Annals of Mathematics, **169** (2009), 903–991

Ricci curvature for metric-measure spaces via optimal transport

By John Lott and Cédric Villani*

Abstract

We define a notion of a measured length space X having nonnegative N-Ricci curvature, for $N \in [1, \infty)$, or having ∞-Ricci curvature bounded below by K, for $K \in \mathbb{R}$. The definitions are in terms of the displacement convexity of certain functions on the associated Wasserstein metric space $P_2(X)$ of probability measures. We show that these properties are preserved under measured Gromov-Hausdorff limits. We give geometric and analytic consequences.

This paper has dual goals. One goal is to extend results about optimal transport from the setting of smooth Riemannian manifolds to the setting of length spaces. A second goal is to use optimal transport to give a notion for a measured length space to have Ricci curvature bounded below. We refer to [11] and [44] for background material on length spaces and optimal transport, respectively. Further bibliographic notes on optimal transport are in Appendix F. In the present introduction we motivate the questions that we address and we state the main results.

To start on the geometric side, there are various reasons to try to extend notions of curvature from smooth Riemannian manifolds to more general spaces. A fairly general setting is that of *length spaces*, meaning metric spaces (X, d) in which the distance between two points equals the infimum of the lengths of curves joining the points. In the rest of this introduction we assume that X is a compact length space. Alexandrov gave a good notion of a length space having "*curvature bounded below by K*", with K a real number, in terms of the geodesic triangles in X. In the case of a Riemannian manifold M with the induced length structure, one recovers the Riemannian notion of having sectional curvature bounded below by K. Length spaces with Alexandrov curvature bounded below by K behave nicely with respect to the Gromov-Hausdorff topology on compact metric spaces (modulo isometries); they form a closed subset.

*The research of the first author was supported by NSF grant DMS-0306242 and the Clay Mathematics Institute.

In view of Alexandrov's work, it is natural to ask whether there are metric space versions of other types of Riemannian curvature, such as Ricci curvature. This question takes substance from *Gromov's precompactness theorem* for Riemannian manifolds with Ricci curvature bounded below by K, dimension bounded above by N and diameter bounded above by D [23, Th. 5.3]. The precompactness indicates that there could be a notion of a length space having "Ricci curvature bounded below by K", special cases of which would be Gromov-Hausdorff limits of manifolds with lower Ricci curvature bounds.

Gromov-Hausdorff limits of manifolds with Ricci curvature bounded below have been studied by various authors, notably Cheeger and Colding [15], [16], [17], [18]. One feature of their work, along with the earlier work of Fukaya [21], is that it turns out to be useful to add an auxiliary Borel probability measure ν and consider *metric-measure spaces* (X, d, ν). (A compact Riemannian manifold M has a canonical measure ν given by the normalized Riemannian density $\frac{\mathrm{dvol}_M}{\mathrm{vol}(M)}$.) There is a measured Gromov-Hausdorff topology on such triples (X, d, ν) (modulo isometries) and one again has precompactness for Riemannian manifolds with Ricci curvature bounded below by K, dimension bounded above by N and diameter bounded above by D. Hence the question is whether there is a good notion of a measured length space (X, d, ν) having "*Ricci curvature bounded below by K*". Whatever definition one takes, one would like the set of such triples to be closed in the measured Gromov-Hausdorff topology. One would also like to derive some nontrivial consequences from the definition, and of course in the case of Riemannian manifolds one would like to recover classical notions. We refer to [16, App. 2] for further discussion of the problem of giving a "synthetic" treatment of Ricci curvature.

Our approach is in terms of a metric space $(P(X), W_2)$ that is canonically associated to the original metric space (X, d). Here $P(X)$ is the space of Borel probability measures on X and W_2 is the so-called *Wasserstein distance* of order 2. The square of the Wasserstein distance $W_2(\mu_0, \mu_1)$ between $\mu_0, \mu_1 \in P(X)$ is defined to be the infimal cost to transport the total mass from the measure μ_0 to the measure μ_1, where the cost to transport a unit of mass between points $x_0, x_1 \in X$ is taken to be $d(x_0, x_1)^2$. A transportation scheme with infimal cost is called an *optimal transport.* The topology on $P(X)$ coming from the metric W_2 turns out to be the weak-$*$ topology. We will write $P_2(X)$ for the metric space $(P(X), W_2)$, which we call the *Wasserstein space.* If (X, d) is a length space then $P_2(X)$ turns out also to be a length space. Its geodesics will be called *Wasserstein geodesics.* If M is a Riemannian manifold then we write $P_2^{\mathrm{ac}}(M)$ for the elements of $P_2(M)$ that are absolutely continuous with respect to the Riemannian density.

In the past fifteen years, optimal transport of measures has been extensively studied in the case $X = \mathbb{R}^n$, with motivation coming from the study of certain partial differential equations. A notion which has proved useful is that of *displacement convexity*, i.e. convexity along Wasserstein geodesics, which was introduced by McCann in order to show the existence and uniqueness of minimizers for certain relevant functions on $P_2^{\mathrm{ac}}(\mathbb{R}^n)$, [31].

In the past few years, some regularity results for optimal transport on $\mathbb{R}^n$ have been extended to Riemannian manifolds [19], [32]. This made it possible to study displacement convexity in a Riemannian setting. Otto and Villani [36] carried out Hessian computations for certain functions on $P_2(M)$ using a formal infinite-dimensional Riemannian structure on $P_2(M)$ defined by Otto [35]. These formal computations indicated a relationship between the Hessian of an "entropy" function on $P_2(M)$ and the Ricci curvature of M. Later, a rigorous displacement convexity result for a class of functions on $P_2^{\mathrm{ac}}(M)$, when M has nonnegative Ricci curvature, was proved by Cordero-Erausquin, McCann and Schmuckenschläger [19]. This work was extended by von Renesse and Sturm [40].

Again in the case of Riemannian manifolds, a further circle of ideas relates displacement convexity to log Sobolev inequalities, Poincaré inequalities, Talagrand inequalities and concentration of measure [8], [9], [27], [36].

In this paper we use optimal transport and displacement convexity in order to *define* a notion of a measured length space (X, d, ν) having Ricci curvature bounded below. If N is a finite parameter (playing the role of a dimension) then we will define a notion of (X, d, ν) having nonnegative N-Ricci curvature. We will also define a notion of (X, d, ν) having ∞-Ricci curvature bounded below by $K \in \mathbb{R}$. (The need to input the possibly-infinite parameter N can be seen from the Bishop-Gromov inequality for complete n-dimensional Riemannian manifolds with nonnegative Ricci curvature. It states that $r^{-n} \operatorname{vol}(B_r(m))$ is nonincreasing in r, where $B_r(m)$ is the r-ball centered at m [23, Lemma 5.3.bis]. When we go from manifolds to length spaces there is no *a priori* value for the parameter n. This indicates the need to specify a dimension parameter in the definition of Ricci curvature bounds.)

We now give the main results of the paper, sometimes in a simplified form. For consistency, we assume in the body of the paper that the relevant length space X is compact. The necessary modifications to deal with complete pointed locally compact length spaces are given in Appendix E.

Let $U : [0, \infty) \to \mathbb{R}$ be a continuous convex function with $U(0) = 0$. Given a reference probability measure $\nu \in P(X)$, define the function $U_\nu : P_2(X) \to \mathbb{R} \cup \{\infty\}$ by

$$U_\nu(\mu) = \int_X U(\rho(x))\, d\nu(x) + U'(\infty)\, \mu_s(X), \tag{0.1}$$

where

$$\mu = \rho\nu + \mu_s \tag{0.2}$$

is the Lebesgue decomposition of μ with respect to ν into an absolutely continuous part $\rho\nu$ and a singular part μ_s, and

$$U'(\infty) = \lim_{r\to\infty} \frac{U(r)}{r}. \tag{0.3}$$

If $N \in [1,\infty)$ then we define $\mathcal{DC}_N$ to be the set of such functions U so that the function

$$\psi(\lambda) = \lambda^N U(\lambda^{-N}) \tag{0.4}$$

is convex on $(0,\infty)$. We further define $\mathcal{DC}_\infty$ to be the set of such functions U so that the function

$$\psi(\lambda) = e^{\lambda} U(e^{-\lambda}) \tag{0.5}$$

is convex on $(-\infty,\infty)$. A relevant example of an element of $\mathcal{DC}_N$ is given by

$$U_N(r) = \begin{cases} Nr(1-r^{-1/N}) & \text{if } 1 < N < \infty, \\ r\log r & \text{if } N = \infty. \end{cases} \tag{0.6}$$

Definition 0.7. Given $N \in [1,\infty]$, we say that a compact measured length space (X,d,ν) has nonnegative N-Ricci curvature if for all $\mu_0,\mu_1 \in P_2(X)$ with $\mathrm{supp}(\mu_0) \subset \mathrm{supp}(\nu)$ and $\mathrm{supp}(\mu_1) \subset \mathrm{supp}(\nu)$, there is *some* Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 so that for all $U \in \mathcal{DC}_N$ and all $t \in [0,1]$,

$$U_\nu(\mu_t) \le tU_\nu(\mu_1) + (1-t)U_\nu(\mu_0). \tag{0.8}$$

Given $K \in \mathbb{R}$, we say that (X,d,ν) has ∞-Ricci curvature bounded below by K if for all $\mu_0,\mu_1 \in P_2(X)$ with $\mathrm{supp}(\mu_0) \subset \mathrm{supp}(\nu)$ and $\mathrm{supp}(\mu_1) \subset \mathrm{supp}(\nu)$, there is *some* Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 so that for all $U \in \mathcal{DC}_\infty$ and all $t \in [0,1]$,

$$U_\nu(\mu_t) \le tU_\nu(\mu_1) + (1-t)U_\nu(\mu_0) - \frac{1}{2}\lambda(U)t(1-t)W_2(\mu_0,\mu_1)^2, \tag{0.9}$$

where $\lambda : \mathcal{DC}_\infty \to \mathbb{R} \cup \{-\infty\}$ is as defined in (5.14) below.

Note that the inequalities (0.8) and (0.9) are only assumed to hold along *some* Wasserstein geodesic from μ_0 to μ_1, and not necessarily along all such geodesics. This is what we call *weak displacement convexity.*

Naturally, one wants to know that in the case of a Riemannian manifold, our definitions are equivalent to classical ones. Let M be a smooth compact connected n-dimensional manifold with Riemannian metric g. We let (M,g) denote the corresponding metric space. Given $\Psi \in C^\infty(M)$ with $\int_M e^{-\Psi}\,\mathrm{dvol}_M = 1$, put $d\nu = e^{-\Psi}\,\mathrm{dvol}_M$.

Definition 0.10. For $N \in [1, \infty]$, let the N-Ricci tensor Ric_N of (M, g, ν) be defined by

$$\mathrm{Ric}_N = \begin{cases} \mathrm{Ric} + \mathrm{Hess}\,(\Psi) & \text{if } N = \infty, \\ \mathrm{Ric} + \mathrm{Hess}\,(\Psi) - \frac{1}{N-n} d\Psi \otimes d\Psi & \text{if } n < N < \infty, \\ \mathrm{Ric} + \mathrm{Hess}\,(\Psi) - \infty(d\Psi \otimes d\Psi) & \text{if } N = n, \\ -\infty & \text{if } N < n, \end{cases} \tag{0.11}$$

where by convention $\infty \cdot 0 = 0$.

THEOREM 0.12. (a) *For $N \in [1, \infty)$, the measured length space (M, g, ν) has nonnegative N-Ricci curvature if and only if* $\mathrm{Ric}_N \geq 0$.

(b) (M, g, ν) *has ∞-Ricci curvature bounded below by K if and only if* $\mathrm{Ric}_\infty \geq Kg$.

In the special case when Ψ is constant, and so $\nu = \frac{\mathrm{dvol}_M}{\mathrm{vol}(M)}$, Theorem 0.12 shows that we recover the usual notion of a Ricci curvature bound from our length space definition as soon as $N \geq n$.

The next theorem, which is the main result of the paper, says that our notion of N-Ricci curvature has good behavior under measured Gromov-Hausdorff limits.

THEOREM 0.13. *Let $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ be a sequence of compact measured length spaces with $\lim_{i\to\infty}(X_i, d_i, \nu_i) = (X, d, \nu)$ in the measured Gromov-Hausdorff topology.*

(a) *For any $N \in [1, \infty)$, if each (X_i, d_i, ν_i) has nonnegative N-Ricci curvature then (X, d, ν) has nonnegative N-Ricci curvature.*

(b) *If each (X_i, d_i, ν_i) has ∞-Ricci curvature bounded below by K then (X, d, ν) has ∞-Ricci curvature bounded below by K.*

Theorems 0.12 and 0.13 imply that measured Gromov-Hausdorff limits (X, d, ν) of smooth manifolds $\left(M, g, \frac{\mathrm{dvol}_M}{\mathrm{vol}(M)}\right)$ with lower Ricci curvature bounds fall under our considerations. Additionally, we obtain the following new characterization of such limits (X, d, ν) which happen to be *smooth*, meaning that (X, d) is a smooth n-dimensional Riemannian manifold (B, g_B) and $d\nu = e^{-\Psi}\,\mathrm{dvol}_B$ for some $\Psi \in C^\infty(B)$:

COROLLARY 0.14. (a) *If (B, g_B, ν) is a measured Gromov-Hausdorff limit of Riemannian manifolds with nonnegative Ricci curvature and dimension at most N then* $\mathrm{Ric}_N(B) \geq 0$.

(b) *If (B, g_B, ν) is a measured Gromov-Hausdorff limit of Riemannian manifolds with Ricci curvature bounded below by $K \in \mathbb{R}$ then* $\mathrm{Ric}_\infty(B) \geq K g_B$.

There is a partial converse to Corollary 0.14 (see Corollary 7.45(ii, ii′)).

Finally, if a measured length space has lower Ricci curvature bounds then there are analytic consequences, such as a *log Sobolev inequality.* To state it, we define the *gradient norm* of a Lipschitz function f on X by the formula

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{d(x,y)}. \tag{0.15}$$

THEOREM 0.16. *Suppose that a compact measured length space (X, d, ν) has ∞-Ricci curvature bounded below by $K \in \mathbb{R}$. Suppose that $f \in \mathrm{Lip}(X)$ satisfies $\int_X f^2 \, d\nu = 1$.*

(a) *If $K > 0$ then*

$$\int_X f^2 \log(f^2) \, d\nu \le \frac{2}{K} \int_X |\nabla f|^2 \, d\nu. \tag{0.17}$$

(b) *If $K \le 0$ then*

$$\int_X f^2 \log(f^2) \, d\nu \le 2 \, \mathrm{diam}(X) \sqrt{\int_X |\nabla f|^2 \, d\nu} - \frac{1}{2} K \, \mathrm{diam}(X)^2. \tag{0.18}$$

In the case of Riemannian manifolds, one recovers from (0.17) the log Sobolev inequality of Bakry and Émery [6].

A consequence of (0.17) is a Poincaré inequality.

COROLLARY 0.19. *Suppose that a compact measured length space (X, d, ν) has ∞-Ricci curvature bounded below by $K > 0$. Then for all $h \in \mathrm{Lip}(X)$ with $\int_X h \, d\nu = 0$, we have*

$$\int_X h^2 \, d\nu \le \frac{1}{K} \int_X |\nabla h|^2 \, d\nu. \tag{0.20}$$

In the case of Riemannian manifolds, Corollary 0.19 follows from the Lichnerowicz inequality for the smallest positive eigenvalue of the Laplacian [28].

We now give the structure of the paper. More detailed descriptions appear at the beginnings of the sections.

Section 1 gives basic definitions about length spaces and optimal transport. Section 2 shows that the Wasserstein space of a length space is also a length space, and that Wasserstein geodesics arise from displacement interpolations. Section 3 defines weak displacement convexity and its variations. This is used to prove functional inequalities called the HWI inequalities.

Section 4 proves, modulo the technical results of Appendices B and C, that weak displacement convexity is preserved by measured Gromov-Hausdorff limits. The notion of N-Ricci curvature is defined in Section 5, which contains the proof of Theorem 0.13, along with a Bishop-Gromov-type inequality. Section 6 proves log Sobolev, Talagrand and Poincaré inequalities for measured length spaces, such as Theorem 0.16 and Corollary 0.19, along with a weak Bonnet-

Myers theorem. Section 7 looks at the case of smooth Riemannian manifolds and proves, in particular, Theorem 0.12 and Corollary 0.14.

There are six appendices that contain either technical results or auxiliary results. Appendix A, which is a sequel to Section 2, discusses the geometry of the Wasserstein space of a Riemannian manifold M. It shows that if M has nonnegative sectional curvature then $P_2(M)$ has nonnegative Alexandrov curvature. The tangent cones at absolutely continuous measures are computed, thereby making rigorous the formal Riemannian metric on $P_2(M)$ introduced by Otto.

Appendices B and C are the technical core of Theorem 0.13. Appendix B shows that $U_\nu(\mu)$ is lower semicontinuous in both μ and ν, and is nonincreasing under pushforward of μ and ν. Appendix C shows that a measure $\mu \in P_2(X)$ with $\mathrm{supp}(\mu) \subset \mathrm{supp}(\nu)$ can be weak-$*$ approximated by measures $\{\mu_k\}_{k=1}^{\infty}$ with continuous densities (with respect to ν) so that $U_\nu(\mu) = \lim_{k\to\infty} U_\nu(\mu_k)$.

Appendix D contains formal computations of the Hessian of U_ν. Appendix E explains how to extend the results of the paper from the setting of compact measured length spaces to the setting of complete pointed locally compact measured length spaces. Appendix F has some bibliographic notes on optimal transport and displacement convexity.

The results of this paper were presented at the workshop "Collapsing and metric geometry" in Münster, August 1–7, 2004. After the writing of the paper was essentially completed we learned of related work by Karl-Theodor Sturm [41], [42]. Also, Ludger Rüschendorf kindly pointed out to us that Theorem B.33 was already proven in [29, Ch. 1] by different means. We decided to retain our proof of Theorem B.33 rather than just quoting [29], partly because the method of proof may be of independent interest, partly for completeness and convenience to the reader, and partly because our method of proof is used in the extension of the theorem considered in Appendix E.

We thank MSRI and the UC-Berkeley mathematics department for their hospitality while part of this research was performed. We also thank the anonymous referees for their suggestions.

Contents

1. Notation and basic definitions
   1.1. Convex analysis
   1.2. Geometry of metric spaces
      1.2.1. Gradient norms
      1.2.2. Length spaces
      1.2.3. (Measured) Gromov-Hausdorff convergence
   1.3. Optimal transport: basic definitions
2. Geometry of the Wasserstein space
   2.1. Displacement interpolations
   2.2. The Wasserstein space as a length space

2.3. Wasserstein geodesics as displacement interpolations
2.4. Optimal transport on Riemannian manifolds
3. Functionals on the Wasserstein space
3.1. Weak displacement convexity
3.2. Important examples
3.3. HWI inequalities
4. Weak displacement convexity and measured Gromov-Hausdorff limits
4.1. Gromov-Hausdorff convergence of the Wasserstein space
4.2. Stability of weak displacement convexity
5. N-Ricci curvature for measured length spaces
5.1. Displacement convex classes
5.2. Ricci curvature via weak displacement convexity
5.3. Preservation of N-Ricci curvature bounds
5.4. Bishop-Gromov inequality
5.5. Compact group actions
5.6. Uniform integrability and absolute continuity
6. Log Sobolev, Talagrand and Poincaré inequalities
6.1. The general inequalities
6.2. The case $N = \infty$
6.3. The case $N < \infty$
6.4. Weak Bonnet-Myers theorem
7. The case of Riemannian manifolds
7.1. Formulation of N-Ricci curvature in classical terms
7.2. Geometric corollaries
Appendix A. The Wasserstein space as an Alexandrov space
A.1. Lipschitz functions and optimal transport
A.2. The case of nonnegatively curved manifolds
A.3. Application to the geometric description of $P_2(M)$
Appendix B. Some properties of the functionals U_ν
B.1. The functional U_ν via Legendre transform
B.2. Lower semicontinuity and contraction
Appendix C. Approximation in $P_2(X)$
C.1. Mollifiers
C.2. Approximation by continuous densities
Appendix D. Hessian calculations
Appendix E. The noncompact case
E.1. Pointed spaces
E.2. Wasserstein space
E.3. Functionals
E.4. Approximation arguments
E.5. Stability of N-Ricci curvature bounds
E.6. Tangent Cones
Appendix F. Bibliographic notes on optimal transport
References

1. Notation and basic definitions

In this section we first recall some facts about convex functions. We then define gradient norms, length spaces and measured Gromov-Hausdorff convergence. Finally, we define the 2-Wasserstein metric W_2 on $P(X)$.

1.1. *Convex analysis.* Let us recall a few results from convex analysis. See [44, Ch. 2.1] and references therein for further information.

Given a convex lower semicontinuous function $U : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ (which we assume is not identically ∞), its Legendre transform $U^* : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is defined by

$$U^*(p) = \sup_{r \in \mathbb{R}} \big[pr - U(r)\big]. \tag{1.1}$$

Then U^* is also convex and lower semicontinuous. We will sometimes identify a convex lower semicontinuous function U defined on a closed interval $I \subset \mathbb{R}$ with the convex function defined on the whole of $\mathbb{R}$ by extending U by ∞ outside of I.

Let $U : [0, \infty) \to \mathbb{R}$ be a convex lower semicontinuous function. Then U admits a left derivative $U'_- : (0, \infty) \to \mathbb{R}$ and a right derivative $U'_+ : [0, \infty) \to \{-\infty\} \cup \mathbb{R}$, with $U'_+(0, \infty) \subset \mathbb{R}$. Furthermore, $U'_- \le U'_+$. They agree almost everywhere and are both nondecreasing. We will write

$$U'(\infty) = \lim_{r \to \infty} U'_+(r) = \lim_{r \to \infty} \frac{U(r)}{r} \in \mathbb{R} \cup \{\infty\}. \tag{1.2}$$

If we extend U by ∞ on $(-\infty, 0)$ then its Legendre transform $U^* : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ becomes $U^*(p) = \sup_{r \ge 0}\big[pr - U(r)\big]$. It is nondecreasing in p, infinite on $(U'(\infty), \infty)$ and equals $-U(0)$ on $(-\infty, U'_+(0)]$. Furthermore, it is continuous on $(-\infty, U'(\infty))$. For all $r \in [0, \infty)$, we have $U^*(U'_+(r)) = rU'_+(r) - U(r)$.

1.2. *Geometry of metric spaces.*

1.2.1. *Gradient norms.* Let (X, d) be a compact metric space (with d valued in $[0, \infty)$). The open ball of radius r around $x \in X$ will be denoted by $B_r(x)$ and the sphere of radius r around x will be denoted by $S_r(x)$.

Let $L^\infty(X)$ denote the set of bounded measurable functions on X. (We will consider such a function to be defined everywhere.) Let $\mathrm{Lip}(X)$ denote the set of Lipschitz functions on X. Given $f \in \mathrm{Lip}(X)$, we define the *gradient norm* of f by

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{d(x, y)} \tag{1.3}$$

if x is not an isolated point, and $|\nabla f|(x) = 0$ if x is isolated. Then $|\nabla f| \in L^\infty(X)$.

On some occasions we will use a finer notion of gradient norm:

$$|\nabla^- f|(x) = \limsup_{y\to x} \frac{[f(y)-f(x)]_-}{d(x,y)} = \limsup_{y\to x} \frac{[f(x)-f(y)]_+}{d(x,y)} \tag{1.4}$$

if x is not isolated, and $|\nabla^- f|(x) = 0$ if x is isolated. Here $a_+ = \max(a,0)$ and $a_- = \max(-a,0)$. Clearly $|\nabla^- f|(x) \leq |\nabla f|(x)$. Note that $|\nabla^- f|(x)$ is automatically zero if f has a local minimum at x. In a sense, $|\nabla^- f|(x)$ measures the downward pointing component of f near x.

1.2.2. *Length spaces.* If γ is a curve in X, i.e. a continuous map $\gamma : [0,1] \to X$, then its length is

$$L(\gamma) = \sup_{J\in\mathbb{N}} \sup_{0=t_0\leq t_1\leq\ldots\leq t_J=1} \sum_{j=1}^{J} d\big(\gamma(t_{j-1}),\gamma(t_j)\big). \tag{1.5}$$

Clearly $L(\gamma) \geq d(\gamma(0),\gamma(1))$.

We will assume that X is a *length space*, meaning that the distance between two points $x_0, x_1 \in X$ is the infimum of the lengths of curves from x_0 to x_1. Such a space is path connected.

As X is compact, it is a strictly intrinsic length space, meaning that we can replace infimum by minimum [11, Th. 2.5.23]. That is, for any $x_0, x_1 \in X$, there is a minimal geodesic (possibly nonunique) from x_0 to x_1. We may sometimes write “geodesic” instead of “minimal geodesic”.

By [11, Prop. 2.5.9], any minimal geodesic γ joining x_0 to x_1 can be parametrized uniquely by $t \in [0,1]$ so that

$$d(\gamma(t),\gamma(t')) = |t-t'|d(x_0,x_1). \tag{1.6}$$

We will often assume that the geodesic has been so parametrized.

By definition, a subset $A \subset X$ is *convex* if for any $x_0, x_1 \in A$ there is a minimizing geodesic from x_0 to x_1 that lies entirely in A. It is *totally convex* if for any $x_0, x_1 \in A$, any minimizing geodesic in X from x_0 to x_1 lies in A.

Given $\lambda \in \mathbb{R}$, a function $F : X \to \mathbb{R}\cup\{\infty\}$ is said to be λ-*convex* if for any geodesic $\gamma : [0,1] \to X$ and any $t \in [0,1]$,

$$F(\gamma(t)) \leq tF(\gamma(1)) + (1-t)F(\gamma(0)) - \frac{1}{2}\lambda t(1-t)\, L(\gamma)^2. \tag{1.7}$$

In the case when X is a smooth Riemannian manifold with Riemannian metric g, and $F \in C^2(X)$, this is the same as saying that Hess $F \geq \lambda g$.

1.2.3. (*Measured*) *Gromov-Hausdorff convergence.*

Definition 1.8. Given two compact metric spaces (X_1, d_1) and (X_2, d_2), an ε-Gromov-Hausdorff approximation from X_1 to X_2 is a (not necessarily continuous) map $f : X_1 \to X_2$ so that

(i) For all $x_1, x_1' \in X_1$, $\left| d_2(f(x_1), f(x_1')) - d_1(x_1, x_1') \right| \leq \varepsilon$.

(ii) For all $x_2 \in X_2$, there is an $x_1 \in X_1$ so that $d_2(f(x_1), x_2) \leq \varepsilon$.

An ε-Gromov-Hausdorff approximation $f : X_1 \to X_2$ has an approximate inverse $f' : X_2 \to X_1$, which can be constructed as follows: Given $x_2 \in X_2$, choose $x_1 \in X_1$ so that $d_2(f(x_1), x_2) \leq \varepsilon$ and put $f'(x_2) = x_1$. Then f' is a 3ε-Gromov-Hausdorff approximation from X_2 to X_1. Moreover, for all $x_1 \in X_1$, $d_1(x_1, (f' \circ f)(x_1)) \leq 2\varepsilon$, and for all $x_2 \in X_2$, $d_2(x_2, (f \circ f')(x_2)) \leq \varepsilon$.

Definition 1.9. A sequence of compact metric spaces $\{X_i\}_{i=1}^\infty$ converges to X in the *Gromov-Hausdorff topology* if there is a sequence of ε_i-approximations $f_i : X_i \to X$ with $\lim_{i\to\infty} \varepsilon_i = 0$.

This notion of convergence comes from a metrizable topology on the space of all compact metric spaces modulo isometries. If $\{X_i\}_{i=1}^\infty$ are length spaces that converge to X in the Gromov-Hausdorff topology then X is also a length space [11, Th. 7.5.1].

For the purposes of this paper, we can and will assume that the maps f and f' in Gromov-Hausdorff approximations are Borel. Let $P(X)$ denote the space of Borel probability measures on X. We give $P(X)$ the weak-$*$ topology, i.e. $\lim_{i\to\infty} \mu_i = \mu$ if and only if for all $F \in C(X)$, $\lim_{i\to\infty} \int_X F d\mu_i = \int_X F d\mu$.

Definition 1.10. Given $\nu \in P(X)$, consider the metric-measure space (X, d, ν). A sequence $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ converges to (X, d, ν) in the *measured Gromov-Hausdorff topology* if there are ε_i-approximations $f_i : X_i \to X$, with $\lim_{i\to\infty} \varepsilon_i = 0$, so that $\lim_{i\to\infty} (f_i)_* \nu_i = \nu$ in $P(X)$.

Other topologies on the class of metric-measure spaces are discussed in [23, Ch. $3\frac{1}{2}$].

For later use we note the following generalization of the Arzelà-Ascoli theorem.

LEMMA 1.11 (cf. [22, p. 66], [24, App. A]). *Let $\{X_i\}_{i=1}^\infty$ be a sequence of compact metric spaces converging to X in the Gromov-Hausdorff topology, with ε_i-approximations $f_i : X_i \to X$. Let $\{Y_i\}_{i=1}^\infty$ be a sequence of compact metric spaces converging to Y in the Gromov-Hausdorff topology, with ε_i-approximations $g_i : Y_i \to Y$. For each i, let $f_i' : X \to X_i$ be an approximate inverse to f_i, as in the paragraph following Definition* 1.8. *Let $\{\alpha_i\}_{i=1}^\infty$ be a sequence of maps $\alpha_i : X_i \to Y_i$ that are asymptotically equicontinuous in the sense that for every $\varepsilon > 0$, there are $\delta = \delta(\varepsilon) > 0$ and $N = N(\varepsilon) \in \mathbb{Z}^+$ so that for all $i \geq N$,*

$$d_{X_i}(x_i, x_i') < \delta \quad \Longrightarrow \quad d_{Y_i}(\alpha_i(x_i), \alpha_i(x_i')) < \varepsilon. \tag{1.12}$$

Then after passing to a subsequence, the maps $g_i \circ \alpha_i \circ f'_i : X \to Y$ *converge uniformly to a continuous map* $\alpha : X \to Y$.

In the conclusion of Lemma 1.11 the maps $g_i \circ \alpha_i \circ f'_i$ may not be continuous, but the notion of uniform convergence makes sense nevertheless.

1.3. *Optimal transport*: *basic definitions.* Given $\mu_0, \mu_1 \in P(X)$, we say that a probability measure $\pi \in P(X \times X)$ is a *transference plan* between μ_0 and μ_1 if

$$(p_0)_* \pi = \mu_0, \qquad (p_1)_* \pi = \mu_1, \tag{1.13}$$

where $p_0, p_1 : X \times X \to X$ are projections onto the first and second factors, respectively. In words, π represents a way to transport the mass from μ_0 to μ_1, and $\pi(x_0, x_1)$ is the amount of mass which is taken from a point x_0 and brought to a point x_1.

We will use optimal transport with quadratic cost function (square of the distance). Namely, given $\mu_0, \mu_1 \in P(X)$, we consider the variational problem

$$W_2(\mu_0, \mu_1)^2 = \inf_\pi \int_{X \times X} d(x_0, x_1)^2 d\pi(x_0, x_1), \tag{1.14}$$

where π ranges over the set of all transference plans between μ_0 and μ_1. Any minimizer π for this variational problem is called an *optimal transference plan.*

In (1.14), one can replace the infimum by the minimum [44, Prop. 2.1], i.e. there always exists (at least) one optimal transference plan. Since X has finite diameter, the infimum is obviously finite. The quantity W_2 will be called the *Wasserstein distance of order 2* between μ_0 and μ_1; it defines a metric on $P(X)$. The topology that it induces on $P(X)$ is the weak-$*$ topology [44, Ths. 7.3 and 7.12]. When equipped with the metric W_2, $P(X)$ is a compact metric space, often denoted by $P_2(X)$.

We remark that there is an isometric embedding $X \to P_2(X)$ given by $x \to \delta_x$. This shows that $\mathrm{diam}(P_2(X)) \geq \mathrm{diam}(X)$. Since the reverse inequality follows from the definition of W_2, actually $\mathrm{diam}(P_2(X)) = \mathrm{diam}(X)$.

A *Monge transport* is a transference plan coming from a map $F : X \to X$ with $F_* \mu_0 = \mu_1$, given by $\pi = (\mathrm{Id}\,, F)_* \mu_0$. In general an optimal transference plan does not have to be a Monge transport, although this may be true under some assumptions (as we will recall below).

A function $\phi : X \to [-\infty, \infty)$ is $\frac{d^2}{2}$-*concave* if it is not identically $-\infty$ and it can be written in the form

$$\phi(x) = \inf_{x' \in X} \left(\frac{d(x, x')^2}{2} - \widetilde{\phi}(x') \right) \tag{1.15}$$

for some function $\widetilde{\phi} : X \to [-\infty, \infty)$. Such functions play an important role in the description of optimal transport on Riemannian manifolds.

2. Geometry of the Wasserstein space

In this section, we investigate some features of the Wasserstein space $P_2(X)$ associated to a compact length space (X, d). (Recall that the subscript 2 in $P_2(X)$ means that $P(X)$ is equipped with the 2-Wasserstein metric.) We show that $P_2(X)$ is a length space. We define displacement interpolations and show that every Wasserstein geodesic comes from a displacement interpolation. We then recall some facts about optimal transport on Riemannian manifolds.

2.1. *Displacement interpolations.* We denote by $\mathrm{Lip}([0,1], X)$ the space of Lipschitz continuous maps $c : [0,1] \to X$ with the uniform topology. For any $k > 0$,

$$\begin{aligned} \text{(2.1)} \quad & \mathrm{Lip}_k([0,1], X) \\ & = \Big\{ c \in \mathrm{Lip}([0,1], X) : d(c(t), c(t')) \leq k|t - t'| \text{ for all } t, t' \in [0,1] \Big\} \end{aligned}$$

is a compact subset of $\mathrm{Lip}([0,1], X)$.

Let Γ denote the set of minimizing geodesics on X. It is a closed subspace of $\mathrm{Lip}_{\mathrm{diam}(X)}([0,1], X)$, defined by the equation $L(c) = d(c(0), c(1))$.

For any $t \in [0,1]$, the *evaluation map* $e_t : \Gamma \to X$ defined by

$$e_t(\gamma) = \gamma(t) \tag{2.2}$$

is continuous. Let $E : \Gamma \to X \times X$ be the "endpoints" map given by $E(\gamma) = (e_0(\gamma), e_1(\gamma))$. A *dynamical transference plan* consists of a transference plan π and a Borel measure Π on Γ such that $E_* \Pi = \pi$; it is said to be optimal if π itself is. In other words, the transference plan π tells us how much mass goes from a point x_0 to another point x_1, but does not tell us about the actual path that the mass has to follow. Intuitively, mass should flow along geodesics, but there may be several possible choices of geodesics between two given points and the transport may be divided among these geodesics; this is the information provided by Π.

If Π is an optimal dynamical transference plan then for $t \in [0,1]$, we put

$$\mu_t = (e_t)_* \Pi. \tag{2.3}$$

The one-parameter family of measures $\{\mu_t\}_{t \in [0,1]}$ is called a *displacement interpolation.* In other words, μ_t is what has become of the mass of μ_0 after it has travelled from time 0 to time t according to the dynamical transference plan Π.

