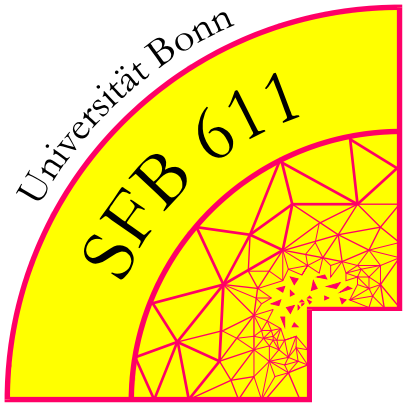
Rheinische Friedrich-Wilhelms-Universität Bonn

On the Geometry of Metric Measure Spaces

Karl-Theodor Sturm

Preprint No. **203**, 2004



Sonderforschungsbereich 611

Singuläre Phänomene und Skalierung in mathematischen Modellen

On the Geometry of Metric Measure Spaces

Karl-Theodor Sturm

Abstract

We introduce and analyze lower ('Ricci') curvature bounds $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ for metric measure spaces (M, d, m). Our definition is based on convexity properties of the relative entropy $\mathrm{Ent}(.|m)$ regarded as a function on the L_2-Wasserstein space of probability measures on the metric space (M, d). Among others, we show that $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ implies estimates for the volume growth of concentric balls. For Riemannian manifolds, $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ if and only if $\mathrm{Ric}_M(\xi, \xi) \geq K \cdot |\xi|^2$ for all $\xi \in TM$.

The crucial point is that our lower curvature bounds are stable under an appropriate notion of $\mathbb{D}$-convergence of metric measure spaces. We define a complete and separable metric $\mathbb{D}$ on the family of all isomorphism classes of normalized metric measure spaces. The metric $\mathbb{D}$ has a natural interpretation, based on the concept of optimal mass transportation.

We also prove that the family of normalized metric measure spaces with doubling constant $\leq C$ is closed under $\mathbb{D}$-convergence. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence.

1 Introduction

The notion of a 'metric space' is one of the basic concepts of mathematics. Metric spaces play a prominent role in many fields of mathematics. In particular, they constitute natural generalizations of manifolds admitting all kinds of singularities and still providing rich geometric structures.

A. D. Alexandrov [Al51] introduced the notion of lower curvature bounds for metric spaces in terms of comparison properties for geodesic triangles. These curvature bounds are equivalent to lower bounds for the *sectional curvature* in the case where the metric spaces are Riemannian manifolds, – and they may be regarded as generalized lower bounds for the 'sectional curvature' for general metric spaces. A fundamental observation is that these lower bounds are *stable* under an appropriate notion of convergence of metric spaces, the so-called *Gromov-Hausdorff convergence*, introduced by M. Gromov [Gr81a]. The family of manifolds with sectional curvature $\geq K$ is, of course, not closed under Gromov-Hausdorff convergence but the family of metric spaces with curvature $\geq K$ in the sense of Alexandrov is closed (for each $K \in \mathbb{R}$). Even more, the family of compact metric spaces with curvature $\geq K$, Hausdorff dimension $\leq N$ and diameter $\leq \Delta$ is compact (for any choice of K, N, Δ).

For many fundamental results in geometric analysis, however, the crucial ingredients are not bounds for the sectional curvature but bounds for the *Ricci curvature:* estimates for heat kernels and Green functions, Harnack inequalities, eigenvalue estimates, isoperimetric inequalities, Sobolev inequalities, — they all depend on lower bounds for the Ricci curvature of the underlying manifolds as pointed out by S. T. Yau and others, see e.g. [LY86], [Ch93], [Da89], [SC02]. The family of Riemannian manifolds with given lower bound for the Ricci curvature is not closed under Gromov-Hausdorff convergence (nor it is closed under any other reasonable notion of convergence). One of the great challenges thus is to establish a generalized notion of lower Ricci curvature bounds for singular spaces. For detailed investigations and a survey of the state of

the art for this problem, we refer to the contributions by J. Cheeger and T. Colding [CC97/00]. Fascinating new developments have been outlined just recently by G. Perelman [Pe02] in the context of his work on the Poincaré conjecture.

Generalizations of lower Ricci curvature bounds should be formulated in the framework of metric measure spaces. These are triples (M, d, m) where (M, d) is a metric space and m is a measure on the Borel σ-algebra of M. We will always require that the metric space (M, d) is complete and separable and that the measure m is locally finite. Recall that for generalizations of sectional curvature bounds only the metric structure (M, d) is required whereas for generalizations of Ricci curvature bounds in addition a reference measure m has to be specified. In a certain sense, this phenomenon is well-known from the discussion of the curvature-dimension condition of D. Bakry and M. Emery [BE85] in the framework of Dirichlet forms and symmetric Markov semigroups. Of course, also the Bakry-Emery condition is a kind of generalized lower bound for the Ricci curvature (together with an upper bound for the dimension). However, it is not given in terms of the basic data (M, d, m) but in terms of the Dirichlet form (or heat semigroup) derived from the original quantities in a highly non-trivial manner.

Metric measure spaces have been studied quite intensively in recent years. Of particular interest is the study of functional inequalities, like Sobolev and Poincaré inequalities, on metric measure spaces and the construction and investigation of function spaces of various types [HK95, HK00], [Ko00], [He01]. To some extent, doubling properties for the volume and scale invariant Poincaré inequalities on metric balls can be regarded as weak replacements of lower Ricci curvature bounds. Among others, they allow to construct Dirichlet forms, Laplacians and heat kernels on given metric measure spaces and to derive (elliptic and parabolic) Harnack inequalities as well as (upper and lower) Gaussian estimates for heat kernels [St98], [Ch99]. On the other hand, however, even in simplest cases doubling constant and Poincaré constant do not characterize spaces with lower bounded Ricci curvature: they always allow at least also metrics which are equivalent to the given ones.

The main results of this paper are:

- We define a *complete and separable metric* $\mathbb{D}$ on the family of all isomorphism classes of normalized metric measure spaces, Theorem 3.6. The metric $\mathbb{D}$ has a natural interpretation, based on the concept of optimal mass transportation.
- The family of normalized metric measure spaces with doubling constant $\leq C$ is *closed under $\mathbb{D}$-convergence*, Theorem 3.15. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is *compact under $\mathbb{D}$-convergence* (for any choice of real numbers C, R), Theorem 3.16.
- We introduce a notion of *lower curvature bounds* $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m)$ for metric measure spaces (M, d, m), based on convexity properties of the relative entropy $\mathrm{Ent}(.|m)$ w.r.t. the reference measure m. Here $\nu \mapsto \mathrm{Ent}(\nu|m)$ is regarded as a function on the L_2-Wasserstein space of probability measures on the metric space (M, d). For Riemannian manifolds, $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ if and only if $\mathrm{Ric}_M(\xi, \xi) \geq K \cdot |\xi|^2$ for all $\xi \in TM$, Theorem 4.9.
- *Local* lower curvature bounds imply *global* lower curvature bounds, Theorem 4.17.
- Lower curvature bounds are *stable under $\mathbb{D}$-convergence*, Theorem 4.20.

- Lower curvature bounds of the form $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ imply *estimates for the volume growth* of concentric balls, for instance, if $K \leq 0$

$$m(B_r(x)) \leq C(x) \cdot \exp(-Kr^2/2)$$

for all $r \geq 1$, Theorem 4.24.

The concept of optimal mass transportation plays a crucial role in our approach. It originates in the classical transportation problems of G. Monge [M1781] and L. V. Kantorovich [Ka42]. The basic quantity for us is the so-called L_2-*Wasserstein distance* between two probability measures μ and ν on a given complete separable metric space (M, d) defined as

$$\mathsf{d}_W(\mu, \nu) := \inf_q \left(\int_{M \times M} \mathsf{d}^2(x, y)\, dq(x, y) \right)^{1/2}$$

where the infimum is taken over all *couplings* q of μ and ν. The latter are probability measures on the product space $M \times M$ whose marginals (i.e. image measures under the projections) are the given measures μ and ν. One choice, of course, is $q = \mu \otimes \nu$ but in most cases this will be a very bad choice if one aims for minimal transportation costs. The L_2-Wasserstein distance can be interpreted as the minimal transportation costs (measured in L_2-sense) for transporting goods from producers at locations distributed according to μ to consumers at locations distributed according to ν.

Two results may be regarded as milestones in the recent development of theory and application of mass transportation concepts; these results have raised an increasing interest in this topic of people from various fields of mathematics including pde's, geometry, fluid mechanics, and probability. See e.g. [Ta95], [Le01], [BG99], [OV00], [DD02], [CMS01], [AT04], [FÜ04a] and in particular the monograph by C. Villani [Vi03] which gives an excellent survey on the whole field. The first of these two results is the polar factorization of Y. Brenier [Br91] and its extension to the Riemannian setting by R. McCann [Mc95, Mc97]. The second one is F. Otto's [JKO98, Ot01] formal Riemannian calculus on the space $\mathcal{P}_2(M)$ of probability measures on M, equipped with the L_2-Wasserstein metric, and his interpretation of the heat equation (and of other nonlinear dissipative evolution equations) as gradient flow(s) of the *relative entropy*

$$\mathrm{Ent}(\nu|m) = \int_M \frac{d\nu}{dm} \log\left(\frac{d\nu}{dm}\right) dm$$

(or related functionals, resp.) on $\mathcal{P}_2(M)$.

It turned out that convexity properties of the function $\nu \mapsto \mathrm{Ent}(\nu|m)$ are intimately related to curvature properties of the underlying metric measure space (M, d, m). If M is a complete Riemannian manifold with Riemannian distance d and if $m = e^{-V} dx$ then M.-K. von Renesse and the author proved (see [RS04] for the case $V = 0$ and [St04] for the general case) that the function $\nu \mapsto \mathrm{Ent}(\nu|m)$ is K-convex[1] on $\mathcal{P}_2(M)$ if and only if

$$\mathrm{Ric}_M(\xi, \xi) + \mathrm{Hess} V(\xi, \xi) \geq K \cdot |\xi|^2$$

for all $\xi \in TM$. An heuristic argument for the 'if'-implication of this equivalence was presented in [OV00], based on the formal Riemannian calculus on $\mathcal{P}_2(M)$. In the particular case $K = 0, V = 0$ the 'if'-implication was proven in [CMS01].

[1] By definition, this means that property (1.1) below holds for each geodesic Γ in $\mathcal{P}_2(M)$.

Having in mind these results, it seems quite natural to say that an arbitrary metric measure space (M, d, m) *has curvature* $\geq K$ if and only if for any pair $\nu_0, \nu_1 \in \mathcal{P}_2(M)$ with $\mathrm{Ent}(\nu_0|m) < \infty$, $\mathrm{Ent}(\nu_1|m) < \infty$ there exists a geodesic $\Gamma : [0,1] \to \mathcal{P}_2(M)$ connecting ν_0 and ν_1 with

$$\mathrm{Ent}(\Gamma(t)|m) \leq (1-t)\mathrm{Ent}(\Gamma(0)|m) + t\mathrm{Ent}(\Gamma(1)|m) - \frac{K}{2}t(1-t)\,\mathsf{d}_W^2(\Gamma(0),\Gamma(1)) \tag{1.1}$$

for all $t \in [0,1]$. In this case, we also briefly write $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$.

A crucial property of this kind of curvature bound is its stability under convergence of metric measure spaces. Of course, this requires to have an appropriate notion of topology or distance on the family of all metric measure spaces. We define the *distance* between two normalized metric measure spaces by

$$\mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m')) := \inf_{\mathrm{q}, \hat{\mathsf{d}}} \left(\int_{M \times M'} \hat{\mathsf{d}}^2(x,y)\, dq(x,y) \right)^{1/2}$$

where the infimum is taken over all couplings q of m and m' and over all couplings $\hat{\mathsf{d}}$ of d and d'. The former are probability measures on $M \times M'$ with marginals m and m'. The latter are pseudo metrics on the disjoint union $M \sqcup M'$ which extend d and d'.
Also the distance $\mathbb{D}$ has an interpretation in terms of mass transportation: In order to realize the distance

$$\mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m'))$$

between two normalized metric measure spaces (M, d, m) and (M', d', m') one first may use isometric transformations of (M, d) and (M', d') to bring the images of m and m' in optimal position to each other. (In the sense of transportation costs, these transformations are for free). Then one has to solve the usual mass transportation problem, trying to minimize the transportation costs in the L_2-sense.

It turns out that $\mathbb{D}$ is a complete separable metric on the family $\mathbb{X}_1$ of all isomorphism classes of normalized metric measure spaces. The family of normalized metric measure spaces with curvature $\geq K$ is closed under $\mathbb{D}$-convergence. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ is closed under $\mathbb{D}$-convergence and the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence (for any choice of real numbers K, C, R).
For various other distances on the family $\mathbb{X}_1$, see the additional chapter $3\frac{1}{2}$ in [Gr99]. A completely different notion of distance between Riemannian manifolds was proposed by A. Kasue [KK94, Ka04], based on the short time asymptotics of the heat kernel. Yet another convergence concept was proposed by K. Kuwae and T. Shioya [KS03] extending the concept of Γ-convergence and Mosco convergence towards a notion of convergence of operators (or Dirichlet forms or heat semigroups) on varying spaces.

A major advantage of our distance $\mathbb{D}$ seems to be that it has a very natural geometric interpretation, namely, in terms of the above-mentioned mass transportation concept. We also expect that it is closely related to more analytic properties of metric measure spaces. Following [JKO98], the heat semigroup on a metric measure space (M, d, m) should be obtained as the gradient flow on $\mathcal{P}_2(M)$ for the relative entropy $\mathrm{Ent}(.|m)$. Curvature bounds of the form $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ should e.g. imply K-contractivity of the heat flow

$$\mathsf{d}_W(\mu p_t, \nu p_t) \leq e^{-Kt}\, \mathsf{d}_W(\mu, \nu),$$

gradient estimates for harmonic functions, isoperimetric inequalities, and volume growth estimates.
The basic concepts and main results of this paper have been presented in [St04a]. In a forthcoming paper [St05], we will study metric measure spaces satisfying a so-called curvature-dimension condition (K, N) being more restrictive than the condition $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$. The additional parameter N plays the role of an upper bound for the dimension. Among others, this will lead to more precise volume growth estimates in the spirit of the Bishop-Gromov volume comparison theorem.

Here in this paper, we will proceed as follows:
In Chapter 2 we give a brief survey on the geometry of metric spaces, recalling the concepts of length and geodesic spaces, the Gromov-Hausdorff distance and the lower curvature bounds in the sense of Alexandrov. We introduce the L_2-Wasserstein space of probability measures on a given metric space and derive some of the basic properties.
Chapter 3 is devoted to the metric $\mathbb{D}$. The first main result states that it indeed defines a (complete and separable) metric on the family of isomorphism classes of normalized metric measure spaces. We collect several simple examples of $\mathbb{D}$-convergence with increasing and decreasing dimensions and we discuss closedness and compactness properties of the families of normalized metric measure spaces with the doubling property.
In Chapter 4 we study metric measure spaces with curvature bounds. First we introduce and discuss the relative entropy, then we present the definition of curvature bounds and analyze their behavior under various transformations (isomorphisms, scaling, weights, subsets, products). The main results are the Globalization Theorem and the Convergence Theorem. Finally, we deduce growth estimates for the volume of concentric balls.

2 On the Geometry of Metric Spaces

2.1 Length and Geodesic Spaces

Let us summarize some definitions and basic results on the geometry of metric spaces. For proofs and further details we refer to [BH99], [Gr99], and [BBI01].
Throughout this paper, a pseudo metric on a set M will be a function $\mathsf{d} : M \times M \to [0, \infty]$ which is symmetric, vanishes on the diagonal and satisfies the triangle inequality. If it does not vanish outside the diagonal and does not take the value $+\infty$ then it is called *metric*. From now on, let (M, d) be a metric space. Open balls in M will be denoted by $B_r(x) = \{y \in M : \mathsf{d}(x, y) < r\}$, their closures by $\bar{B}_r(x) \subset \{y \in M : \mathsf{d}(x, y) \leq r\}$. A curve connecting two points $x, y \in M$ is a continuous map $\gamma : [a, b] \to M$ with $\gamma(a) = x$ and $\gamma(b) = y$. Obviously, then $\mathrm{Length}(\gamma) \geq \mathsf{d}(x, y)$ with the length of γ being defined as

$$\mathrm{Length}(\gamma) = \sup \sum_{k=1}^{n} \mathsf{d}(\gamma(t_{k-1}), \gamma(t_k))$$

where the supremum is taken over all partitions $a = t_0 < t_1 < \cdots < t_n = b$. If $\mathrm{Length}(\gamma) < \infty$ then γ is called rectifiable. In this case we can and will henceforth always assume that (after suitable reparametrization) γ has constant speed, i.e. $\mathrm{Length}(\gamma|_{[s,t]}) = \frac{t-s}{b-a} \cdot \mathrm{Length}(\gamma)$ for all $a < s < t < b$. In general, we will not distinguish between curves and equivalence classes of curves which are reparametrizations of each other. The curve $\gamma : [a, b] \to M$ is called *geodesic* iff $\mathrm{Length}(\gamma) = \mathsf{d}(\gamma(a), \gamma(b))$. A geodesic in this sense is always minimizing.