LEMMA 2.4. *The map $c : [0,1] \to P_2(X)$ given by $c(t) = \mu_t$ has length $L(c) = W_2(\mu_0, \mu_1)$.*

Proof. Given $0 \le t \le t' \le 1$, $(e_t, e_{t'})_* \Pi$ is a particular transference plan from μ_t to $\mu_{t'}$, and so

(2.5)
$$\begin{aligned}
W_2(\mu_t, \mu_{t'})^2 &\le \int_{X \times X} d(x_0, x_1)^2 d\left((e_t, e_{t'})_* \Pi\right)(x_0, x_1) = \int_\Gamma d(\gamma(t), \gamma(t'))^2 d\Pi(\gamma) \\
&= \int_\Gamma (t'-t)^2 L\left(\gamma\right)^2 d\Pi(\gamma) = (t'-t)^2 \int_\Gamma d(\gamma(0), \gamma(1))^2 d\Pi(\gamma) \\
&= (t'-t)^2 \int_{X \times X} d(x_0, x_1)^2 (dE_* \Pi)(x_0, x_1) = (t'-t)^2 W_2(\mu_0, \mu_1)^2.
\end{aligned}$$

Equation (2.5) implies that $L(c) \le W_2(\mu_0, \mu_1)$, and so $L(c) = W_2(\mu_0, \mu_1)$. □

2.2. *The Wasserstein space as a length space.*

PROPOSITION 2.6. *Let (X, d) be a compact length space. Then any two points $\mu_0, \mu_1 \in P_2(X)$ can be joined by a displacement interpolation.*

Proof. The endpoints map E is Borel and surjective. Given $(x_0, x_1) \in X \times X$, $E^{-1}(x_0, x_1)$ is compact. It follows that there is a Borel map $S : X \times X \to \Gamma$ so that $E \circ S = \mathrm{Id}_{X \times X}$ [46, Cor. A.6]. In words, S is a measurable way to join points by minimizing geodesics. Given $\mu_0, \mu_1 \in P_2(X)$, let π be an optimal transference plan between μ_0 and μ_1, and put $\Pi = S_*(\pi)$. The corresponding displacement interpolation joins μ_0 and μ_1. □

COROLLARY 2.7. *If X is a compact length space then $P_2(X)$ is a compact length space.*

Proof. We already know that $P_2(X)$ is compact. Given $\mu_0, \mu_1 \in P_2(X)$, Proposition 2.6 gives a displacement interpolation c from μ_0 to μ_1. By Lemma 2.4, $L(c) = W_2(\mu_0, \mu_1)$, so that $P_2(X)$ is also a length space. □

Remark 2.8. The same argument shows that $(P(X), W_p)$ is a compact length space for all $p \in [1, \infty)$, where W_p is the Wasserstein distance of order p [44, §7.1.1].

Example 2.9. Suppose that $X = A \cup B \cup C$, where A, B and C are subsets of the plane given by $A = \{(x_1, 0) : -2 \le x_1 \le -1\}$, $B = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$ and $C = \{(x_1, 0) : 1 \le x_1 \le 2\}$. Let μ_0 be the one-dimensional Hausdorff measure of A and let μ_1 be the one-dimensional Hausdorff measure of C. Then there is an uncountable number of Wasserstein geodesics from μ_0 to μ_1, given by the whims of a switchman at the point $(-1, 0)$.

2.3. *Wasserstein geodesics as displacement interpolations.* The next result states that every Wasserstein geodesic arises from a displacement interpolation.

Proposition 2.10. *Let (X,d) be a compact length space and let $\{\mu_t\}_{t\in[0,1]}$ be a geodesic path in $P_2(X)$. Then there exists some optimal dynamical transference plan Π such that $\{\mu_t\}_{t\in[0,1]}$ is the displacement interpolation associated to Π.*

Proof. Let $\{\mu_t\}_{t\in[0,1]}$ be a Wasserstein geodesic. Up to reparametrization, we can assume that for all $t,t'\in[0,1]$,

$$W_2(\mu_t,\mu_{t'}) = |t-t'|\, W_2(\mu_0,\mu_1). \tag{2.11}$$

Let $\pi^{(0)}_{x_0,x_{1/2}}$ be an optimal transference plan from μ_0 to $\mu_{1/2}$, and let $\pi^{(1/2)}_{x_{1/2},x_1}$ be an optimal transference plan from $\mu_{1/2}$ to μ_1. Consider the measure obtained by "gluing together" $\pi^{(0)}_{x_0,x_{1/2}}$ and $\pi^{(1/2)}_{x_{1/2},x_1}$:

$$M^{(1)} = \frac{d\pi^{(0)}_{x_0,x_{1/2}}\, d\pi^{(1/2)}_{x_{1/2},x_1}}{d\mu_{1/2}(x_{1/2})} \tag{2.12}$$

on $X\times X\times X$.

The precise meaning of this expression is just as in the "gluing lemma" stated in [44, Lemma 7.6]: Decompose $\pi^{(0)}$ with respect to the projection $p_1 : X\times X\to X$ on the second factor as $\pi^{(0)} = \sigma^{(0)}_{x_{1/2}}\mu_{1/2}(x_{1/2})$, where for $\mu_{1/2}$-almost all $x_{1/2}$, $\sigma^{(0)}_{x_{1/2}}\in P(p_1^{-1}(x_{1/2}))$ is a probability measure on $p_1^{-1}(x_{1/2})$. Decompose $\pi^{(1/2)}$ with respect to the projection $p_0 : X\times X\to X$ on the first factor as $\pi^{(1/2)} = \sigma^{(1/2)}_{x_{1/2}}\mu_{1/2}(x_{1/2})$, where for $\mu_{1/2}$-almost all $x_{1/2}$, $\sigma^{(1/2)}_{x_{1/2}}\in P(p_0^{-1}(x_{1/2}))$. Then, for $F\in C(X\times X\times X)$,

$$\int_{X\times X\times X} F\,dM^{(1)} \equiv \int_X \int_{p_1^{-1}(x_{1/2})\times p_0^{-1}(x_{1/2})} \cdot F(x_0,x_{1/2},x_1)\,d\sigma^{(0)}_{x_{1/2}}(x_0)\,d\sigma^{(1/2)}_{x_{1/2}}(x_1)\,d\mu_{1/2}(x_{1/2}). \tag{2.13}$$

The formula

$$d\pi_{x_0,x_1} = \int_X M^{(1)}_{x_0,x_{1/2},x_1} \tag{2.14}$$

defines a transference plan from μ_0 to μ_1 with cost

$$\begin{aligned} \int_{X\times X} d(x_0,x_1)^2 d\pi_{x_0,x_1} &\le \int_{X\times X\times X} (d(x_0,x_{1/2})+d(x_{1/2},x_1))^2 \frac{d\pi^{(0)}_{x_0,x_{1/2}}\, d\pi^{(1/2)}_{x_{1/2},x_1}}{d\mu_{1/2}(x_{1/2})} \\ &\le \int_{X\times X\times X} 2(d(x_0,x_{1/2})^2+d(x_{1/2},x_1)^2) \frac{d\pi^{(0)}_{x_0,x_{1/2}}\, d\pi^{(1/2)}_{x_{1/2},x_1}}{d\mu_{1/2}(x_{1/2})} \\ &= 2\left(\int_{X\times X} d(x_0,x_{1/2})^2 d\pi^{(0)}_{x_0,x_{1/2}} + \int_{X\times X} d(x_{1/2},x_1)^2 d\pi^{(1/2)}_{x_{1/2},x_1}\right) \\ &= 2\left(W_2(\mu_0,\mu_{\frac12})^2 + W_2(\mu_{\frac12},\mu_1)^2\right) = W_2(\mu_0,\mu_1)^2. \end{aligned} \tag{2.15}$$

Thus π is an optimal transference plan and we must have equality everywhere in (2.15). Let

$$B^{(1)} = \Big\{(x_0, x_{1/2}, x_1) \in X \times X \times X : d(x_0, x_{1/2}) = d(x_{1/2}, x_1) = \frac{1}{2} d(x_0, x_1)\Big\}; \tag{2.16}$$

then $M^{(1)}$ is supported on $B^{(1)}$. For $t \in \{0, \frac{1}{2}, 1\}$, define $e_t : B^{(1)} \to X$ by $e_t(x_0, x_{1/2}, x_1) = x_t$. Then $(e_t)_* M^{(1)} = \mu_t$.

We can repeat the same procedure using a decomposition of the interval $[0,1]$ into 2^i subintervals. For any $i \geq 1$, define

$$B^{(i)} = \Big\{(x_0, x_{2^{-i}}, x_{2\cdot 2^{-i}}, \ldots, x_{1-2^{-i}}, x_1) \in X^{2^i+1} : d(x_0, x_{2^{-i}}) = d(x_{2^{-i}}, x_{2\cdot 2^{-i}}) = \ldots = d(x_{1-2^{-i}}, x_1) = 2^{-i} d(x_0, x_1)\Big\}. \tag{2.17}$$

For $0 \leq j \leq 2^i - 1$, choose an optimal transference plan $\pi^{(j\cdot 2^{-i})}_{x_{j\cdot 2^{-i}}, x_{(j+1)\cdot 2^{-i}}}$ from $\mu_{j\cdot 2^{-i}}$ to $\mu_{(j+1)\cdot 2^{-i}}$. Then as before, we obtain a probability measure $M^{(i)}$ on $B^{(i)}$ by

$$M^{(i)}_{x_0, x_{2^{-i}}, \ldots, x_1} = \frac{d\pi^{(0)}(x_0, x_{2^{-i}}) d\pi^{(2^{-i})}(x_{2^{-i}}, x_{2\cdot 2^{-i}}) \ldots d\pi^{(1-2^{-i})}(x_{1-2^{-i}}, x_1)}{d\mu_{2^{-i}}(x_{2^{-i}}) \ldots d\mu_{1-2^{-i}}(x_{1-2^{-i}})}. \tag{2.18}$$

The formula

$$d\pi_{x_0, x_1} = \int_{X^{2^i-1}} M^{(i)}_{x_0, x_{2^{-i}}, \ldots, x_1} \tag{2.19}$$

defines a transference plan from μ_0 to μ_1. For $t = j \cdot 2^{-i}$, $0 \leq j \leq 2^i$, define $e_t : B^{(i)} \to X$ by $e_t(x_0, \ldots, x_1) = x_t$; then $(e_t)_* M^{(i)} = \mu_t$.

Let S be as in the proof of Proposition 2.6. Given $(x_0, \ldots, x_1) \in B^{(i)}$, define a map $p_{x_0, \ldots, x_1} : [0,1] \to X$ as the concatenation of the paths $S(x_0, x_{2^{-i}})$, $S(x_{2^{-i}}, x_{2\cdot 2^{-i}})$, $\ldots$, and $S(x_{1-2^{-i}}, x_1)$. As $p_{x_0, \ldots, x_1}$ is a normalized continuous curve from x_0 to x_1 of length $d(x_0, x_1)$, it is a geodesic. For each i, the linear functional on $C(\Gamma)$ given by

$$F \to \int_{X^{2^i+1}} F(p_{x_0, \ldots, x_1}) dM^{(i)}_{x_0, \ldots, x_1} \tag{2.20}$$

defines a probability measure $R^{(i)}$ on the compact space Γ. Let $R^{(\infty)}$ be the limit of a weak-$*$ convergent subsequence of $\{R^{(i)}\}_{i=1}^{\infty}$. It is also a probability measure on Γ.

For any $t \in \frac{\mathbb{N}}{2^{\mathbb{N}}} \cap [0,1]$ and $f \in C(X)$, we have $\int_K (e_t)^* f dR^{(i)} = \int_X f d\mu_t$ for large i. Then $\int_K (e_t)^* f dR^{(\infty)} = \int_X f d\mu_t$ for all $f \in C(X)$, or equivalently, $(e_t)_* R^{(\infty)} = \mu_t$. But as in the proof of Lemma 2.4, $(e_t)_* R^{(\infty)}$ is weak-$*$ continuous in t. It follows that $(e_t)_* R^{(\infty)} = \mu_t$ for all $t \in [0,1]$. $\square$

2.4. *Optimal transport on Riemannian manifolds.* In the rest of this section we discuss the case when X is a smooth, compact, connected Riemannian manifold M with Riemannian metric g. (The results are also valid if g is only C^3-smooth.)

Given $\mu_0, \mu_1 \in P_2(M)$ which are absolutely continuous with respect to dvol_M, we know that there is a unique Wasserstein geodesic c joining μ_0 to μ_1 [32, Th. 9]. Furthermore, for each $t \in [0,1]$, $c(t)$ is absolutely continuous with respect to dvol_M [19, Prop. 5.4]. Thus it makes sense to talk about the length space $P_2^{\mathrm{ac}}(M)$ of Borel probability measures on M that are absolutely continuous with respect to the Riemannian density, equipped with the metric W_2. It is a dense totally convex subset of $P_2(M)$. Note that if M is other than a point then $P_2^{\mathrm{ac}}(M)$ is an incomplete metric space and is neither open nor closed in $P_2(M)$.

An optimal transference plan in $P_2^{\mathrm{ac}}(M)$ turns out to be a Monge transport; that is, $c(t) = (F_t)_* \mu_0$ for a family of Monge transports $\{F_t\}_{t \in [0,1]}$ of M. For each $m \in M$, $\{F_t(m)\}_{t \in [0,1]}$ is a minimizing geodesic. Furthermore, there is a $\frac{d^2}{2}$-concave function ϕ on M so that for almost all $m \in M$, $F_t(m) = \exp_m(-t \nabla \phi(m))$ [19, Th. 3.2 and Cor. 5.2]. This function ϕ, just as any $\frac{d^2}{2}$-concave function on a compact Riemannian manifold, is Lipschitz [32, Lemma 2] and has a Hessian almost everywhere [19, Prop. 3.14]. If we only want the Wasserstein geodesic to be defined for an interval $[0, r^{-1}]$ then we can use the same formula with ϕ being $\frac{rd^2}{2}$-concave.

3. Functionals on the Wasserstein space

This section is devoted to the study of certain functions on the Wasserstein space $P_2(X)$. We first define the functional U_ν on $P_2(X)$. We then define λ-displacement convexity of the functional, along with its variations: (weak) λ-(a.c.) displacement convexity. We give relations among these various notions of displacement convexity. We define the H-functionals $H_{N,\nu}$. Finally, under certain displacement convexity assumptions, we prove HWI functional inequalities.

The notion of λ-displacement convexity is more conventional than that of weak λ-displacement convexity. However, the "weak" notion turns out to be more useful when considering measured Gromov-Hausdorff limits. We will see that the "weak" hypothesis is sufficient for proving functional inequalities.

3.1. *Weak displacement convexity.* All of our results will involve a distinguished reference measure, which is not *a priori* canonically given. So by "measured length space" we will mean a triple (X, d, ν), where (X, d) is a compact length space and ν is a Borel probability measure on X. These assumptions automatically imply that ν is a regular measure.

We write

$$P_2(X,\nu) = \big\{\mu \in P_2(X) : \mathrm{supp}(\mu) \subset \mathrm{supp}(\nu)\big\}. \tag{3.1}$$

We denote by $P_2^{\mathrm{ac}}(X,\nu)$ the elements of $P_2(X,\nu)$ that are absolutely continuous with respect to ν.

Definition 3.2. Let U be a continuous convex function on $[0,\infty)$ with $U(0)=0$. Given $\mu,\nu \in P_2(X)$, we define the functional $U_\nu : P_2(X) \to \mathbb{R}\cup\{\infty\}$ by

$$U_\nu(\mu) = \int_X U(\rho(x))\, d\nu(x) \; + \; U'(\infty)\, \mu_s(X), \tag{3.3}$$

where

$$\mu = \rho\nu + \mu_s \tag{3.4}$$

is the Lebesgue decomposition of μ with respect to ν into an absolutely continuous part $\rho\nu$ and a singular part μ_s.

Remark 3.5. If $U'(\infty) = \infty$, then finiteness of $U_\nu(\mu)$ implies that μ is absolutely continuous with respect to ν. This is not true if $U'(\infty) < \infty$.

Lemma 3.6. $U_\nu(\mu) \ge U_\nu(\nu) = U(1)$.

Remark 3.7. The lemma says that as a function of μ, U_ν is minimized at ν. If μ is absolutely continuous with respect to ν then the lemma is just Jensen's inequality in the form

$$\int_X U(\rho(x))\, d\nu(x) \ge U\left(\int_X \rho(x)\, d\nu(x)\right). \tag{3.8}$$

The general case could be proved using this particular case together with an approximation argument such as Theorem C.12. However, we give a direct proof below.

Proof of Lemma 3.6. As U is convex, for any $\alpha \in (0,1)$ we have

$$U(\alpha r + 1 - \alpha) \le \alpha U(r) + (1-\alpha) U(1), \tag{3.9}$$

or

$$U(r) - U(1) \ge \frac{1}{\alpha}\left[U(\alpha r + 1 - \alpha) - U(1)\right]. \tag{3.10}$$

Then

$$\int_X U(\rho)\, d\nu - U(1) \ge \int_X \frac{U(\alpha\rho + 1 - \alpha) - U(1)}{\alpha\rho - \alpha} (\rho - 1)\, d\nu, \tag{3.11}$$

where we take the integrand of the right-hand-side to vanish at points $x \in X$ where $\rho(x) = 1$. We break up the right-hand-side of (3.11) according to whether $\rho(x) \leq 1$ or $\rho(x) > 1$. From monotone convergence, for $\rho \leq 1$ we have

(3.12)
$$\lim_{\alpha\to 0^+} \int_X \frac{U(\alpha\rho+1-\alpha)-U(1)}{\alpha\rho-\alpha}(\rho-1)1_{\rho\leq 1}\,d\nu = U'_-(1)\int_X(\rho-1)1_{\rho\leq 1}\,d\nu,$$

while for $\rho > 1$ we have

(3.13)
$$\lim_{\alpha\to 0^+} \int_X \frac{U(\alpha\rho+1-\alpha)-U(1)}{\alpha\rho-\alpha}(\rho-1)1_{\rho> 1}\,d\nu = U'_+(1)\int_X(\rho-1)1_{\rho> 1}\,d\nu.$$

Then

$$\begin{aligned}(3.14)\quad \int_X U(\rho)\,d\nu - U(1) &\geq U'_-(1)\int_X(\rho-1)\,d\nu \\ &\quad +(U'_+(1)-U'_-(1))\int_X(\rho-1)1_{\rho>1}\,d\nu \\ &\geq U'_-(1)\int_X(\rho-1)\,d\nu \geq U'(\infty)\int_X(\rho-1)\,d\nu \\ &= -U'(\infty)\mu_s(X).\end{aligned}$$

As $U_\nu(\nu) = U(1)$, the lemma follows. □

Definition 3.15. Given a compact measured length space (X,d,ν) and a number $\lambda\in\mathbb{R}$, we say that U_ν is

- λ-*displacement convex* if for all Wasserstein geodesics $\{\mu_t\}_{t\in[0,1]}$ with $\mu_0,\mu_1 \in P_2(X,\nu)$, we have

  $$(3.16)\qquad U_\nu(\mu_t) \leq tU_\nu(\mu_1) + (1-t)U_\nu(\mu_0) - \frac{1}{2}\lambda t(1-t)W_2(\mu_0,\mu_1)^2$$

  for all $t\in[0,1]$;
- *weakly* λ-*displacement convex* if for all $\mu_0,\mu_1 \in P_2(X,\nu)$, there is *some* Wasserstein geodesic from μ_0 to μ_1 along which (3.16) is satisfied;
- (*weakly*) λ-a.c. *displacement convex* if the condition is satisfied when we just assume that $\mu_0,\mu_1 \in P_2^{\mathrm{ac}}(X,\nu)$.

Remark 3.17. In Definition 3.15 we assume that $\mathrm{supp}(\mu_0)\subset\mathrm{supp}(\nu)$ and $\mathrm{supp}(\mu_1)\subset\mathrm{supp}(\nu)$, but we do not assume that $\mathrm{supp}(\mu_t)\subset\mathrm{supp}(\nu)$ for $t\in(0,1)$.

Remark 3.18. If U_ν is λ-displacement convex and $\mathrm{supp}(\nu) = X$ then the function $t\to U_\nu(\mu_t)$ is λ-convex on $[0,1]$; i.e. for all $0\leq s\leq s'\leq 1$ and $t\in[0,1]$,

(3.19)
$$U_\nu(\mu_{ts'+(1-t)s}) \leq tU_\nu(\mu_{s'}) + (1-t)U_\nu(\mu_s) - \frac{1}{2}\lambda t(1-t)(s'-s)^2W_2(\mu_0,\mu_1)^2.$$

This is not *a priori* the case if we only assume that U_ν is weakly λ-displacement convex.

We may sometimes write "displacement convex" instead of 0-displacement convex. In short, *weakly* means that we require a condition to hold only for *some* geodesic between two measures, as opposed to all geodesics, and a.c. means that we only require the condition to hold when the two measures are absolutely continuous.

There are obvious implications

$$
\begin{array}{ccc}
\lambda\text{-displacement convex} & \Longrightarrow & \text{weakly } \lambda\text{-displacement convex} \\
\Downarrow & & \Downarrow \\
\lambda\text{-a.c. displacement convex} & \Longrightarrow & \text{weakly } \lambda\text{-a.c. displacement convex.}
\end{array}
\tag{3.20}
$$

The next proposition reverses the right vertical implication in (3.20).

PROPOSITION 3.21. *Let U be a continuous convex function on $[0, \infty)$ with $U(0) = 0$. Let (X, d, ν) be a compact measured length space. Then U_ν is weakly λ-displacement convex if and only if it is weakly λ-*a.c. *displacement convex.*

Proof. We must show that if U_ν is weakly λ-a.c. displacement convex then it is weakly λ-displacement convex. That is, for $\mu_0, \mu_1 \in P_2(X, \nu)$, we must show that there is some Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which

$$
U_\nu(\mu_t) \le tU_\nu(\mu_0) + (1-t)U_\nu(\mu_1) - \frac{1}{2}\lambda t(1-t)W_2(\mu_0, \mu_1)^2. \tag{3.22}
$$

We may assume that $U_\nu(\mu_0) < \infty$ and $U_\nu(\mu_1) < \infty$, as otherwise (3.22) is trivially true for any Wasserstein geodesic from μ_0 to μ_1. From Theorem C.12 in Appendix C, there are sequences $\{\mu_{k,0}\}_{k=1}^\infty$ and $\{\mu_{k,1}\}_{k=1}^\infty$ in $P_2^{\mathrm{ac}}(X, \nu)$ (in fact with continuous densities) so that $\lim_{k\to\infty} \mu_{k,0} = \mu_0$, $\lim_{k\to\infty} \mu_{k,1} = \mu_1$, $\lim_{k\to\infty} U_\nu(\mu_{k,0}) = U_\nu(\mu_0)$ and $\lim_{k\to\infty} U_\nu(\mu_{k,1}) = U_\nu(\mu_1)$. Let $c_k : [0,1] \to P_2(X)$ be a minimal geodesic from $\mu_{k,0}$ to $\mu_{k,1}$ such that for all $t \in [0,1]$,

$$
U_\nu(c_k(t)) \le tU_\nu(\mu_{k,1}) + (1-t)U_\nu(\mu_{k,0}) - \frac{1}{2}\lambda t(1-t)W_2(\mu_{k,0}, \mu_{k,1})^2. \tag{3.23}
$$

After taking a subsequence, we may assume that the geodesics $\{c_k\}_{k=1}^\infty$ converge uniformly (i.e. in $C([0,1], P_2(X))$) to a geodesic $c : [0,1] \to P_2(X)$ from μ_0 to μ_1 [11, Th. 2.5.14 and Prop. 2.5.17]. The lower semicontinuity of U_ν, Theorem B.33(i) in Appendix B, implies that

$$
U_\nu(c(t)) \le \liminf_{k\to\infty} U_\nu(c_k(t)).
$$

The proposition follows. $\square$

In fact, the proof of Proposition 3.21 gives the following slightly stronger result.

LEMMA 3.24. *Let U be a continuous convex function on $[0, \infty)$ with $U(0) = 0$. Let (X, d, ν) be a compact measured length space. Suppose that for* *all* $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(X, \nu)$ *with continuous densities, there is some Wasserstein geodesic from* μ_0 *to* μ_1 *along which* (3.16) *is satisfied. Then* U_ν *is weakly* λ*-displacement convex.*

The next lemma gives sufficient conditions for the horizontal implications in (3.20) to be reversed. We recall the definition of total convexity from Section 1.2.2.

LEMMA 3.25. (i) *Suppose that* X *has the property that for each minimizing geodesic* $c : [0,1] \to P_2(X)$, *there is some* $\delta_c > 0$ *so that the minimizing geodesic between* $c(t)$ *and* $c(t')$ *is unique whenever* $|t - t'| \leq \delta_c$. *Suppose that* $\mathrm{supp}(\nu) = X$. *If* U_ν *is weakly* λ*-displacement convex then it is* λ*-displacement convex.*

(ii) *Suppose that* $P_2^{\mathrm{ac}}(X, \nu)$ *is totally convex in* $P_2(X)$. *Suppose that* X *has the property that for each minimizing geodesic* $c : [0,1] \to P_2^{\mathrm{ac}}(X, \nu)$, *there is some* $\delta_c > 0$ *so that the minimizing geodesic between* $c(t)$ *and* $c(t')$ *is unique whenever* $|t - t'| \leq \delta_c$. *Suppose that* $\mathrm{supp}(\nu) = X$. *If* U_ν *is weakly* λ-a.c. *displacement convex then it is* λ-a.c. *displacement convex.*

Proof. For part (i), suppose that U_ν is weakly λ-displacement convex. Let $c : [0,1] \to P_2(X)$ be a minimizing geodesic. We want to show the λ-convexity of U_ν along c. By assumption, for all $0 \leq s \leq s' \leq 1$ there is some geodesic from $c(s)$ to $c(s')$ so that (3.19) is satisfied for all $t \in [0,1]$. If $|s - s'| \leq \delta_c$ then this geodesic must be $c\big|_{[s,s']}$. It follows that the function $s \to U_\nu(c(s))$ is λ-convex on each interval $[s, s']$ with $|s - s'| \leq \delta_c$, and hence on $[0,1]$. This proves part (i).

The same argument works for (ii) provided that we restrict to absolutely continuous measures. $\square$

3.2. *Important examples.* The following functionals will play an important role.

Definition 3.26. Put

$$U_N(r) = \begin{cases} Nr(1 - r^{-1/N}) & \text{if } 1 < N < \infty, \\ r \log r & \text{if } N = \infty. \end{cases} \tag{3.27}$$

Definition 3.28. Let $H_{N,\nu} : P_2(X) \to [0, \infty]$ be the functional associated to U_N, via Definition 3.3. More explicitly:

- For $N \in (1, \infty)$,

$$H_{N,\nu} = N - N \int_X \rho^{1 - \frac{1}{N}} \, d\nu, \tag{3.29}$$

where $\rho\nu$ is the absolutely continuous part in the Lebesgue decomposition of μ with respect to ν.

• For $N = \infty$, the functional $H_{\infty,\nu}$ is defined as follows: if μ is absolutely continuous with respect to ν, with $\mu = \rho\nu$, then

$$H_{\infty,\nu}(\mu) = \int_X \rho \log \rho \, d\nu, \tag{3.30}$$

while if μ is not absolutely continuous with respect to ν then $H_{\infty,\nu}(\mu) = \infty$.

To verify that $H_{N,\nu}$ is indeed the functional associated to U_N, we note that $U_N'(\infty) = N$ and write

$$\begin{aligned} N \int_X \rho \left(1 - \rho^{-\frac{1}{N}}\right) d\nu + N\mu_s(X) &= N \int_X \rho \left(1 - \rho^{-\frac{1}{N}}\right) d\nu \\ &\quad + N\left(1 - \int_X \rho \, d\nu\right) \\ &= N - N \int_X \rho^{1-\frac{1}{N}} \, d\nu. \end{aligned} \tag{3.31}$$

Of course, the difference of treatment of the singular part of ν according to whether N is finite or not reflects the fact that U_N grows at most linearly when $N < \infty$, but superlinearly when $N = \infty$. Theorem B.33 in Appendix B ensures that $H_{N,\nu}$ is lower semicontinuous on $P_2(X)$.

Remark 3.32. Formally extending (3.27) to the case $N = 1$ would give $U_1(r) = r - 1$, which does not satisfy the condition $U(0) = 0$. This could be ameliorated by instead considering the function $U(r) = r$. However, the corresponding entropy functional U_ν is identically one, which is not of much use. We will deal with the case $N = 1$ separately.

Remark 3.33. The quantity $H_{\infty,\nu}(\mu)$ is variously called the Boltzmann H-functional, the negative entropy or the relative Kullback information of μ with respect to ν. As a function of μ, $H_{N,\nu}(\mu)$ attains a minimum when $\mu = \nu$, which can be considered to be the measure with the least information content with respect to ν. In some sense, $H_{N,\nu}(\mu)$ is a way of measuring the nonuniformity of μ with respect to ν.

3.3. HWI *inequalities.*

Definition 3.34. Let (X, d, ν) be a compact measured length space. Let U be a continuous convex function on $[0, \infty)$, with $U(0) = 0$, which is C^2-regular on $(0, \infty)$. Given $\mu \in P_2^{\mathrm{ac}}(X, \nu)$ with $\rho = \frac{d\mu}{d\nu}$ a positive Lipschitz function on X, define the "generalized Fisher information" I_U by

$$I_U(\mu) = \int_X U''(\rho)^2 \, |\nabla^- \rho|^2 \, d\mu = \int_X \rho \, U''(\rho)^2 \, |\nabla^- \rho|^2 \, d\nu. \tag{3.35}$$

(See Remark 3.56 about the terminology.)

The following estimates generalize the ones that underlie the HWI inequalities in [36].

PROPOSITION 3.36. *Let (X, d, ν) be a compact measured length space. Let U be a continuous convex function on $[0, \infty)$ with $U(0) = 0$. Given $\mu \in P_2(X, \nu)$, let $\{\mu_t\}_{t \in [0,1]}$ be a Wasserstein geodesic from $\mu_0 = \mu$ to $\mu_1 = \nu$. Given $\lambda \in \mathbb{R}$, suppose that* (3.16) *is satisfied. Then*

$$\frac{\lambda}{2} W_2(\mu, \nu)^2 \leq U_\nu(\mu) - U_\nu(\nu). \tag{3.37}$$

Now suppose in addition that U is C^2-regular on $(0, \infty)$ and that $\mu \in P_2^{\mathrm{ac}}(X, \nu)$ is such that $\rho = \frac{d\mu}{d\nu}$ is a positive Lipschitz function on X. Suppose that $U_\nu(\mu) < \infty$ and $\mu_t \in P_2^{\mathrm{ac}}(X, \nu)$ for all $t \in [0, 1]$. Then

$$U_\nu(\mu) - U_\nu(\nu) \leq W_2(\mu, \nu) \sqrt{I_U(\mu)} - \frac{\lambda}{2} W_2(\mu, \nu)^2. \tag{3.38}$$

Proof. Consider the function $\phi(t) = U_\nu(\mu_t)$. Then $\phi(0) = U_\nu(\mu)$ and $\phi(1) = U_\nu(\nu)$. By assumption,

$$\phi(t) \leq t\phi(1) + (1 - t)\phi(0) - \frac{1}{2}\lambda t(1 - t) W_2(\mu, \nu)^2. \tag{3.39}$$

If $\phi(0) - \phi(1) < \frac{1}{2}\lambda W_2(\mu, \nu)^2$ then as

$$\phi(t) - \phi(1) \leq (1 - t) \left(\phi(0) - \phi(1) - \frac{1}{2}\lambda t W_2(\mu, \nu)^2 \right),$$

we conclude that $\phi(t) - \phi(1)$ is negative for t close to 1, which contradicts Lemma 3.6. Thus $\phi(0) - \phi(1) \geq \frac{1}{2}\lambda W_2(\mu, \nu)^2$, which proves (3.37).

To prove (3.38), put $\rho_t = \frac{d\mu_t}{d\nu}$. Then $\phi(t) = \int_X U(\rho_t)\, d\nu$. From (3.39), for $t > 0$ we have

$$\phi(0) - \phi(1) \leq -\frac{\phi(t) - \phi(0)}{t} - \frac{1}{2}\lambda(1 - t) W_2(\mu, \nu)^2. \tag{3.40}$$

To prove the inequality (3.38), it suffices to prove that

$$\liminf_{t \to 0} \left(-\frac{\phi(t) - \phi(0)}{t} \right) \leq W_2(\mu, \nu) \sqrt{I_U(\mu)}. \tag{3.41}$$

The convexity of U implies that

$$U(\rho_t) - U(\rho_0) \geq U'(\rho_0)(\rho_t - \rho_0). \tag{3.42}$$

Integrating with respect to ν and dividing by $-t < 0$, we infer

$$-\frac{1}{t} [\phi(t) - \phi(0)] \leq -\frac{1}{t} \int_X U'(\rho_0(x))[d\mu_t(x) - d\mu_0(x)]. \tag{3.43}$$

By Proposition 2.10, $\mu_t = (e_t)_* \Pi$, where Π is a certain probability measure on the space Γ of minimal geodesics in X. In particular,

(3.44)
$$-\frac{1}{t}\int_X U'(\rho_0(x))[d\mu_t(x) - d\mu_0(x)] = -\frac{1}{t}\int_\Gamma [U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]\, d\Pi(\gamma).$$

Since U' is nondecreasing and $td(\gamma(0),\gamma(1)) = d(\gamma(0),\gamma(t))$, we have

(3.45)
$$\begin{aligned} &-\frac{1}{t}\int_\Gamma \big[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))\big]\, d\Pi(\gamma) \\ &\le -\frac{1}{t}\int_\Gamma 1_{\rho_0(\gamma(t))\le\rho_0(\gamma(0))} \big[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))\big]\, d\Pi(\gamma) \\ &= \int_\Gamma \frac{U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))}{\rho_0(\gamma(t)) - \rho_0(\gamma(0))} \frac{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]_-}{d(\gamma(0),\gamma(t))} d(\gamma(0),\gamma(1)) d\Pi(\gamma), \end{aligned}$$

where strictly speaking we define the integrand of the last term to be zero when $\rho_0(\gamma(t)) = \rho_0(\gamma(0))$. Applying the Cauchy-Schwarz inequality, we can bound the last term above by

$$\sqrt{\int_\Gamma \frac{[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]^2}{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]^2} \frac{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]_-^2}{d(\gamma(0),\gamma(t))^2} d\Pi(\gamma)} \sqrt{\int_\Gamma d(\gamma(0),\gamma(1))^2 d\Pi(\gamma)}. \tag{3.46}$$

The second square root is just $W_2(\mu_0, \mu_1)$. To conclude the argument, it suffices to show that

(3.47)
$$\liminf_{t\to 0} \int_\Gamma \frac{[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]^2}{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]^2} \frac{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]_-^2}{d(\gamma(0),\gamma(t))^2} d\Pi(\gamma) \le I_U(\mu).$$

The continuity of ρ_0 implies that $\lim_{t\to 0}\rho_0(\gamma(t)) = \rho_0(\gamma(0))$. So that

$$\lim_{t\to 0} \frac{[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]^2}{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]^2} = U''(\rho_0(\gamma(0)))^2. \tag{3.48}$$

On the other hand, the definition of the gradient implies

$$\limsup_{t\to 0} \frac{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]_-^2}{d(\gamma(0),\gamma(t))^2} \le |\nabla^-\rho_0|^2(\gamma(0)). \tag{3.49}$$

As ρ_0 is a positive Lipschitz function on X, and U' is C^1-regular on $(0,\infty)$, $U' \circ \rho_0$ is also Lipschitz on X. Then $\frac{[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]^2}{d(\gamma(0),\gamma(t))^2}$ is uniformly bounded on Γ, with respect to t, and dominated convergence implies that

(3.50)

$$\liminf_{t\to 0}\int_\Gamma \frac{[U'(\rho_0(\gamma(t))) - U'(\rho_0(\gamma(0)))]^2}{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]^2}\frac{[\rho_0(\gamma(t)) - \rho_0(\gamma(0))]_-^2}{d(\gamma(0),\gamma(t))^2}d\Pi(\gamma)$$
$$\leq \int_\Gamma U''(\rho_0(\gamma(0)))^2|\nabla^-\rho_0|^2(\gamma(0))\,d\Pi(\gamma) = \int_X U''(\rho_0(x))^2|\nabla^-\rho_0|^2(x)\,d\mu(x).$$

This concludes the proof of the inequality on the right-hand-side of (3.38). $\square$

Remark 3.51. Modulo the notational burden caused by the nonsmooth setting, the proof of Proposition 3.36 is somewhat simpler than the "standard" Euclidean proof because we used a convexity inequality to avoid computing $\phi'(0)$ explicitly (compare with [44, p. 161]).

Particular cases 3.52. Taking $U = U_N$, with $\mu = \rho\nu$ and $\rho \in \mathrm{Lip}(X)$ a positive function, define

$$I_{N,\nu}(\mu) = \begin{cases} \left(\dfrac{N-1}{N}\right)^2 \displaystyle\int_X \frac{|\nabla^-\rho|^2}{\rho^{\frac{2}{N}+1}}\,d\nu & \text{if } 1 < N < \infty, \\[2ex] \displaystyle\int_X \frac{|\nabla^-\rho|^2}{\rho}\,d\nu & \text{if } N = \infty. \end{cases} \tag{3.53}$$

Proposition (3.36) implies the following inequalities:

- If $\lambda > 0$ then

(3.54)

$$\frac{\lambda}{2}W_2(\mu,\nu)^2 \leq H_{N,\nu}(\mu) \leq W_2(\mu,\nu)\sqrt{I_{N,\nu}(\mu)} - \frac{\lambda}{2}W_2(\mu,\nu)^2 \leq \frac{1}{2\lambda}I_{N,\nu}(\mu).$$

- If $\lambda \leq 0$ then

$$H_{N,\nu}(\mu) \leq \mathrm{diam}(X)\sqrt{I_{N,\nu}(\mu)} - \frac{\lambda}{2}\,\mathrm{diam}(X)^2. \tag{3.55}$$

Remark 3.56. $I_{\infty,\nu}(\mu)$ is the classical Fisher information of μ relative to the reference measure ν, which is why we call I_U a "generalized Fisher information".

4. Weak displacement convexity and measured Gromov-Hausdorff limits

In this section we first show that if a sequence of compact metric spaces converges in the Gromov-Hausdorff topology then their associated Wasserstein spaces also converge in the Gromov-Hausdorff topology. Assuming the results of Appendices B and C, we show that weak displacement convexity of U_ν is preserved by measured Gromov-Hausdorff limits. Finally, we define the notion of weak λ-displacement convexity for a family $\mathcal{F}$ of functions U.

4.1. *Gromov-Hausdorff convergence of the Wasserstein space.*

PROPOSITION 4.1. *If $f : (X_1, d_1) \to (X_2, d_2)$ is an ε-Gromov-Hausdorff approximation then $f_* : P_2(X_1) \to P_2(X_2)$ is an $\widetilde{\varepsilon}$-Gromov-Hausdorff approximation, where*

$$\widetilde{\varepsilon} = 4\varepsilon + \sqrt{3\varepsilon(2\,\mathrm{diam}(X_2) + 3\varepsilon)}. \tag{4.2}$$

COROLLARY 4.3. *If a sequence of compact metric spaces $\{(X_i, d_i)\}_{i=1}^{\infty}$ converges in the Gromov-Hausdorff topology to a compact metric space (X, d) then $\{P_2(X_i)\}_{i=1}^{\infty}$ converges in the Gromov-Hausdorff topology to $P_2(X)$.*

Proof of Proposition 4.1. Given $\mu_1, \mu_1' \in P_2(X_1)$, let π_1 be an optimal transference plan for μ_1 and μ_1'. Put $\pi_2 = (f \times f)_* \pi_1$. Then π_2 is a transference plan for $f_*\mu_1$ and $f_*\mu_1'$. We have

$$\begin{aligned} W_2(f_*\mu_1, f_*\mu_1')^2 &\le \int_{X_2 \times X_2} d_2(x_2, y_2)^2 d\pi_2(x_2, y_2) \\ &= \int_{X_1 \times X_1} d_2(f(x_1), f(y_1))^2 d\pi_1(x_1, y_1). \end{aligned} \tag{4.4}$$

As

$$\begin{aligned} &|d_2(f(x_1), f(y_1))^2 - d_1(x_1, y_1)^2| \\ &\quad = |d_2(f(x_1), f(y_1)) - d_1(x_1, y_1)| \big(d_2(f(x_1), f(y_1)) + d_1(x_1, y_1)\big), \end{aligned} \tag{4.5}$$

we have

$$|d_2(f(x_1), f(y_1))^2 - d_1(x_1, y_1)^2| \le \varepsilon(2\,\mathrm{diam}(X_1) + \varepsilon) \tag{4.6}$$

and

$$|d_2(f(x_1), f(y_1))^2 - d_1(x_1, y_1)^2| \le \varepsilon(2\,\mathrm{diam}(X_2) + \varepsilon). \tag{4.7}$$

It follows that

$$W_2(f_*\mu_1, f_*\mu_1')^2 \le W_2(\mu_1, \mu_1')^2 + \varepsilon(2\,\mathrm{diam}(X_1) + \varepsilon) \tag{4.8}$$

and

$$W_2(f_*\mu_1, f_*\mu_1')^2 \le W_2(\mu_1, \mu_1')^2 + \varepsilon(2\,\mathrm{diam}(X_2) + \varepsilon). \tag{4.9}$$

From this last inequality,

$$W_2(f_*\mu_1, f_*\mu_1') \le W_2(\mu_1, \mu_1') + \sqrt{\varepsilon(2\,\mathrm{diam}(X_2) + \varepsilon)}. \tag{4.10}$$

We now exchange the roles of X_1 and X_2. We correspondingly apply (4.8) instead of (4.9), to the map f' and the measures $f_*\mu_1$ and $f_*\mu_1'$, and use the fact that f' is a 3ε-Gromov-Hausdorff approximation, to obtain

$$W_2(f'_*(f_*\mu_1), f'_*(f_*\mu_1')) \leq W_2(f_*\mu_1, f_*\mu_1') + \sqrt{3\varepsilon(2\operatorname{diam}(X_2) + 3\varepsilon)}. \tag{4.11}$$

Since $f' \circ f$ is an admissible Monge transport between μ_1 and $(f' \circ f)_*\mu_1$, or between μ_1' and $(f' \circ f)_*\mu_1'$, which moves points by a distance at most 2ε, we have

$$W_2((f' \circ f)_*\mu_1, \mu_1) \leq 2\varepsilon, \qquad W_2((f' \circ f)_*\mu_1', \mu_1') \leq 2\varepsilon. \tag{4.12}$$

Thus by (4.11) and the triangle inequality,

$$W_2(\mu_1, \mu_1') \leq W_2(f_*\mu_1, f_*\mu_1') + 4\varepsilon + \sqrt{3\varepsilon(2\operatorname{diam}(X_2) + 3\varepsilon)}. \tag{4.13}$$

Equations (4.10) and (4.13) show that condition (i) of Definition 1.8 is satisfied.

Finally, given $\mu_2 \in P_2(X_2)$, consider the Monge transport $f \circ f'$ from μ_2 to $(f \circ f')_*\mu_2$. Then $W_2(\mu_2, f_*(f'_*\mu_2)) \leq \varepsilon$. Thus condition (ii) of Definition 1.8 is satisfied as well. □

Remark 4.14. The map f_* is generally discontinuous. In fact, it is continuous if and only if f is continuous.

4.2. *Stability of weak displacement convexity.*

THEOREM 4.15. *Let $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ be a sequence of compact measured length spaces so that $\lim_{i\to\infty}(X_i, d_i, \nu_i) = (X, d, \nu_\infty)$ in the measured Gromov-Hausdorff topology. Let U be a continuous convex function on $[0, \infty)$ with $U(0) = 0$. Given $\lambda \in \mathbb{R}$, suppose that for all i, U_{ν_i} is weakly λ-displacement convex for (X_i, d_i, ν_i). Then U_{ν_∞} is weakly λ-displacement convex for (X, d, ν).*

Proof. By Lemma 3.24, it suffices to show that for any $\mu_0, \mu_1 \in P_2(X)$ with continuous densities with respect to ν_∞, there is a Wasserstein geodesic joining them along which inequality (3.16) holds for U_{ν_∞}. We may assume that $U_{\nu_\infty}(\mu_0) < \infty$ and $U_{\nu_\infty}(\mu_1) < \infty$, as otherwise any Wasserstein geodesic works.

Write $\mu_0 = \rho_0\nu_\infty$ and $\mu_1 = \rho_1\nu_\infty$. Let $f_i : X_i \to X$ be an ε_i-approximation, with $\lim_{i\to\infty}\varepsilon_i = 0$ and $\lim_{i\to\infty}(f_i)_*\nu_i = \nu_\infty$. If i is sufficiently large then $\int_X \rho_0 d(f_i)_*\nu_i > 0$ and $\int_X \rho_1 d(f_i)_*\nu_i > 0$. For such i, put $\mu_{i,0} = \frac{(f_i^*\rho_0)\nu_i}{\int_X \rho_0 d(f_i)_*\nu_i}$ and $\mu_{i,1} = \frac{(f_i^*\rho_1)\nu_i}{\int_X \rho_1 d(f_i)_*\nu_i}$. Then $(f_i)_*\mu_{i,0} = \frac{\rho_0 (f_i)_*\nu_i}{\int_X \rho_0 d(f_i)_*\nu_i}$ and $(f_i)_*\mu_{i,1} = \frac{\rho_1 (f_i)_*\nu_i}{\int_X \rho_1 d(f_i)_*\nu_i}$. Now choose geodesics $c_i : [0,1] \to P_2(X_i)$ with $c_i(0) = \mu_{i,0}$ and $c_i(1) = \mu_{i,1}$ so that for all $t \in [0,1]$, we have

$$U_{\nu_i}(c_i(t)) \leq tU_{\nu_i}(\mu_{i,1}) + (1-t)U_{\nu_i}(\mu_{i,0}) - \frac{1}{2}\lambda t(1-t)W_2(\mu_{i,0}, \mu_{i,1})^2. \tag{4.16}$$

From Lemma 1.11 and Corollary 4.3, after passing to a subsequence, the maps $(f_i)_* \circ c_i : [0,1] \to P_2(X)$ converge uniformly to a continuous map $c : [0,1] \to P_2(X)$. As $W_2(c_i(t), c_i(t')) = |t - t'| W_2(\mu_{i,0}, \mu_{i,1})$, it follows that $W_2(c(t), c(t')) = |t - t'| W_2(\mu_0, \mu_1)$. Thus c is a Wasserstein geodesic. The problem is to pass to the limit in (4.16) as $i \to \infty$.

Given $F \in C(X)$, the fact that $\rho_0 \in C(X)$ implies that

$$\lim_{i\to\infty} \int_X F d(f_i)_* \mu_{i,0} = \lim_{i\to\infty} \int_X F \rho_0 \frac{d(f_i)_*\nu_i}{\int_X \rho_0 d(f_i)_*\nu_i} = \int_X F \rho_0 \, d\nu_\infty. \tag{4.17}$$

Thus $\lim_{i\to\infty} (f_i)_* \mu_{i,0} = \mu_0$. Similarly, $\lim_{i\to\infty} (f_i)_* \mu_{i,1} = \mu_1$. It follows from Corollary 4.3 that

$$\lim_{i\to\infty} W_2(\mu_{i,0}, \mu_{i,1}) = W_2(\mu_0, \mu_1). \tag{4.18}$$

Next,

$$U_{\nu_i}(\mu_{i,0}) = \int_{X_i} U\left(\frac{f_i^* \rho_0}{\int_X \rho_0 d(f_i)_*\nu_i} \right) d\nu_i = \int_X U\left(\frac{\rho_0}{\int_X \rho_0 d(f_i)_*\nu_i} \right) d(f_i)_*\nu_i. \tag{4.19}$$

As

$$\lim_{i\to\infty} U\left(\frac{\rho_0}{\int_X \rho_0 d(f_i)_*\nu_i} \right) = U(\rho_0) \tag{4.20}$$

uniformly on X, it follows that

$$\lim_{i\to\infty} \int_X U\left(\frac{\rho_0}{\int_X \rho_0 d(f_i)_*\nu_i} \right) d(f_i)_*\nu_i = \lim_{i\to\infty} \int_X U(\rho_0) d(f_i)_*\nu_i = \int_X U(\rho_0) \, d\nu_\infty. \tag{4.21}$$

Thus $\lim_{i\to\infty} U_{\nu_i}(\mu_{i,0}) = U_{\nu_\infty}(\mu_0)$. Similarly, $\lim_{i\to\infty} U_{\nu_i}(\mu_{i,1}) = U_{\nu_\infty}(\mu_1)$.

It follows from Theorem B.33(ii) in Appendix B that

$$U_{(f_i)_*\nu_i}((f_i)_* c_i(t)) \le U_{\nu_i}(c_i(t)). \tag{4.22}$$

Then, for any $t \in [0,1]$, we can combine this with the lower semicontinuity of $(\mu, \nu) \to U_\nu(\mu)$ (Theorem B.33(i) in Appendix B) to obtain

$$U_{\nu_\infty}(c(t)) \le \liminf_{i\to\infty} U_{(f_i)_*\nu_i}((f_i)_* c_i(t)) \le \liminf_{i\to\infty} U_{\nu_i}(c_i(t)). \tag{4.23}$$

Combining this with (4.18) and the preceding results, we can take $i \to \infty$ in (4.16) and find

$$U_{\nu_\infty}(c(t)) \le t U_{\nu_\infty}(\mu_1) + (1-t) U_{\nu_\infty}(\mu_0) - \frac{1}{2} \lambda t (1-t) W_2(\mu_0, \mu_1)^2. \tag{4.24}$$

This concludes the proof. $\square$

Definition 4.25. Let $\mathcal{F}$ be a family of continuous convex functions U on $[0, \infty)$ with $U(0) = 0$. Given a function $\lambda : \mathcal{F} \to \mathbb{R} \cup \{-\infty\}$, we say that a compact measured length space (X, d, ν) is weakly λ-displacement convex for the family $\mathcal{F}$ if for any $\mu_0, \mu_1 \in P_2(X, \nu)$, one can find a Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 so that for each $U \in \mathcal{F}$, U_ν satisfies

$$U_\nu(\mu_t) \leq tU_\nu(\mu_1) + (1-t)U_\nu(\mu_0) - \frac{1}{2}\lambda(U)t(1-t)W_2(\mu_0, \mu_1)^2 \tag{4.26}$$

for all $t \in [0, 1]$.

There is also an obvious definition of "weakly λ-a.c. displacement convex for the family $\mathcal{F}$", in which one just requires the condition to hold when $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(X, \nu)$. Note that in Definition 4.25, the same Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ is supposed to work for all of the functions $U \in \mathcal{F}$. Hence if (X, d, ν) is weakly λ-displacement convex for the family $\mathcal{F}$ then it is weakly $\lambda(U)$-displacement convex for each $U \in \mathcal{F}$, but the converse is not *a priori* true.

The proof of Theorem 4.15 establishes the following result.

THEOREM 4.27. *Let $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ be a sequence of compact measured length spaces with $\lim_{i\to\infty}(X_i, d_i, \nu_i) = (X, d, \nu_\infty)$ in the measured Gromov-Hausdorff topology. Let $\mathcal{F}$ be a family of continuous convex functions U on $[0, \infty)$ with $U(0) = 0$. Given a function $\lambda : \mathcal{F} \to \mathbb{R} \cup \{-\infty\}$, suppose that each (X_i, d_i, ν_i) is weakly λ-displacement convex for the family $\mathcal{F}$. Then (X, d, ν_∞) is weakly λ-displacement convex for the family $\mathcal{F}$.*

For later use, we note that the proof of Proposition 3.21 establishes the following result.

PROPOSITION 4.28. *Let $\mathcal{F}$ be a family of continuous convex functions U on $[0, \infty)$ with $U(0) = 0$. Given a function $\lambda : \mathcal{F} \to \mathbb{R} \cup \{-\infty\}$, (X, d, ν) is weakly λ-displacement convex for the family $\mathcal{F}$ if and only if it is weakly λ-*a.c. *displacement convex for the family $\mathcal{F}$.*

5. N-Ricci curvature for measured length spaces

This section deals with N-Ricci curvature and its basic properties. We first define certain classes $\mathcal{DC}_N$ of convex functions U. We use these to define the notions of a measured length space (X, d, ν) having nonnegative N-Ricci curvature, or ∞-Ricci curvature bounded below by $K \in \mathbb{R}$. We show that these properties pass to totally convex subsets of X. We prove that the Ricci curvature definitions are preserved by measured Gromov-Hausdorff limits. We show that nonnegative N-Ricci curvature for $N < \infty$ implies a Bishop-Gromov-type inequality. We show that in certain cases, lower Ricci curvature bounds are preserved upon quotienting by compact group actions. Finally, we show that under the assumption of nonnegative N-Ricci curvature with $N < \infty$, any two measures that are absolutely continuous with respect to ν can be joined by a Wasserstein geodesic all of whose points are absolutely continuous measures with respect to ν.

5.1. *Displacement convex classes.* We first define a suitable class of convex functions, introduced by McCann [31]. Consider a continuous convex function $U : [0, \infty) \to \mathbb{R}$ with $U(0) = 0$. We define the nonnegative function

$$p(r) = rU'_+(r) - U(r), \tag{5.1}$$

with $p(0) = 0$. If one thinks of U as defining an internal energy for a continuous medium then p can be thought of as a pressure. By analogy, if U is C^2-regular on $(0, \infty)$ then we define the "iterated pressure"

$$p_2(r) = rp'(r) - p(r). \tag{5.2}$$

Definition 5.3. For $N \in [1, \infty)$, we define $\mathcal{DC}_N$ to be the *set of all continuous convex functions* U *on* $[0, \infty)$, with $U(0) = 0$, such that the function

$$\psi(\lambda) = \lambda^N U(\lambda^{-N}) \tag{5.4}$$

is convex on $(0, \infty)$.

We further define $\mathcal{DC}_\infty$ to be the *set of all continuous convex functions* U *on* $[0, \infty)$, with $U(0) = 0$, such that the function

$$\psi(\lambda) = e^\lambda U(e^{-\lambda}) \tag{5.5}$$

is convex on $(-\infty, \infty)$.

We note that the convexity of U implies that ψ is nonincreasing in λ, as $\frac{U(x)}{x}$ is nondecreasing in x. Below are some useful facts about the classes $\mathcal{DC}_N$.

LEMMA 5.6. *If $N \leq N'$ then $\mathcal{DC}_{N'} \subset \mathcal{DC}_N$.*

Proof. If $N' < \infty$, let ψ_N and $\psi_{N'}$ denote the corresponding functions. Then $\psi_N(\lambda) = \psi_{N'}\left(\lambda^{N/N'}\right)$. The conclusion follows from the fact that the function $x \to x^{N/N'}$ is concave on $[0, \infty)$, along with the fact that the composition of a nonincreasing convex function and a concave function is convex. The case $N' = \infty$ is similar. $\square$

LEMMA 5.7. *For $N \in [1, \infty]$,*

(a) *If U is a continuous convex function on $[0, \infty)$ with $U(0) = 0$ then $U \in \mathcal{DC}_N$ if and only if the function $r \longmapsto p(r)/r^{1-\frac{1}{N}}$ is nondecreasing on $(0, \infty)$.*

(b) *If U is a continuous convex function on $[0, \infty)$ that is C^2-regular on $(0, \infty)$ with $U(0) = 0$, then $U \in \mathcal{DC}_N$ if and only if $p_2 \geq -\frac{p}{N}$.*

Proof. Suppose first that U is a continuous convex function on $[0,\infty)$ and $N \in [1,\infty)$. Putting $r(\lambda) = \lambda^{-N}$, one can check that

$$\psi'_{-}(\lambda) = -Np(r)/r^{1-\frac{1}{N}}. \tag{5.8}$$

Then ψ is convex if and only if ψ'_{-} is nondecreasing, which is the case if and only if the function $r \longmapsto p(r)/r^{1-\frac{1}{N}}$ is nondecreasing (since the map $\lambda \to \lambda^{-N}$ is nonincreasing). Next, suppose that U is C^2-regular on $(0,\infty)$. One can check that

$$\psi''(\lambda) = N^2 r^{\frac{2}{N}-1}\left(p_2(r) + \frac{p(r)}{N}\right). \tag{5.9}$$

Then ψ is convex if and only if $\psi'' \geq 0$, which is the case if and only if $p_2 \geq -\frac{p}{N}$.

The proof in the case $N = \infty$ is similar. □

LEMMA 5.10. *Given $U \in \mathcal{DC}_\infty$, either U is linear or there exist $a, b > 0$ such that $U(r) \geq a\, r \log r \,-\, br$.*

Proof. The function U can be reconstructed from ψ by the formula

$$U(x) = x\psi(\log(1/x)). \tag{5.11}$$

As ψ is convex and nonincreasing, either ψ is constant or there are constants $a, b > 0$ such that $\psi(\lambda) \geq -a\lambda - b$ for all $\lambda \in \mathbb{R}$. In the first case, U is linear. In the second case, $U(x) \geq -ax\log(1/x) - bx$, as required. □

5.2. *Ricci curvature via weak displacement convexity.* We recall from Definition 4.25 the notion of a compact measured length space (X, d, ν) being weakly λ-displacement convex for a family of convex functions $\mathcal{F}$.

Definition 5.12. Given $N \in [1,\infty]$, we say that a compact measured length space (X, d, ν) has nonnegative N-Ricci curvature if it is weakly displacement convex for the family $\mathcal{DC}_N$.

By Lemma 5.6, if $N \leq N'$ and X has nonnegative N-Ricci curvature then it has nonnegative N'-Ricci curvature. In the case $N = \infty$, we can define a more precise notion.

Definition 5.13. Given $K \in \mathbb{R}$, define $\lambda : \mathcal{DC}_\infty \to \mathbb{R} \cup \{-\infty\}$ by

$$\lambda(U) = \inf_{r>0} K\frac{p(r)}{r} = \begin{cases} K \lim_{r\to 0^+} \frac{p(r)}{r} & \text{if } K > 0, \\ 0 & \text{if } K = 0, \\ K \lim_{r\to\infty} \frac{p(r)}{r} & \text{if } K < 0, \end{cases} \tag{5.14}$$

where p is given by (5.1). We say that a compact measured length space (X, d, ν) has ∞-Ricci curvature bounded below by K if it is weakly λ-displacement convex for the family $\mathcal{DC}_\infty$.

If $K \leq K'$ and (X, d, ν) has ∞-Ricci curvature bounded below by K' then it has ∞-Ricci curvature bounded below by K.

The next proposition shows that our definitions localize on totally convex subsets.

PROPOSITION 5.15. *Suppose that a closed set $A \subset X$ is totally convex. Given $\nu \in P_2(X)$ with $\nu(A) > 0$, put $\nu' = \frac{1}{\nu(A)} \nu\big|_A \in P_2(A)$.*

(a) *If (X, d, ν) has nonnegative N-Ricci curvature then (A, d, ν') has nonnegative N-Ricci curvature.*
(b) *If (X, d, ν) has ∞-Ricci curvature bounded below by K then (A, d, ν') has ∞-Ricci curvature bounded below by K.*

Proof. By Proposition 2.10, $P_2(A)$ is a totally convex subset of $P_2(X)$. Given $\mu \in P_2(A) \subset P_2(X)$, let $\mu = \rho\nu + \mu_s$ be its Lebesgue decomposition with respect to ν. Then $\mu = \rho'\nu' + \mu_s$ is the Lebesgue decomposition of μ with respect to ν', where $\rho' = \nu(A)\rho\big|_A$. Given a continuous convex function $U : [0, \infty) \to \mathbb{R}$ with $U(0) = 0$, define

$$\widetilde{U}(r) = \frac{U(\nu(A)r)}{\nu(A)}. \tag{5.16}$$

Then $\widetilde{U}'(\infty) = U'(\infty)$, and $U \in \mathcal{DC}_N$ if and only if $\widetilde{U} \in \mathcal{DC}_N$. Now

$$\begin{aligned} U_{\nu'}(\mu) &= \int_A U(\rho')\, d\nu' + U'(\infty)\mu_s(A) \\ &= \frac{1}{\nu(A)} \int_A U(\nu(A)\rho)\, d\nu + U'(\infty)\mu_s(A) \\ &= \int_X \widetilde{U}(\rho)\, d\nu + \widetilde{U}'(\infty)\mu_s(X) = \widetilde{U}_\nu(\mu). \end{aligned} \tag{5.17}$$

As $P_2(A, \nu') \subset P_2(X, \nu)$, part (a) follows.

Letting $\widetilde{p}$ denote the pressure of $\widetilde{U}$, one finds that

$$\frac{\widetilde{p}(r)}{r} = \frac{p(\nu(A)r)}{\nu(A)r}. \tag{5.18}$$

Then with reference to Definition 5.13, $\lambda(\widetilde{U}) = \lambda(U)$. Part (b) follows. □

5.3. *Preservation of N-Ricci curvature bounds.* The next theorem can be considered to be the main result of this paper.

THEOREM 5.19. *Let $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ be a sequence of compact measured length spaces with $\lim_{i\to\infty}(X_i, d_i, \nu_i) = (X, d, \nu)$ in the measured Gromov-Hausdorff topology.*

• *If each (X_i, d_i, ν_i) has nonnegative N-Ricci curvature then (X, d, ν) has nonnegative N-Ricci curvature.*

• *If each* (X_i, d_i, ν_i) *has* ∞*-Ricci curvature bounded below by* K*, for some* $K \in \mathbb{R}$*, then* (X, d, ν) *has* ∞*-Ricci curvature bounded below by* K*.*

Proof. If $N < \infty$ then the theorem follows from Theorem 4.27 with the family $\mathcal{F} = \mathcal{DC}_N$ and $\lambda = 0$. If $N = \infty$ then it follows from Theorem 4.27 with the family $\mathcal{F} = \mathcal{DC}_\infty$ and λ given by Definition 5.13. □

In what we have presented so far, the concept of (X, d, ν) having nonnegative N-Ricci curvature, or having ∞-Ricci curvature bounded below by K, may seem somewhat abstract. In Section 7 we will show that in the setting of Riemannian manifolds, it can be expressed in terms of classical tensors related to the Ricci tensor.

5.4. *Bishop-Gromov inequality.* We first show that a weak displacement convexity assumption implies that the measure ν either is a delta function or is nonatomic.

PROPOSITION 5.20. *Let* (X, d, ν) *be a compact measured length space. For all* $N \in (1, \infty]$*, if* $H_{N,\nu}$ *is weakly* λ*-displacement convex then* ν *either is a delta function or is nonatomic.*

Proof. We will assume that $\nu(\{x\}) \in (0, 1)$ for some $x \in X$ and derive a contradiction.

Suppose first that $N \in (1, \infty)$. Put $\mu_0 = \delta_x$ and $\mu_1 = \frac{\nu - \nu(\{x\})\delta_x}{1 - \nu(\{x\})}$. By the hypothesis and Proposition 2.10, there is a displacement interpolation $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which (3.16) is satisfied with $U_\nu = H_{N,\nu}$. Now $H_{N,\nu}(\mu_0) = N - N(\nu(\{x\}))^{1/N}$ and $H_{N,\nu}(\mu_1) = N - N(1 - \nu(\{x\}))^{1/N}$. Hence

$$
\begin{aligned}
(5.21) \qquad H_{N,\nu}(\mu_t) \leq\, & N - (1-t)N(\nu(\{x\}))^{1/N} \\
& - tN(1-\nu(\{x\}))^{1/N} - \frac{1}{2}\lambda t(1-t)W_2(\mu_0, \mu_1)^2.
\end{aligned}
$$

Put $D = \operatorname{diam}(X)$. As we have a displacement interpolation, it follows that if $t > 0$ then $\operatorname{supp}(\mu_t) \subset \overline{B_{tD}(x)}$ and $\mu_t(\{x\}) = 0$. When $\mu_t = \rho_t \nu + (\mu_t)_s$ is the Lebesgue decomposition of μ_t with respect to ν, Hölder's inequality implies that

$$
\begin{aligned}
(5.22) \qquad \int_X \rho_t^{1-\frac{1}{N}}\, d\nu &= \int_{\overline{B_{tD}(x)} - \{x\}} \rho_t^{1-\frac{1}{N}}\, d\nu \\
&\leq \left(\int_{\overline{B_{tD}(x)} - \{x\}} \rho_t \, d\nu \right)^{1-\frac{1}{N}} \nu\left(\overline{B_{tD}(x)} - \{x\}\right)^{\frac{1}{N}} \\
&\leq \nu\left(\overline{B_{tD}(x)} - \{x\}\right)^{\frac{1}{N}}.
\end{aligned}
$$

Then

$$H_{N,\nu}(\mu_t) \geq N - N \left(\nu \left(\overline{B_{tD}(x)} \right) - \nu(\{x\}) \right)^{1/N}. \tag{5.23}$$

As $\lim_{t \to 0^+} \nu \left(\overline{B_{tD}(x)} \right) = \nu(\{x\})$, we obtain a contradiction to (5.21) when t is small.

If $N = \infty$ then $H_{\infty,\nu}(\mu_0) = \log \frac{1}{\nu(\{x\})}$ and $H_{\infty,\nu}(\mu_1) = \log \frac{1}{1-\nu(\{x\})}$. Hence

$$H_{\infty,\nu}(\mu_t) \leq (1-t) \log \frac{1}{\nu(\{x\})} + t \log \frac{1}{1-\nu(\{x\})} - \frac{1}{2} \lambda t(1-t) W_2(\mu_0, \mu_1)^2. \tag{5.24}$$

In particular, μ_t is absolutely continuous with respect to ν. Write $\mu_t = \rho_t \nu$. Jensen's inequality implies that for $t > 0$,

(5.25)

$$\begin{aligned}
&\int_{\overline{B_{tD}(x)} - \{x\}} \rho_t \log(\rho_t) \frac{d\nu}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \\
&\geq \left(\int_{\overline{B_{tD}(x)} - \{x\}} \rho_t \frac{d\nu}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \right) \cdot \log \left(\int_{\overline{B_{tD}(x)} - \{x\}} \rho_t \frac{d\nu}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \right) \\
&= \frac{1}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \log \left(\frac{1}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \right).
\end{aligned}$$

Then

$$\begin{aligned}
H_{\infty,\nu}(\mu_t) &= \int_X \rho_t \log(\rho_t)\, d\nu = \int_{\overline{B_{tD}(x)} - \{x\}} \rho_t \log(\rho_t)\, d\nu \\
&\geq \log \left(\frac{1}{\nu \left(\overline{B_{tD}(x)} - \{x\} \right)} \right).
\end{aligned} \tag{5.26}$$

As $\lim_{t \to 0^+} \nu \left(\overline{B_{tD}(x)} \right) = \nu(\{x\})$, we obtain a contradiction to (5.24) when t is small. □

We now prove a Bishop-Gromov-type inequality.

PROPOSITION 5.27. *Let (X, d, ν) be a compact measured length space. Assume that $H_{N,\nu}$ is weakly displacement convex on $P_2(X)$, for some $N \in (1, \infty)$. Then for all $x \in \mathrm{supp}(\nu)$ and all $0 < r_1 \leq r_2$,*

$$\nu(B_{r_2}(x)) \leq \left(\frac{r_2}{r_1} \right)^N \nu(B_{r_1}(x)). \tag{5.28}$$

Proof. From Proposition 5.20 we may assume that ν is nonatomic, as the theorem is trivially true when $\nu = \delta_x$. Put $\mu_0 = \delta_x$ and $\mu_1 = \frac{1_{B_{r_2}(x)}}{\nu(B_{r_2}(x))}\nu$. By the hypothesis and Proposition 2.10, there is a displacement interpolation $\{\mu_t\}_{t\in[0,1]}$ from μ_0 to μ_1 along which (3.16) is satisfied with $U_\nu = H_{N,\nu}$ and $\lambda = 0$. Now $H_{N,\nu}(\mu_0) = N$ and $H_{N,\nu}(\mu_1) = N - N(\nu(B_{r_2}(x)))^{1/N}$. Hence

$$H_{N,\nu}(\mu_t) \leq N - tN(\nu(B_{r_2}(x)))^{1/N}. \tag{5.29}$$

Let $\mu_t = \rho_t\nu + (\mu_t)_s$ be the Lebesgue decomposition of μ_t with respect to ν. As we have a displacement interpolation, ρ_t vanishes outside of $B_{tr_2}(x)$. Then from Hölder's inequality,

$$H_{N,\nu}(\mu_t) \geq N - N(\nu(B_{tr_2}(x)))^{1/N}. \tag{5.30}$$

The theorem follows by taking $t = \frac{r_1}{r_2}$. □

THEOREM 5.31. *If a compact measured length space (X, d, ν) has nonnegative N-Ricci curvature for some $N \in [1,\infty)$ then for all $x \in \operatorname{supp}(\nu)$ and all $0 < r_1 \leq r_2$,*

$$\nu(B_{r_2}(x)) \leq \left(\frac{r_2}{r_1}\right)^N \nu(B_{r_1}(x)). \tag{5.32}$$

Proof. If $N \in (1,\infty)$ then the theorem follows from Proposition 5.27. If $N = 1$ then (X, d, ν) has nonnegative N'-Ricci curvature for all $N' \in (1,\infty)$. The theorem now follows by replacing N in (5.32) by N' and taking $N' \to 1$. □

COROLLARY 5.33. *Given $N \in [1,\infty)$ and $D \geq 0$, the space of compact measured length spaces (X, d, ν) with nonnegative N-Ricci curvature*, $\operatorname{diam}(X, d) \leq D$ *and* $\operatorname{supp}(\nu) = X$ *is sequentially compact in the measured Gromov-Hausdorff topology.*

Proof. Let $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ be a sequence of such spaces. Using the Bishop-Gromov inequality of Theorem 5.31, along with the fact that $\operatorname{supp}(\nu_i) = X_i$, it follows as in [23, Th. 5.3] that after passing to a subsequence we may assume that $\{(X_i, d_i)\}_{i=1}^\infty$ converges in the Gromov-Hausdorff topology to a compact length space (X, d), necessarily with $\operatorname{diam}(X, d) \leq D$. Let $f_i : X_i \to X$ be Borel ε_i-approximations, with $\lim_{i\to\infty} \varepsilon_i = 0$. From the compactness of $P_2(X)$, after passing to a subsequence we may assume that $\lim_{i\to\infty}(f_i)_*\nu_i = \nu$ for some $\nu \in P_2(X)$. From Theorem 5.19, (X, d, ν) has nonnegative N-Ricci curvature.