A metric space $(M,\,\mathsf{d})$ is called *length space* (or length metric space) iff for all $x, y \in M$

$$\mathsf{d}(x,y) = \inf_{\gamma}\ \mathrm{Length}(\gamma)$$

where the infimum is taken over all curves γ in M which connect x and y. A metric space $(M,\,\mathsf{d})$ is called *geodesic space* (or geodesic metric space) iff each pair of points $x, y \in M$ is connected by a geodesic. (This geodesic is not required to be unique.)

Lemma 2.1. *A complete metric space $(M,\,\mathsf{d})$ is a length space (or geodesic space) if and only if for each pair of points $x_0, x_1 \in M$ and for each $\epsilon > 0$ (or for $\epsilon = 0$, resp.) there exists a point $y \in M$ satisfying for each $i = 0, 1$*

$$\mathsf{d}(x_i, y) \le \frac{1}{2}\,\mathsf{d}(x_0, x_1) + \epsilon. \tag{2.1}$$

Any such point y will be called ϵ-midpoint of x_0 and x_1. In the case $\epsilon = 0$ it will be called midpoint of x_0 and x_1.

Remark 2.2. Given $x_0, x_1 \in M$ then each ϵ-midpoint $y \in M$ satisfies

$$\mathsf{d}^2(x_0, y) + \mathsf{d}^2(y, x_1) \le \frac{1}{2}\,\mathsf{d}^2(x_0, x_1) + \epsilon' \tag{2.2}$$

with $\epsilon' = 2\epsilon\,\mathsf{d}(x_0, x_1) + 2\epsilon^2$. Vice versa, each $y \in M$ which satisfies (2.2) is an ϵ-midpoint with $\epsilon = \sqrt{(\,\mathsf{d}(x_0, x_1)/2)^2 - \epsilon'/2} - \mathsf{d}(x_0, x_1)/2 - \sqrt{\epsilon'/2}$.
Indeed, if y is an ϵ-midpoint then

$$\mathsf{d}^2(x_0, y) + \mathsf{d}^2(y, x_1) \le 2\left[\frac{1}{2}\,\mathsf{d}(x_0, x_1) + \epsilon\right]^2 = \frac{1}{2}\,\mathsf{d}^2(x_0, x_1) + \epsilon'$$

with ϵ' chosen as above. Conversely, if y satisfies (2.2) then

$$\begin{aligned}
\frac{1}{2}\,\mathsf{d}^2(x_0, x_1) + \epsilon' &\ge \mathsf{d}^2(x_0, y) + \mathsf{d}^2(y, x_1) \\
&= \frac{1}{2}[\,\mathsf{d}(x_0, y) + \mathsf{d}(y, x_1)]^2 + \frac{1}{2}[\,\mathsf{d}(x_0, y) - \mathsf{d}(y, x_1)]^2 \\
&\ge \frac{1}{2}\,\mathsf{d}^2(x_0, x_1) + \frac{1}{2}[\,\mathsf{d}(x_0, y) - \mathsf{d}(y, x_1)]^2.
\end{aligned}$$

Hence, $|\,\mathsf{d}(x_0, y) - \mathsf{d}(y, x_1)| \le \sqrt{2\epsilon'}$ and

$$2\,\mathsf{d}(x_i, y) - \sqrt{2\epsilon'} \le \mathsf{d}(x_0, y) + \mathsf{d}(y, x_1) \le \sqrt{\mathsf{d}^2(x_0, x_1) + 2\epsilon'}$$

for $i = 0, 1$.

Lemma 2.3. *If $(M,\,\mathsf{d})$ is a complete length space then:*

(i) The closure of $B_r(x)$ is $\{y \in M : \mathsf{d}(x, y) \le r\}$.

(ii) M is locally compact if and only if each closed ball in M is compact.

(iii) If M is locally compact then it is a geodesic space.

Recall that the Hausdorff distance between two subsets A_1, A_2 of a metric space (M, d) is given by

$$\mathsf{d}^{\mathsf{H}}(A_1, A_2) = \inf\{\epsilon > 0 : A_1 \subset B_\epsilon(A_2), A_2 \subset B_\epsilon(A_1)\}$$

where $B_\epsilon(A) := \{x \in M : \inf_{y\in A} \mathsf{d}(x, y) < \epsilon\}$ denotes the ϵ-neighborhood of $A \subset M$. The *Gromov-Hausdorff distance* between two metric spaces (M_1, d_1) and (M_2, d_2) is defined by

$$\mathsf{D}^{\mathsf{GH}}((M_1, \mathsf{d}_1), (M_2, \mathsf{d}_2)) = \inf \ \mathsf{d}^{\mathsf{H}}(j_1(M_1), j_2(M_2))$$

where the inf is taken over all metric spaces (M, d) and over all isometric embeddings $j_1 : M_1 \hookrightarrow M$, $j_2 : M_2 \hookrightarrow M$.

Proposition 2.4.

(i) The Gromov-Hausdorff metric D^{GH} is a pseudo metric on the family X of isometry classes of metric spaces.

(ii) Let X_l (and X_{lcl}) denote the family of isometry classes of (locally compact)complete length spaces. Then X_l and X_{lcl} are closed under D^{GH}-convergence.

(iii) Let X_c (and X_f) denote the family of isometry classes of compact (or finite, resp.) metric spaces. Then $(\mathsf{X}_c, \mathsf{D}^{\mathsf{GH}})$ is a complete separable metric space. The family X_f is dense in X_c.

2.2 Alexandrov Spaces

Now let us briefly discuss metric spaces with lower curvature bounds in the sense of A.D. Alexandrov [Al51]. The latter are generalizations of lower bounds for the sectional curvature for Riemannian manifolds. The results of this section will not be used in the sequel. The focus in this paper is on generalizations of lower bounds for the Ricci curvature. Partly, however, there will be some analogy to Alexandrov's generalizations of lower bounds for the sectional curvature. We summarize some of the basic properties of these metric spaces and refer to [BGP92], [GP97], [Gr99], [BBI01], [Pl02] for further details.

Given any $K \in \mathbb{R}$ we say that a metric space (M, d) has *curvature* $\geq K$ iff for each quadruple of points $z, x_1, x_2, x_3 \in M$

$$\sphericalangle_K(z; x_1, x_2) + \sphericalangle_K(z; x_2, x_3) + \sphericalangle_K(z; x_3, x_1) \leq 2\pi. \tag{2.3}$$

Here for any triple of points $z, x, y \in M$ with $K \cdot [\,\mathsf{d}(z, x) + \mathsf{d}(x, y) + \mathsf{d}(y, z)]^2 < (2\pi)^2$ we denote by $\sphericalangle_K(z; x, y)$ the angle at $\bar{z}$ of a triangle $\Delta(\bar{z}, \bar{x}, \bar{y})$ with side lengths $\bar{z}\bar{x} = \mathsf{d}(z, x), \bar{z}\bar{y} = \mathsf{d}(z, y), \bar{x}\bar{y} = \mathsf{d}(x, y)$ in the simply connected two-dimensional space of constant curvature K, i.e.

$$\sphericalangle_K(z; x, y) = \arccos\left(\frac{\cos(\sqrt{K}\,\mathsf{d}(x, y)) - \cos(\sqrt{K}\,\mathsf{d}(z, x)) \cdot \cos(\sqrt{K}\,\mathsf{d}(z, y))}{\sin(\sqrt{K}\,\mathsf{d}(z, x)) \cdot \sin(\sqrt{K}\,\mathsf{d}(z, y))}\right) \tag{2.4}$$

(with appropriate interpretations/modifications if $K < 0$ or $K = 0$). If $K \cdot [\,\mathsf{d}(z, x) + \mathsf{d}(x, y) + \mathsf{d}(y, z)]^2 \geq (2\pi)^2$ we put $\sphericalangle_K(z; x, y) := -\infty$. We define

$$\underline{\mathsf{curv}}(M, \mathsf{d}) = \sup\{K \in \mathbb{R} : (M, \mathsf{d}) \text{ has curvature } \geq K\}.$$

Complete length spaces with curvature $\geq K$ and finite Hausdorff dimension are called *Alexandrov spaces* with curvature $\geq K$. For complete geodesic spaces there are several alternative (but

equivalent) ways to define this curvature bound: via triangle comparison, angle monotonicity, convexity properties of the distance. For instance, one can interpret it as a weak formulation of

$$\operatorname{Hess}\frac{1}{K}\cos(\sqrt{K}\,\mathsf{d}(z,\cdot)) \geq -\cos(\sqrt{K}\,\mathsf{d}(z,\cdot)) \tag{2.5}$$

for all $z \in M$ (with appropriate modification in the case $K \leq 0$, e.g. $\operatorname{Hess}\mathsf{d}^2(z,\cdot)/2 \leq 1$ if $K = 0$).