It remains to show that $\operatorname{supp}(\nu) = X$. Given $x \in X$, the measured Gromov-Hausdorff convergence of $\{(X_i, d_i, \nu_i)\}_{i=1}^\infty$ to (X, d, ν) implies that there is a sequence of points $x_i \in X_i$ with $\lim_{i\to\infty} f_i(x_i) = x$ so that for all $r > 0$ and $\varepsilon \in (0, r)$, we have $\limsup_{i\to\infty} \nu_i(\overline{B_{r-\varepsilon}(x_i)}) \leq \nu(B_r(x))$. By Theorem 5.31,

$(r-\varepsilon)^{-N}\nu_i(B_{r-\varepsilon}(x_i)) \geq \mathrm{diam}(X_i,d_i)^{-N}$. Then $\nu(B_r(x)) \geq \left(\frac{r}{\mathrm{diam}(X,d)}\right)^N$, which proves the claim. □

Remark 5.34. Corollary 5.33 shows that it is consistent in some sense to restrict to the case $\mathrm{supp}(\nu) = X$, at least when N is finite; see also Theorem 5.53. The analog of Corollary 5.33 does not hold in the case $N = \infty$, as can be seen by taking $X = [-1,1]$, $d(x,y) = |x-y|$, $\nu = \frac{e^{-tx^2}dx}{\int_{-1}^{1} e^{-tx^2}dx}$ and letting t go to infinity.

5.5. *Compact group actions.* In this section we show that in certain cases, lower Ricci curvature bounds are preserved upon quotienting by a compact group action.

THEOREM 5.35. *Let (X,d,ν) be a compact measured length space. Suppose that any two $\mu_0,\mu_1 \in P_2^{\mathrm{ac}}(X,\nu)$ are joined by a unique Wasserstein geodesic, that lies in $P_2^{\mathrm{ac}}(X,\nu)$. Suppose that a compact topological group G acts continuously and isometrically on X, preserving ν. Let $p : X \to X/G$ be the quotient map and let $d_{X/G}$ be the quotient metric. We have the following implications*:

a. *For $N \in [1,\infty)$, if (X,d,ν) has nonnegative N-Ricci curvature then $(X/G, d_{X/G}, p_*\nu)$ has nonnegative N-Ricci curvature.*
b. *If (X,d,ν) has ∞-Ricci curvature bounded below by K then $(X/G, d_{X/G}, p_*\nu)$ has ∞-Ricci curvature bounded below by K.*

The proof of this theorem will be an easy consequence of the following lemma, which does not involve the length space structure.

LEMMA 5.36. *The map $p_* : P_2(X) \to P_2(X/G)$ restricts to an isometric isomorphism between the set $P_2(X)^G$ of G-invariant elements in $P_2(X)$, and $P_2(X/G)$.*

Proof. Let dh be the normalized Haar measure on G. The map p_* : $P_2(X) \to P_2(X/G)$ restricts to an isomorphism $p_* : P_2(X)^G \to P_2(X/G)$; the problem is to show that it is an isometry.

Let $\widetilde{\pi}$ be a transference plan between $\widetilde{\mu}_0, \widetilde{\mu}_1 \in P_2(X)^G$. Then $\widetilde{\pi}' = \int_G g \cdot \widetilde{\pi}\, dh(g)$ is also a transference plan between $\widetilde{\mu}_0$ and $\widetilde{\mu}_1$, with

$$\int_{X\times X} d_X(\widetilde{x},\widetilde{y})^2 d\widetilde{\pi}'(\widetilde{x},\widetilde{y}) = \int_G \int_{X\times X} d_X(\widetilde{x}g,\widetilde{y}g)^2 d\pi(\widetilde{x},\widetilde{y}) dh(g) = \int_{X\times X} d_X(\widetilde{x},\widetilde{y})^2 d\pi(\widetilde{x},\widetilde{y}). \tag{5.37}$$

Thus there is a G-invariant optimal transference plan $\widetilde{\pi}$ between $\widetilde{\mu}_0$ and $\widetilde{\mu}_1$. As $\pi = (p \times p)_* \widetilde{\pi}$ is a transference plan between $p_* \widetilde{\mu}_0$ and $p_* \widetilde{\mu}_1$, with

$$\begin{aligned}(5.38) \quad \int_{(X/G)\times(X/G)} d_{X/G}(x,y)^2 d\pi(x,y) &= \int_{X \times X} d_{X/G}(p(\widetilde{x}), p(\widetilde{y}))^2 d\widetilde{\pi}(\widetilde{x}, \widetilde{y}) \\ &\leq \int_{X \times X} d_X(\widetilde{x}, \widetilde{y})^2 d\widetilde{\pi}(\widetilde{x}, \widetilde{y}),\end{aligned}$$

it follows that the map $p_* : P_2(X)^G \to P_2(X/G)$ is metrically nonincreasing.

Conversely, let $s : (X/G) \times (X/G) \to X \times X$ be a Borel map such that $(p \times p) \circ s = \mathrm{Id}$ and $d_X \circ s = d_{X/G}$. That is, given $x, y \in X/G$, the map s picks points $\widetilde{x} \in p^{-1}(x)$ and $\widetilde{y} \in p^{-1}(y)$ in the corresponding orbits so that the distance between $\widetilde{x}$ and $\widetilde{y}$ is minimized among all pairs of points in $p^{-1}(x) \times p^{-1}(y)$. (The existence of s follows from applying [46, Cor. A.6] to the restriction of $p \times p$ to $\{(\widetilde{x}, \widetilde{y}) \in X \times X : d_X(\widetilde{x}, \widetilde{y}) = d_{X/G}(p(\widetilde{x}), p(\widetilde{y}))\}$. The restriction map is a surjective Borel map with compact preimages.) Given an optimal transference plan π between $\mu_0, \mu_1 \in P_2(X/G)$, define a measure $\widetilde{\pi}$ on $X \times X$ by saying that for all $\widetilde{F} \in C(X \times X)$,

$$(5.39) \qquad \int_{X \times X} \widetilde{F} d\widetilde{\pi} = \int_G \int_{(X/G)\times(X/G)} \widetilde{F}(s(x,y) \cdot (g,g)) d\pi(x,y) dh(g).$$

Then for $F \in C((X/G) \times (X/G))$,

$$\begin{aligned}(5.40) \quad & \int_{(X/G)\times(X/G)} F d(p \times p)_* \widetilde{\pi} \\ &= \int_{X \times X} (p \times p)^* F d\widetilde{\pi} \\ &= \int_G \int_{(X/G)\times(X/G)} ((p \times p)^* F)\,(s(x,y) \cdot (g,g)) d\pi(x,y) dh(g) \\ &= \int_G \int_{(X/G)\times(X/G)} F\left((p \times p)(s(x,y) \cdot (g,g))\right) d\pi(x,y) dh(g) \\ &= \int_{(X/G)\times(X/G)} F(x,y) d\pi(x,y).\end{aligned}$$

Thus $(p \times p)_* \widetilde{\pi} = \pi$. As $\widetilde{\pi}$ is G-invariant, it is a transference plan between $(p_*)^{-1}(\mu_0), (p_*)^{-1}(\mu_1) \in P_2(X)^G$. Now

$$\begin{aligned}(5.41) \quad \int_{X \times X} d_X(\widetilde{x}, \widetilde{y})^2 d\widetilde{\pi}(\widetilde{x}, \widetilde{y}) &= \int_G \int_{(X/G)\times(X/G)} d_X(s(x,y) \cdot (g,g))^2 d\pi(x,y) dh(g) \\ &= \int_{(X/G)\times(X/G)} d_{X/G}(x,y)^2 d\pi(x,y).\end{aligned}$$

Thus p_* and $(p_*)^{-1}$ are metrically nonincreasing, which shows that p_* defines an isometric isomorphism between $P_2(X)^G$ and $P_2(X/G)$. $\square$

Proof of Theorem 5.35. The proofs of parts a. and b. of the theorem are similar, so we will be content with proving just part a.

First, $(X/G, d_{X/G})$ is a length space. (Given $x, y \in X/G$, let $\widetilde{x} \in p^{-1}(x)$ and $\widetilde{y} \in p^{-1}(y)$ satisfy $d_X(\widetilde{x}, \widetilde{y}) = d_{X/G}(x, y)$. If c is a geodesic from $\widetilde{x}$ to $\widetilde{y}$ then $p \circ c$ is a geodesic from x to y.)

Given $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(X/G, p_*\nu)$, write $\mu_0 = \rho_0 p_* \nu$ and $\mu_1 = \rho_1 p_* \nu$. Put $\widetilde{\mu}_0 = (p^*\rho_0)\nu$ and $\widetilde{\mu}_1 = (p^*\rho_1)\nu$. From Lemma 5.36, $W_2(\widetilde{\mu}_0, \widetilde{\mu}_1) = W_2(\mu_0, \mu_1)$. By hypothesis, there is a Wasserstein geodesic $\{\widetilde{\mu}_t\}_{t\in[0,1]}$ from $\widetilde{\mu}_0$ to $\widetilde{\mu}_1$ so that for all $U \in \mathcal{DC}_N$, equation (3.16) is satisfied along $\{\widetilde{\mu}_t\}_{t\in[0,1]}$, with $\lambda = 0$. The geodesic $\{\widetilde{\mu}_t\}_{t\in[0,1]}$ is G-invariant, as otherwise by applying an appropriate element of G we would obtain two distinct Wasserstein geodesics between $\widetilde{\mu}_0$ and $\widetilde{\mu}_1$. Put $\mu_t = p_*\widetilde{\mu}_t$. It follows from the above discussion that $\{\mu_t\}_{t\in[0,1]}$ is a curve with length $W_2(\mu_0, \mu_1)$, and so is a Wasserstein geodesic. As $\widetilde{\mu}_t \in P_2^{\mathrm{ac}}(X, \nu)$, we have $\mu_t \in P_2^{\mathrm{ac}}(X/G, p_*\nu)$. Write $\mu_t = \rho_t p_* \nu$. Then $\widetilde{\mu}_t = (p^*\rho_t)\nu$. As

(5.42)
$$U_{p_*\nu}(\mu_t) = \int_{X/G} U(\rho_t) dp_*\nu = \int_X p^* U(\rho_t)\, d\nu = \int_X U(p^*\rho_t)\, d\nu = U_\nu(\widetilde{\mu}_t),$$

equation (3.16) is satisfied along $\{\mu_t\}_{t\in[0,1]}$, with $\lambda = 0$. Along with Proposition 3.21, this concludes the proof of part a. $\square$

5.6. *Uniform integrability and absolute continuity.* In what has been done so far, it would be logically consistent to make our definition of nonnegative N-Ricci curvature to mean weak displacement convexity of just $H_{N,\nu}$, and not necessarily all of $\mathcal{DC}_N$. The reasons to require weak displacement convexity for $\mathcal{DC}_N$ are first that we can, in the sense of being consistent with the classical definitions in the case of a Riemannian manifold, and second that we thereby obtain a useful absolute continuity property for the measures appearing in a Wasserstein geodesic joining two absolutely continuous measures. This last property will feed into Proposition 3.36, when proving Theorem 6.1.

Lemma 5.43. *Let $\{\mu_i\}_{i=1}^m$ be a finite subset of $P_2^{\mathrm{ac}}(X, \nu)$, with densities $\rho_i = \frac{d\mu_i}{d\nu}$. If $N < \infty$ then there is a function $U \in \mathcal{DC}_N$ such that*

$$\lim_{r\to\infty} \frac{U(r)}{r} = \infty \tag{5.44}$$

and

$$\sup_{1\le i\le m} \int_X U(\rho_i(x))\, d\nu(x) < \infty. \tag{5.45}$$

Proof. As a special case of the Dunford-Pettis theorem [13, Th. 2.12], there is an increasing function $\Phi : (0, \infty) \to \mathbb{R}$ such that

$$\lim_{r\to\infty} \frac{\Phi(r)}{r} = \infty \tag{5.46}$$

and

$$\sup_{1 \le i \le m} \int_X \Phi(\rho_i(x))\, d\nu(x) < \infty. \tag{5.47}$$

We may assume that Φ is identically zero on $[0, 1]$.

Consider the function $\psi : (0, \infty) \to \mathbb{R}$ given by

$$\psi(\lambda) = \lambda^N \Phi(\lambda^{-N}). \tag{5.48}$$

Then $\psi \equiv 0$ on $[1, \infty)$, and $\lim_{\lambda \to 0^+} \psi(\lambda) = \infty$. Let $\widetilde{\psi}$ be the lower convex hull of ψ on $(0, \infty)$, i.e. the supremum of the linear functions bounded above by ψ. Then $\widetilde{\psi} \equiv 0$ on $[1, \infty)$ and $\widetilde{\psi}$ is nonincreasing. We claim that $\lim_{\lambda \to 0^+} \widetilde{\psi}(\lambda) = \infty$. If not, suppose that $\lim_{\lambda \to 0^+} \widetilde{\psi}(\lambda) = M < \infty$. Let $a = \sup_{\lambda \ge 0} \frac{M+1-\psi(\lambda)}{\lambda} < \infty$ (because this quantity is ≤ 0 when λ is small enough). Then $\psi(\lambda) \ge M + 1 - a\lambda$, so $\lim_{\lambda \to 0^+} \widetilde{\psi}(\lambda) \ge M + 1$, which is a contradiction.

Now set

$$U(r) = r\widetilde{\psi}(r^{-1/N}). \tag{5.49}$$

Since $\widetilde{\psi} \le \psi$ and $\Phi(r) = r\psi(r^{-1/N})$, we see that $U \le \Phi$. Hence

$$\sup_{1 \le i \le m} \int_X U(\rho_i(x))\, d\nu(x) < \infty. \tag{5.50}$$

Since $\lim_{\lambda \to 0^+} \widetilde{\psi}(\lambda) = \infty$, we also know that

$$\lim_{r \to \infty} \frac{U(r)}{r} = \infty. \tag{5.51}$$

Clearly U is continuous with $U(0) = 0$. As $\widetilde{\psi}$ is convex and nonincreasing, it follows that U is convex. Hence $U \in \mathcal{DC}_N$. $\square$

THEOREM 5.52. *If (X, d, ν) has nonnegative N-Ricci curvature for some $N \in [1, \infty)$ then $P_2^{\mathrm{ac}}(X, \nu)$ is a convex subset of $P_2(X)$.*

Proof. Given $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(X, \nu)$, put $\rho_0 = \frac{d\mu_0}{d\nu}$ and $\rho_1 = \frac{d\mu_1}{d\nu}$. By Lemma 5.43, there is a $U \in \mathcal{DC}_N$ with $U'(\infty) = \infty$ such that $U_\nu(\mu_0) < \infty$ and $U_\nu(\mu_1) < \infty$. As (X, d, ν) has nonnegative N-Ricci curvature, there is a Wasserstein geodesic $\{\mu_t\}_{t \in [0,1]}$ from μ_0 to μ_1 so that (3.16) is satisfied with $\lambda = 0$. In particular, $U_\nu(\mu_t) < \infty$ for all $t \in [0, 1]$. As $U'(\infty) = \infty$, it follows that $\mu_t \in P_2^{\mathrm{ac}}(X, \nu)$ for each t. $\square$

We now clarify the relationship between (X, d, ν) having nonnegative N-Ricci curvature and the analogous statement for $\mathrm{supp}(\nu)$. We recall the notion of a subset $A \subset X$ being convex or totally convex, from Section 1.2.2, and we note that $d\big|_A$ defines a length space structure on a closed subset A if and only if A is convex in X.

THEOREM 5.53. a. *Given* $N \in [1, \infty)$, *suppose that a compact measured length space* (X, d, ν) *has nonnegative* N*-Ricci curvature. Then* $\operatorname{supp}(\nu)$ *is a convex subset of* X (*although not necessarily totally convex*) *and* $(\operatorname{supp}(\nu), d|_{\operatorname{supp}(\nu)}, \nu)$ *has nonnegative* N*-Ricci curvature. Conversely, if* $\operatorname{supp}(\nu)$ *is a convex subset of* X *and* $(\operatorname{supp}(\nu), d|_{\operatorname{supp}(\nu)}, \nu)$ *has nonnegative* N*-Ricci curvature then* (X, d, ν) *has nonnegative* N*-Ricci curvature.*

b. *Given* $K \in \mathbb{R}$, *the analogous statement holds when one replaces "nonnegative* N*-Ricci curvature" by "*∞*-Ricci curvature bounded below by* K*".*

Proof. a. Let (X, d, ν) be a compact measured length space with nonnegative N-Ricci curvature. Let μ_0 and μ_1 be elements of $P_2(X, \nu)$. By Theorem C.12 in Appendix C, there are sequences $\{\mu_{k,0}\}_{k=1}^\infty$ and $\{\mu_{k,1}\}_{k=1}^\infty$ in $P_2^{\mathrm{ac}}(X, \nu)$ (in fact with continuous densities) such that $\lim_{k\to\infty} \mu_{k,0} = \mu_0$, $\lim_{k\to\infty} \mu_{k,1} = \mu_1$ and for all $U \in \mathcal{DC}_N$, $\lim_{k\to\infty} U_\nu(\mu_{k,0}) = U_\nu(\mu_0)$ and $\lim_{k\to\infty} U_\nu(\mu_{k,1}) = U_\nu(\mu_1)$. From the definition of nonnegative N-Ricci, for each k there is a Wasserstein geodesic $\{\mu_{k,t}\}_{t\in[0,1]}$ such that

$$U_\nu(\mu_{k,t}) \le t\, U_\nu(\mu_{k,1}) + (1-t)\, U_\nu(\mu_{k,0}) \tag{5.54}$$

for all $U \in \mathcal{DC}_N$ and $t \in [0,1]$. By repeating the proof of Theorem 5.52, each $\mu_{k,t}$ is absolutely continuous with respect to ν. In particular, it is supported in $\operatorname{supp}(\nu)$. By the same reasoning as in the proof of Proposition 3.21, after passing to a subsequence we may assume that as $k \to \infty$, the geodesics $\{\mu_{k,t}\}_{t\in[0,1]}$ converge uniformly to a Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ that satisfies

$$U_\nu(\mu_t) \le t\, U_\nu(\mu_1) + (1-t)\, U_\nu(\mu_0). \tag{5.55}$$

For each $t \in [0,1]$, the measure μ_t is the weak-$*$ limit of the probability measures $\{\mu_{k,t}\}_{k=1}^\infty$, which are all supported in the closed set $\operatorname{supp}(\nu)$. Hence μ_t is also supported in $\operatorname{supp}(\nu)$. To summarize, we have shown that $\{\mu_t\}_{t\in[0,1]}$ is a Wasserstein geodesic lying in $P_2(X, \nu)$ that satisfies (5.55) for all $U \in \mathcal{DC}_N$ and $t \in [0,1]$.

We now check that $\operatorname{supp}(\nu)$ is convex. Let x_0 and x_1 be any two points in $\operatorname{supp}(\nu)$. Applying the reasoning above to $\mu_0 = \delta_{x_0}$ and $\mu_1 = \delta_{x_1}$, one obtains the existence of a Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ joining δ_{x_0} to δ_{x_1} such that each μ_t is supported in $\operatorname{supp}(\nu)$. By Proposition 2.10, there is an optimal dynamical transference plan $\Pi \in P(\Gamma)$ such that $\mu_t = (e_t)_*\Pi$ for all $t \in [0,1]$. For each $t \in [0,1]$, we know that $\gamma(t) \in \operatorname{supp}(\nu)$ holds for Π-almost all γ. It follows that for Π-almost all γ, we have $\gamma(t) \in \operatorname{supp}(\nu)$ for all $t \in \mathbb{Q} \cap [0,1]$. As $\gamma \in \Gamma$ is continuous, this is the same as saying that for Π-almost all γ, the geodesic γ is entirely contained in $\operatorname{supp}(\nu)$. Also, for Π-almost all γ we have $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Thus x_0 and x_1 are indeed joined by a geodesic path contained in $\operatorname{supp}(\nu)$.

This proves the direct implication in part a. The converse is immediate.

b. The proof of part b. follows the same lines as that of part a. We construct the approximants $\{\mu_{k,0}\}_{k=1}^{\infty}$ and $\{\mu_{k,1}\}_{k=1}^{\infty}$, with continuous densities, and the geodesics $\{\mu_{k,t}\}_{t\in[0,1]}$. As $H_{\infty,\nu}(\mu_{0,k}) < \infty$ and $H_{\infty,\nu}(\mu_{1,k}) < \infty$, we can apply inequality (3.16) with $U = H_\infty$ and $\lambda = K$, to deduce that $H_{\infty,\nu}(\mu_{t,k}) < \infty$ for all $t \in [0,1]$. This implies that $\mu_{t,k}$ is absolutely continuous with respect to ν. The rest of the argument is similar to that of part a. □

Remark 5.56. Corollary 5.33 and Theorem 5.53.a together show that in the case $N < \infty$, we do not lose much by assuming that $X = \mathrm{supp}(\nu)$.

6. Log Sobolev, Talagrand and Poincaré inequalities

In this section we study several functional inequalities with geometric content that are associated to optimal transport and concentration of measure: log Sobolev inequalities, Talagrand inequalities and Poincaré inequalities. We refer to [4] and [44, Ch. 9] for concise surveys about previous work on these inequalities.

We first write some general functional inequalities. In the case of ∞-Ricci curvature bounded below by K, we make explicit the ensuing log Sobolev inequalities, Talagrand inequalities and Poincaré inequalities. We then write out explicit functional inequalities in the case of nonnegative N-Ricci curvature. Finally, we prove a weak Bonnet-Myers theorem, following [36, §6].

6.1. *The general inequalities.* We recall the generalized Fisher information of (3.53), where $\rho \in \mathrm{Lip}(X)$ is positive and $\mu = \rho\nu$ is the corresponding measure.

THEOREM 6.1. *Suppose that (X, d, ν) has ∞-Ricci curvature bounded below by $K > 0$. Then for all $\mu \in P_2(X, \nu)$,*

$$\frac{K}{2} W_2(\mu,\nu)^2 \le H_{\infty,\nu}(\mu). \tag{6.2}$$

If now $\mu \in P_2^{\mathrm{ac}}(X,\nu)$ and its density $\rho = \frac{d\mu}{d\nu}$ is a positive Lipschitz function on X then

$$H_{\infty,\nu}(\mu) \le W_2(\mu,\nu)\sqrt{I_{\infty,\nu}(\mu)} - \frac{K}{2} W_2(\mu,\nu)^2 \le \frac{1}{2K} I_{\infty,\nu}(\mu). \tag{6.3}$$

If on the other hand (X, d, ν) has ∞-Ricci curvature bounded below by $K \le 0$ then

$$H_{\infty,\nu}(\mu) \le \mathrm{diam}(X)\sqrt{I_{\infty,\nu}(\mu)} - \frac{K}{2}\,\mathrm{diam}(X)^2. \tag{6.4}$$

If (X, d, ν) has nonnegative N-Ricci curvature then

$$H_{N,\nu}(\mu) \le \mathrm{diam}(X)\sqrt{I_{N,\nu}(\mu)}. \tag{6.5}$$

Proof. We wish to apply Proposition 3.36 to the cases described in particular cases 3.52. Under the assumption that $U_\nu(\mu) < \infty$, we have to show that there is a Wasserstein geodesic as in the statement of Proposition 3.36 with $\mu_t \in P_2^{\mathrm{ac}}(X, \nu)$ for all $t \in [0,1]$. If $N = \infty$ then there is some Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ from μ to ν which in particular satisfies equation (3.16) with $U_\nu = H_{\infty,\nu}$ and $\lambda = K$. Hence $H_{\infty,\nu}(\mu_t) < \infty$ for all $t \in [0,1]$ and the claim follows from the fact that $U'_\infty(\infty) = \infty$. If $N \in [1,\infty)$ then the claim follows from Theorem 5.52. □

We now express the conclusion of Theorem 6.1 in terms of more standard inequalities, starting with the case $N = \infty$.

6.2. *The case $N = \infty$.*

Definition 6.6. Suppose that $K > 0$.

• We say that ν satisfies a *log Sobolev inequality* with constant K, LSI(K), if for all $\mu \in P_2^{\mathrm{ac}}(X, \nu)$ whose density $\rho = \frac{d\mu}{d\nu}$ is Lipschitz and positive, we have

$$H_{\infty,\nu}(\mu) \leq \frac{1}{2K}\, I_{\infty,\nu}(\mu). \tag{6.7}$$

• We say that ν satisfies a *Talagrand inequality* with constant K, T(K), if for all $\mu \in P_2(X, \nu)$,

$$W_2(\mu, \nu) \leq \sqrt{\frac{2H_{\infty,\nu}(\mu)}{K}}. \tag{6.8}$$

• We say that ν satisfies a *Poincaré inequality* with constant K, P(K), if for all $h \in \mathrm{Lip}(X)$ with $\int_X h\, d\nu = 0$, we have

$$\int_X h^2\, d\nu \leq \frac{1}{K} \int_X |\nabla^- h|^2\, d\nu. \tag{6.9}$$

Remark 6.10. Here we used the gradient norm defined in (1.4), instead of the one defined in (1.3). Accordingly, our log Sobolev inequality and Poincaré inequalities are slightly stronger statements than those discussed by some other authors.

All of these inequalities are associated with concentration of measure [4], [8], [9], [26], [27]. For example, T(K) implies a Gaussian-type concentration of measure. The following chain of implications, none of which is an equivalence, is well-known in the context of smooth Riemannian manifolds:

$$[\mathrm{Ric} \geq K] \Longrightarrow \mathrm{LSI}(K) \Longrightarrow \mathrm{T}(K) \Longrightarrow \mathrm{P}(K). \tag{6.11}$$

In the context of length spaces, we see from Theorem 6.1 that having ∞-Ricci curvature bounded below by $K > 0$ implies LSI(K) and T(K). The next corollary makes the statement of the log Sobolev inequality more explicit.

COROLLARY 6.12. *Suppose that* (X, d, ν) *has* ∞*-Ricci curvature bounded below by* $K \in \mathbb{R}$. *If* $f \in \mathrm{Lip}(X)$ *satisfies* $\int_X f^2 \, d\nu = 1$ *then*

$$\int_X f^2 \log(f^2) \, d\nu \leq 2W_2(f^2\nu, \nu) \sqrt{\int_X |\nabla^- f|^2 \, d\nu} - \frac{K}{2} W_2(f^2\nu, \nu)^2. \tag{6.13}$$

In particular, if $K > 0$ *then*

$$\int_X f^2 \log(f^2) \, d\nu \leq \frac{2}{K} \int_X |\nabla^- f|^2 \, d\nu, \tag{6.14}$$

while if $K \leq 0$ *then*

$$\int_X f^2 \log(f^2) \, d\nu \leq 2 \operatorname{diam}(X) \sqrt{\int_X |\nabla^- f|^2 \, d\nu} - \frac{K}{2} \operatorname{diam}(X)^2. \tag{6.15}$$

Proof. For any $\varepsilon > 0$, put $\rho_\varepsilon = \frac{f^2+\varepsilon}{1+\varepsilon}$. From Theorem 6.1,

$$\int_X \rho_\varepsilon \log(\rho_\varepsilon) \, d\nu \leq W_2(\rho_\varepsilon\nu, \nu) \sqrt{\int_X \frac{|\nabla^- \rho_\varepsilon|^2}{\rho_\varepsilon} \, d\nu} - \frac{K}{2} W_2(\rho_\varepsilon\nu, \nu)^2. \tag{6.16}$$

As

$$\frac{|\nabla^- \rho_\varepsilon|^2}{\rho_\varepsilon} = \frac{1}{1+\varepsilon} \frac{4f^2}{f^2+\varepsilon} |\nabla^- f|^2, \tag{6.17}$$

the corollary follows by taking $\varepsilon \to 0$. □

We now recall the standard fact that LSI(K) implies P(K).

THEOREM 6.18. *Let* (X, d, ν) *be a compact measured length space satisfying* LSI(K) *for some* $K > 0$. *Then it also satisfies* P(K).

Proof. Suppose that $h \in \mathrm{Lip}(X)$ satisfies $\int_X h \, d\nu = 0$. For $\varepsilon \in [0, \frac{1}{\|h\|_\infty})$, put $f_\varepsilon = \sqrt{1 + \varepsilon h} > 0$. As $2 f_\varepsilon \nabla^- f_\varepsilon = \varepsilon \nabla^- h$, it follows that

$$\lim_{\varepsilon \to 0^+} \left(\frac{1}{\varepsilon^2} \int_X |\nabla^- f_\varepsilon|^2 \, d\nu \right) = \frac{1}{4} \int_X |\nabla^- h|^2 \, d\nu. \tag{6.19}$$

As the Taylor expansion of $x \log(x) - x + 1$ around $x = 1$ is $\frac{1}{2}(x-1)^2 + \ldots$,

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon^2} \int_X f_\varepsilon^2 \log(f_\varepsilon^2) \, d\nu = \frac{1}{2} \int_X h^2 \, d\nu. \tag{6.20}$$

Then the conclusion follows from (6.14). □

Remark 6.21. If $Q(h) = \int_X |\nabla h|^2 \, d\nu$ defines a quadratic form on $\mathrm{Lip}(X)$, which in addition is closable in $L^2(X, \nu)$, then there is a (nonpositive) self-adjoint Laplacian $\triangle_\nu$ associated to Q. In this case, P(K) implies that $-\triangle_\nu \geq K$ on the orthogonal complement of the constant functions.

We do not claim to show that there are such Laplacians on (X, d, ν) in general. In the case of a limit space arising from a sequence of manifolds with Ricci curvature bounded below, Cheeger and Colding used additional structure on the limit space in order to show the Laplacian does exist [18].

As mentioned above, in the case of smooth Riemannian manifolds there are stronger implications: $\mathrm{T}(K)$ implies $\mathrm{P}(K)$, and $\mathrm{LSI}(K)$ implies $\mathrm{T}(K)$. We will show elsewhere that the former is always true, while the latter is true under the additional assumption of a lower bound on the Alexandrov curvature:

THEOREM 6.22. *Let (X, d, ν) be a compact measured length space.*

(i) *If ν satisfies $\mathrm{T}(K)$ for some $K > 0$, then it also satisfies $\mathrm{P}(K)$.*

(ii) *If X is a finite-dimensional Alexandrov space with Alexandrov curvature bounded below, and ν satisfies $\mathrm{LSI}(K)$ for some $K > 0$, then it also satisfies $\mathrm{T}(K)$.*

Remark 6.23. The Alexandrov curvature bound in (ii) essentially serves as a regularity assumption. One can ask whether it can be weakened.

Remark 6.24. We have only discussed *global* Poincaré inequalities. There is also a notion of a metric-measure space admitting a *local* Poincaré inequality, as considered for example in [18]. If a measured length space (X, d, ν) has nonnegative N-Ricci curvature, with $N < \infty$, then it admits a local Poincaré inequality, at least if one assumes almost-everywhere uniqueness of geodesics. We will discuss this in detail elsewhere.

6.3. *The case $N < \infty$.* We now write an analog of Corollary 6.12 in the case $N < \infty$. Suppose that (X, d, ν) has nonnegative N-Ricci curvature. Then if ρ is a positive Lipschitz function on X, (6.5) says that

$$N - N \int_X \rho^{1-\frac{1}{N}} \, d\nu \leq \frac{N-1}{N} \operatorname{diam}(X) \sqrt{\int_X \frac{|\nabla^- \rho|^2}{\rho^{\frac{2}{N}+1}} \, d\nu}. \tag{6.25}$$

If $N > 2$, put $f = \rho^{\frac{N-2}{2N}}$. Then $\int_X f^{\frac{2N}{N-2}} \, d\nu = 1$ and one finds that (6.25) is equivalent to

$$1 - \int_X f^{\frac{2(N-1)}{N-2}} \, d\nu \leq \frac{2(N-1)}{N(N-2)} \operatorname{diam}(X) \sqrt{\int_X |\nabla^- f|^2 \, d\nu}. \tag{6.26}$$

As in the proof of Corollary 6.12, equation (6.26) holds for all $f \in \mathrm{Lip}(X)$ satisfying $\int_X f^{\frac{2N}{N-2}} \, d\nu = 1$. From the Hölder inequality,

$$\int_X f^{\frac{2(N-1)}{N-2}} \, d\nu \leq \left(\int_X f \, d\nu \right)^{\frac{2}{N+2}} \left(\int_X f^{\frac{2N}{N-2}} \, d\nu \right)^{\frac{N}{N+2}} = \left(\int_X f \, d\nu \right)^{\frac{2}{N+2}}. \tag{6.27}$$

Then (6.26) implies

$$1 \leq \frac{2(N-1)}{N(N-2)} \operatorname{diam}(X) \sqrt{\int_X |\nabla^- f|^2 \, d\nu} + \left(\int_X f \, d\nu \right)^{\frac{2}{N+2}}. \tag{6.28}$$

Writing (6.28) in a homogeneous form, one sees that its content is as follows: for a function F on X, bounds on $\| \nabla^- F \|_2$ and $\| F \|_1$ imply a bound on $\| F \|_{\frac{2N}{N-2}}$. This is of course an instance of Sobolev embedding.

If $N = 2$, putting $f = \log(\frac{1}{\rho})$, one finds that $\int_X e^{-f} \, d\nu = 1$ implies

$$1 - \int_X e^{-\frac{f}{2}} \, d\nu \leq \frac{1}{4} \operatorname{diam}(X) \sqrt{\int_X |\nabla^- f|^2 \, d\nu}. \tag{6.29}$$

6.4. *Weak Bonnet-Myers theorem.* The classical Bonnet-Myers theorem says that if M is a smooth connected complete N-dimensional Riemannian manifold with $\mathrm{Ric}_M \geq K g_M > 0$, then $\operatorname{diam}(M) \leq \pi \sqrt{\frac{N-1}{K}}$.

We cannot give an immediate generalization of this theorem to a measured length space (X, d, ν), as we have not defined what it means to have N-Ricci curvature bounded below by K for $N < \infty$ and $K > 0$. However, it does make sense to state a weak version of the Bonnet-Myers theorem under the assumptions that (X, d, ν) has nonnegative N-Ricci curvature and has ∞-Ricci curvature bounded below by $K > 0$.

THEOREM 6.30. *There is a constant $C > 0$ with the following property. Let (X, d, ν) be a compact measured length space with nonnegative N-Ricci curvature, and ∞-Ricci curvature bounded below by $K > 0$. Suppose that* $\operatorname{supp}(\nu) = X$. *Then*

$$\operatorname{diam}(X) \leq C \sqrt{\frac{N}{K}}. \tag{6.31}$$

Proof. From Theorem 5.31, ν satisfies the growth estimate

$$\frac{\nu(B_r(x))}{\nu(B_{\alpha r}(x))} \leq \alpha^{-N}, \qquad 0 < \alpha \leq 1. \tag{6.32}$$

From Theorem 6.1, ν satisfies $\mathrm{T}(K)$. The result follows by repeating verbatim the proof of [36, Th. 4] with $R = 0$, $n = N$ and $\rho = K$. $\square$

Remark 6.33. The remark at the end of [36, §6] shows that $C = 7.7$ is admissible.

7. The case of Riemannian manifolds

In this section we look at the case of a smooth Riemannian manifold (M, g) equipped with a smooth measure ν. We define the tensor Ric_N and show the equivalence of lower bounds on Ric_N to various displacement convexity conditions. In particular, we show that the measured length space (M, g, ν) has nonnegative N-Ricci curvature if and only if $\mathrm{Ric}_N \geq 0$, and that it has ∞-Ricci curvature bounded below by K if and only if $\mathrm{Ric}_\infty \geq Kg$.