Example 2.5. Let M be a complete Riemannian manifold with Riemannian distance d and dimension $n \geq 2$. Then $\underline{\mathsf{curv}}(M, \mathsf{d})$ is the greatest lower bound for the *sectional curvature* of M.

Example 2.6. If M is an interval $\subset \mathbb{R}$ or if M is a circle of length L then $\underline{\mathsf{curv}}(M, \mathsf{d}) = +\infty$. This example is regarded as pathological.

Similarly, one can define metric spaces of curvature $\leq K$ and a number $\overline{\mathsf{curv}}(M, \mathsf{d})$ (which coincides with the least upper bound for the sectional curvature if M is a Riemannian manifold). However, in this paper we concentrate on lower curvature bounds.

Proposition 2.7. *For each complete length space (M, d) the following properties hold:*

(i) Scaling*: $\underline{\mathsf{curv}}(M, \alpha\,\mathsf{d}) = \alpha^{-2}\,\underline{\mathsf{curv}}(M, \mathsf{d})$ for all $\alpha \in \mathbb{R}_+$.*

(ii) Products*: If $(M, \mathsf{d}) = \bigotimes_{i=1}^n (M_i, \mathsf{d}_i)$ with complete length spaces (M_1, d_1), …, (M_n, d_n) consisting of more than one point then*

$$\underline{\mathsf{curv}}(M, \mathsf{d}) = \inf\{\underline{\mathsf{curv}}(M_1, \mathsf{d}_1), \ldots, \underline{\mathsf{curv}}(M_n, \mathsf{d}_n), 0\}.$$

(iii) Local/global*: If $M = \bigcup_{i\in I} M_i$ with open subsets $M_i \subset M$ then*

$$\underline{\mathsf{curv}}(M, \mathsf{d}) = \inf_{i\in I} \underline{\mathsf{curv}}(M_i, \mathsf{d}_i)$$

('Topogonov's globalization theorem').

(iv) Convergence*: If $((M_n, \mathsf{d}_n))_{n\in\mathbb{N}}$ is a sequence of complete length spaces with Gromov-Hausdorff distance $\mathsf{D}^{\mathsf{GH}}((M_n, \mathsf{d}_n), (M, \mathsf{d})) \to 0$ as $n \to \infty$ then*

$$\underline{\mathsf{curv}}(M, \mathsf{d}) \geq \limsup_{n\to\infty} \underline{\mathsf{curv}}(M_n, \mathsf{d}_n).$$

In particular, for each $K \in \mathbb{R}$ the set

$$\mathsf{X}_c(K) = \{(M, \mathsf{d}) \in \mathsf{X}_c : \underline{\mathsf{curv}}(M, \mathsf{d}) \geq K\}$$

is a closed subset of $(\mathsf{X}_c, \mathsf{D}^{\mathsf{GH}})$.

(v) Compactness*: For each $K \in \mathbb{R}, N \in \mathbb{N}$ and $\Delta \in \mathbb{R}_+$ the set $\mathsf{X}_c(K, N, \Delta)$ of all compact length spaces (M, d) with curvature $\geq K$, Hausdorff dimension $\leq N$ and diameter $\leq \Delta$ is compact w.r.t. D^{GH} ('Gromov's compactness theorem').*

Definition 2.8. A geodesic space (M, d) is called *nonbranching* iff for each quadruple of points z, x_0, x_1, x_2 with z being the midpoint of x_0 and x_1 as well as the midpoint of x_0 and x_2 it follows that $x_1 = x_2$.

Remark 2.9. If a geodesic space has curvature $\geq K$ for some $K \in \mathbb{R}$ then it is nonbranching.

2.3 The L_2-Wasserstein Space

Probability measures on metric spaces will play an important role throughout this paper. We collect some definitions and the basic facts on the L_2-Wasserstein distance. For further reading we recommend [Du89], [KR57], [RR98], [Vi03] and [Wa69].
For the rest of this chapter, let (M, d) be a *complete separable* metric space. A *measure* ν on M will always mean a measure on $(M, \mathcal{B}(M))$ with $\mathcal{B}(M)$ being the Borel σ-algebra of M (generated by the open balls in M). Recall that $\mathrm{supp}[\nu]$, the *support* of ν, is the smallest closed set $M_0 \subset M$ such that $\nu(M \backslash M_0) = 0$. The *push forward* of ν under a measurable map $f : M \to M'$ into another metric space M' is the probability measure $f_*\nu$ on M' given by

$$(f_*\nu)(A) := \nu(f^{-1}(A))$$

for all measurable $A \subset M'$. Given two measures μ, ν on M we say that a measure q on $M \times M$ is a *coupling* of μ and ν iff its marginals are μ and ν, that is, iff

$$q(A \times M) = \mu(A), \quad q(M \times A) = \nu(A)$$

for all measurable sets $A \subset M$. (This in particular implies that the total masses coincide: $\nu(M) = q(M \times M) = \mu(M)$.) If μ and ν are probability measures then for instance one such coupling is the product measure $\mu \times \nu$.
The L_2-*Wasserstein distance* between μ and ν is defined as

$$\mathsf{d}_W(\mu, \nu) = \inf \left\{ \left(\int_{M \times M} \mathsf{d}^2(x, y) dq(x, y) \right)^{1/2} : q \text{ is a coupling of } \mu \text{ and } \nu \right\}. \tag{2.6}$$

Note that $\mathsf{d}_W(\mu, \nu) = +\infty$ whenever $\mu(M) \neq \nu(M)$. We denote by $\mathcal{P}_2(M, \mathsf{d})$ or briefly $\mathcal{P}_2(M)$ the space of all probability measures ν on M with finite second moments:

$$\int_M \mathsf{d}^2(o, x) d\nu(x) < \infty$$

for some (hence all) $o \in M$. The pair $(\mathcal{P}_2(M), \mathsf{d}_W)$ is called L_2-*Wasserstein space* over (M, d).

Proposition 2.10.

(i) $(\mathcal{P}_2(M), \mathsf{d}_W)$ *is a complete separable metric space.*

The map $x \mapsto \delta_x$ *defines an isometric and totally geodesic embedding of* (M, d) *into* $(\mathcal{P}_2(M), \mathsf{d}_W)$.

The set of all normalized configurations $\mu = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$ *with* $n \in \mathbb{N}$ *and* $x_1, \ldots, x_n \in M$ *is dense in* $\mathcal{P}_2(M)$.

(ii) d_W*-convergence implies weak convergence (in the sense of measures). More precisely. if* $(\mu_n)_{n \in \mathbb{N}}$ *is a sequence in* $\mathcal{P}_2(M)$ *then* $\mathsf{d}_W(\mu_n, \mu) \to 0$ *if and only if* $\mu_n \to \mu$ *weakly and*

$$\lim_{R \to \infty} \sup_n \int_{M \backslash B_R(o)} \mathsf{d}^2(o, x) d\mu_n(x) = 0 \tag{2.7}$$

for some (hence each) point $o \in M$. *Note that obviously (2.7) is always satisfied if* (M, d) *is bounded.*

(iii) $(\mathcal{P}_2(M), \mathsf{d}_W)$ *is a compact space or a length space if and only if* (M, d) *is so.*

(iv) If M is a length space with more than one point then

$$\underline{\mathsf{curv}}(\mathcal{P}_2(M),\, \mathsf{d}_W) = 0 \quad \Longleftrightarrow \quad \underline{\mathsf{curv}}(M,\, \mathsf{d}) \geq 0$$

and

$$\underline{\mathsf{curv}}(\mathcal{P}_2(M),\, \mathsf{d}_W) = -\infty \quad \Longleftrightarrow \quad \underline{\mathsf{curv}}(M,\, \mathsf{d}) < 0.$$

Proof. (i), (ii): [RR98], [Vi03].
$(iii)_a$ The 'only if' statements follow from the fact that M is isometrically embedded in $\mathcal{P}_2(M)$.
$(iii)_b$ Compactness of M implies compactness of $\mathcal{P}_2(M)$ according to (ii) and Prohorov's theorem.
$(iii)_c$ Assume that $(M,\, \mathsf{d})$ is a length space and let $\epsilon > 0$ and $\mu, \nu \in \mathcal{P}_2(M)$ be given. We have to prove that there exists an ϵ-midpoint η of μ and ν. Choose $n \in \mathbb{N}$ and $x_1, \ldots, x_n, y_1, \ldots, y_n \in M$ such that $\mathsf{d}_W(\mu, \bar{\mu}) \leq \epsilon/3$, $\mathsf{d}_W(\nu, \bar{\nu}) \leq \epsilon/3$ and $\mathsf{d}_W^2(\bar{\mu}, \bar{\nu}) = \frac{1}{n}\sum_{i=1}^n \mathsf{d}^2(x_i, y_i)$ where $\bar{\mu} := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}, \bar{\nu} := \frac{1}{n}\sum_{i=1}^n \delta_{y_i}$ For each $i = 1, \ldots, n$ let z_i be an $\epsilon/3$-midpoint of x_i and y_i and put $\eta := \frac{1}{n}\sum_{i=1}^n \delta_{z_i}$. Then

$$\begin{aligned}
\mathsf{d}_W(\bar{\mu}, \eta) \;&\leq\; \left(\frac{1}{n}\sum_{i=1}^n \mathsf{d}^2(x_i, z_i)\right)^{1/2} \;\leq\; \left(\frac{1}{n}\sum_{i=1}^n \left[\frac{1}{2}\,\mathsf{d}(x_i, y_i) + \frac{\epsilon}{3}\right]^2\right)^{1/2} \\
&\leq\; \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^n \mathsf{d}^2(x_i, y_i)\right)^{1/2} + \frac{\epsilon}{3} \;\leq\; \frac{1}{2}\,\mathsf{d}_W(\bar{\mu}, \bar{\nu}) + \frac{\epsilon}{3}
\end{aligned}$$

and thus $\mathsf{d}_W(\mu, \eta) \leq \frac{1}{2}\,\mathsf{d}_W(\mu, \nu) + \epsilon$. Similarly, $\mathsf{d}_W(\nu, \eta) \leq \frac{1}{2}\,\mathsf{d}_W(\mu, \nu) + \epsilon$. This proves the claim.
$(iv)_a$ Assume that $(M,\, \mathsf{d})$ has curvature ≥ 0. Then for each $n \in \mathbb{N}$ the space $M^n = M \times \cdots \times M$ has curvature ≥ 0 (Proposition 2.7 (ii)). According to [St99], the latter is equivalent to

$$\sum_{i,j=1}^l \lambda_i \lambda_j\, \mathsf{d}^2(y_i, y_j) \leq 2 \sum_{i=1}^l \lambda_i\, \mathsf{d}^2(y_i, y_0) \tag{2.8}$$

for all $l \in \mathbb{N}$, all $\lambda_1, \ldots, \lambda_l \in \mathbb{R}_+$ with $\sum_{i=1}^l \lambda_i = 1$ and all $y_0, y_1, \ldots, y_l \in M^n$. In order to prove that $(\mathcal{P}_2(M),\, \mathsf{d}_W)$ has curvature ≥ 0, let $l \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_l \in \mathbb{R}_+$ and $\nu_0, \nu_1, \ldots, \nu_l \in \mathcal{P}_2(M)$ be given. For $\epsilon > 0$ choose $n \in \mathbb{N}$ and $y_0 = (y_{01}, \ldots, y_{0n}), \ldots, y_l = (y_{l1}, \ldots, y_{ln}) \in M^n$ such that $\mathsf{d}_W(\nu_i, \bar{\nu}_i) \leq \epsilon$ for all $i = 0, 1, \ldots, l$ and $\mathsf{d}_W^2(\bar{\nu}_i, \bar{\nu}_0) = \frac{1}{n}\sum_{k=1}^n \mathsf{d}^2(y_{ik}, y_{0k}) = \frac{1}{n}\,\mathsf{d}^2(y_i, y_0)$ for all $i = 1, \ldots, l$ where we put

$$\bar{\nu}_i = \frac{1}{n}\sum_{k=1}^n \delta_{y_{ik}}.$$

Then $\mathsf{d}_W^2(\bar{\nu}_i, \bar{\nu}_j) \leq \frac{1}{n}\sum_{k=1}^n \mathsf{d}^2(y_{ik}, y_{jk}) = \frac{1}{n}\,\mathsf{d}^2(y_i, y_j)$ for all $i, j = 1, \ldots, l$ and thus by (2.8)

$$\begin{aligned}
\sum_{i,j=1}^l \lambda_i \lambda_j\, \mathsf{d}_W^2(\bar{\nu}_i, \bar{\nu}_j) &\leq \frac{1}{n}\sum_{i,j=1}^l \lambda_i \lambda_j\, \mathsf{d}^2(y_i, y_j) \\
\leq\; \frac{2}{n}\sum_{i=1}^l \lambda_i\, \mathsf{d}^2(y_i, y_0) \;&=\; 2\sum_{i=1}^l \lambda_i\, \mathsf{d}_W^2(\bar{\nu}_i, \bar{\nu}_0).
\end{aligned}$$

In the limit $\epsilon \to 0$ this yields

$$\sum_{i,j=1}^l \lambda_i \lambda_j\, \mathsf{d}_W^2(\nu_i, \nu_j) \leq 2 \sum_{i=1}^l \lambda_i\, \mathsf{d}_W^2(\nu_i, \nu_0)$$

which (again by [St99]) proves the claim.
$(iv)_b$ Since $(M,\,\mathsf{d})$ is isometrically and totally geodesically embedded into $(\mathcal{P}_2(M),\,\mathsf{d}_W)$ it is obvious that $\underline{\mathsf{curv}}(\mathcal{P}_2(M),\,\mathsf{d}_W) \le \underline{\mathsf{curv}}(M,\,\mathsf{d})$.
$(iv)_c$ Assume that $\underline{\mathsf{curv}}(M,\,\mathsf{d}) < 0$. Choose $K < 0$ such that $\underline{\mathsf{curv}}(M,\,\mathsf{d}) < K$. Then there exist points $x_0, x_1, x_2, x_3 \in M$ with

$$\sphericalangle_K(x_0; x_1, x_2) + \sphericalangle_K(x_0; x_2, x_3) + \sphericalangle_K(x_0; x_3, x_1) > 2\pi.$$

Choose a point $z \in M$ 'far away' from the x_i, say $\mathsf{d}(z, x_i) \ge 3\,\mathsf{d}(x_i, x_j)$ for all $i, j = 0, 1, 2, 3$. [This is always possible since M is a length space and the x_i for $i = 1, 2, 3$ can be replaced by points x_i' lying arbitrarily close to x_0 on approximate geodesics connecting x_0 and x_i.]
For $t \in]0, 1]$ and $i = 0, 1, 2, 3$ define $\mu_i := t\delta_{x_i} + (1-t)\delta_z$. Then

$$\mathsf{d}_W^2(\mu_i, \mu_j) = t\,\mathsf{d}^2(x_i, x_j)$$

for all $i, j = 0, 1, 2, 3$ and thus according to formula (2.4)

$$\sphericalangle_K(x_0; x_i, x_j) = \sphericalangle_{K/t}(\mu_0; \mu_i, \mu_j).$$

Therefore,

$$\sphericalangle_{K/t}(\mu_0; \mu_1, \mu_2) + \sphericalangle_{K/t}(\mu_0; \mu_2, \mu_3) + \sphericalangle_{K/t}(\mu_0; \mu_3, \mu_1) > 2\pi$$

which implies

$$\underline{\mathsf{curv}}(\mathcal{P}(M),\,\mathsf{d}_W) < K/t.$$

Since the latter holds for the chosen $K < 0$ and all arbitrarily small $t > 0$, it proves the claim.
$(iv)_d$ Finally, it remains to prove that $\underline{\mathsf{curv}}(\mathcal{P}_2(M),\,\mathsf{d}_W) \le 0$ if M has more than one point. Assume that $\underline{\mathsf{curv}}(\mathcal{P}_2(M),\,\mathsf{d}_W) \ge K$ for some $K > 0$ and that $x_0, x_1 \in M$ with $x_0 \ne x_1$. Given $\epsilon > 0$ let y be an ϵ-midpoint of x_0 and x_1. Put $\nu_0 = \delta_{x_0}, \nu_1 = \delta_{x_1}, \mu = \delta_y$ and $\eta = \frac{1}{2}\delta_{x_0} + \frac{1}{2}\delta_{x_1}$. Then $\mathsf{d}_W(\nu_0, \nu_1) = \mathsf{d}(x_0, x_1)$, $\mathsf{d}_W(\eta, \mu) \ge \frac{1}{2}\,\mathsf{d}(x_0, x_1)$, $\mathsf{d}_W(\nu_i, \eta) = \frac{1}{\sqrt{2}}\,\mathsf{d}(x_0, x_1)$ and $\mathsf{d}_W(\nu_i, \mu) \le \frac{1}{2}\,\mathsf{d}(x_0, x_1) + \epsilon$ for $i = 0, 1$. In particular, μ is an ϵ-midpoint of ν_0 and ν_1.
Our curvature assumption on $\mathcal{P}_2(M)$ implies (via quadruple comparison for $(\mu; \nu_0, \nu_1, \eta)$ or via triangle comparison for (ν_0, ν_1, η)) that

$$\begin{aligned} &2\cos\Big(\frac{\sqrt{K}}{2}\,\mathsf{d}_W(\nu_0, \nu_1)\Big) \cdot \cos\big(\sqrt{K}\,\mathsf{d}_W(\eta, \mu)\big) \\ &\quad \le \cos\big(\sqrt{K}\,\mathsf{d}_W(\eta, \nu_0)\big) + \cos\big(\sqrt{K}\,\mathsf{d}_W(\eta, \nu_1)\big) + \epsilon' \end{aligned}$$

with some $\epsilon' \to 0$ as $\epsilon \to 0$. Therefore,

$$\begin{aligned} &2\cos\Big(\frac{\sqrt{K}}{2}\,\mathsf{d}(x_0, x_1)\Big) \cdot \cos\Big(\frac{\sqrt{K}}{2}\,\mathsf{d}(x_0, x_1)\Big) \\ &\quad \le \cos\Big(\sqrt{\frac{K}{2}}\,\mathsf{d}(x_0, x_1)\Big) + \cos\Big(\sqrt{\frac{K}{2}}\,\mathsf{d}(x_0, x_1)\Big). \end{aligned}$$