We use this, along with Theorem 5.19, to characterize measured Ricci limit spaces that happen to be smooth. We give some consequences concerning their metric structure. We then show that for Riemannian manifolds, lower N-Ricci curvature bounds are preserved under taking compact quotients. Finally, we use displacement convexity to give a "synthetic" proof of a part of the Ricci O'Neill theorem from [30].

7.1. *Formulation of N-Ricci curvature in classical terms.* Let (M, g) be a smooth compact connected n-dimensional Riemannian manifold. Let Ric denote its Ricci tensor.

Given $\Psi \in C^\infty(M)$ with $\int_M e^{-\Psi}\,\mathrm{dvol}_M = 1$, put $d\nu = e^{-\Psi}\,\mathrm{dvol}_M$.

Definition 7.1. For $N \in [1, \infty]$, the *N-Ricci tensor* of (M, g, ν) is

$$\mathrm{Ric}_N = \begin{cases} \mathrm{Ric} + \mathrm{Hess}\,(\Psi) & \text{if } N = \infty, \\ \mathrm{Ric} + \mathrm{Hess}\,(\Psi) - \frac{1}{N-n} d\Psi \otimes d\Psi & \text{if } n < N < \infty, \\ \mathrm{Ric} + \mathrm{Hess}\,(\Psi) - \infty (d\Psi \otimes d\Psi) & \text{if } N = n, \\ -\infty & \text{if } N < n, \end{cases} \tag{7.2}$$

where by convention $\infty \cdot 0 = 0$.

The expression for Ric_∞ is the Bakry-Émery tensor [6]. The expression for Ric_N with $n < N < \infty$ was considered in [30], [39]. The statement $\mathrm{Ric}_N \geq Kg$ is equivalent to the statement that the operator $L = \triangle - (\nabla\Psi) \cdot \nabla$ satisfies Bakry's curvature-dimension condition $\mathrm{CD}(K, N)$ [5, Prop. 6.2].

Given $K \in \mathbb{R}$, we recall the definition of $\lambda : \mathcal{DC}_\infty \to \mathbb{R} \cup \{-\infty\}$ from Definition 5.13.

THEOREM 7.3. a. *For $N \in (1, \infty)$, the following are equivalent*:

(1) $\mathrm{Ric}_N \geq 0$.

(2) *The measured length space (M, g, ν) has nonnegative N-Ricci curvature.*

(3) *For all $U \in \mathcal{DC}_N$, U_ν is weakly displacement convex on $P_2(M)$.*

(4) *For all $U \in \mathcal{DC}_N$, U_ν is weakly* a.c. *displacement convex on $P_2^{\mathrm{ac}}(M)$.*

(5) *$H_{N,\nu}$ is weakly* a.c. *displacement convex on $P_2^{\mathrm{ac}}(M)$.*

b. *For any $K \in \mathbb{R}$, the following are equivalent*:

(1) $\mathrm{Ric}_\infty \geq Kg$.

(2) *The measured length space (M, g, ν) has ∞-Ricci curvature bounded below by K.*

(3) *For all* $U \in \mathcal{DC}_\infty$, U_ν *is weakly* $\lambda(U)$*-displacement convex on* $P_2(M)$.

(4) *For all* $U \in \mathcal{DC}_\infty$, U_ν *is weakly* $\lambda(U)$-a.c. *displacement convex on* $P_2^{\mathrm{ac}}(M)$.

(5) $H_{\infty,\nu}$ *is weakly* K-a.c. *displacement convex on* $P_2^{\mathrm{ac}}(M)$.

For both parts (a) and (b), the nontrivial implications are (1) $\Rightarrow$ (2) and (5) $\Rightarrow$ (1). The proof that (1) $\Rightarrow$ (2) will be along the lines of [19, Th. 6.2], with some differences. One ingredient is the following lemma.

Lemma 7.4. *Let* $\phi : M \to \mathbb{R}$ *be a* $\frac{d^2}{2}$*-concave function. We recall that* ϕ *is necessarily Lipschitz and hence* $(\nabla\phi)(y)$ *exists for almost all* $y \in M$. *For such* y, *define*

$$F_t(y) \equiv \exp_y(-t\nabla\phi(y)). \tag{7.5}$$

Assume furthermore that $y \in M$ *is such that*

(i) ϕ *admits a Hessian at* y (*in the sense of Alexandrov*),

(ii) F_t *is differentiable at* y *for all* $t \in [0,1)$ *and*

(iii) $dF_t(y)$ *is nonsingular for all* $t \in [0,1)$.

Then $D(t) \equiv \det^{\frac{1}{n}}(dF_t(y))$ *satisfies the differential inequality*

$$\frac{D''(t)}{D(t)} \leq -\frac{1}{n}\operatorname{Ric}(F_t'(y), F_t'(y)) \qquad t \in (0,1). \tag{7.6}$$

Proof. Let $\{e_i\}_{i=1}^n$ be an orthonormal basis of T_yM. For each i, let $J_i(t)$ be defined by

$$J_i(t) = (dF_t)_y(e_i). \tag{7.7}$$

Then $\{J_i(t)\}_{i=1}^n$ is a Jacobi field with $J_i(0) = e_i$. Next, we note that $d\phi$ is differentiable at y, and that $d(d\phi)_y$ coincides with $\operatorname{Hess}_y(\phi)$, up to identification. This is not so obvious (indeed, the existence of a Hessian only means the existence of a second-order Taylor expansion) but can be shown as a consequence of the semiconcavity of ϕ, as in [19, Prop. 4.1 (b)]. (The case of a convex function in $\mathbb{R}^n$ is treated in [1, Ths. 3.2 and 7.10].) It follows that

$$J_i'(0) = -\operatorname{Hess}(\phi)(y)e_i. \tag{7.8}$$

Let now $W(t)$ be the $(n \times n)$-matrix with

$$W_{ij}(t) = \langle J_i(t), J_j(t)\rangle; \tag{7.9}$$

then $\det^{\frac{1}{n}}(dF_t)(y) = \det^{\frac{1}{2n}} W(t)$.

Since $W(t)$ is nonsingular for $t \in [0,1)$, $\{J_i(t)\}_{i=1}^n$ is a basis of $T_{F_t(y)}M$. Define a matrix $R(t)$ by $J_i'(t) = \sum_j R(t)_i^j J_j(t)$. It follows from the equation

$$\frac{d}{dt}\left(\langle J_i'(t), J_j(t)\rangle - \langle J_i(t), J_j'(t)\rangle\right) = 0 \tag{7.10}$$

and the self-adjointness of Hess $(\phi)(y)$ that $RW - WR^T = 0$ for all $t \in [0,1)$, or equivalently, $R = WR^T W^{-1}$. (More intrinsically, the linear operator on $T_{F_t(y)}M$ defined by R satisfies $R = R^*$, where R^* is the dual defined using the inner product on $T_{F_t(y)}M$.)

Next,

$$W' = RW + WR^T. \tag{7.11}$$

Applying the Jacobi equation to

$$W''_{ij} = \langle J''_i(t), J_j(t)\rangle + \langle J_i(t), J''_j(t)\rangle + 2\langle J'_i(t), J'_j(t)\rangle \tag{7.12}$$

gives

$$W'' = -2\,\mathrm{Riem}(\cdot, F'_t(y), \cdot, F'_t(y)) + 2RWR^T. \tag{7.13}$$

Now

$$\frac{d}{dt}\det^{\frac{1}{2n}} W(t) = \frac{1}{2n}\det^{\frac{1}{2n}} W(t)\,\mathrm{Tr}\left(W'W^{-1}\right) \tag{7.14}$$

and

$$\begin{aligned} \frac{d^2}{dt^2}\det^{\frac{1}{2n}} W(t) = {} & \frac{1}{4n^2}\det^{\frac{1}{2n}} W(t)\left(\mathrm{Tr}\left(W'W^{-1}\right)\right)^2 \\ & - \frac{1}{2n}\det^{\frac{1}{2n}} W(t)\,\mathrm{Tr}\left((W'W^{-1})^2\right) \\ & + \frac{1}{2n}\det^{\frac{1}{2n}} W(t)\,\mathrm{Tr}\left(W''W^{-1}\right). \end{aligned} \tag{7.15}$$

Then by (7.11) and (7.13),

$$D^{-1}\frac{d^2 D}{dt^2} = \frac{1}{n^2}(\mathrm{Tr}(R))^2 - \frac{2}{n}\,\mathrm{Tr}(R^2) - \frac{1}{n}\,\mathrm{Ric}(F'_t(y), F'_t(y)) + \frac{1}{n}\,\mathrm{Tr}(R^2). \tag{7.16}$$

As R is self-adjoint,

$$\frac{1}{n}(\mathrm{Tr}(R))^2 - \mathrm{Tr}(R^2) \le 0, \tag{7.17}$$

from which the conclusion follows. □

Proof of Theorem 7.3, part (a). To show (1) ⇒ (2), suppose that $\mathrm{Ric}_N \ge 0$. By the definition of Ric_N, we must have $n < N$, or $n = N$ and Ψ is constant. Suppose first that $n < N$. We can write

$$\mathrm{Ric}_N = \mathrm{Ric} - (N-n)e^{\frac{\Psi}{N-n}}\,\mathrm{Hess}\left(e^{-\frac{\Psi}{N-n}}\right). \tag{7.18}$$

Given $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(M)$, let $\{\mu_t\}_{t\in[0,1]}$ be the unique Wasserstein geodesic from μ_0 to μ_1. From Proposition 4.28, in order to prove (2) it suffices to show that for all such μ_0 and μ_1, and all $U \in \mathcal{DC}_N$, the inequality (3.16) is satisfied with $\lambda = 0$.

We recall facts from Subsection 2.4 about optimal transport on Riemannian manifolds. In particular, μ_t is absolutely continuous with respect to dvol_M for all t, and takes the form $(F_t)_*\mu_0$, where $F_t(y) = \exp_y(-t\nabla\phi(y))$ for some $\frac{d^2}{2}$-concave function ϕ. Put $\eta_t = \frac{d\mu_t}{\mathrm{dvol}_M}$. Using the nonsmooth change-of-variables formula proven in [19, Cor. 4.7] (see also [31, Th. 4.4]), we can write

(7.19)
$$\begin{aligned} U_\nu(\mu_t) &= \int_M U(e^{\Psi(m)}\eta_t(m))e^{-\Psi(m)}\,\mathrm{dvol}_M(m) \\ &= \int_M U\left(e^{\Psi(F_t(y))}\frac{\eta_0(y)}{\det(dF_t)(y)}\right)e^{-\Psi(F_t(y))}\det(dF_t)(y)\,\mathrm{dvol}_M(y). \end{aligned}$$

Putting

$$C(y,t) = e^{-\frac{\Psi(F_t(y))}{N}}\det{}^{\frac{1}{N}}(dF_t)(y), \tag{7.20}$$

we can write

$$U_\nu(\mu_t) = \int_M C(y,t)^N U\left(\eta_0(y)C(y,t)^{-N}\right)\mathrm{dvol}_M(y). \tag{7.21}$$

Suppose that we can show that $C(y,t)$ is concave in t for almost all $y \in M$. Then for $y \in \mathrm{supp}(\mu_0)$, as the map

$$\lambda \to \eta_0^{-1}(y)\lambda^N U\left(\eta_0(y)\lambda^{-N}\right) \tag{7.22}$$

is nonincreasing and convex, and the composition of a nonincreasing convex function with a concave function is convex, the integrand of (7.21) is convex in t. Hence $U_\nu(\mu_t)$ will be convex in t.

To show that $C(y,t)$ is concave in t, fix y. Put

$$C_1(t) = e^{-\frac{\Psi(F_t(y))}{N-n}} \tag{7.23}$$

and

$$C_2(t) = \det{}^{\frac{1}{n}}(dF_t)(y), \tag{7.24}$$

so that $C(y,t) = C_1(t)^{\frac{N-n}{N}}C_2(t)^{\frac{n}{N}}$. We have

$$\begin{aligned} NC^{-1}\frac{d^2C}{dt^2} &= (N-n)C_1^{-1}\frac{d^2C_1}{dt^2} \\ &\quad + nC_2^{-1}\frac{d^2C_2}{dt^2} - \frac{n(N-n)}{N}\left(C_1^{-1}\frac{dC_1}{dt} - C_2^{-1}\frac{dC_2}{dt}\right)^2 \\ &\le (\mathrm{Ric} - \mathrm{Ric}_N)\left(F_t'(y), F_t'(y)\right) + nC_2^{-1}\frac{d^2C_2}{dt^2}. \end{aligned} \tag{7.25}$$

We may assume that the function ϕ has a Hessian at y [19, Th. 4.2(a)], and that dF_t is well-defined and nonsingular at y for all $t \in [0,1)$ [19, Claim 4.3(a–b)]. Then Lemma 7.4 shows that

$$nC_2^{-1}\frac{d^2C_2}{dt^2} \le -\mathrm{Ric}(F_t'(y), F_t'(y)). \tag{7.26}$$

So $NC^{-1}(t)\,C''(t) \leq -\operatorname{Ric}_N(F_t'(y), F_t'(y)) \leq 0$. This shows that (M, g, ν) is weakly displacement convex for the family $\mathcal{DC}_N$.

The proof in the case $N = n$ follows the same lines, if we replace C_1 by 1 and C_2 by C.

We now prove the implication (5) ⇒ (1). Putting $U = U_N$ in (7.21), we obtain

$$H_{N,\nu}(\mu_t) = N - N \int_M C(y,t)\eta_0(y)^{1-\frac{1}{N}} \operatorname{dvol}_M(y). \tag{7.27}$$

Suppose first that $n < N$ and $H_{N,\nu}$ is weakly a.c. displacement convex. Given $m \in M$ and $v \in T_mM$, we want to show that $\operatorname{Ric}_N(v,v) \geq 0$. Choose a smooth function ϕ, defined in a neighborhood of m, so that $v = -(\nabla\phi)(m)$, $\operatorname{Hess}(\phi)(m)$ is proportionate to $g(m)$ and

$$\frac{1}{N-n} v\Psi = \frac{1}{n}(\triangle\phi)(m). \tag{7.28}$$

Consider the geodesic segment $t \to \exp_m(tv)$. Then

$$C_1^{-1}(0)C_1'(0) = -\frac{1}{N-n} v\Psi \tag{7.29}$$

and

$$\begin{aligned} C_2^{-1}(0)C_2'(0) &= \frac{1}{2n}\operatorname{Tr}(W'(0)W^{-1}(0)) = \frac{1}{n}\operatorname{Tr}(R(0)) \\ &= -\frac{1}{n}\operatorname{Tr}(\operatorname{Hess}(\phi)(m)) = -\frac{1}{n}(\triangle\phi)(m). \end{aligned} \tag{7.30}$$

Hence by construction, $C_1^{-1}(0)C_1'(0) = C_2^{-1}(0)C_2'(0)$. From (7.25),

$$NC^{-1}(0)C''(0) = (\operatorname{Ric} - \operatorname{Ric}_N)(v,v) + nC_2^{-1}(0)C_2''(0). \tag{7.31}$$

As $R(0)$ is a multiple of the identity, (7.16) now implies that

$$NC^{-1}(0)C''(0) = -\operatorname{Ric}_N(v,v). \tag{7.32}$$

For small numbers $\varepsilon_1, \varepsilon_2 > 0$, consider a smooth probability measure μ_0 with support in an ε_1-ball around m. Put $\mu_1 = (F_{\varepsilon_2})_*\mu_0$ where F_t is defined by $F_t(y) = \exp_y(-t\nabla\phi(y))$. If ε_2 is small enough then $\varepsilon_2\phi$ is $\frac{d^2}{2}$-concave. As μ_0 is absolutely continuous, F_{ε_2} is the unique optimal transport between μ_0 and $(F_{\varepsilon_2})_*\mu_0$. As a consequence, $\mu_t \equiv (F_{t\varepsilon_2})_*\mu_0$ is the unique Wasserstein geodesic from μ_0 to μ_1. Since $\varepsilon_1 \to 0$ and then $\varepsilon_2 \to 0$, if $H_{N,\nu}$ is to satisfy (3.16) for all such μ_0 then we must have $C''(0) \leq 0$. Hence $\operatorname{Ric}_N(v,v) \geq 0$. Since v was arbitrary, this shows that $\operatorname{Ric}_N \geq 0$.

Now suppose that $N = n$ and $H_{N,\nu}$ is weakly a.c. displacement convex. Given $m \in M$ and $v \in T_mM$, we want to show that $v\Psi = 0$ and $\operatorname{Ric}(v,v) \geq 0$. Choose a smooth function ϕ, defined in a neighborhood of m, so that $v =$

$-(\nabla\phi)(m)$, and $\mathrm{Hess}\,(\phi)(m)$ is proportionate to $g(m)$. We must again have $C''(0) \leq 0$, where now $C(t) = e^{-\frac{\Psi(F_t(y))}{n}} \det^{\frac{1}{n}}(dF_t)(y)$. By direct computation,

$$\frac{C''(0)}{C(0)} = -\frac{1}{n}(\mathrm{Ric} + \mathrm{Hess}\,(\Psi))(v,v) + \frac{(v\Psi)^2}{n^2} + \frac{2(v\Psi)(\triangle\phi)(m)}{n^2}. \tag{7.33}$$

If $v\Psi \neq 0$ then we can make $C''(0) > 0$ by an appropriate choice of $\triangle\phi$. Hence Ψ must be constant and then we must have $\mathrm{Ric}(v,v) \geq 0$.

Finally, if $N < n$ then (7.25) gives

$$\begin{aligned} N\frac{C''(0)}{C(0)} = &-(\mathrm{Ric} + \mathrm{Hess}\,(\Psi))(v,v) \\ &+ \frac{(v\Psi)^2}{N-n} - \frac{n(N-n)}{N}\left(-\frac{v\Psi}{N-n} + \frac{(\triangle\phi)(m)}{n}\right)^2. \end{aligned} \tag{7.34}$$

One can always choose $(\triangle\phi)(m)$ to make $C''(0)$ positive and so $H_{N,\nu}$ cannot be weakly a.c. displacement convex. □

Proof of Theorem 7.3, *part* (b). We first show (1) ⇒ (2). Suppose that $\mathrm{Ric}_\infty \geq Kg$. Given $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(M)$, we again use (7.19), with $U \in \mathcal{DC}_\infty$. Putting

$$C(y,t) = -\Psi(F_t(y)) + \log\det(dF_t)(y), \tag{7.35}$$

we have

$$U_\nu(\mu_t) = \int_M e^{C(y,t)} U\left(\eta_0(y) e^{-C(y,t)}\right) \mathrm{dvol}_M(y). \tag{7.36}$$

As in the proof of (a), the condition $\mathrm{Ric}_\infty \geq Kg$ implies that

$$\frac{d^2C}{dt^2} \leq -K|F_t'(y)|^2 = -K|\nabla\phi|^2(y), \tag{7.37}$$

where the last equality comes from the constant speed of the geodesic $t \to F_t(y)$. By assumption, the map

$$\lambda \to \eta_0^{-1}(y) e^\lambda U(\eta_0(y) e^{-\lambda}) \tag{7.38}$$

is nonincreasing and convex in λ, with derivative $-\frac{p(\eta_0(y)e^{-\lambda})}{\eta_0(y)e^{-\lambda}}$. It follows that the composition

$$t \to \eta_0^{-1}(y) e^{C(y,t)} U\left(\eta_0(y) e^{-C(y,t)}\right) \tag{7.39}$$

is $\lambda(U)|\nabla\phi|^2(y)$-convex in t. Then

$$\begin{aligned} e^{C(y,t)} U\left(\eta_0(y) e^{-C(y,t)}\right) \leq\ & t e^{C(y,1)} U\left(\eta_0(y) e^{-C(y,1)}\right) \\ &+ (1-t) e^{C(y,0)} U\left(\eta_0(y) e^{-C(y,0)}\right) \\ &- \frac{1}{2}\lambda(U)|\nabla\phi|^2(y)\eta_0(y) t(1-t). \end{aligned} \tag{7.40}$$

Integrating with respect to $\mathrm{dvol}_M(y)$ and using the fact that

$$W_2(\mu_0,\mu_1)^2 = \int_M |\nabla\phi|^2(y)\eta_0(y)\,\mathrm{dvol}_M(y) \tag{7.41}$$

show that (3.16) is satisfied with $\lambda = \lambda(U)$. The implication (1) $\Rightarrow$ (2) now follows from Proposition 4.28.

The proof that (5) $\Rightarrow$ (1) is similar to the proof in part (a). □

The case $N = 1$ is slightly different because $H_{1,\nu}$ is not defined. However, the rest of Theorem 7.3.a carries through.

THEOREM 7.42. a. *The following are equivalent*:

(1) $\mathrm{Ric}_1 \geq 0$.

(2) *The measured length space* (M, g, ν) *has nonnegative* 1*-Ricci curvature.*

(3) *For all* $U \in \mathcal{DC}_1$, U_ν *is weakly displacement convex on* $P_2(M)$.

(4) *For all* $U \in \mathcal{DC}_1$, U_ν *is weakly* a.c. *displacement convex on* $P_2^{\mathrm{ac}}(M)$.

Proof. The proofs of (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) are as in the proof of Theorem 7.3.a. It remains to show that (4) $\Rightarrow$ (1). Since $\mathcal{DC}_N \subset \mathcal{DC}_1$ for all $N > 1$, condition (4) implies that U_ν is weakly a.c. displacement convex on $P_2^{\mathrm{ac}}(M)$ for all $U \in \mathcal{DC}_N$. So by Theorem 7.3.a, M satisfies $\mathrm{Ric}_N \geq 0$ for all $N > 1$. It follows that $n \leq 1$. If $n = 0$, i.e. M is a point, then $\mathrm{Ric}_1 \geq 0$ holds automatically. If $n = 1$, i.e. M is a circle, then taking $N \to 1^+$ shows that $\mathrm{Ric}_1 \geq 0$, i.e. Ψ is constant. □

Remark 7.43. In the Riemannian case there is a unique Wasserstein geodesic joining $\mu_0, \mu_1 \in P_2^{\mathrm{ac}}(M)$. Hence we could add two more equivalences to Theorem 7.3. Namely, a.(4) is equivalent to saying that for all $U \in \mathcal{DC}_N$, U_ν is a.c. displacement convex on $P_2^{\mathrm{ac}}(M)$, and b.(4) is equivalent to saying that for all $U \in \mathcal{DC}_\infty$, U_ν is $\lambda(U)$-a.c. displacement convex on $P_2^{\mathrm{ac}}(M)$.

Remark 7.44. Theorem 7.3 also holds under weaker regularity assumptions. For example, if Ψ is a continuous function on Euclidean $\mathbb{R}^n$ then $\left(\mathbb{R}^n, e^{-\Psi}\,\mathrm{dvol}_{\mathbb{R}^n}\right)$ has ∞-Ricci curvature bounded below by zero if and only if Ψ is convex.

7.2. *Geometric corollaries.* We have shown that our abstract notion of a lower Ricci curvature bound is stable under measured Gromov-Hausdorff convergence (Theorem 5.19) and, in the Riemannian setting, coincides with a classical notion (Theorem 7.3). This subsection is devoted to various geometric applications.

We first give a characterization of the *smooth* elements in the set of measured Gromov-Hausdorff limits of manifolds with Ricci curvature bounded below.

COROLLARY 7.45. *Let (B, g_B) be a smooth compact connected Riemannian manifold, equipped with the Riemannian density* dvol_B, *and let* Ψ *be a* C^2*-regular function on* B *which is normalized by an additive constant so that* $e^{-\Psi}\,\mathrm{dvol}_B$ *is a probability measure on* B. *We have the following implications*:

(i) *If* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *is a measured Gromov-Hausdorff limit of Riemannian manifolds with nonnegative Ricci curvature and dimension at most* N *then* $\mathrm{Ric}_N(B) \geq 0$.

(i′) *If* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *is a measured Gromov-Hausdorff limit of Riemannian manifolds with Ricci curvature bounded below by* $K \in \mathbb{R}$ *then* $\mathrm{Ric}_\infty(B) \geq K g_B$.

(ii) *As a partial converse, if* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *has* $\mathrm{Ric}_N(B) \geq 0$ *with* $N \geq \dim(B) + 2$ *an integer then* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *is a measured Gromov-Hausdorff limit of Riemannian manifolds with nonnegative Ricci curvature and dimension at most* N.

(ii′) *If* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *has* $\mathrm{Ric}_\infty(B) \geq K g_B$ *then* $(B, g_B, e^{-\Psi}\,\mathrm{dvol}_B)$ *is a measured Gromov-Hausdorff limit of Riemannian manifolds* M_i *with* $\mathrm{Ric}(M_i) \geq (K - \frac{1}{i}) g_{M_i}$.

Proof. Parts (i) and (i′) are a direct consequence of Theorems 5.19, 7.3 and 7.42. Part (ii) follows from the warped product construction of [30, Th. 3.1]. The proof of (ii′) is similar. □

Remark 7.46. In Corollary 7.45(ii′), if $K \neq 0$ then one can use a rescaling argument to transform the condition $\mathrm{Ric}(M_i) \geq \left(K - \frac{1}{i}\right) g_{M_i}$ into the more stringent condition $\mathrm{Ric}(M_i) \geq K\, g_{M_i}$.

The next two corollaries give some consequences of Corollary 7.45 for the *metric* structure of smooth limit spaces, i.e. for the aspects of the limit metric-measure spaces that are independent of the measure. In general, one cannot change the conclusion of Corollary 7.45(i) to obtain a lower bound on Ric instead of Ric_N. However, one does obtain such a lower bound in the noncollapsing case.

COROLLARY 7.47. a. *Suppose that* (X, d) *is a Gromov-Hausdorff limit of* n*-dimensional Riemannian manifolds with nonnegative Ricci curvature. If* (X, d) *has Hausdorff dimension* n, *and* ν_H *is its normalized* n*-dimensional Hausdorff measure, then* (X, d, ν_H) *has nonnegative* n*-Ricci curvature.*

b. *If in addition* (X, d) *happens to be a smooth* n*-dimensional Riemannian manifold* (B, g_B) *then* $\mathrm{Ric}(B) \geq 0$.

Proof. a. If $\{M_i\}_{i=1}^\infty$ is a sequence of n-dimensional Riemannian manifolds with nonnegative Ricci curvature and $\{f_i\}_{i=1}^\infty$ is a sequence of ε_i-approximations

$f_i : M_i \to X$, with $\lim_{i\to\infty} \varepsilon_i = 0$, then $\lim_{i\to\infty}(f_i)_* \operatorname{dvol}_{M_i} = \nu_H$ in the weak-$*$ topology [16, Ths. 3.1 and 5.9]. (This also shows that the n-dimensional Hausdorff measure on X can be normalized to be a probability measure.) Then part a. follows from Theorems 5.19 and 7.3.

b. If $(X, d) = (B, g_B)$ then $\nu_H = \frac{\operatorname{dvol}_B}{\operatorname{vol}(B)}$ and the claim follows from Theorem 7.3, along with the definition of Ric_n. □

Remark 7.48. A special case of Corollary 7.47.a occurs when (X, d) is an n-dimensional Gromov-Hausdorff limit of a sequence of n-dimensional Riemannian manifolds with nonnegative sectional curvature. In this case, (X, d) has nonnegative Alexandrov curvature and (X, d, ν_H) has nonnegative n-Ricci curvature. More generally, we expect that for an n-dimensional compact length space (X, d) with Alexandrov curvature bounded below, equipped with the normalized n-dimensional Hausdorff measure ν_H,

1. If (X, d) has nonnegative Alexandrov curvature then (X, d, ν_H) has nonnegative n-Ricci curvature, and
2. For $n > 1$, if (X, d) has Alexandrov curvature bounded below by $\frac{K}{n-1}$ then (X, d, ν_H) has ∞-Ricci curvature bounded below by K.

It is possible that the proof of Theorem 7.3 can be adapted to show this.

As mentioned above, in the collapsing case the lower bound in the conclusion of Corollary 7.45(i) (or Corollary 7.45(i′)) would generally fail if we replaced Ric_N (or $\operatorname{Ric}_\infty$) by Ric. However, one does obtain a lower bound on the average scalar curvature of B.

Corollary 7.49. *If $\left(B, g_B, e^{-\Psi} \operatorname{dvol}_B\right)$ is a smooth n-dimensional measured Gromov-Hausdorff limit of Riemannian manifolds (of arbitrary dimension), each with Ricci curvature bounded below by $K \in \mathbb{R}$, then the scalar curvature S of (B, g_B) satisfies*

$$\frac{\int_B S \operatorname{dvol}_B}{\operatorname{vol}(B)} \geq nK. \tag{7.50}$$

Proof. From Corollary 7.45(iii), $\operatorname{Ric}(B) + \operatorname{Hess}(\Psi) \geq K g_B$. Tracing gives $S + \triangle\Psi \geq nK$. Integrating gives $\int_B S \operatorname{dvol}_B \geq nK \operatorname{vol}(B)$. □

Next, we show that for Riemannian manifolds, lower N-Ricci curvature bounds are preserved upon taking quotients by compact Lie group actions.

Corollary 7.51. *Let M be a compact connected Riemannian manifold. Let G be a compact Lie group that acts isometrically on M, preserving a function $\Psi \in C^\infty(M)$ that satisfies $\int_M e^{-\Psi} \operatorname{dvol}_M = 1$. Let $p : M \to M/G$ be the quotient map.*

a. *For $N \in [1, \infty)$, if $\left(M, e^{-\Psi} \operatorname{dvol}_M\right)$ has $\operatorname{Ric}_N \geq 0$ then*

$$(M/G, d_{M/G}, p_*(e^{-\Psi} \operatorname{dvol}_M))$$

has nonnegative N-Ricci curvature.

b. *If $\left(M, e^{-\Psi} \operatorname{dvol}_M\right)$ has $\operatorname{Ric}_\infty \geq K g_M$ then $(M/G, d_{M/G}, p_*(e^{-\Psi} \operatorname{dvol}_M))$ has ∞-Ricci curvature bounded below by K.*

Proof. This follows from Theorem 5.35. □

Corollary 7.51 provides many examples of singular spaces with lower Ricci curvature bounds. Of course, the main case is when Ψ is constant.

We conclude this section by giving a "synthetic" proof of a part of the Ricci-O'Neill theorem of [30, Th. 2].

COROLLARY 7.52. *Let $p : M \to B$ be a Riemannian submersion of compact connected manifolds, with fibers Z_b. Choose $N \geq \dim(M)$ and $\Psi_M \in C^\infty(M)$ with $\int_M e^{-\Psi_M} \operatorname{dvol}_M = 1$; if $N = \dim(M)$ then we assume that Ψ_M is constant. Define $\Psi_B \in C^\infty(B)$ by $p_*\left(e^{-\Psi_M} \operatorname{dvol}_M\right) = e^{-\Psi_B} \operatorname{dvol}_B$. Suppose that the fiber parallel transport of the Riemannian submersion preserves the fiberwise measures $e^{-\Psi_M}\big|_Z \operatorname{dvol}_Z$ up to multiplicative constants. (That is, if $\gamma : [0,1] \to B$ is a smooth path in B, let $P_\gamma : Z_{\gamma(0)} \to Z_{\gamma(1)}$ denote the fiber transport diffeomorphism.) Then we assume that there is a constant $C_\gamma > 0$ so that*

$$P_\gamma^*\left(e^{-\Psi_M}\big|_{Z_{\gamma(1)}} \operatorname{dvol}_{Z_{\gamma(1)}}\right) = C_\gamma e^{-\Psi_M}\big|_{Z_{\gamma(0)}} \operatorname{dvol}_{Z_{\gamma(0)}}. \tag{7.53}$$

With these assumptions,

a. *If $\operatorname{Ric}_N(M) \geq 0$ then $\operatorname{Ric}_N(B) \geq 0$.*

b. *For any $K \in \mathbb{R}$, if $\operatorname{Ric}_\infty(M) \geq K g_M$ then $\operatorname{Ric}_\infty(B) \geq K g_B$.*

Proof. Put $\nu_M = e^{-\Psi_M} \operatorname{dvol}_M$ and $\nu_B = e^{-\Psi_B} \operatorname{dvol}_B$. We can decompose ν_M with respect to p as $\sigma(b)\nu_B(b)$, with $\sigma(b) \in P_2^{\mathrm{ac}}(Z_b)$. From the assumptions, the family $\{\sigma(b)\}_{b \in B}$ of vertical densities is invariant under fiber parallel transport.

To prove part (a), let $\{\mu_t\}_{t\in[0,1]}$ be a Wasserstein geodesic in $P_2^{\mathrm{ac}}(B)$. Define $\{\mu'_t\}_{t\in[0,1]}$ in $P_2^{\mathrm{ac}}(M)$ by $\mu'_t \equiv \sigma(b)\mu_t(b)$. By construction, the corresponding densities satisfy $\rho'_t = p^*\rho_t$. Thus $H_{N,\nu_M}(\mu'_t) = H_{N,\nu_B}(\mu_t)$. Furthermore, $\{\mu'_t\}_{t\in[0,1]}$ is a Wasserstein geodesic; if $(F_t)_{t\in[0,1]}$ is an optimal Monge transport from μ_0 to μ_1 then its horizontal lift is an optimal Monge transport from μ'_0 to μ'_1, with generating function $\phi_M = p^*\phi_B$. From Theorem 7.3(a) and Remark 7.43, H_{N,ν_M} is a.c. displacement convex on $P_2^{\mathrm{ac}}(M)$. In particular, (3.16) is satisfied along $\{\mu'_t\}_{t\in[0,1]}$ with $U_\nu = H_{N,\nu_M}$ and $\lambda = 0$. Then the same equation is satisfied along $\{\mu_t\}_{t\in[0,1]}$ with $U_\nu = H_{N,\nu_B}$ and $\lambda = 0$. Thus

H_{N,ν_B} is a.c. displacement convex on $P_2^{\mathrm{ac}}(B)$. Theorem 7.3(a) now implies that $\mathrm{Ric}_N(B) \geq 0$.

The proof of part (b) is similar. □

Remark 7.54. In fact, for any $N \in [1, \infty]$ and any $K \in \mathbb{R}$, if $\mathrm{Ric}_N(M) \geq K g_M$ then $\mathrm{Ric}_N(B) \geq K g_B$. This was proven in [30, Th. 2] in the cases $N = \infty$ and $N = \dim(M)$ by explicit tensor calculations. (The paper [30] writes Ric_q for what we write as Ric_N, where $q = N - n$.) The same method of proof works for all N.

Remark 7.55. Suppose that M is a compact connected Riemannian manifold on which a compact Lie group G acts isometrically, with all orbits of the same orbit type. Put $B = M/G$. If $\Psi_M \in C^\infty(M)$ is a G-invariant function that satisfies $\int_M e^{-\Psi_M} \, \mathrm{dvol}_M = 1$, and $\left(M, g_M, e^{-\Psi_M} \, \mathrm{dvol}_M\right)$ has $\mathrm{Ric}_N(M) \geq 0$, then Corollaries 7.51 and 7.52 overlap in saying that $\left(B, g_B, e^{-\Psi_B} \, \mathrm{dvol}_B\right)$ has $\mathrm{Ric}_N(B) \geq 0$. There is a similar statement when $\mathrm{Ric}_\infty(M) \geq K g_M$.

Remark 7.56. There is an obvious analogy between the Ricci-O'Neill theorem and O'Neill's theorem that sectional curvature is nondecreasing under pushforward by a Riemannian submersion. There is also a "synthetic" proof of O'Neill's theorem, obtained by horizontally lifting a geodesic hinge from B and using triangle comparison results, along with the fact that p is distance nonincreasing.

Appendix A. The Wasserstein space as an Alexandrov space

This section is concerned with the geometry of the Wasserstein space $P_2(M)$ of a Riemannian manifold M. Otto introduced a formal infinite-dimensional Riemannian metric on $P_2(M)$ and showed that $P_2(\mathbb{R}^n)$ formally has nonnegative sectional curvature [35]. We make such results rigorous by looking at $P_2(M)$ as an Alexandrov space.

We first give a general lower bound on Wasserstein distances in terms of Lipschitz functions. We show that if M is a compact Riemannian manifold with nonnegative sectional curvature then $P_2(M)$ has nonnegative Alexandrov curvature. Using the above-mentioned lower bound on Wasserstein distances, we compute the tangent cones of $P_2(M)$ at the absolutely continuous measures.

A.1. *Lipschitz functions and optimal transport.* In general, one can estimate Wasserstein distances from above by choosing particular transference plans. The next lemma provides a way to estimate these distances from below by using Lipschitz functions.

LEMMA A.1. *If X is a compact length space and $\{\mu_t\}_{t\in[0,1]}$ is a Wasserstein geodesic then for all $f \in \mathrm{Lip}(X)$,*

$$\left| \int_X f \, d\mu_1 - \int_X f \, d\mu_0 \right|^2 \leq W_2(\mu_0, \mu_1)^2 \int_0^1 \left(\int_X |\nabla f|^2 d\mu_t \right) dt. \tag{A.2}$$

Proof. By Proposition 2.10, the Wasserstein geodesic arises as the displacement interpolation associated to some optimal dynamical transference plan Π. We have

(A.3)
$$\int_X f d\mu_1 - \int_X f d\mu_0 = \int_X f d(e_1)_*\Pi - \int_X f d(e_0)_*\Pi = \int_\Gamma ((e_1)^* f - (e_0)^* f) d\Pi$$
$$= \int_\Gamma (f(\gamma(1)) - f(\gamma(0))) d\Pi(\gamma).$$

As $f \circ \gamma \in \mathrm{Lip}([0,1])$,

$$|f(\gamma(1)) - f(\gamma(0))| = \left| \int_0^1 \frac{df(\gamma(t))}{dt} dt \right| \leq \int_0^1 \left| \frac{df(\gamma(t))}{dt} \right| dt \leq \int_0^1 |\nabla f|(\gamma(t)) L(\gamma) dt. \tag{A.4}$$

Then

$$\left| \int_X f d\mu_1 - \int_X f d\mu_0 \right| \leq \int_\Gamma \int_0^1 |\nabla f|(\gamma(t)) L(\gamma) dt d\Pi(\gamma). \tag{A.5}$$

From the Cauchy-Schwarz inequality,

(A.6)
$$\left| \int_X f d\mu_1 - \int_X f d\mu_0 \right|^2 \leq \int_\Gamma \int_0^1 |\nabla f|^2(\gamma(t)) dt d\Pi(\gamma) \int_\Gamma \int_0^1 L(\gamma)^2 dt d\Pi(\gamma)$$
$$= \int_0^1 \int_\Gamma |\nabla f|^2(\gamma(t)) dt d\Pi(\gamma) \int_\Gamma L(\gamma)^2 d\Pi(\gamma).$$

Now, $W_2(\mu_0, \mu_1)^2 = \int_\Gamma L(\gamma)^2 d\Pi(\gamma)$. To conclude the proof, we note that

$$\int_\Gamma |\nabla f|^2(\gamma(t)) d\Pi(\gamma) = \int_\Gamma |\nabla f|^2(e_t(\gamma)) d\Pi(\gamma) = \int_X |\nabla f|^2 d(e_t)_*\Pi = \int_X |\nabla f|^2 d\mu_t. \tag{A.7}$$
□

A.2. *The case of nonnegatively curved manifolds.*

THEOREM A.8. *A smooth compact connected manifold M has nonnegative sectional curvature if and only if $P_2(M)$ has nonnegative Alexandrov curvature.*

Proof. Suppose that M has nonnegative sectional curvature. We first show that $P_2^{\mathrm{ac}}(M)$ has nonnegative Alexandrov curvature. Let μ_0, μ_1, μ_2 and μ_3 be points in $P_2^{\mathrm{ac}}(M)$ with $\mu_i \neq \mu_0$ for $1 \leq i \leq 3$. For $1 \leq i \neq j \leq 3$, let $\widetilde{\angle} \mu_i \mu_0 \mu_j$ denote the comparison angle at μ_0 of the triangle formed by μ_i, μ_0 and μ_j

[11, Def. 3.6.25]. For $1 \le i \le 3$, let $c_i : [0,1] \to P_2^{\mathrm{ac}}(M)$ be a Wasserstein geodesic from μ_0 to μ_i. Let ϕ_i be the corresponding $\frac{d^2}{2}$-concave function on M that generates c_i. That is, $c_i(t) = (F_{i,t})_* \mu_0$, with $F_{i,t}(m) = \exp_m(-t \nabla \phi_i(m))$. Then $W_2(\mu_0, \mu_i)^2 = \int_M d(m, F_{i,t}(m))^2 d\mu_0(m) = \int_M |\nabla \phi_i|^2 d\mu_0$.

Let us define a particular transference plan from μ_i to μ_j as a Monge transport $F_{j,1} \circ F_{i,1}^{-1}$. It gives an upper bound on $W_2(\mu_i, \mu_j)^2$ by

$$W_2(\mu_i, \mu_j)^2 \le \int_M d(F_{i,1}(m), F_{j,1}(m))^2 d\mu_0(m). \tag{A.9}$$

For almost all m, the nonnegative curvature of M, applied to the hinge at m formed by the geodesic segments $t \to F_{i,t}(m)$ and $t \to F_{j,t}(m)$, implies

$$d(F_{i,1}(m), F_{j,1}(m))^2 \le |\nabla \phi_i(m)|^2 + |\nabla \phi_j(m)|^2 - 2 \langle \nabla \phi_i(m), \nabla \phi_j(m) \rangle. \tag{A.10}$$

Integrating (A.10) with respect to μ_0 yields

$$W_2(\mu_i, \mu_j)^2 \le W_2(\mu_0, \mu_i)^2 + W_2(\mu_0, \mu_j)^2 - 2 \int_M \langle \nabla \phi_i(m), \nabla \phi_j(m) \rangle d\mu_0(m). \tag{A.11}$$

Thus $\widetilde{\angle} \mu_i \mu_0 \mu_j \le \theta_{ij}$, where $\theta_{ij} \in [0, \pi]$ is defined by

$$\cos \theta_{ij} = \frac{\int_M \langle \nabla \phi_i, \nabla \phi_j \rangle d\mu_0}{\sqrt{\int_M |\nabla \phi_i|^2 d\mu_0 \int_M |\nabla \phi_j|^2 d\mu_0}}. \tag{A.12}$$

It follows from the geometry of an inner product space that

$$\theta_{12} + \theta_{23} + \theta_{31} \le 2\pi. \tag{A.13}$$

Thus

$$\widetilde{\angle} \mu_1 \mu_0 \mu_2 + \widetilde{\angle} \mu_2 \mu_0 \mu_3 + \widetilde{\angle} \mu_3 \mu_0 \mu_1 \le 2\pi, \tag{A.14}$$

which implies that $P_2^{\mathrm{ac}}(M)$ has nonnegative Alexandrov curvature [11, Prop. 10.1.1].

As $P_2(M)$ is the completion of $P_2^{\mathrm{ac}}(M)$, the fact that (A.14) can be written solely in terms of distances implies that it also holds for $P_2(M)$. Thus $P_2(M)$ has nonnegative Alexandrov curvature, and as the embedding $M \to P_2(M)$ by delta functions defines a totally geodesic subspace of $P_2(M)$, it follows that M has nonnegative Alexandrov curvature. Thus M has nonnegative sectional curvature. □

Remark A.15. The fact that $P_2(M)$ has nonnegative Alexandrov curvature ensures the existence and uniqueness of the gradient flow of a Lipschitz λ-concave function on $P_2(M)$ [38, Appendix]. (The conventions of [38] are such that the function increases along the flowlines of its gradient flow; some other authors have the convention that a function decreases along the flowlines of its gradient flow, and hence consider λ-convex functions.) Other approaches to geometrizing $P_2(M)$, with a view toward defining gradient flows, appear in [2], [3], [14].

Now suppose that M has nonnegative sectional curvature. Let $c_0, c_1 : [0,1] \to P_2(M)$ be nontrivial Wasserstein geodesics, with $c_0(0) = c_1(0) = \mu$. Theorem A.8 implies that the comparison angle $\widetilde{\angle} c_0(s_0)\mu c_1(s_1)$ is monotonically nonincreasing as s_0 and s_1 increase, separately in s_0 and s_1 [11, Def. 4.3.1, Ths. 4.3.5 and 10.1.1]. Then there is a well-defined angle $\angle(c_0, c_1)$ that c_0 and c_1 form at μ, in the sense of [11, Def. 3.6.26], given by

$$\angle(c_0, c_1) = \lim_{s_0, s_1 \to 0^+} \widetilde{\angle} c_0(s_0)\mu c_1(s_1). \tag{A.16}$$

Proposition A.17. *Let $c_0, c_1 : [0,1] \to P_2^{\mathrm{ac}}(M)$ be nontrivial Wasserstein geodesics, with $c_0(0) = c_1(0) = \mu$. If ϕ_0 and ϕ_1 are the $\frac{d^2}{2}$-concave functions that generate c_0 and c_1, respectively, then*

$$\cos \angle(c_0, c_1) = \frac{\int_M \langle \nabla \phi_0, \nabla \phi_1 \rangle d\mu}{\sqrt{\int_M |\nabla \phi_0|^2 d\mu \int_M |\nabla \phi_1|^2 d\mu}}, \tag{A.18}$$

Proof. Applying (A.12) (and the sentence preceding it) to the triangle $\triangle c_0(s_0)\mu c_1(s_1)$ gives

$$\cos \widetilde{\angle} c_0(s_0)\mu c_1(s_1) \geq \frac{\int_M \langle \nabla \phi_0, \nabla \phi_1 \rangle d\mu}{\sqrt{\int_M |\nabla \phi_0|^2 d\mu \int_M |\nabla \phi_1|^2 d\mu}}, \tag{A.19}$$

and so

$$\cos \angle(c_0, c_1) \geq \frac{\int_M \langle \nabla \phi_0, \nabla \phi_1 \rangle d\mu}{\sqrt{\int_M |\nabla \phi_0|^2 d\mu \int_M |\nabla \phi_1|^2 d\mu}}. \tag{A.20}$$

From the monotonicity of $\widetilde{\angle} c_0(s_0)\mu c_1(s_1)$, it suffices to show that

$$\lim_{s \to 0} \cos \widetilde{\angle} c_0(s)\mu c_1(s) \leq \frac{\int_M \langle \nabla \phi_0, \nabla \phi_1 \rangle d\mu}{\sqrt{\int_M |\nabla \phi_0|^2 d\mu \int_M |\nabla \phi_1|^2 d\mu}}. \tag{A.21}$$

This amounts to showing a lower bound on $W_2(c_0(s), c_1(s))$.

Let $\{\mu_{t,s}\}_{t \in [0,1]}$ be a Wasserstein geodesic from $c_0(s)$ to $c_1(s)$. From Lemma A.1, for any $f \in C^1(M)$,

$$\left(\int_M f dc_1(s) - \int_M f dc_0(s) \right)^2 \leq W_2(c_0(s), c_1(s))^2 \int_0^1 \int_M |\nabla f|^2 d\mu_{t,s} dt. \tag{A.22}$$

In terms of the Monge transport maps $F_{0,t}$ and $F_{1,t}$,

$$\begin{aligned} \int_M f dc_1(s) - \int_M f dc_0(s) &= \int_M f d(F_{1,s})_* \mu - \int_M f d(F_{0,s})_* \mu \\ &= \int_M ((F_{1,s})^* f - (F_{0,s})^* f) d\mu. \end{aligned} \tag{A.23}$$

Thus
(A.24)
$$\left(\frac{\int_M ((F_{1,s})^* f - (F_{0,s})^* f) d\mu}{s} \right)^2 \leq \frac{W_2(c_0(s), c_1(s))^2}{s^2} \int_0^1 \int_M |\nabla f|^2 d\mu_{t,s} dt.$$

Since $\{\mu_{t,s}\}_{t \in [0,1]}$ is minimizing between its endpoints, we must have

$$W_2(\mu, \mu_{t,s}) \leq W_2(\mu, c_0(s)) + W_2(\mu, c_1(s)) \tag{A.25}$$

for all $t \in [0,1]$. (Otherwise the length of $\{\mu_{t,s}\}_{t \in [0,1]}$ would have to be greater than $W_2(\mu, c_0(s)) + W_2(\mu, c_1(s))$. Then there would be a path from $c_0(s)$ to $c_1(s)$ that is shorter than $\{\mu_{t,s}\}_{t\in[0,1]}$, obtained by going from $c_0(s)$ to μ along c_0 and then from μ to $c_1(s)$ along c_1.) If $\pi_{t,s}$ is an optimal transference plan between μ and $\mu_{t,s}$ then $\int_0^1 \pi_{t,s} dt$ is a transference plan between μ and $\int_0^1 \mu_{t,s} dt$, showing that
(A.26)
$$W_2 \left(\mu, \int_0^1 \mu_{t,s} dt \right)^2 \leq \int_0^1 W_2 \left(\mu, \mu_{t,s} \right)^2 dt \leq \left(W_2(\mu, c_0(s)) + W_2(\mu, c_1(s)) \right)^2.$$

Thus $\lim_{s \to 0} \int_0^1 \mu_{t,s} dt = \mu$ in the weak-$*$ topology. As $|\nabla f|^2 \in C(M)$, taking $s \to 0$ in (A.24) gives

$$\left(\int_M \langle \nabla f, \nabla \phi_0 - \nabla \phi_1 \rangle d\mu \right)^2 \leq \left(\lim_{s \to 0} \frac{W_2(c_0(s), c_1(s))^2}{s^2} \right) \int_M |\nabla f|^2 d\mu. \tag{A.27}$$

We claim that in fact (A.27) holds for any $f \in \mathrm{Lip}(M)$. To see this, let $e^{v\triangle}$ be the heat operator on M. Given $f \in \mathrm{Lip}(M)$, for any $v > 0$, $e^{v\triangle} f \in C^1(M)$. It follows from spectral theory that $\lim_{v \to 0} \nabla e^{v\triangle} f = \nabla f$ in the Hilbert space of square-integrable vector fields on M. Then $\lim_{v \to 0} \left| \nabla e^{v\triangle} f \right|^2 = |\nabla f|^2$ in $L^1(M, \mathrm{dvol}_M)$. There is a uniform bound on the L^∞-norm of $\nabla e^{v\triangle} f \equiv e^{v\triangle_1} df$ for $v \in [0,1]$. Writing $\mu = \rho \, \mathrm{dvol}_M$ with $\rho \in L^1(M, \mathrm{dvol}_M)$, for any $N \in \mathbb{Z}^+$ we have

(A.28)
$$\begin{aligned}
&\left| \int_M \left| \nabla e^{v\triangle} f \right|^2 \rho \, \mathrm{dvol}_M - \int_M |\nabla f|^2 \rho \, \mathrm{dvol}_M \right| \\
&= \left| \int_{\rho^{-1}([0,N])} \left(\left| \nabla e^{v\triangle} f \right|^2 - |\nabla f|^2 \right) \rho \, \mathrm{dvol}_M \right. \\
&\qquad \left. + \int_{\rho^{-1}((N,\infty))} \left(\left| \nabla e^{v\triangle} f \right|^2 - |\nabla f|^2 \right) \rho \, \mathrm{dvol}_M \right| \\
&\leq N \parallel \left| \nabla e^{v\triangle} f \right|^2 - |\nabla f|^2 \parallel_1 + \left(\parallel \nabla e^{v\triangle} f \parallel_\infty^2 + \parallel \nabla f \parallel_\infty^2 \right) \int_{\rho^{-1}((N,\infty))} \rho \, \mathrm{dvol}_M \, .
\end{aligned}$$

For any $\varepsilon > 0$, by taking N large we can make

$$\left(\parallel \nabla e^{v\triangle} f \parallel_\infty^2 + \parallel \nabla f \parallel_\infty^2 \right) \int_{\rho^{-1}((N,\infty))} \rho \, \mathrm{dvol}_M$$

less than ε. Then by taking v small, we can make $N \parallel \left|\nabla e^{v\triangle} f\right|^2 - |\nabla f|^2 \parallel_1$ less than ε. It follows that

$$\lim_{v\to 0}\int_M \left|\nabla e^{v\triangle} f\right|^2 \rho\, \mathrm{dvol}_M = \int_M |\nabla f|^2 \rho\, \mathrm{dvol}_M\,. \tag{A.29}$$

By a similar argument,

$$\lim_{v\to 0}\int_M \left\langle \nabla e^{v\triangle} f, \nabla\phi_0 - \nabla\phi_1\right\rangle \rho\, \mathrm{dvol}_M = \int_M \langle \nabla f, \nabla\phi_0 - \nabla\phi_1\rangle\, \rho\, \mathrm{dvol}_M\,. \tag{A.30}$$

Thus (A.27) holds for f.

In particular, taking $f = \phi_0 - \phi_1$ in (A.30) gives

$$\lim_{s\to 0}\frac{W_2(c_0(s), c_1(s))^2}{s^2} \geq \int_M |\nabla\phi_0 - \nabla\phi_1|^2 d\mu, \tag{A.31}$$

or

$$\begin{aligned}\lim_{s\to 0}\frac{W_2(c_0(s), c_1(s))^2}{s^2} \geq{}& \frac{W_2(\mu, c_0(s))^2}{s^2} + \frac{W_2(\mu, c_1(s))^2}{s^2} \\ &- 2\frac{W_2(\mu, c_0(s))}{s}\frac{W_2(\mu, c_1(s))}{s}\frac{\int_M \langle\nabla\phi_0, \nabla\phi_1\rangle d\mu}{\sqrt{\int_M |\nabla\phi_0|^2 d\mu \int_M |\nabla\phi_1|^2 d\mu}}.\end{aligned} \tag{A.32}$$

Equation (A.18) follows. □

A.3. *Application to the geometric description of* $P_2(M)$. Let us recall some facts about a finite-dimensional Alexandrov space Y with curvature bounded below [11], [12]. Let n be the dimension of Y. A point $y \in Y$ is a *regular point* if its tangent cone is isometric to $\mathbb{R}^n$. The complement of the regular points is the set S of *singular points*. The regular points $Y - S$ form a dense totally convex subset of Y, but need not be open or closed in Y; see [34, pp. 632–633] for simple but relevant examples. The existence of a Riemannian metric on Y was studied in [33], [34], [37]. We recall the results of [37]: There is a dense open totally convex subset Y^δ of Y, containing $Y - S$, which is a topological manifold with DC (=difference of concave) transition maps; there is a Riemannian metric g on Y^δ which in local charts is bounded, measurable and of bounded variation, with the restriction of g to $Y - S$ being continuous; the Christoffel symbols exist as measures in local charts on Y^δ; the lengths of curves in $Y - S$ can be computed using g.

There is an evident analogy between $Y - S \subset Y$ and $P_2^{\mathrm{ac}}(M) \subset P_2(M)$. The arguments of the above papers do not directly extend to infinite-dimensional Alexandrov spaces. Nevertheless, in order to make a zeroth order approximation to a Riemannian geometry on $P_2(M)$, it makes sense to look at the tangent cones. We recall that for a finite-dimensional Riemannian manifold, the tangent cone at a point p is isometric to T_pM equipped with the inner product coming from the Riemannian metric at p.

PROPOSITION A.33. *Let M be a smooth compact connected Riemannian manifold. If M has nonnegative sectional curvature then for each absolutely continuous measure $\mu \in P_2^{\mathrm{ac}}(M)$, the tangent cone of $P_2(M)$ at μ is an inner product space.*

Proof. Given $\mu \in P_2^{\mathrm{ac}}(M)$, we consider the space Σ'_μ of equivalence classes of geodesic segments emanating from μ, with the equivalence relation identifying two segments if they form a zero angle at μ[11, §9.1.8] (which in the case of curvature bounded below means that one segment is contained in the other). The metric on Σ'_μ is the angle. By definition, the space of directions Σ_μ is the metric completion of Σ'_μ. The tangent cone K_μ is the union of $\Sigma_\mu \times \mathbb{R}^+$ and a "vertex" point, with the metric described in [11, §10.9].

We first note that any Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ emanating from μ is of the form $\mu_t = (F_t)_*\mu$, with F_t as in Subsection 2.4 [44, Th. 2.47]. It follows that we can apply the angle calculation in Proposition A.17 to all Wasserstein geodesics emanating from μ. (One can also use the fact that any such Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ has $\mu_t \in P_2^{\mathrm{ac}}(M)$ for $t \in [0,1)$ [7, Lemma 22]; see [44, Prop. 5.9(iii)] for the $\mathbb{R}^n$-case.)

Now to identify Σ'_μ, consider the space $\mathcal{S}$ of Lipschitz functions on M that are rd^2-concave for some $r > 0$. In terms of the function ϕ, we can identify the geodesic segments from μ with $\mathcal{S}' = \mathcal{S}/\mathbb{R}$, where $\mathbb{R}$ acts additively on $\mathcal{S}$. There is an action of $\mathbb{R}^+$ on $\mathcal{S}'$ by multiplication. As the angle between geodesic segments is given by (A.18), we can identify Σ'_μ with the corresponding quotient of the space of Lipschitz functions on M that are rd^2-concave for some $r > 0$.

We can approximate a Lipschitz function on M with respect to the quadratic form $Q(\phi) = \int_M |\nabla\phi|^2 d\mu$ by functions that are rd^2-concave for various $r > 0$, for example by flowing the Lipschitz function for a short time under the heat equation on M. Hence when considering the metric completion of Σ'_μ, it doesn't matter whether we start with Lipschitz functions on M that are rd^2-concave for some $r > 0$ or arbitrary Lipschitz functions on M. It follows that K_μ is the inner product space constructed by starting with $\mathrm{Lip}(M)$, quotienting by the kernel of Q and taking the metric completion with respect to Q. □

The tangent cone constructed in Proposition A.33 agrees with the formal infinite-dimensional Riemannian metric on $P_2(M)$ considered by Otto [35]. Proposition A.33 can be considered as a way of making this formal Riemannian metric rigorous, and Theorem A.8 as a rigorous version of Otto's formal argument that his Riemannian metric on $P_2(\mathbb{R}^n)$ has nonnegative sectional curvature.

Appendix B. Some properties of the functionals U_ν

The goal of this section is to gather several results about functionals of the form U_ν, as defined in Definition 3.2. (We will generalize slightly to allow X to be a compact Hausdorff space, but the same definition makes sense.)

We show that $U_\nu(\mu)$ is lower semicontinuous in μ and ν. Such a lower semicontinuity in μ is well known in the setting of the weak topology on L^p functions, but we need to prove it in the weak-$*$ topology on Borel measures. To do so, we derive a Legendre-type formula for $U_\nu(\mu)$; this Legendre formula is also well-known in certain cases, e.g. $U(r) = r \log r$, but it is not so easy to find a precise reference for general nonlinearities.

We will also show that $U_\nu(\mu)$ is nonincreasing under "pushforward". For notation, if U is a convex function then U^* is its Legendre transform and U' is its right-derivative.

B.1. *The functional U_ν via Legendre transform.* We start by recalling, without proof, a consequence of Lusin's theorem.

THEOREM B.1. *Let X be a compact Hausdorff space. Let μ be a Borel probability measure on X. Then for all $f \in L^\infty(X)$ there is a sequence $\{f_k\}_{k=1}^\infty$ of continuous functions on X such that*

(i) $\inf f \le \inf f_k \le \sup f_k \le \sup f$ *and*

(ii) $\lim_{k\to\infty} f_k(x) = f(x)$ *for μ-almost all $x \in X$.*

We now prove a useful Legendre-type representation formula.

THEOREM B.2. *Let X be a compact Hausdorff space. Let $U : [0,\infty) \to \mathbb{R}$ be a continuous convex function with $U(0) = 0$. Given $\mu, \nu \in P(X)$, we have*

(B.3)

$$\begin{aligned} U_\nu(\mu) &= \sup \left\{ \int_X \varphi \, d\mu - \int_X U^*(\varphi) \, d\nu : \quad \varphi \in L^\infty(X), \ \varphi \le U'(\infty) \right\} \\ &= \sup_{M \in \mathbb{Z}^+} \sup \left\{ \int_X \varphi \, d\mu - \int_X U^*(\varphi) \, d\nu : \quad \varphi \in C(X), \ \varphi \le U'(M) \right\}. \end{aligned}$$

Remark B.4. The reason to add the condition $\varphi \le U'(M)$ is to ensure that $U^*(\phi)$ is continuous on X. This will be used in the proof of Theorem B.33(i).

Proof of Theorem B.2. As an initial reduction, if $U'(0) = U'(\infty)$ then U is linear and the result of the theorem is easy to check. If $U'(0) < U'(\infty)$, choose $c \in (U'(0), U'(\infty))$. Replacing $U(r)$ by $U(r) - cr$, we can reduce to the case when $U'(0) < 0$ and $U'(\infty) > 0$.

Let $\mu = \rho\nu + \mu_s$ be the Lebesgue decomposition of μ with respect to ν. Let S be a measurable set such that $\mu_s(S) = \mu_s(X)$ and $\nu(S) = 0$. Without loss of generality, we may assume that $\rho < \infty$ everywhere on $X - S$, and we set $\rho = \infty$ on S.

We will prove that

$$U_\nu(\mu) \geq \sup\left\{\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu : \quad \varphi \in L^\infty(X),\ \varphi \leq U'(\infty)\right\} \tag{B.5}$$

and

$$U_\nu(\mu) \leq \sup_{M\in\mathbb{Z}^+} \sup\left\{\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu : \varphi \in C(X),\ U'\left(\frac{1}{M}\right) \leq \varphi \leq U'(M)\right\}. \tag{B.6}$$

As the right-hand side of (B.6) is clearly less than or equal to the right-hand-side of (B.5), this will imply that

$$\begin{aligned} U_\nu(\mu) &= \sup\left\{\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu : \varphi \in L^\infty(X),\ \varphi \leq U'(\infty)\right\} \\ &= \sup_{M\in\mathbb{Z}^+} \sup\left\{\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu : \varphi \in C(X),\ U'\left(\frac{1}{M}\right) \leq \varphi \leq U'(M)\right\}, \end{aligned} \tag{B.7}$$

which in turn implies the theorem.

The proof of (B.5) is obtained by a direct argument: for any $\varphi \in L^\infty(X)$ with $\varphi \leq U'(\infty)$, we will show that

$$U_\nu(\mu) \geq \int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu. \tag{B.8}$$

We may assume that $U(\rho) \in L^1(X, \nu)$, as otherwise there is nothing to prove. For all $x \in X$, we have

$$U(\rho(x)) - \varphi(x)\rho(x) \geq -U^*(\varphi(x)). \tag{B.9}$$

Also, $\varphi\rho \in L^1(X, \nu)$. Integrating (B.9) with respect to ν gives

$$\int_X U(\rho)\, d\nu - \int_X \varphi\rho\, d\nu \geq -\int_X U^*(\varphi)\, d\nu. \tag{B.10}$$

On the other hand, since $\varphi \leq U'(\infty)$, we also have

$$U'(\infty)\mu_s(X) \geq \int_X \varphi\, d\mu_s. \tag{B.11}$$

Adding (B.10) and (B.11) gives (B.8).

To prove (B.6), it suffices to show the existence of a sequence $\{\varphi_M\}_{M=1}^\infty$ in $C(X)$ such that

$$\begin{cases} U'\left(\frac{1}{M}\right) \leq \varphi_M \leq U'(M) \text{ and} \\ U_\nu(\mu) \leq \liminf_{M\to\infty}\left(\int_X \varphi_M\, d\mu - \int_X U^*(\varphi_M)\, d\nu\right). \end{cases} \tag{B.12}$$

For $M \geq 1$, we define

$$\rho_M = \max\left(\frac{1}{M}, \min(\rho, M)\right). \tag{B.13}$$

It is clear that

(i) $M^{-1} \leq \rho_M \leq M$;

(ii) for all $x \in X$, $\lim_{M\to\infty} \rho_M(x) = \rho(x)$;

(iii) if $0 < \rho(x) < \infty$ then $\rho_M(x) = \rho(x)$ for M large enough.

Choose $\varepsilon > 0$ so that U is nonincreasing on $[0, \varepsilon)$ and nondecreasing on $[\varepsilon, \infty)$. Monotone convergence implies that

$$\int_{\rho^{-1}[\varepsilon,\infty)} U(\rho)\, d\nu = \lim_{M\to\infty} \int_{\rho^{-1}[\varepsilon,\infty)} U(\rho_M)\, d\nu \tag{B.14}$$

and

$$\int_{\rho^{-1}[0,\varepsilon)} U(\rho)\, d\nu = \lim_{M\to\infty} \int_{\rho^{-1}[0,\varepsilon)} U(\rho_M)\, d\nu. \tag{B.15}$$

Hence

$$\int_X U(\rho)\, d\nu = \lim_{M\to\infty} \int_X U(\rho_M)\, d\nu. \tag{B.16}$$

Define now a function $\overline{\varphi}_M : X \to \mathbb{R}$ by

$$\overline{\varphi}_M = U'(\rho_M). \tag{B.17}$$

Since U' is nondecreasing, we have

$$U'\left(\frac{1}{M}\right) \leq \overline{\varphi}_M \leq U'(M). \tag{B.18}$$

We also have the pointwise equality

$$U(\rho_M) = \overline{\varphi}_M \rho_M - U^*(\overline{\varphi}_M). \tag{B.19}$$

All of the functions appearing in this identity are integrable with respect to ν, so that

$$\int_X U(\rho_M)\, d\nu = \int_X \overline{\varphi}_M \rho_M\, d\nu - \int_X U^*(\overline{\varphi}_M)\, d\nu. \tag{B.20}$$

Our first goal is to prove that

$$U_\nu(\mu) \leq \liminf_{M\to\infty} \left(\int_X \overline{\varphi}_M\, d\mu - \int_X U^*(\overline{\varphi}_M)\, d\nu\right). \tag{B.21}$$

If this is true then we have shown that the sequence $\{\overline{\varphi}_M\}_{M=1}^\infty$ satisfies all of the properties required in (B.12), except maybe continuity. We split (B.21) into two parts:

$$U'(\infty)\mu_s(X) = \lim_{M\to\infty} \int_X \overline{\varphi}_M \, d\mu_s \tag{B.22}$$

and

$$\int_X U(\rho)\, d\nu \le \liminf_{M\to\infty} \left[\int_X \overline{\varphi}_M \, \rho \, d\nu - \int_X U^*(\overline{\varphi}_M)\, d\nu \right]. \tag{B.23}$$

To prove (B.22), we write

$$\int_X \overline{\varphi}_M \, d\mu_s = \int_S \overline{\varphi}_M \, d\mu_s = U'(M)\mu_s(S) = U'(M)\mu_s(X) \stackrel{M\to\infty}{\longrightarrow} U'(\infty)\mu_s(X). \tag{B.24}$$

To prove (B.23), we note that for large M,

$$U'(\rho_M)(\rho_M - \rho) \le 0. \tag{B.25}$$

Indeed, if $1/M \le \rho \le M$ then $\rho_M = \rho$; if $\rho > M$ then $\rho > \rho_M$ and $U'(\rho_M) \ge 0$; while if $\rho < 1/M$ then $\rho < \rho_M$ and $U'(\rho_M) \le 0$. Thus

$$\int_X \overline{\varphi}_M \, \rho_M \, d\nu \le \int_X \overline{\varphi}_M \, \rho \, d\nu. \tag{B.26}$$

Combining this with (B.16) and (B.20), we find

$$\begin{aligned} \int_X U(\rho)\, d\nu &= \lim_{M\to\infty} \int_X U(\rho_M)\, d\nu \\ &= \lim_{M\to\infty} \left[\int_X \overline{\varphi}_M \, \rho_M \, d\nu - \int_X U^*(\overline{\varphi}_M)\, d\nu \right] \\ &\le \liminf_{M\to\infty} \left[\int_X \overline{\varphi}_M \rho \, d\nu - \int_X U^*(\overline{\varphi}_M)\, d\nu \right]. \end{aligned} \tag{B.27}$$

This proves (B.21). To conclude the proof of the theorem, it suffices to show that for any $M \in \mathbb{Z}^+$ there is a *continuous* function φ_M such that $U'(1/M) \le \varphi_M \le U'(M)$ and

$$\left| \left(\int_X \varphi_M \, d\mu - \int_X U^*(\varphi_M)\, d\nu \right) - \left(\int_X \overline{\varphi}_M \, d\mu - \int_X U^*(\overline{\varphi}_M)\, d\nu \right) \right| \le \frac{1}{M}. \tag{B.28}$$

Fix M. By Theorem B.1, there is a sequence $\{\psi_k\}_{k=1}^\infty$ of continuous functions such that $U'(1/M) \le \inf \overline{\varphi}_M \le \inf \psi_k \le \sup \psi_k \le \sup \overline{\varphi}_M \le U'(M)$ (in particular $\{\psi_k\}_{k=1}^\infty$ is uniformly bounded) and $\lim_{k\to\infty} \psi_k(x) = \overline{\varphi}_M(x)$ $(\mu + \nu)$-almost everywhere.

At this point we note that

$$U^*(p) = \sup_{r\ge 0} [pr - U(r)] \tag{B.29}$$

is bounded below by $-U(0) = 0$ and is a nondecreasing function of p. Also,

$$0 \le U^*(U'(1/M)) \le U^*(U'(M)) = MU'(M) - U(M) < \infty. \tag{B.30}$$

Thus U^* is bounded on $[U'(1/M), U'(M)]$. Since it is also lower semicontinuous and convex, it is actually continuous on that interval. So $\{U^*(\psi_k)\}_{k=1}^\infty$ converges ν-almost everywhere to $U^*(\overline{\varphi}_M)$. By dominated convergence,

$$\lim_{k\to\infty} \int_X U^*(\psi_k)\, d\nu = \int_X U^*(\overline{\varphi}_M)\, d\nu. \tag{B.31}$$

Also by dominated convergence,

$$\lim_{k\to\infty} \int_X \psi_k\, d\mu = \int_X \overline{\varphi}_M\, d\mu. \tag{B.32}$$

We can conclude by choosing $\varphi_M = \psi_k$ for some large k. □

B.2. *Lower semicontinuity and contraction.* The following theorem is an easy consequence of the duality formulas established above.