Now choosing $x_0, x_1 \in M$ with sufficiently small $\mathsf{d}(x_0, x_1)$ leads to a contradiction. □

Let us recall that a *Markov kernel* on M is a map $Q : M \times \mathcal{B}(M) \to [0, 1]$ (where $\mathcal{B}(M)$ denotes the Borel σ-algebra of M) with the following properties:

- for each $x \in M$ the map $Q(x, \cdot) : \mathcal{B}(M) \to [0, 1]$ is a probability measure on M, usually denoted by $Q(x, dy)$;

- for each $A \in \mathcal{B}(M)$ the function $Q(\cdot, A) : M \to [0,1]$ is measurable.

Lemma 2.11.

(i) For each pair $\mu, \nu \in \mathcal{P}_2(M)$ there exists a coupling q (called 'optimal coupling') such that

$$\mathsf{d}_W^2(\mu,\nu) = \int_{M\times M} \mathsf{d}^2(x,y)\, dq(x,y)$$

and there exist Markov kernels Q, Q' on M ('optimal transport kernels') such that

$$dq(x,y) = Q(x,dy)d\mu(x) = Q'(y,dx)d\nu(x).$$

(In general, neither q nor Q, Q' are unique.)

(ii) For each geodesic $\Gamma : [0,1] \to \mathcal{P}_2(M)$, each $l \in \mathbb{N}$ and each partition $0 = t_0 < t_1 < \cdots < t_l = 1$ there exists a probability measure $\hat{q}$ on M^{l+1} with the following properties:

- *the projection on the i-th factor is $\Gamma(t_i)$ (for all $i = 0, 1, \ldots, l$);*
- *for $\hat{q}$-almost every $x(x_0, \ldots, x_l) \in M^{l+1}$ and every $i, j = 0, 1, \ldots, l$*

$$\mathsf{d}(x_i, x_j) = |t_i - t_j| \cdot \mathsf{d}(x_0, x_l). \tag{2.9}$$

In particular, for every pair $i, j \in \{0, 1, \ldots, l\}$ the projection on the i-th and j-th factor is an optimal coupling of $\Gamma(t_i)$ and $\Gamma(t_j)$.
In the case $l = 2$ and $t = \frac{1}{2}$, (2.9) states that for $\hat{q}$-a.e. $(x_0, x_1, x_2) \in M^3$ the point x_1 is a midpoint of x_0 and x_2.

(iii) If M is a nonbranching geodesic space then in the previous situation for $\hat{q}$-almost every (x_0, x_1, x_2) and $(y_0, y_1, y_2) \in M^3$

$$x_1 = y_1 \quad \Rightarrow \quad (x_0, x_2) = (y_0, y_2).$$

Proof. (i) For the existence of optimal coupling, see [RR98] or [Du89], 11.8.2.
The existence of optimal transport kernels is a straightforward application of disintegration of measures on Polish spaces (or of the existence of regular conditional probabilities), namely, Q is the disintegration of q w.r.t μ.

(ii) We assume $l = 2$ and $t = \frac{1}{2}$. (The general case follows by iterated application and appropriate modifications.)
Let q_1 be an optimal coupling of $\Gamma(0)$ and $\Gamma(\frac{1}{2})$ and let q_2 be an optimal coupling of $\Gamma(\frac{1}{2})$ and $\Gamma(1)$. Then there exists a probability measure $\hat{q}$ on $M \times M \times M$ such that its projection on the first two factors is q_1 and the projection on last two factors is q_2 ([Du89], section 11.8). Hence, for $i = 1, 2, 3$ the projection of $\hat{q}$ on the i-th factor is $\Gamma(\frac{i-1}{2})$ and for $i = 1, 2$

$$\mathsf{d}_W^2(\Gamma(\frac{i-1}{2}), \Gamma(\frac{i}{2})) = \int_{M^3} \mathsf{d}^2(x_{i-1}, x_i) d\hat{q}(x_0, x_1, x_2).$$

Then

$$\begin{aligned}
\mathsf{d}_W(\Gamma(0),\Gamma(1)) &\leq \left[\int \mathsf{d}^2(x_0,x_2)d\hat{q}(x_0,x_1,x_2)\right]^{\frac{1}{2}} \\
&\overset{(*)}{\leq} \left[\int [\,\mathsf{d}(x_0,x_1)+\,\mathsf{d}(x_1,x_2)]^2\, d\hat{q}(x_0,x_1,x_2)\right]^{\frac{1}{2}} \\
&\overset{(**)}{\leq} \left[\int \mathsf{d}^2(x_0,x_1)d\hat{q}(x_0,x_1,x_2)\right]^{\frac{1}{2}} \\
&\quad + \left[\int \mathsf{d}^2(x_1,x_2)d\hat{q}(x_0,x_1,x_2)\right]^{\frac{1}{2}} \\
&= \mathsf{d}_W(\Gamma(0),\Gamma(\frac{1}{2}))+\,\mathsf{d}_W(\Gamma(\frac{1}{2}),\Gamma(1)).
\end{aligned}$$

Since $\Gamma(\frac{1}{2})$ is a midpoint of $\Gamma(0)$ and $\Gamma(1)$, the previous inequalities (*) and (**) have to be equalities. From equality in (*) we conclude that $\hat{q}$-almost surely the point x_1 lies on some geodesic connecting x_0 and x_2. Equality in (**) implies that $\hat{q}$-almost surely the point x_1 is a midpoint of x_0 and x_2.

(*iii*) Let $\eta = \Gamma(\frac{1}{2})$ be the distribution of the midpoints and let Q be a disintegration of $\hat{q}$ w.r.t. η, i.e.

$$d\hat{q}(x,z,y) = Q(z,d(x,y))d\eta(z).$$

We have to prove that for η-a.e. $z \in M$ the probability measure $Q(z,\cdot)$ is a Dirac measure (sitting on some $(x,y) \in M \times M$). Denote the marginals of $Q(z,\cdot)$ by $p_1(z,\cdot)$ and $p_2(z,\cdot)$. Then

$$\begin{aligned}
\int \mathsf{d}^2(x,y)Q(z,d(x,y)) &= \int [2\,\mathsf{d}^2(x,z)+2\,\mathsf{d}^2(z,y)]Q(z,d(x,y)) \\
&= \int\int [2\,\mathsf{d}^2(x,z)+2\,\mathsf{d}^2(z,y)]\, p_1(z,dx)\, p_2(z,dy) \\
&\overset{(***)}{\geq} \int\int \mathsf{d}^2(x,y)\, p_1(z,dx)\, p_2(z,dy).
\end{aligned}$$

The optimality of $\hat{q}$ implies that for η-a.e. $z \in M$ the measure $Q(z,\cdot)$ is an optimal coupling of $p_1(z,\cdot)$ and $p_2(z,\cdot)$. Hence, there has to be equality in $(***)$ which in turn implies that for $p_1(z,.)$-a.e. $x \in M$ and $p_2(z,.)$-a.e. $y \in M$ the point z is a midpoint of x and y. Since M is nonbranching this implies that both $p_1(z,.)$ and $p_2(z,.)$ are Dirac measures. Thus $Q(z,\cdot)$ is also a Dirac measure. This proves the claim. □

Remark 2.12.
(i) Couplings q of μ and ν are also called *transportation plans* from μ to ν. If μ is the distribution of locations at which a good is produced and ν is the distribution of locations where it is consumed, then each coupling q of μ and ν gives a plan how to transport the products to the consumer. More precisely, for each x the kernel $Q(x,dy)$ determines how to distribute goods produced at the location x to various consumers at location y.