THEOREM B.33. *Let X and U satisfy the assumptions of Theorem* B.2. *Then*

(i) *$U_\nu(\mu)$ is a lower semicontinuous function of $(\mu, \nu) \in P(X) \times P(X)$. That is, if $\{\mu_k\}_{k=1}^\infty$ and $\{\nu_k\}_{k=1}^\infty$ are sequences in $P(X)$ with $\lim_{k\to\infty} \mu_k = \mu$ and $\lim_{k\to\infty} \nu_k = \nu$ in the weak-$*$ topology then*

$$U_\nu(\mu) \le \liminf_{k\to\infty} U_{\nu_k}(\mu_k). \tag{B.34}$$

(ii) *$U_\nu(\mu)$ is nonincreasing under pushforward. That is, if Y is a compact Hausdorff space and $f : X \to Y$ is a Borel map then*

$$U_{f_*\nu}(f_*\mu) \le U_\nu(\mu). \tag{B.35}$$

Proof. Using the second representation formula in Theorem B.2, we can write

$$U_\nu(\mu) = \sup_{(L_1, L_2)\in\mathcal{L}} [L_1(\mu) + L_2(\nu)], \tag{B.36}$$

where $\mathcal{L}$ is a certain subset of $C(X) \oplus C(X)$, and L_1 and L_2 define continuous linear functionals on the space of measures $C(X)^*$. As the supremum of a set of lower semicontinuous functions (in particular linear functions) is lower semicontinuous, it follows that U_ν is lower semicontinuous in (μ, ν).

To prove part (ii), we use the first representation formula in Theorem B.2 to obtain

(B.37)

$$\begin{aligned} U_{f_*\nu}(f_*\mu) &= \sup \left\{ \int_Y \varphi\, d(f_*\mu) - \int_Y U^*(\varphi)\, d(f_*\nu);\;\; \varphi \in L^\infty(Y),\ \varphi \le U'(\infty) \right\} \\ &= \sup \left\{ \int_X (\varphi \circ f)\, d\mu - \int_X U^*(\varphi \circ f)\, d\nu;\;\; \varphi \in L^\infty(Y),\ \varphi \le U'(\infty) \right\}. \end{aligned}$$

If $\varphi \in L^\infty(Y)$ and $\varphi \leq U'(\infty)$ then $\varphi \circ f \in L^\infty(X)$ and $\varphi \circ f \leq U'(\infty)$. So the above supremum is bounded above by

$$U_\nu(\mu) = \sup \left\{ \int_X \psi \, d\mu - \int_X U^*(\psi) \, d\nu; \quad \psi \in L^\infty(X), \, \psi \leq U'(\infty) \right\}. \tag{B.38}$$

□

Appendix C. Approximation in $P_2(X)$

In this section we show how to effectively approximate a measure $\mu \in P_2(X, \nu)$ by measures $\{\mu_k\}_{k=1}^\infty$ whose densities, with respect to ν, are continuous. The approximation will be such that $\lim_{k\to\infty} \mu_k = \mu$ in the weak-$*$ topology and $\lim_{k\to\infty} U_\nu(\mu_k) = U_\nu(\mu)$.

We first construct a mollification operator on measures, in terms of a partition of unity for X. We then use finer and finer partitions of unity to construct the sequence $\{\mu_k\}_{k=1}^\infty$.

C.1. *Mollifiers.* Let (X, d, ν) be a compact metric space equipped with a reference Borel probability measure ν. Let $K : X \times X \to [0, \infty)$ be a symmetric continuous kernel satisfying

$$\forall x \in \mathrm{supp}(\nu), \quad \int_X K(x, y) \, d\nu(y) = 1. \tag{C.1}$$

For $\rho \in L^1(X, \nu)$, define $K\rho \in C(X)$ by

$$(K\rho)(x) = \int_X K(x, y) \, \rho(y) \, d\nu(y). \tag{C.2}$$

Note that $\int_X K\rho \, d\nu = \int_X \rho \, d\nu$ and the map $\rho \to 1_{\mathrm{supp}(\nu)} K\rho$ is a bounded operator on $L^1(X, \nu)$ with norm 1. For $\mu \in P_2(X, \nu)$, define $K\mu \in P_2(X, \nu)$ by saying that for $f \in C(\mathrm{supp}(\nu))$,

$$\int_{\mathrm{supp}(\nu)} f d(K\mu) = \int_X (Kf) d\mu. \tag{C.3}$$

More explicitly,

$$K\mu = \left(\int_X K(\cdot, y) d\mu(y) \right) \nu. \tag{C.4}$$

In particular, $K\mu \in P_2^{\mathrm{ac}}(X, \nu)$ is the product of a continuous function on X with ν. The notation is consistent, in the sense that if $\rho \in L^1(X, \nu)$ then $K(\rho \, d\nu) = K(\rho) \, d\nu$. Moreover, taking $f = 1$ in (C.3), it follows that if μ is a probability measure then $K\mu$ is a probability measure.

THEOREM C.5. *Let (X,d) be a compact metric space equipped with a reference Borel probability measure ν. Then there is a sequence $\{K_I\}_{I=1}^{\infty}$ of continuous nonnegative kernels with the following properties*:

(i) $\forall x, y \in X, \quad K_I(x,y) = K_I(y,x)$.

(ii) $\forall x \in \mathrm{supp}(\nu), \quad \int_X K_I(x,y)\, d\nu(y) = 1$.

(iii) *There is a sequence $\{\varepsilon_I\}_{I=1}^{\infty}$ converging to zero so that $K_I(x,y) = 0$ whenever $d(x,y) \geq \varepsilon_I$.*

(iv) *For all $\mu \in P_2(X,\nu)$, $\lim_{I\to\infty} K_I\mu = \mu$ in the weak-$*$ topology.*

Proof. Let $\mathcal{C} \equiv \{U_j\}$ be a finite open cover of X. Let $\{\phi_j\}$ be a subordinate partition of unity. Put

$$K_{\mathcal{C}}(x,y) = \sum_{j:\int_X \phi_j\, d\nu > 0} \frac{\phi_j(x)\phi_j(y)}{\int_X \phi_j\, d\nu}. \tag{C.6}$$

Then $K_{\mathcal{C}}(x,y) = K_{\mathcal{C}}(y,x)$. If $x \in \mathrm{supp}(\nu)$ and $\phi_j(x) > 0$ then $\int_X \phi_j\, d\nu > 0$, so that $\int_X K_{\mathcal{C}}(x,y)\, d\nu(y) = 1$. Properties (i) and (ii) are ensured if our sequence is made of such kernels. Moreover, (iii) will be satisfied if each ϕ_j has support in a small ball of radius $\varepsilon_I/2$.

Given $\delta > 0$, let $\mathcal{C}_\delta$ denote a finite cover of X by δ-balls. Given $f \in C(\mathrm{supp}(\nu))$ and $\varepsilon > 0$, suppose that $\delta > 0$ is such that $|f(x') - f(x)| \leq \varepsilon$ whenever $x', x \in \mathrm{supp}(\nu)$ satisfy $d(x',x) \leq 2\delta$. With such a cover $\mathcal{C}_\delta$, if $x \in \mathrm{supp}(\nu)$ has $\phi_j(x) > 0$ then

$$\left| \frac{\int_X f\phi_j\, d\nu}{\int_X \phi_j\, d\nu} - f(x) \right| \leq \varepsilon. \tag{C.7}$$

As

$$(K_{\mathcal{C}_\delta} f)(x) = \sum_{j:\int_X \phi_j\, d\nu > 0} \phi_j(x) \frac{\int_X f\phi_j\, d\nu}{\int_X \phi_j\, d\nu} = \sum_{j:\phi_j(x) > 0} \phi_j(x) \frac{\int_X f\phi_j\, d\nu}{\int_X \phi_j\, d\nu} \tag{C.8}$$

and

$$f(x) = \sum_j \phi_j(x) f(x) = \sum_{j:\phi_j(x)>0} \phi_j(x) f(x), \tag{C.9}$$

it follows that

$$|(K_{\mathcal{C}_\delta} f)(x) - f(x)| \leq \varepsilon. \tag{C.10}$$

Thus $\lim_{I\to\infty} K_{\mathcal{C}_{1/I}} f = f$ in $C(\mathrm{supp}(\nu))$.

Now put $K_I = K_{\mathcal{C}_{1/I}}$. For any $f \in C(\mathrm{supp}(\nu))$,

$$\lim_{I\to\infty} \int_X f dK_I\mu = \lim_{I\to\infty} \int_X (K_I f) d\mu = \int_X f d\mu. \tag{C.11}$$

This proves (iv). $\square$

C.2. *Approximation by continuous densities.*

THEOREM C.12. *Let U be a continuous convex function on $[0,\infty)$ with $U(0)=0$. Let (X,d) be a compact metric space equipped with a Borel probability measure ν. Let $\mu \in P_2(X,\nu)$ satisfy $U_\nu(\mu) < \infty$. Then there is a sequence $\{f_k\}_{k=1}^\infty$ in $C(X)$ such that $\lim_{k\to\infty} f_k\nu = \mu$ in the weak-$*$ topology and $\lim_{k\to\infty} U_\nu(f_k\nu) = U_\nu(\mu)$.*

Proof. We introduce a sequence of mollifying kernels K_I as above. We will prove that $\{K_I\mu\}_{I=1}^\infty$ does the job. Each $K_I\mu$ is the product of a continuous function on X with ν. Theorem C.5(iv) gives $\lim_{I\to\infty} K_I\mu = \mu$. By Theorem B.33(i), $U_\nu(\mu) \le \liminf_{I\to\infty} U_\nu(K_I\mu)$. Hence it suffices to show that $U_\nu(K_I\mu) \le U_\nu(\mu)$ for all I.

Before giving the general proof, it might be enlightening first to consider two "extreme" cases.

First particular case: *μ is absolutely continuous.* Assume that $\mu = \rho\nu$. We write $K_I\mu(x)$ for the density of $K_I\mu$ with respect to ν. By Jensen's inequality, for $x \in \mathrm{supp}(\nu)$,

(C.13)
$$U(K_I\mu(x)) = U\left(\int_X K_I(x,y)\,\rho(y)\,d\nu(y)\right) \le \int_X K_I(x,y)U(\rho(y))\,d\nu(y).$$

Then

$$U_\nu(K_I\mu) = \int_X U(K_I\mu(x))\,d\nu(x) \le \int_X U(\rho)\,d\nu. \tag{C.14}$$

Second particular case: *μ is completely singular.* With the notation used before, $\mu = \mu_s$. We write $K_I\mu_s(x)$ for the density of $K_I\mu_s$ with respect to ν. Then

$$\int_X U((K_I\mu_s)(x))\,d\nu(x) \le U'(\infty)\int_X (K_I\mu_s)(x)\,d\nu(x) = U'(\infty)\mu_s(X). \tag{C.15}$$

General case. Now we introduce the Lebesgue decomposition $\mu = \rho\nu + \mu_s$. In view of the first particular case above, we may assume that $\mu_s(X) > 0$. If $U'(\infty) = \infty$ then $U_\nu(\mu) = \infty$, so we can restrict to the case $U'(\infty) < \infty$.

We write $K_I\mu(x)$ for the density of $K_I\mu$ with respect to ν, and similarly for $K_I\mu_s(x)$. For all $\theta \in (0,1)$, there is a pointwise inequality

$$U(K_I\mu) = U(K_I\rho + K_I\mu_s) \le \theta U\left(\frac{K_I\rho}{\theta}\right) + (1-\theta)U\left(\frac{K_I\mu_s}{1-\theta}\right). \tag{C.16}$$

Then

(C.17)

$$U_\nu(K_I\mu) = \int_X U(K_I\mu(x))\, d\nu(x) \le \theta \int_X U\left(\frac{K_I\rho}{\theta}\right) d\nu + U'(\infty) \int_X K_I\mu_s(x)\, d\nu(x)$$
$$= \theta \int_X U\left(\frac{K_I\rho}{\theta}\right) d\nu + U'(\infty)\mu_s(X).$$

As $K_I\rho \in C(X)$,

$$\lim_{\theta\to 1^-} \left(\theta \int_X U\left(\frac{K_I\rho}{\theta}\right) d\nu\right) = \int_X U(K_I\rho)\, d\nu. \tag{C.18}$$

As in the proof of the first particular case,

$$\int_X U(K_I\rho)\, d\nu \le \int_X U(\rho)\, d\nu. \tag{C.19}$$

Combining (C.17), (C.18) and (C.19) gives $U_\nu(K_I\mu) \le U_\nu(\mu)$. □

Appendix D. Hessian calculations

How can one check, in practice, the displacement convexity of a given functional on $P_2(X)$, say in the case when X is a smooth compact connected Riemannian manifold M? We provide below some explicit computations to that effect, following [36]. The computations are purely formal and we do not rigorously justify them, ignoring all regularity issues. Although formal, these computations motivate the definition of nonnegative N-Ricci curvature in terms of displacement convexity.

Denote by dvol_M the Riemannian density on M, and introduce a reference measure

$$d\nu = e^{-\Psi}\, \mathrm{dvol}_M, \tag{D.1}$$

with $\Psi \in C^\infty(M)$ satisfying $\int_M e^{-\Psi}\, \mathrm{dvol}_M = 1$.

The direction vector along a curve $\{\mu_t\}$ in the space of probability measures $P(M)$ can be "represented" as

$$\frac{\partial\mu}{\partial t} = -\nabla \cdot (\mu\nabla\Phi), \tag{D.2}$$

where $\Phi \equiv \Phi(t)$ is a function on M that is defined up to constants. The meaning of (D.2) is that

$$\frac{d}{dt}\int_M \xi d\mu = \int_M \nabla\xi \cdot \nabla\Phi d\mu \tag{D.3}$$

for all $\xi \in C^\infty(M)$. Thus we can parametrize the tangent space $T_\mu P(M)$ by the function Φ. Otto's formal inner product on $T_\mu P(M)$ is given by the quadratic form $\int_M \langle\nabla\Phi, \nabla\Phi\rangle d\mu$. The function Φ here is related to the function ϕ of Section 2.4 by $\Phi = -\phi$.

With this Riemannian metric, the geodesic equation in $P(M)$ becomes

$$\partial_t \Phi + \frac{1}{2}|\nabla \Phi|^2 = 0. \tag{D.4}$$

It has the solution

$$\Phi(t)(x) = \inf_{y \in X} \left[\Phi(0)(y) + \frac{d(x,y)^2}{2t} \right]. \tag{D.5}$$

The corresponding length metric on $P(M)$ is formally the Wasserstein metric.

Let U be a continuous convex function on $[0, \infty)$ that is C^2-regular on $(0, \infty)$. Put

$$E(\mu) = \int_M U\left(\frac{d\mu}{d\nu}\right) d\nu. \tag{D.6}$$

Recall that

$$p(r) = rU'(r) - U(r), \qquad p_2(r) = rp'(r) - p(r). \tag{D.7}$$

Then along a curve $\{\mu_t\}$ in $P(M)$, the derivative of $E(\mu_t)$ is given by

$$\begin{aligned} \frac{dE}{dt} &= \int_M U'\left(\frac{d\mu}{d\nu}\right) \partial_t \frac{d\mu}{d\nu}\, d\nu \\ &= \int_M \nabla U'\left(\frac{d\mu}{d\nu}\right) \cdot \nabla \Phi d\mu \\ &= \int_M \nabla \Phi \cdot \left(\frac{d\mu}{d\nu} U''\left(\frac{d\mu}{d\nu}\right) \nabla \frac{d\mu}{d\nu}\right) d\nu \\ &= \int_M \nabla \Phi \cdot \nabla p\left(\frac{d\mu}{d\nu}\right) d\nu \\ &= \int_M (-\triangle \Phi + \nabla \Psi \cdot \nabla \Phi) p\left(\frac{d\mu}{d\nu}\right) d\nu. \end{aligned} \tag{D.8}$$

Parametrizing $T_\mu P(M)$ by $\{\Phi\}$, equation (D.8) shows in particular that grad E is represented by the function $U'\left(\frac{d\mu}{d\nu}\right)$.

To compute the Hessian of E, we assume that $\{\mu_t\}$ is a geodesic curve in $P_2(M)$. Then

$$\begin{aligned} \frac{d^2E}{dt^2} = \int_M &\left(\triangle\left(\frac{1}{2}|\nabla\Phi|^2\right) - \nabla\Psi \cdot \nabla\left(\frac{1}{2}|\nabla\Phi|^2\right)\right) p\left(\frac{d\mu}{d\nu}\right) d\nu \\ &+ \nabla\left((-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\, p'\left(\frac{d\mu}{d\nu}\right)\right) \cdot \nabla\Phi d\mu. \end{aligned} \tag{D.9}$$

Now

(D.10)

$$\begin{aligned}
&\int_M \nabla \left((-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\, p'\left(\frac{d\mu}{d\nu}\right)\right)\cdot\nabla\Phi d\mu \\
&= \int_M \nabla \left((-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\, p'\left(\frac{d\mu}{d\nu}\right)\right)\cdot\nabla\Phi \frac{d\mu}{d\nu}\, d\nu \\
&= \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\, p'\left(\frac{d\mu}{d\nu}\right)\left(-\nabla\cdot\left(\nabla\Phi\frac{d\mu}{d\nu}\right) + \nabla\Psi\cdot\nabla\Phi\frac{d\mu}{d\nu}\right) d\nu \\
&= \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\, p'\left(\frac{d\mu}{d\nu}\right)\frac{d\mu}{d\nu}\, d\nu \\
&\quad - \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\,\nabla\Phi\cdot\nabla p\left(\frac{d\mu}{d\nu}\right)\, d\nu \\
&= \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\, p'\left(\frac{d\mu}{d\nu}\right)\frac{d\mu}{d\nu}\, d\nu \\
&\quad + \int_M [\nabla(-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\cdot\nabla\Phi + (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\cdot\triangle\Phi \\
&\quad - \nabla\Psi\cdot(-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\,\nabla\Phi]\, p\left(\frac{d\mu}{d\nu}\right)\, d\nu \\
&= \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\, p'\left(\frac{d\mu}{d\nu}\right)\frac{d\mu}{d\nu}\, d\nu \\
&\quad + \int_M \left[\nabla(-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\cdot\nabla\Phi - (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\right] p\left(\frac{d\mu}{d\nu}\right)\, d\nu.
\end{aligned}$$

Combining (D.9) and (D.10) gives

$$\begin{aligned}
\frac{d^2E}{dt^2} = &\int_M \Big[\triangle\left(\frac{1}{2}|\nabla\Phi|^2\right) - \nabla\Psi\cdot\nabla\left(\frac{1}{2}|\nabla\Phi|^2\right) \\
&\qquad + \nabla(-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)\cdot\nabla\Phi \\
&\qquad - (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\Big] p\left(\frac{d\mu}{d\nu}\right)\, d\nu \\
&+ \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\, p'\left(\frac{d\mu}{d\nu}\right)\frac{d\mu}{d\nu}\, d\nu \\
= &\int_M \left[|\operatorname{Hess}\Phi|^2 + \nabla\Phi\cdot(\operatorname{Ric} + \operatorname{Hess}\Psi)\nabla\Phi\right] p\left(\frac{d\mu}{d\nu}\right)\, d\nu \\
&+ \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\left(-p\left(\frac{d\mu}{d\nu}\right) + p'\left(\frac{d\mu}{d\nu}\right)\frac{d\mu}{d\nu}\right) d\nu \\
= &\int_M \left[|\operatorname{Hess}\Phi|^2 + \nabla\Phi\cdot(\operatorname{Ric} + \operatorname{Hess}\Psi)\nabla\Phi\right] p\left(\frac{d\mu}{d\nu}\right)\, d\nu \\
&+ \int_M (-\triangle\Phi + \nabla\Psi\cdot\nabla\Phi)^2\, p_2\left(\frac{d\mu}{d\nu}\right)\, d\nu.
\end{aligned} \tag{D.11}$$

In particular, if $U \in \mathcal{DC}_N$ with $n < N$ then from Lemma 5.7(b),

(D.12)
$$\begin{aligned}
\frac{d^2 E}{dt^2} &\geq \int_M \Big[|\operatorname{Hess} \Phi|^2 + \nabla \Phi \cdot (\operatorname{Ric} + \operatorname{Hess} \Psi) \nabla \Phi \\
&\qquad\qquad - \frac{1}{N} \left(-\triangle \Phi + \nabla \Psi \cdot \nabla \Phi \right)^2 \Big] p \left(\frac{d\mu}{d\nu} \right) d\nu \\
&\geq \int_M \Big[\frac{1}{n} (\triangle \Phi)^2 + \nabla \Phi \cdot (\operatorname{Ric} + \operatorname{Hess} \Psi) \nabla \Phi \\
&\qquad\qquad - \frac{1}{N} \left(-\triangle \Phi + \nabla \Psi \cdot \nabla \Phi \right)^2 \Big] p \left(\frac{d\mu}{d\nu} \right) d\nu \\
&\geq \int_M \Big[\nabla \Phi \cdot (\operatorname{Ric} + \operatorname{Hess} \Psi) \nabla \Phi - \frac{1}{N-n} \left(\nabla \Psi \cdot \nabla \Phi \right)^2 \Big] p \left(\frac{d\mu}{d\nu} \right) d\nu \\
&= \int_M \Big[\nabla \Phi \cdot \Big(\operatorname{Ric} + \operatorname{Hess} \Psi - \frac{1}{N-n} \nabla \Psi \otimes \nabla \Psi \Big) \nabla \Phi \Big] p \left(\frac{d\mu}{d\nu} \right) d\nu \\
&= \int_M \Big[\nabla \Phi \cdot \operatorname{Ric}_N (\nabla \Phi) \Big] p \left(\frac{d\mu}{d\nu} \right) d\nu.
\end{aligned}$$

The same conclusion applies if $N = n$ and $\nabla \Psi = 0$, in which case $\operatorname{Ric}_N = \operatorname{Ric}$.

If $\operatorname{Ric}_N \geq K g_M$ then

$$\frac{d^2 E}{dt^2} \geq K \int_M |\nabla \Phi|^2 \frac{p\left(\frac{d\mu}{d\nu}\right)}{\frac{d\mu}{d\nu}} d\mu. \tag{D.13}$$

If $K = 0$ then (D.13) gives $\frac{d^2 E}{dt^2} \geq 0$. That is, E is formally convex on $P_2(M)$, no matter what the value of N is.

If $N = \infty$ then (D.13) gives

$$\frac{d^2 E}{dt^2} \geq \left(\inf_{r > 0} K \frac{p(r)}{r} \right) \int_M |\nabla \Phi|^2 d\mu = \lambda(U) \int_M |\nabla \Phi|^2 d\mu. \tag{D.14}$$

As $\int_M |\nabla \Phi|^2 d\mu$ is the square of the speed of the geodesic, it follows that E is formally $\lambda(U)$-convex on $P_2(M)$.

Appendix E. The noncompact case

In the preceding part of the paper we worked with compact spaces X. We now discuss how to adapt our arguments to (possibly noncompact) pointed metric spaces. To avoid expanding the size of this section too much, we sometimes simplify the proofs by slightly restricting the generality of the discussion, and we give details mainly for the case of nonnegative N-Ricci curvature with $N < \infty$.

E.1. *Pointed spaces.* In this section we will always assume our metric spaces have distinguished basepoints. In other words, the objects under study will be complete pointed metric spaces, which are pairs $(X, \star)$ where X is a complete metric space and $\star \in X$. A map f between pointed spaces $(X_1, \star_1)$ and $(X_2, \star_2)$ is said to be a pointed map if $f(\star_1) = f(\star_2)$. In this setting, the analog of Gromov-Hausdorff convergence is the following:

Definition E.1. Let $\{(X_i, \star_i)\}_{i=1}^\infty$ be a sequence of complete pointed metric spaces. It converges to a complete pointed metric space $(X, \star)$ in the *pointed Gromov-Hausdorff topology*, by means of pointed approximations $f_i : X_i \to X$, if for every $R > 0$ there is a sequence $\{\varepsilon_{R,i}\}_{i=1}^\infty$ of positive numbers converging to zero so that

1. For all $x_i, y_i \in B_R(\star_i)$, we have $|d_X(f_i(x_i), f_i(y_i)) - d_{X_i}(x_i, y_i)| \le \varepsilon_{R,i}$.
2. For all $x \in B_R(\star)$, there is some $x_i \in B_R(\star_i)$ so that $d_X(f_i(x_i), x) \le \varepsilon_{R,i}$.

A more usual definition would involve approximations defined just on balls in X_i, instead of all of X_i. It is convenient for us to assume that f_i is defined on all of X_i, as will be seen when defining maps between Wasserstein spaces. The notions of convergence are equivalent.

Next, a pointed metric-measure space is a complete pointed metric space $(X, \star)$ equipped with a nonnegative nonzero Radon measure ν. We do not assume that ν has finite mass. In this context, a pointed map $f : X_1 \to X_2$ will be assumed to be Borel, with the preimage of a compact set being precompact. (This ensures that the pushforward of a Radon measure is a Radon measure.)

Definition E.2. Let $\{(X_i, \star_i, \nu_i)\}_{i=1}^\infty$ be a sequence of complete pointed locally compact metric-measure spaces. It is said to converge to a complete pointed locally compact metric-measure space $(X, \star, \nu)$ in the *pointed measured Gromov-Hausdorff topology* if $\{(X_i, \star_i)\}_{i=1}^\infty$ converges to $(X, \star)$ in the pointed Gromov-Hausdorff topology by means of pointed approximations $f_i : X_i \to X$ which additionally satisfy $\lim_{i\to\infty} (f_i)_* \nu_i = \nu$ in the weak-$*$ topology on $C_c(X)^*$.

The pointed measured Gromov-Hausdorff topology was used, for example, in [18]. In what follows we will examine its compatibility with the Wasserstein space.

E.2. *Wasserstein space.* If X is a complete pointed metric space, possibly noncompact, let $P_2(X)$ be the space of probability measures μ on X with a finite second moment, namely

$$P_2(X) = \left\{ \mu \in P(X) : \int_X d(\star, x)^2 d\mu(x) < \infty \right\}. \tag{E.3}$$

One can still introduce the Wasserstein distance W_2 by the same formula as in (1.14). Then W_2 is a well-defined metric on $P_2(X)$ [44, Th. 7.3]. The metric space $(P_2(X), W_2)$ will be called the 2-Wasserstein space associated to X. It does not depend on the choice of basepoint $\star \in X$.

We will assume that X is a complete separable metric space, in which case $P_2(X)$ is also a complete separable metric space. Put

$$\text{(E.4)} \qquad \big(1 + d(\star, \cdot)^2\big) C_b(X) = \left\{ f \in C(X) : \sup_{x \in X} \frac{|f(x)|}{1 + d(\star, x)^2} < \infty \right\}.$$

Then $(1 + d(\star, \cdot)^2) C_b(X)$ is a Banach space with norm

$$\text{(E.5)} \qquad \parallel f \parallel = \sup_{x \in X} \frac{|f(x)|}{1 + d(\star, x)^2},$$

and the underlying topological vector space is independent of the choice of basepoint $\star$. The dual space $\big((1 + d(\star, \cdot)^2) C_b(X)\big)^*$ contains $P_2(X)$ as a subset. As such, $P_2(X)$ inherits a subspace topology from the weak-$*$ topology on $\big((1 + d(\star, \cdot)^2) C_b(X)\big)^*$, which turns out to coincide with the topology on $P_2(X)$ arising from the metric W_2 [44, Th. 7.12]. (If X is noncompact then $P_2(X)$ is not a closed subset of $\big((1 + d(\star, \cdot)^2) C_b(X)\big)^*$.) A subset $S \subset P_2(X)$ is relatively compact if and only if it satisfies the "tightness" condition that for every $\varepsilon > 0$, there is some $R > 0$ so that for all $\mu \in S$, $\int_{X - B_R(\star)} d(\star, x)^2 d\mu(x) < \varepsilon$ [44, Th. 7.12(ii)]. Applying this to a ball in $P_2(X)$ around $\delta_\star$, it follows that if X is noncompact then $P_2(X)$ is not locally compact, while if X is compact then $P_2(X)$ is compact.

If X is a complete locally compact length space then for all $k > 0$, $\mathrm{Lip}_k([0, 1], X)$ is locally compact, with the set of geodesic paths between any two given points in X forming a compact subset. Then the proof of Proposition 2.6 carries through to show that $P_2(X)$ is a length space. Finally, if X is pointed then there is a distinguished basepoint in $P_2(X)$, namely the Dirac mass $\delta_\star$.

The next proposition is an extension of Proposition 2.10.

Proposition E.6. *Let $(X, \star)$ be a complete pointed locally compact length space and let $\{\mu_t\}_{t \in [0,1]}$ be a geodesic path in $P_2(X)$. Then there exists some optimal dynamical transference plan Π such that $\{\mu_t\}_{t \in [0,1]}$ is the displacement interpolation associated to Π.*

Proof. We can go through the proof of Proposition 2.10, constructing the probability measures $R^{(i)}$ with support on the locally compact space Γ. For each i, we have $(e_0)_* R^{(i)} = \mu_0$ and $(e_1)_* R^{(i)} = \mu_1$. In order to construct a weak-$*$ accumulation point $R^{(\infty)}$, as a probability measure on Γ, it suffices to show that for each $\varepsilon > 0$ there is a compact set $K \subset \Gamma$ so that for all i, $R^{(i)}(\Gamma - K) < \varepsilon$.

Let $E : \Gamma \to X \times X$ be the endpoints map. Given $r > 0$, put $K = E^{-1}(\overline{B_r(\star)} \times \overline{B_r(\star)})$, a compact subset of Γ. As

$$\Gamma - K = E^{-1}\left(\left((X - \overline{B_r(\star)}) \times X\right) \cup \left(X \times (X - \overline{B_r(\star)})\right)\right), \tag{E.7}$$

we have

$$\begin{aligned} R^{(i)}(\Gamma - K) &\leq (E_* R^{(i)})((X - \overline{B_r(\star)}) \times X) + (E_* R^{(i)})(X \times (X - \overline{B_r(\star)})) \\ &= \mu_0(X - \overline{B_r(\star)}) + \mu_1(X - \overline{B_r(\star)}). \end{aligned} \tag{E.8}$$

Taking r sufficiently large, we can ensure that $\mu_0(X - \overline{B_r(\star)}) + \mu_1(X - \overline{B_r(\star)}) < \varepsilon$. □

Using Proposition E.6, we show that geodesics with endpoints in a given compact subset of $P_2(X)$ will all lie in a compact subset of $P_2(X)$.

Proposition E.9. *For any compact set $K \subset P_2(X)$, there is a compact set $K' \subset P_2(X)$ with the property that for any $\mu_0, \mu_1 \in K$, if $\{\mu_t\}_{t\in[0,1]}$ is a Wasserstein geodesic from μ_0 to μ_1 then $\mu_t \in K'$ for all $t \in [0,1]$.*

Proof. Given $\mu_0, \mu_1 \in K$, let $\{\mu_t\}_{t\in[0,1]}$ be a Wasserstein geodesic from μ_0 to μ_1. Then

$$\begin{aligned} \int_{X - B_R(\star)} d(\star, x)^2 d\mu_t(x) &= \int_\Gamma d(\star, \gamma(t))^2 1_{\gamma(t) \notin B_R(\star)} d\Pi(\gamma) \\ &= \int_\Gamma d(\star, \gamma(t))^2 1_{\gamma(t) \notin B_R(\star)} 1_{\max(d(\star,\gamma(0)), d(\star,\gamma(1))) \geq \frac{R}{2}} d\Pi(\gamma), \end{aligned} \tag{E.10}$$

where Π comes from Proposition E.6. We break up the integral in the last term of (E.10) into two pieces according to whether $d(\star, \gamma(0)) \leq d(\star, \gamma(1))$ or $d(\star, \gamma(1)) < d(\star, \gamma(0))$. If $d(\star, \gamma(0)) \leq d(\star, \gamma(1))$ then

$$\begin{aligned} d(\star, \gamma(t)) &\leq d(\star, \gamma(0)) + d(\gamma(0), \gamma(t)) \leq d(\star, \gamma(0)) + d(\gamma(0), \gamma(1)) \\ &\leq 2d(\star, \gamma(0)) + d(\star, \gamma(1)) \leq 3d(\star, \gamma(1)). \end{aligned} \tag{E.11}$$

Then the contribution to the last term of (3.11), when $d(\star, \gamma(0)) \leq d(\star, \gamma(1))$, is bounded above by

$$9 \int_\Gamma d(\star, \gamma(1))^2 1_{d(\star,\gamma(1)) \geq \frac{R}{2}} d\Pi(\gamma) = 9 \int_{X - B_{R/2}(\star)} d(\star, x)^2 d\mu_1(x). \tag{E.12}$$

Adding the contribution when $d(\star, \gamma(1)) < d(\star, \gamma(0))$ gives

$$\begin{aligned} \int_{X - B_R(\star)} d(\star, x)^2 d\mu_t(x) \leq 9 &\int_{X - B_{R/2}(\star)} d(\star, x)^2 d\mu_0(x) \\ +9 &\int_{X - B_{R/2}(\star)} d(\star, x)^2 d\mu_1(x). \end{aligned} \tag{E.13}$$

For any $\varepsilon > 0$ we can choose $R > 0$ so that the right-hand-side of (E.13) is bounded above by ε, uniformly in $\mu_0, \mu_1 \in K$. This proves the proposition. □

Remark E.14. Although we will consider optimal transport between elements of $P_2(X)$, there are also interesting issues concerning "optimal transport" in a generalized sense, with possibly infinite cost, on the whole of $P(X)$. For example, one has McCann's theorem about existence of "generalized optimal transport" between arbitrary probability measures on $\mathbb{R}^n$ [44, Th. 2.32].

E.3. *Functionals.* In the nonpointed part of the paper we dealt with a compact measured length space (X, d, ν), with the background measure ν lying in $P_2(X)$. When generalizing to the case when X is pointed and possibly noncompact, one's first inclination might be to again have ν lie in $P_2(X)$. This is indeed the appropriate choice for some purposes, such as to extend the functional analytic results of Sections 3.3 and 6. However, requiring ν to lie in $P_2(X)$ would rule out such basic cases as $X = \mathbb{R}^N$ with the Lebesgue measure. Additionally, it would preclude the tangent cone construction for a compact space with N-Ricci curvature bounded below. Because of this, in what follows we will allow ν to have infinite mass, at the price of some additional complications

Let ν be a nonnegative nonzero Radon measure on X. Let U be a continuous convex function on $[0, \infty)$ with $U(0) = 0$. One would like to define the functional U_ν as in Definition 3.2, but this requires some care. Even if we use (3.3) to define $U_\nu(\mu)$ for $\mu = \rho\nu$ and $\rho \in C_c(X)$, in general there is no way to extend U_ν to a lower semicontinuous function on $P_2(X)$. A way to circumvent this difficulty is to impose a growth assumption on ν.

Definition E.15. For $k > 0$, we define $M_{-k}(X)$ to be the space of nonnegative Radon measures ν on X such that

$$\int_X (1 + d(\star, x)^2)^{-\frac{k}{2}} \, d\nu(x) < \infty. \tag{E.16}$$

Equivalently, ν is a nonnegative Radon measure on X that lies in the dual space of $(1 + d(\star, \cdot)^2)^{-\frac{k}{2}} C_b(X)$. We further define $M_{-\infty}(X)$ by the condition $\int_X e^{-cd(x,*)^2} \, d\nu(x) < \infty$, where c is a fixed positive constant.

Proposition E.17. *Let X be an arbitrary metric space. Given $N \in [1, \infty]$, suppose that $U \in \mathcal{DC}_N$ and $\nu \in M_{-2(N-1)}(X)$. Then U_ν is a well-defined functional on $P_2(X)$, with values in $\mathbb{R} \cup \{\infty\}$.*

Proof. Suppose first that $N < \infty$. From the definition of $\mathcal{DC}_N$, there is a constant $A > 0$ so that

$$\lambda^N U(\lambda^{-N}) \geq -A\lambda - A, \tag{E.18}$$

and so

$$U(\rho) \geq -A\left(\rho + \rho^{1-\frac{1}{N}}\right). \tag{E.19}$$

Of course, ρ lies in $L^1(X,\nu)$. To prove that $U_\nu(\mu)$ is well-defined, it suffices to show that $\rho^{1-1/N}$ also lies in $L^1(X,\nu)$. For that we use Hölder's inequality:

(E.20)

$$\begin{aligned}\int_X \rho(x)^{1-\frac{1}{N}}\,d\nu(x) &= \int_X \left((1+d(\star,x)^2)\rho(x)\right)^{1-\frac{1}{N}}(1+d(\star,x)^2)^{-1+\frac{1}{N}}\,d\nu(x)\\ &\leq \left(\int_X (1+d(\star,x)^2)\rho(x)\,d\nu(x)\right)^{1-\frac{1}{N}}\\ &\quad\cdot\left(\int_X (1+d(\star,x)^2)^{-(N-1)}\,d\nu(x)\right)^{\frac{1}{N}}.\end{aligned}$$

Now suppose that $N=\infty$. From Lemma 5.10, if U is nonlinear then there are constants $a,b>0$ so that

$$U(\rho) \geq a\rho\log\rho - b\rho. \tag{E.21}$$

Thus it is sufficient to show that $(\rho\log\rho)_- \in L^1(X,\nu)$. Applying Jensen's inequality with the probability measure $\frac{e^{-c\,d(\star,\cdot)^2}\,d\nu}{\int_X e^{-c\,d(\star,\cdot)^2}\,d\nu}$ gives

(E.22)

$$\begin{aligned}&\int_X \rho(x)\log(\rho(x))\,d\nu(x)\\ &= \int_X \rho(x)e^{cd(\star,x)^2}\log\left(\rho(x)e^{cd(\star,x)^2}\right)e^{-cd(\star,x)^2}\,d\nu(x) - c\int_X d(\star,x)^2\rho(x)\,d\nu(x)\\ &= \left(\int_X e^{-cd(\star,x)^2}\,d\nu(x)\right)\left(\int_X \rho(x)e^{cd(\star,x)^2}\log\left(\rho(x)e^{cd(\star,x)^2}\right)\frac{e^{-cd(\star,x)^2}\,d\nu(x)}{\int_X e^{-cd(\star,x)^2}\,d\nu(x)}\right)\\ &\quad - c\int_X d(\star,x)^2\rho(x)\,d\nu(x)\\ &\geq \left(\int_X e^{-cd(\star,x)^2}\,d\nu(x)\right)\left(\frac{\int_X \rho\,d\nu}{\int_X e^{-cd(\star,x)^2}\,d\nu(x)}\right)\log\left(\frac{\int_X \rho\,d\nu}{\int_X e^{-cd(\star,x)^2}\,d\nu(x)}\right)\\ &\quad - c\int_X d(\star,x)^2\rho(x)\,d\nu(x).\end{aligned}$$

This concludes the argument. □

E.4. *Approximation arguments.* Now we check that the technical results in Appendices B and C go through to the pointed locally compact case.

There is an obvious way to generalize the conclusion of Theorem B.2 by introducing the quantity

$$\sup_{\varphi\in(1+d(\star,\cdot)^2)L^\infty(X),\ \varphi\leq U'(\infty)}\left(\int_X \varphi\,d\mu - \int_X U^*(\varphi)\,d\nu\right). \tag{E.23}$$

We claim that this quantity is not $-\infty$ as long as ν lies in $M_{-2(N-1)}(X)$. To prove this, it suffices to exhibit a φ such that $\int \varphi\, d\mu > -\infty$ and $\int U^*(\varphi)\, d\nu < \infty$. It turns out that $\varphi(x) = -c\, d(\star, x)^2$ will do the job, taking into account the following lemma:

LEMMA E.24. *If $U \in \mathcal{DC}_N$ with $N < \infty$ then as $p \to -\infty$,*

$$U^*(p) = O\left((-p)^{1-N}\right). \tag{E.25}$$

If $U \in \mathcal{DC}_\infty$ then as $p \to -\infty$,

$$U^*(p) = O\left(e^p\right). \tag{E.26}$$

Proof. Suppose first that $U \in \mathcal{DC}_N$ with $N < \infty$. Then for p sufficiently negative, using (E.19) we have

$$0 \le U^*(p) = \sup_{r \ge 0} \left[pr - U(r)\right] \le \sup_{r \ge 0} \left[pr + 2Ar^{1-\frac{1}{N}}\right] = \text{const.}(-p)^{1-N}. \tag{E.27}$$

The case $N = \infty$ is similar. $\square$

The analog of Theorem B.2 becomes

$$\begin{aligned} U_\nu(\mu) &= \sup_{\varphi \in (1+d(\star,\cdot)^2)L^\infty(X),\, \varphi \le U'(\infty)} \left(\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu\right) \\ &= \sup_{M \in \mathbb{Z}^+} \sup_{\varphi \in (1+d(\star,\cdot)^2)C_b(X),\, \varphi \le U'(M)} \left(\int_X \varphi\, d\mu - \int_X U^*(\varphi)\, d\nu\right). \end{aligned} \tag{E.28}$$

The proof is similar to the proof of Theorem B.2, with the following modifications. Given $R > 0$ and $M \in \mathbb{Z}^+$ large numbers, we construct $\overline{\phi}_{R,M}$ on $B_R(\star)$ as in the proof of Theorem B.2. We extend $\overline{\phi}_{R,M}$ to X by setting it equal to $-d(\star,\cdot)^2$ on $X - B_R(\star)$. Then after passing to an appropriate subsequence of $\{\overline{\phi}_{R,M}\}$, the analog of (B.21) holds. As in the proof of Lusin's theorem, we can find a sequence $\{\psi_k\}$ of uniformly bounded continuous functions that converges pointwise to $\frac{\overline{\phi}_{R,M}}{1+d(\star,\cdot)^2}$. Considering the function $(1 + d(\star,\cdot)^2)\psi_k$ for large k proves the theorem.

Then Theorem B.33(i) extends, where μ lies in $P_2(X)$ and ν is a measure on X that lies in the dual space of $(1 + d(\star,\cdot)^2)^{-(N-1)}C_b(X)$, which we endow with the weak-$*$ topology.

Theorem B.33(ii) is a bit more subtle because we have to be careful about how the pushforward map f_* acts on the measures at spatial infinity. The discussion is easier when $N < \infty$, so from now on we restrict to this case. Then the statement in Theorem B.33(ii) goes through as soon as we require that the map f be a pointed Borel map satisfying

$$A^{-1}d_X(\star_X, x) - A \le d_Y(\star_Y, f(x)) \le A d_X(\star_X, x) + A \tag{E.29}$$

for some constant $A > 0$. This condition ensures that f_* maps $P_2(X)$ to $P_2(Y)$, and maps measures on X that lie in the dual space of $(1+d(\star_X,\cdot)^2)^{-(N-1)}C_b(X)$ to measures on Y that lie in the dual space of $(1+d(\star_Y,\cdot)^2)^{-(N-1)}C_b(Y)$.

Finally, we wish to extend Theorem C.12 to the pointed locally compact setting, replacing $C(X)$ by $C_c(X)$. Given $\delta > 0$, let $\{x_j\}$ be a maximal $\frac{\delta}{2}$-separated net in X. Then $\mathcal{C} = \{B_\delta(x_j)\}$ is an open cover of X. It is locally finite, as X is locally compact. If $\{\phi_j\}$ is a subordinate partition of unity then we define the operator $K_{\mathcal{C}}$ as in Appendix C. Given $\mu \in P_2(X,\nu)$, we claim that $K_{\mathcal{C}}\mu \in P_2(X,\nu)$. To see this, we write

(E.30)
$$\begin{aligned}\int_X (1+d(\star,x)^2)dK_{\mathcal{C}}\mu(x) &= \int_X (1+d(\star,x)^2) \sum_{j:\int_X \phi_j\, d\nu>0} \phi_j(x)\frac{\int_X \phi_j d\mu}{\int_X \phi_j\, d\nu}\, d\nu(x)\\ &= \int_X \sum_{j:\int_X \phi_j\, d\nu>0} \frac{\int_X (1+d(\star,\cdot)^2)\phi_j\, d\nu}{\int_X \phi_j\, d\nu}\phi_j(y)d\mu(y).\end{aligned}$$

From the choice of $\{x_j\}$ and $\{\phi_j\}$, there is a constant $C < \infty$ so that

$$\frac{\int_X (1+d(\star,\cdot)^2)\phi_j\, d\nu}{\int_X \phi_j\, d\nu} \le C(1+d(\star,x_j)^2) \tag{E.31}$$

for all j with $\int_X \phi_j\, d\nu > 0$. There is another constant $C' < \infty$ so that

$$(1+d(\star,x_j)^2)\phi_j(y) \le C'(1+d(\star,y)^2)\phi_j(y) \tag{E.32}$$

for all j and all $y \in X$, from which the claim follows.

Next, we claim that $U_\nu(K_{\mathcal{C}}\mu) \le U_\nu(\mu)$. We use the fact that $K_{\mathcal{C}}\mu$ is the product of a continuous function on X with ν. As in (C.16)–(C.18), for each $R > 0$ we have

(E.33)
$$\int_{B_R(\star)} U(K_{\mathcal{C}}\mu(x))\, d\nu(x) \le \int_{B_R(\star)} U(K_{\mathcal{C}}\rho)\, d\nu + U'(\infty)\int_{B_R(\star)} K_{\mathcal{C}}\mu_s(x)\, d\nu(x).$$

Taking $R \to \infty$ and applying the arguments of the particular cases in the proof of Theorem C.12 gives the claim.

For $R > 0$ and $\varepsilon > 0$, let $\phi_{R,\varepsilon} : X \to [0,1]$ be a continuous function which is one on $B_R(\star)$ and vanishes outside of $B_{R+\varepsilon}(\star)$. We can find sequences $\{\delta_k\}$, $\{\mathcal{C}_k\}$, $\{R_k\}$ and $\{\varepsilon_k\}$ so that $\lim_{k\to\infty} \frac{\phi_{R_k,\varepsilon_k}K_{\mathcal{C}_k}\mu}{\int_X \phi_{R_k,\varepsilon_k} d(K_{\mathcal{C}_k}\mu)} = \mu$ in $P_2(X,\nu)$ and

$$\limsup_{k\to\infty} U_\nu\left(\frac{\phi_{R_k,\varepsilon_k}K_{\mathcal{C}_k}\mu}{\int_X \phi_{R_k,\varepsilon_k} d(K_{\mathcal{C}_k}\mu)}\right) \le \limsup_{k\to\infty} U_\nu(K_{\mathcal{C}_k}\mu). \tag{E.34}$$

From the previously-shown lower semicontinuity of U_ν, we know that

$$U_\nu(\mu) \le \liminf_{k\to\infty} U_\nu\left(\frac{\phi_{R_k,\varepsilon_k}K_{\mathcal{C}_k}\mu}{\int_X \phi_{R_k,\varepsilon_k} d(K_{\mathcal{C}_k}\mu)}\right). \tag{E.35}$$

We have already show that $U_\nu(K_{\mathcal{C}_k}\mu) \leq U_\nu(\mu)$. Hence

$$U_\nu(\mu) = \lim_{k\to\infty} U_\nu \left(\frac{\phi_{R_k,\varepsilon_k} K_{\mathcal{C}_k}\mu}{\int_X \phi_{R_k,\varepsilon_k} d(K_{\mathcal{C}_k}\mu)} \right).$$

This proves the desired extension of Theorem C.12.

E.5. *Stability of N-Ricci curvature bounds.* We now define the notion of a complete, pointed, measured, locally compact, length space $(X, \star, \nu)$ having nonnegative N-Ricci curvature as in Definition 5.12, provided that $\nu \in M_{-2(N-1)}(X)$. Note that this notion is independent of the choice of basepoint.

Most of the geometric inequalities discussed in Sections 3.3, 5 and 6 have evident extensions to the pointed case. When discussing HWI, log Sobolev, Talagrand and Poincaré inequalities we assume that $\nu \in P_2(X)$. If X is a smooth Riemannian manifold such that the reference measure ν lies in $M_{-2(N-1)}(X)$ then there is an analog of Theorem 7.3, expressing the condition of having nonnegative N-Ricci curvature in terms of classical tensors.

Remark E.36. For $n > 2$, if $X = \mathbb{R}^n$ is endowed with the Lebesgue measure ν then $\nu \in M_{-2(n-1)}(\mathbb{R}^n)$ and (X, ν) will have nonnegative n-Ricci curvature. In case of $X = \mathbb{R}^2$, endowed with the Lebesgue measure ν, it is not true that $\nu \in M_{-2(n-1)}(\mathbb{R}^n)$. The borderline case $n = 2$ merits further study; see also Corollary E.44.

The issue of showing that Ricci bounds are preserved by pointed measured Gromov-Hausdorff convergence is more involved than in the nonpointed case. The following definition seems to be useful.

Definition E.37. A sequence $\{(X_i, \star_i)\}_{i=1}^\infty$ of pointed metric spaces converges to the pointed metric space $(X, \star)$ in the *proper* pointed Gromov-Hausdorff topology if

1. It converges in the pointed Gromov-Hausdorff topology, by means of pointed approximations $f_i : X_i \to X$,
2. There is a function $\widehat{R} : (0,\infty) \to (0,\infty)$ with $\widehat{R}(R) > R$ for all R,
3. There are nondecreasing functions $G_i : (0,\infty) \to (0,\infty)$, each increasing linearly at infinity, and
4. There is a constant $A > 0$ such that
   1. For all $R > 0$, we have $R < \liminf_{i\to\infty} G_i(\widehat{R}(R))$, and
   2. For all $x_i \in X_i$,

$$G_i(d_i(\star_i, x_i)) \leq d(\star, f_i(x_i)) \leq A d_i(\star_i, x_i) + A. \tag{E.38}$$

Here are the main motivations for the definition. The condition that $d(\star, f_i(x_i)) \leq A d_i(\star_i, x_i) + A$ ensures that $(f_i)_*$ maps $P_2(X_i)$ to $P_2(X)$. Condition 3. and (E.38) ensure that $(f_i)_*$ maps a measure on X_i lying in the dual space of $(1+d_i(\star_i,\cdot)^2)^{-(N-1)}C_b(X_i)$ to a measure on X lying in the dual space of $(1+d(\star,\cdot)^2)^{-(N-1)}C_b(X)$. The condition $G_i(d_i(\star_i,x_i)) \le d(\star, f_i(x_i))$ implies that f_i is metrically proper; for example, it cannot map an unbounded sequence in X_i to a bounded sequence in X. The conditions $R < \liminf_{i\to\infty} G_i(\widehat{R}(R))$ and $G_i(d_i(\star_i,x_i)) \le d(\star, f_i(x_i))$ imply that for any $R>0$, we have $f_i^{-1}(B_R(\star)) \subset B_{\widehat{R}(R)}(\star_i)$ for sufficiently large i. It then follows that in fact, $f_i^{-1}(B_R(\star)) \subset B_{R+\varepsilon_{R,i}}(\star_i)$ for large i.

Definition E.39. A sequence of pointed metric-measure spaces $\{(X_i,\star_i,\nu_i)\}_{i=1}^\infty$ converges to $(X,\star,\nu)$ in the proper pointed M_{-k}-measured Gromov-Hausdorff topology if $\lim_{i\to\infty}(X_i,\star_i) = (X,\star)$ in the proper pointed Gromov-Hausdorff topology, by means of pointed Borel approximations $f_i : X_i \to X$ as above, and in addition $\lim_{i\to\infty}(f_i)_*\nu_i = \nu$ in the weak-$*$ topology on the dual space of $(1+d(\star,\cdot)^2)^{-\frac{k}{2}}C_b(X)$.

Now we can prove the stability of Ricci curvature bounds with respect to the proper pointed measured Gromov-Hausdorff topology. Again, for simplicity we restrict to the case of nonnegative N-Ricci curvature with $N<\infty$.

THEOREM E.40. *Let $\{(X_i,\star_i,\nu_i)\}_{i=1}^\infty$ be a sequence of complete pointed measured, locally compact length spaces with $\lim_{i\to\infty}(X_i,\star_i,\nu_i) = (X,\star,\nu_\infty)$ in the proper pointed $M_{-2(N-1)}$-measured Gromov-Hausdorff topology. If each (X_i,ν_i) has nonnegative N-Ricci curvature then (X,ν_∞) has nonnegative N-Ricci curvature.*

Proof. Given $\mu_0,\mu_1 \in P_2(X,\nu_\infty)$, we wish to show that there is a geodesic joining them along which (3.16) holds for U_{ν_∞}, with $\lambda=0$.

We first show that the claim is true if $\mu_0=\rho_0\nu_\infty$ and $\mu_1=\rho_1\nu_\infty$, with ρ_0 and ρ_1 being compactly-supported continuous functions on X. We will follow the proof of Theorem 4.15. This involved constructing a limiting geodesic using the Arzelà-Ascoli theorem, which in turn used the compactness of $P_2(X)$. If X is noncompact then $P_2(X)$ is not locally compact. Nevertheless, we will show that the needed arguments can be carried out in a compact subset of $P_2(X)$.

By assumption, there is some $R>0$ so that ρ_0 and ρ_1 have support in $B_R(\star)$. Put $\mu_{i,0}=(f_i^*\rho_0)\nu_i$ and $\mu_{i,1}=(f_i^*\rho_1)\nu_i$. From the definition of proper pointed Gromov-Hausdorff convergence, for large i we know that $f_i^*\rho_0$ and $f_i^*\rho_1$ have support in $B_{R+\varepsilon_{R,i}}(\star_i)$. Choose Wasserstein geodesics c_i as in the proof of Theorem 4.15. If γ is a geodesic joining two points of $B_{R+\varepsilon_{R,i}}(\star_i)$ then $\gamma([0,1]) \subset B_{2R+2\varepsilon_{R,i}}(\star_i)$, and so Proposition E.6 implies that each $c_i(t)$ has support in $B_{2R+2\varepsilon_{R,i}}(\star_i)$. Then $(f_i)_*(c_i(t))$ has support in $B_{2R+3\varepsilon_{R,i}}(\star)$.

Hence for large i, each measure $(f_i)_*(c_i(t))$ has support in $B_{2R+1}(\star)$. As the elements of $P_2(X)$ with support in $B_{2R+1}(\star)$ form a relatively compact subset of $P_2(X)$, we can now carry out the arguments of the proof of Theorem 4.15.

This proves the theorem when $\mu_0, \mu_1 \in P_2(X, \nu_\infty)$ have compactly supported continuous densities. To handle the general case, we will use the arguments of Proposition 3.21. Again, the main issue is to show that one can carry out the arguments in a compact subset of $P_2(X)$.

Let $r_0 > 0$ be such that $\mu_0(B_{r_0}(\star)) > 0$ and $\mu_1(B_{r_0}(\star)) > 0$. For $r > r_0$, put $\mu_{0,r} = \frac{1_{B_r(\star)}}{\mu_0(B_r(\star))}\mu_0$ and $\mu_{1,r} = \frac{1_{B_r(\star)}}{\mu_1(B_r(\star))}\mu_1$. Let $\{\mu_{\delta,0,r}\}$ and $\{\mu_{\delta,1,r}\}$ be mollifications of $\mu_{0,r}$ and $\mu_{1,r}$, respectively, using a maximal δ-separated net as discussed in Section E.4. Then

$$\begin{aligned}
\text{(E.41)} \quad & \int_{X-B_R(\star)} d(\star,x)^2 d\mu_{\delta,0,r}(x) \\
&= \frac{1}{\mu_0(B_r(\star))} \int_{X-B_R(\star)} d(\star,x)^2 dK_{\mathcal{C}}(1_{B_r(\star)}\mu_0)(x) \\
&\leq \frac{1}{\mu_0(B_r(\star))} \int_{X-B_{R-10\delta}(\star)} (d(\star,x)+10\delta)^2\, 1_{B_r(\star)}(x) d\mu_0(x) \\
&\leq \frac{1}{\mu_0(B_{r_0}(\star))} \int_{X-B_{R-10\delta}(\star)} (d(\star,x)+10\delta)^2\, d\mu_0(x).
\end{aligned}$$

For small δ, we obtain

$$\text{(E.42)} \quad \int_{X-B_R(\star)} d(\star,x)^2 d\mu_{\delta,0,r}(x) \leq \frac{2}{\mu_0(B_{r_0}(\star))} \int_{X-B_{R/2}(\star)} d(\star,x)^2 d\mu_0(x).$$

Similarly,

$$\text{(E.43)} \quad \int_{X-B_R(\star)} d(\star,x)^2 d\mu_{\delta,1,r}(x) \leq \frac{2}{\mu_1(B_{r_0}(\star))} \int_{X-B_{R/2}(\star)} d(\star,x)^2 d\mu_1(x).$$

As the right-hand-sides of (E.42) and (E.43) can be made arbitrarily small by taking R sufficiently large, it follows that $\bigcup_{r>r_0}\bigcup_{i=1}^{\infty}\{\mu_{i^{-1},0,r}, \mu_{i^{-1},1,r}\}$ is relatively compact in $P_2(X)$. With an appropriate choice of i_j for j large, we have $\lim_{j\to\infty}\mu_{i_j^{-1},0,j} = \mu_0$ and $\lim_{j\to\infty}\mu_{i_j^{-1},1,j} = \mu_1$. Using Proposition E.9, the argument in the proof of Proposition 3.21 can now be applied to show that there is a geodesic from μ_0 to μ_1 along which (3.16) holds for U_{ν_∞}, with $\lambda = 0$. □

E.6. *Tangent Cones.* We now give an application of Theorem E.40 that just involves the *pointed measured Gromov-Hausdorff topology* introduced in Definition E.2.

Corollary E.44. *Let* $\{(X_i, \star_i, \nu_i)\}_{i=1}^{\infty}$ *be a sequence of complete, pointed, measured, locally compact length spaces. Suppose that* $\lim_{i\to\infty}(X_i,\star_i,\nu_i) = (X,\star,\nu)$ *in the pointed, measured Gromov-Hausdorff topology of Definition* E.2. *If* $N \in (2,\infty)$ *and each* (X_i,ν_i) *has nonnegative* N*-Ricci curvature then* (X,ν) *has nonnegative* N*-Ricci curvature.*

Proof of Corollary E.44. If X is compact then the result follows from Theorem 5.19, so we will assume that X is noncompact. Let $\{f_i\}_{i=1}^\infty$ be a sequence of approximations as in Definition E.1. Given $R_i > 0$, let $\widehat{f}_i : X_i \to X$ be an arbitrary Borel map such that $\widehat{f}_i(x_i) = f_i(x_i)$ if $d_i(\star_i, x_i) < R_i$ and $d(\star, \widehat{f}_i(x_i)) = d(\star_i, x_i)$ if $d_i(\star_i, x_i) \geq R_i$. (For example, if $\gamma : [0, \infty) \to X$ is a normalized ray with $\gamma(0) = \star$ then we can put $\widehat{f}_i(x_i) = \gamma(d(\star_i, x_i))$ when $d_i(\star_i, x_i) \geq R_i$.) After passing to a subsequence of $\{f_i\}_{i=1}^\infty$ (which we relabel as $\{f_i\}_{i=1}^\infty$) and replacing f_i by $\widehat{f}_i$ for an appropriate choice of R_i, we can assume that $\lim_{i\to\infty}(X_i, \star_i, \nu_i) = (X, \star, \nu)$ in the *proper* pointed measured Gromov-Hausdorff topology, with $\widehat{R}(R) = 3R$ and $G_i(r) = \frac{r}{2}$.

As each (X_i, ν_i) has nonnegative N-Ricci curvature, and $\lim_{i\to\infty}(\widehat{f}_i)_* \nu_i = \nu$ in $C_c(X)^*$, the Bishop-Gromov inequality of Theorem 5.31 (as extended to the noncompact case) implies that there are constants $C, r_0 > 0$ so that for all i, whenever $r \geq r_0$ we have $\nu_i(B_r(\star_i)) \leq Cr^N$. As $N > 2$, it follows from dominated convergence that $\lim_{i\to\infty}(\widehat{f}_i)_* \nu_i = \nu$ in the weak-$*$ topology on the dual space of $(1 + d(\star, \cdot)^2)^{-(N-1)} C_b(X)$. The claim now follows from Theorem E.40. □

Example E.45. We apply Corollary E.44 to tangent cones. Suppose that (X, d, ν) is a complete measured locally compact length space with nonnegative N-Ricci curvature for some $N \in (2, \infty)$. Suppose that $\mathrm{supp}(\nu) = X$. For $i \geq 1$, put $(X_i, d_i) = (X, i \cdot d)$. Given $\star \in X$, let $\star_i$ be the same point in X_i. Using Theorem 5.31 (as extended to the noncompact case), after passing to a subsequence we may assume that $\{(X_i, \star_i)\}_{i=1}^\infty$ converges in the pointed Gromov-Hausdorff topology to a tangent cone $(T_\star X, o)$; see Corollary 5.33. Let ν_i be the pushforward from X to X_i, via the identity map, of the measure $\frac{\nu}{\nu(B_{i^{-1}}(\star))}$. After passing to a further subsequence, we can assume that $\{(X_i, \star_i, \nu_i)\}_{i=1}^\infty$ converges in the pointed measured Gromov-Hausdorff topology to a pointed measured length space $(T_\star X, o, \nu_\infty)$, where ν_∞ is a nonnegative Radon measure on $T_\star X$ that is normalized so that $\nu_\infty(B_1(o)) = 1$; see [16, §1]. From Corollary E.44, $(T_\star X, \nu_\infty)$ has nonnegative N-Ricci curvature.

Appendix F. Bibliographic notes on optimal transport

The following notes are by no means exhaustive, but may provide some entry points to the literature.

Wasserstein was one of many authors who discovered, rediscovered or studied optimal transportation metrics [45]. He was interested in the case when the cost coincides with the distance. Tanaka [43] may have been the first to take advantage of geometric properties of W_2, in his study of the Boltzmann equation. Accordingly, other names could be used for W_2, such as Monge-Kantorovich distance or Tanaka distance. The terminology "Wasserstein distance" was used by Otto and coworkers, and naturally gave rise to the term "Wasserstein space". Otto studied this metric space from a geometric point of view and showed that $P_2(\mathbb{R}^n)$ can be equipped with a formal infinite-dimensional Riemannian metric, thereby allowing insightful computations [35]. He also showed that his Riemannian metric formally has nonnegative sectional curvature. Otto's motivation came from partial differential equations, and in particular from earlier work by Brenier in fluid mechanics [10]. The formal gradient flow of a "free energy" functional on the Wasserstein space was considered by Jordan, Kinderlehrer and Otto [25] and Otto [35].

The notion of displacement convexity was introduced by McCann [31], and later refined to the notion of λ-uniform displacement convexity. A formal differential calculus on $P_2(M)$, when M is a smooth Riemannian manifold, was described by Otto and Villani [36]. It was "shown" that the entropy functional $\int_M \rho \log \rho \, \mathrm{dvol}_M$ is displacement convex on a manifold with nonnegative Ricci curvature. Appendix D of the present paper follows up on the calculations in [36].

Simultaneously, a rigorous theory of optimal transport on manifolds was initiated by McCann [32] and further developed by Cordero-Erausquin, McCann and Schmuckenschläger [19]. In particular, these authors prove the implication (1) $\Rightarrow$ (4) of Theorem 7.3(a) of the present paper when Ψ is constant and $N = n$. The paper [19] was extended by von Renesse and Sturm [40], whose paper contains a proof of the implications (1) $\Leftrightarrow$ (5) of Theorem 7.3(b) of the present paper when Ψ is constant, and also indicates that the condition (5) may make sense for some metric-measure spaces. In a more recent contribution, which was done independently of the present paper, Cordero-Erausquin, McCann and Schmuckenschläger [20] prove the implication (1) $\Rightarrow$ (5) of Theorem 7.3(b) for general Ψ.

Connections between optimal transport and the theory of log Sobolev inequalities and Poincaré inequalities were established by Otto and Villani [36] and developed by many authors. This was the starting point for Section 6 of the present paper. More information can be found in [44].

A proof of a weak Bonnet-Myers theorem, based only on Riemannian growth control and concentration estimates, was given by Ledoux [26] as a special case of a more general result about the control of the diameter of manifolds satisfying a log Sobolev inequality. The simplified proof used in the present paper, based on a Talagrand inequality, is taken from [36].

University of Michigan, Ann Arbor, MI
E-mail address: lott@umich.edu

UMPA (UMR CNRS 5669), École Normale Supérieure de Lyon, Lyon, France
E-mail address: cvillani@umpa.ens-lyon.fr

References

[1] G. Alberti and L. Ambrosio, A geometrical approach to monotone functions in $\mathbb{R}^n$, *Math. Z.* **230** (1999), 259–316.

[2] L. Ambrosio, N. Gigli, and G. Savarè, Gradient flows in metric spaces and in the space of probability measures, *Lectures in Mathematics ETH Zürich*, Birkhaüser, Basel (2005)

[3] ———, Gradient flows with metric and differentiable structures, and applications to the Wasserstein space, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **15** (2004), 327–343.

[4] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, Sur les inégalités de Sobolev logarithmiques, *Panoramas et Synthèses* **10**, Société Mathématique de France, Paris (2000).

[5] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes, in *Lectures on Probability Theory* (*Saint-Flour*, 1992), *Lecture Notes in Math.* **1581** (1994), 1–114, Springer-Verlag, New York.

[6] D. Bakry and M. Émery, Diffusions hypercontractives, in *Séminaire de Probabilités* XIX, *Lecture Notes in Math.* **1123** (1985), 177–206, Springer-Verlag, New York.

[7] P. Bernard and B. Buffoni, Optimal mass transportation and Mather theory, *J. of European Math. Soc.* **9** (2007), 85–121.

[8] S. Bobkov, I. Gentil, and M. Ledoux, Hypercontractivity of Hamilton-Jacobi equations, *J. Math. Pures Appl.* **80** (2001), 669–696.

[9] S. Bobkov and F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.* **163** (1999), 1–28.

[10] Y. Brenier, Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations, *Comm. Pure Appl. Math.* **52** (1999), 411-452.

[11] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry*, *Graduate Studies in Mathematics* **33**, Amer. Math. Soc., Providence, RI, 2001.

[12] Y. Burago, M. Gromov, and G. Perelman, A. D. Aleksandrov spaces with curvatures bounded below, *Russian Math. Surveys* **47** (1992), 1–58.

[13] G. Buttazzo, M. Giaquinta, and S. Hildebrandt, One-dimensional variational problems, *Oxford Lecture Series in Mathematics and its Applications* **15**, Oxford University Press, New York (1998).

[14] J. A. Carrillo, R. J. McCann, and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media, *Arch. Rational Mech. Anal.* **179** (2006), 217–263.

[15] J. Cheeger and T. Colding, Lower bounds on Ricci curvature and the almost rigidity of warped products, *Ann. of Math.* **144** (1996), 189–237.

[16] ———, On the structure of spaces with Ricci curvature bounded below I, *J. Diff. Geom.* **46** (1997), 406–480.

[17] ———, On the structure of spaces with Ricci curvature bounded below II, *J. Diff. Geom.* **54** (2000), 13–35.

[18] ———, On the structure of spaces with Ricci curvature bounded below III, *J. Diff. Geom.* **54** (2000), 37–74.

[19] D. Cordero-Erausquin, R. McCann, and M. Schmuckenschläger, A Riemannian interpolation inequality à la Borell, Brascamp and Lieb, *Invent. Math.* **146** (2001), 219–257.

[20] ———, Prékopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport, *Ann. Fac. Sci. Toulouse Math.* **15** (2006), 613–635.

[21] K. Fukaya, Collapsing of Riemannian manifolds and eigenvalues of Laplace operator, *Invent. Math.* **87** (1987), 517–547.

[22] M. Gromov, Groups of polynomial growth and expanding maps, *Publ. Math. de l'IHES* **53** (1981), 53–78.

[23] ———, Metric structures for Riemannian and non-Riemannian spaces, *Prog. in Math.* **152** (1999), Birkhäuser, Boston.

[24] K. Grove and P. Petersen, Manifolds near the boundary of existence, *J. Diff. Geom.* **33** (1991), 379–394.

[25] R. Jordan, D. Kinderlehrer, and F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.* **29** (1998), 1–17.

[26] M. Ledoux, Concentration of measure and logarithmic Sobolev inequalities, *Séminaire de Probabilités* XXXIII, *Lecture Notes in Mathematics* **1709** (1999), 120–216, Springer-Verlag, New York.

[27] M. Ledoux, *The Concentration of Measure Phenomenon*, Amer. Math. Soc., Providence, RI, 2001.

[28] A. Lichnerowicz, Géométrie des groupes de transformations, *Travaux et Recherches Mathématiques* III, Dunod, Paris (1958).

[29] F. Liese and I. Vajda, Convex statistical distances, *Teubner-Texte zur Mathematik* **95**, BSB B. G. Teubner Verlagsgesellschaft, Leipzig (1987).

[30] J. Lott, Some geometric properties of the Bakry-Émery-Ricci tensor, *Comment. Math. Helv.* **78** (2003), 865–883.

[31] R. J. McCann, A convexity principle for interacting gases, *Adv. Math.* **128** (1997), 153–179.

[32] ———, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.* **11** (2001), 589–608.

[33] Y. Otsu, Differential geometric aspects of Alexandrov spaces, in *Comparison Geometry*, *MSRI Publ.* **30** (1997), 135–148, Cambridge Univ. Press, Cambridge, UK.

[34] Y. Otsu and T. Shioya, The Riemannian structure of Alexandrov spaces, *J. Diff. Geom.* **39** (1994), 629–658.

[35] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Comm. Partial Diff. Eq.* **26** (2001), 101–174.

[36] F. Otto and C. Villani, Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173** (2000), 361–400.

[37] G. Perelman, DC structure on Alexandrov space with curvature bounded below, unpublished preprint.

[38] G. Perelman and A. Petrunin, Quasigeodesics and gradient curves in Alexandrov spaces, unpublished preprint.

[39] Z. Qian, Estimates for weighted volumes and applications, *Quart. J. Math. Oxford* **48** (1997), 235–242.

[40] M.-K. von Renesse and K.-T. Sturm, Transport inequalities, gradient estimates, entropy, and Ricci curvature, *Comm. Pure Appl. Math.* **58** (2005), 923–940.

[41] K.-T. Sturm, Convex functionals of probability measures and nonlinear diffusions on manifolds, *J. Math. Pures. Appl.* **84** (2005), 149–168.

[42] ———, On the geometry of metric measure spaces, *Acta Math.* **196** (2006), 65–131.

[43] H. Tanaka, An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27** (1973), 47–52.

[44] C. Villani, *Topics in Optimal Transportation*, *Graduate Studies in Mathematics* **58**, Amer. Math. Soc., Providence, RI, 2003.

[45] L. Wasserstein, Markov processes over denumerable products of spaces describing large systems of automata, *Problems of Information Transmission* **5** (1969), 47–82.

[46] R. Zimmer, *Ergodic Theory and Semisimple Groups*, *Monographs in Mathematics* **81** (1984), Birkhäuser, Basel.

(Received February 1, 2005)
(Revised November 21, 2005)