(ii) The interpretation of Lemma 2.11 (ii) is that for each geodesic in $\mathcal{P}_2(M)$ the mass is transported along geodesics of the underlying space M. (iii) states that 'the paths of optimal mass transportation do not cross each other halfway'.

(iii) If M is a complete Riemannian manifold with Riemannian volume m then for each pair

$\mu, \nu \in \mathcal{P}_2(M)$ with $\mu \ll m$ there exists an *optimal transport map* $F_1 : M \to M$ such that $dq(x,y) = Q(x,dy)d\mu(x)$ with

$$Q(x,dy) = d\delta_{F_1(x)}(y)$$

is the unique optimal coupling of μ and ν.
More precisely, there exists a function $\varphi : M \to \mathbb{R}$ such that for μ-a.e. $x \in M$ and $t \in [0,1]$

$$F_t(x) = \exp_x(-t\nabla\varphi(x))$$

exists and the unique geodesic Γ in $\mathcal{P}_2(M)$ connecting $\mu = \Gamma(0)$ and $\nu = \Gamma(1)$ is given by

$$\Gamma(t) := (F_t)_*\mu,$$

the push forward of μ under F_t, [CMS01].

3 Metric Measure Spaces

3.1 The Metric $\mathbb{D}$

Throughout this paper, a *metric measure space* will always be a triple (M, d, m) where

- (M, d) is a complete separable metric space,
- m is a measure on $(M, \mathcal{B}(M))$ which is locally finite in the sense that $m(B_r(x)) < \infty$ for all $x \in M$ and all sufficiently small $r > 0$.

A metric measure space (M, d, m) is called *normalized* iff $m(M) = 1$. It is called compact or locally compact or geodesic iff the metric space (M, d) is compact or locally compact or geodesic, resp.
Two metric measure spaces (M, d, m) and (M', d', m') are called *isomorphic* iff there exists an isometry $\psi : M_0 \to M_0'$ between the supports $M_0 := \mathrm{supp}[m] \subset M$ and $M_0' := \mathrm{supp}[m'] \subset M'$ such that

$$\psi_* m = m'.$$

The *variance* of a metric measure space (M, d, m) is defined as

$$\mathbb{V}ar(M, \mathsf{d}, m) = \inf \int_{M'} \mathsf{d}'^2(z,x) dm'(x). \tag{3.1}$$

where the inf is taken over all metric measure spaces (M', d', m') which are isomorphic to (M, d, m) and over all $z \in M'$. Note that a normalized metric measure space (M, d, m) has finite variance if and only if

$$\int_M \mathsf{d}^2(z,x) dm(x) < \infty \tag{3.2}$$

for some (hence all) $z \in M$. Similarly, the *radius* of (M, d, m) is defined as

$$\begin{aligned} \mathrm{rad}(M, \mathsf{d}, m) = \inf\{R > 0 : \quad & \exists (M', \mathsf{d}', m') \text{ isom. to } (M, \mathsf{d}, m), \\ & \exists z \in M' : \quad \mathrm{supp}[m'] \subset \overline{B'}_R(z)\}. \end{aligned}$$

The *diameter* of a metric measure space (M, d, m) is defined as the diameter of the metric space $(\mathrm{supp}[m], d)$:

$$\mathrm{diam}(M, \mathsf{d}, m) = \sup\{d(x,y) : x, y \in \mathrm{supp}[m]\}.$$

Example 3.1. Let $M = \mathbb{R}^2$ with Euclidean distance d and $m = \frac{1}{3}(\delta_{x_1} + \delta_{x_2} + \delta_{x_3})$ where x_1, x_2, x_3 are the vertices of an equilateral triangle of sidelength 1. Then

$$\mathbb{V}ar(M, \mathsf{d}, m) = \frac{1}{4} \quad \text{whereas} \quad \inf_{z \in M} \int \mathsf{d}^2(z, x) dm(x) = \frac{1}{3}.$$

(Hint: embed supp$[m]$ isometrically into a graph or into a hyperbolic space with curvature close to $-\infty$.)

The family of all isomorphism classes of metric measure spaces will be denoted by $\mathbb{X}$. For each $\lambda \in \mathbb{R}_+$, let $\mathbb{X}_\lambda$ denote the family of isomorphism classes of metric measure spaces (M, d, m) with finite variances and total mass $m(M) = \lambda$. Moreover, for $\Delta \in \mathbb{R}_+$, let $\mathbb{X}_\lambda(\Delta)$ denote the family of isomorphism classes of metric measure spaces (M, d, m) with diameter $\leq \Delta$ and total mass $m(M) = \lambda$. If $\lambda\Delta \neq 0$, the map

$$(M, \mathsf{d}, m) \mapsto (M, \Delta\, \mathsf{d}, \lambda m)$$

defines a bijection between $\mathbb{X}_1$ and $\mathbb{X}_\lambda$ and also a bijection between $\mathbb{X}_1(1)$ and $\mathbb{X}_\lambda(\Delta)$.
For $\lambda > 0$, the family $\mathbb{X}_\lambda$ contains a unique element with $\mathbb{V}ar(M, \mathsf{d}, m) = 0$, namely, $m = \lambda \cdot \delta_o$ for some $o \in M$. Here a priori M is an arbitrary nonempty set. But without restriction it contains just one point, say $M = \{o\}$. The family $\mathbb{X}_0$ is pathological: it contains only one element, the 'empty space'.

Definition 3.2.

(i) Given two metric measure spaces (M, d, m) and (M', d', m') we say that a measure q on the product space $M \times M'$ is a *coupling of m and m'* iff

$$q(A \times M') = m(A), \quad q(M \times A') = m'(A') \tag{3.3}$$

for all measurable sets $A \subset M, A' \subset M'$. We say that a pseudo metric $\hat{\mathsf{d}}$ on the disjoint union $M \sqcup M'$ is a *coupling of* d *and* d' iff

$$\hat{\mathsf{d}}(x, y) = \mathsf{d}(x, y), \quad \hat{\mathsf{d}}(x', y') = \mathsf{d}'(x', y') \tag{3.4}$$

for all $x, y \in \text{supp}[m] \subset M$ and all $x', y' \in \text{supp}[m'] \subset M'$.

(ii) We define the *distance* $\mathbb{D}$ between two metric measure spaces by

$$\begin{aligned} \mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m')) \;=\; \inf\{ & \Big(\int_{M \times M'} \hat{\mathsf{d}}^2(x, y) dq(x, y) \Big)^{1/2} : \\ & \hat{\mathsf{d}} \text{ is a coupling of } \mathsf{d} \text{ and } \mathsf{d}', \\ & q \text{ is a coupling of } m \text{ and } m' \}. \end{aligned}$$

Remark 3.3. (i) Note that the integrals involved in the definition of $\mathbb{D}$ are well-defined since each coupling $\hat{\mathsf{d}}$ is a function on $(M \sqcup M') \times (M \sqcup M') = (M \times M) \sqcup (M \times M') \sqcup (M' \times M) \sqcup (M' \times M')$ and each coupling q is a measure on $M \times M'$.
(ii) In the definition of the distance $\mathbb{D}$ we may restrict ourselves to take the infimum over all *complete separable metrics* $\hat{\mathsf{d}}$ on $M \sqcup M'$ which are couplings of d and d'. Indeed, given any (pseudo metric) coupling $\hat{\mathsf{d}}$ of d and d' and any $\epsilon > 0$ we obtain a complete separable metric $\hat{\mathsf{d}}_\epsilon$ which is a coupling of d and d' as follows:

$$\hat{\mathsf{d}}_\epsilon = \begin{cases} \mathsf{d}, & \text{on } (M \times M) \sqcup (M' \times M') \\ \mathsf{d} + \epsilon, & \text{on } (M \times M') \sqcup (M' \times M). \end{cases}$$

(iii) One easily verifies that

$$\mathbb{D}((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) = \inf\,\hat{\mathsf{d}}_W(\psi_* m,\psi'_* m') \tag{3.5}$$

where the inf is taken over all metric spaces $(\hat{M},\,\hat{\mathsf{d}})$ with isometric embeddings $\psi : M_0 \hookrightarrow \hat{M}$, $\psi' : M'_0 \hookrightarrow \hat{M}$ of the supports M_0 and M'_0 of m and m', resp. Here $\hat{\mathsf{d}}_W$ denotes the L_2-Wasserstein distance for measures on $\hat{M}$ as introduced in Chapter 2. In other words,

$$\mathbb{D}((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) = \inf\left(\int_{M\times M'} \hat{\mathsf{d}}^2(\psi(x),\psi'(x'))\,dq(x,x')\right)^{1/2} \tag{3.6}$$

where the inf now is taken over all metric spaces $(\hat{M},\,\hat{\mathsf{d}})$ with isometric embeddings $\psi : M_0 \hookrightarrow \hat{M}$, $\psi' : M'_0 \hookrightarrow \hat{M}$ and over all couplings q of m and m'.
Indeed, the set $\hat{M}$ can always be chosen as the disjoint union of M_0 and M'_0 (the supports of m and m'), i.e.

$$\hat{M} = M_0 \sqcup M'_0$$

and ψ, ψ' can be chosen as identities. Hence,

$$\mathbb{D}((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) \;=\; \inf_{\hat{\mathsf{d}}}\,\hat{\mathsf{d}}_W(m,m')$$

where the $\inf_{\hat{\mathsf{d}}}$ is taken over all metrics (or, equivalently, over all pseudo metrics) on $M_0 \sqcup M'_0$ the restriction of which coincide on M_0 with d and on M'_0 with d'.

Let us summarize some elementary properties of $\mathbb{D}$.

Lemma 3.4.

(i) If $(M,\,\mathsf{d}) = (M',\,\mathsf{d}')$ then $\mathbb{D}((M,\,\mathsf{d},m),(M,\,\mathsf{d},m')) \le \mathsf{d}_W(m,m')$. In general, there will be no equality.

(ii) If $m(M) \neq m'(M')$ then $\mathbb{D}((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) = +\infty$.

(iii) For all $\alpha,\beta \in \mathbb{R}_+$

$$\mathbb{D}((M,\alpha\,\mathsf{d},\beta m),(M',\alpha\,\mathsf{d}',\beta m')) = \alpha\sqrt{\beta}\cdot\mathbb{D}((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m'))$$

(iv) If $m = \sum_n m_n$ and $m' = \sum_n m'_n$ then

$$\mathbb{D}^2((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) \le \sum_n \mathbb{D}^2((M,\,\mathsf{d},m_n),(M',\,\mathsf{d}',m'_n)).$$

In particular, with $M_n = \mathrm{supp}[m_n], M'_n = \mathrm{supp}[m'_n]$

$$\mathbb{D}^2((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) \le \sum_n \mathbb{D}^2((M_n,\,\mathsf{d}|_{M_n},m_n),(M'_n,\,\mathsf{d}'|_{M'_n},m'_n)).$$

Now let us concentrate on normalized metric measure spaces.

Lemma 3.5.

(i) If $m(M) = 1$ and $m' = \delta_o$ for some $o \in M'$ then

$$\mathbb{D}^2((M,\,\mathsf{d},m),(M',\,\mathsf{d}',m')) = \mathbb{V}ar(M,\,\mathsf{d},m).$$

(ii) *The family $\mathbb{X}_{1,*}$ of isomorhpism classes of (M, d, m) with finite supports M_0, say $\{x_1, \dots, x_n\}$, and uniform distribution $m = \frac{1}{n}\sum \delta_{x_i}$ ('normalized configurations') is dense in $\mathbb{X}_1$.*

(iii) *For each $(M, \mathsf{d}, m) \in \mathbb{X}_1$ let $X_1, X_2, \dots$ be an independent sequence of random variables $X_i : \Omega \to M$ (defined on some probability space Ω with values in M) with distribution m and let*

$$m_n(\omega, \cdot) := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)}$$

be their empirical distributions. Then for m-almost every $\omega \in \Omega$

$$(M, \mathsf{d}, m_n(\omega, \cdot)) \to (M, \mathsf{d}, m)$$

in $(\mathbb{X}_1, \mathbb{D})$ as $n \to \infty$.

(iv) *If $m = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $m' = \frac{1}{n}\sum_{i=1}^n \delta_{x'_i}$ then*

$$\mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m')) \le \sup_{i,j} |\, \mathsf{d}_{ij} - \mathsf{d}'_{ij}|$$

where $\mathsf{d}_{ij} := \mathsf{d}(x_i, x_j)$ and $\mathsf{d}'_{ij} := \mathsf{d}'(x'_i, x'_j)$.

Proof. (i) is obvious.
(ii) Given $(M, \mathsf{d}, m) \in \mathbb{X}_1$ we have $m \in \mathcal{P}_2(M, \mathsf{d})$ by (3.2). Then by Proposition 2.10(i)

$$\forall \epsilon > 0 : \exists n \in \mathbb{N},\ x_1, \dots, x_n \in M : \qquad \mathsf{d}_W(m, \bar{m}) \le \epsilon$$

where $\bar{m} := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Hence, $(M, \mathsf{d}, \bar{m}) \in \mathbb{X}_{1,*}$ and $\mathbb{D}((M, \mathsf{d}, m), (M, \mathsf{d}, \bar{m})) \le \mathsf{d}_W(m, \bar{m}) \le \epsilon$.
(iii) follows from the Empirical Law of Large Numbers or Varadarajan's Theorem, e.g. [Du89], Theorem 11.4.1.
(iv) Assume (without restriction) that $M = \{x_1, \dots, x_n\}$, $M' = \{x'_1, \dots, x'_n\}$, and $|\, \mathsf{d}_{ij} - \mathsf{d}'_{ij}| \le \epsilon$ for all i, j (with d_{ij}, d'_{ij} as above). Define $\hat{\mathsf{d}}$ on $M \times M'$ by

$$\hat{\mathsf{d}}(x_i, x'_j) := \inf_{k=1,\dots,n} \big[\, \mathsf{d}(x_i, x_k) + \mathsf{d}'(x'_k, x'_j)\big] + \epsilon$$

and analogously on $M' \times M$. As usual, put $\hat{\mathsf{d}} := \mathsf{d}$ on $M \times M$ and $\hat{\mathsf{d}} := \mathsf{d}'$ on $M' \times M'$. Moreover, put

$$m = \frac{1}{n}\sum_{i=1}^{n} \delta_{(x_i, x'_i)}.$$

Then $\hat{\mathsf{d}}$ is a coupling of d and d' and q is a coupling of m and m'. Thus

$$\mathbb{D}^2((M, \mathsf{d}, m), (M', \mathsf{d}', m')) \le \int_{M \times M'} \hat{\mathsf{d}}^2(x, y) dq(x, y) = \epsilon^2.$$

□

Theorem 3.6. *$(\mathbb{X}_1, \mathbb{D})$ is a complete separable metric space.*

Proof.

(i) Obviously, $\mathbb{D}$ is well-defined and symmetric on $\mathbb{X}_1 \times \mathbb{X}_1$ with values in $\mathbb{R}_+$.

(ii) The density of $\mathbb{X}_{1,*}$ in $\mathbb{X}_1$ follows from Lemma 3.5 (ii).

(iii) According to (ii), separability of $\mathbb{X}_1$ will follow from separability of $\mathbb{X}_{1,*}$. The latter is the disjoint union $\dot\cup_{n\in\mathbb{N}}\tilde{\mathcal{K}}(n)$ where $\tilde{\mathcal{K}}(n) := \{(M, \mathsf{d}, m) \in \mathbb{X}_{1,*} : \mathrm{supp}[m] \text{ has } n \text{ points}\}$. But $\tilde{\mathcal{K}}(n)$ can be identified with

$$\mathcal{K}(n) = \{D = (D_{ij})_{i,j} \in \mathbb{R}_+^{n\times n} : \quad \forall i,j,k \in \{1,\dots,n\} : \\ D_{ij} = D_{ji},\ D_{ij} + D_{jk} \geq D_{ik} \text{ and } D_{ij} = 0 \Leftrightarrow i = j\}.$$

Now each of the $\mathcal{K}(n)$ is separable (as a subset of $\mathbb{R}^{n\times n}$), hence, $\tilde{\mathcal{K}}(n)$ is separable (Lemma 3.5(iv)) and thus finally $\mathbb{X}_{1,*}$ is separable.

(iv) In order to prove the triangle inequality let three metric measure spaces $(M_i, \mathsf{d}_i, m_i) \in \mathbb{X}_1, i = 1,2,3$, be given. Without restriction, we may assume $M_i = \mathrm{supp}[m_i]$ for $i = 1,2,3$. Then for each $\epsilon > 0$ there exist a complete separable metric d_{12} on $M_1 \sqcup M_2$ and a complete separable metric d_{23} on $M_2 \sqcup M_3$ such that

$$\mathbb{D}((M_1, \mathsf{d}_1, m_1), (M_2, \mathsf{d}_2, m_2)) \geq \mathsf{d}_{12}^W(m_1, m_2) - \epsilon,$$

$$\mathbb{D}((M_2, \mathsf{d}_2, m_2), (M_3, \mathsf{d}_3, m_3)) \geq \mathsf{d}_{23}^W(m_2, m_3) - \epsilon,$$

and d_{ij} restricted to M_i coincides with d_i, restricted to M_j coincides with d_j for $(i,j) = (1,2)$ or $(2,3)$. (Here for typographical reasons, we use not a lower but an upper index to indicate the Wasserstein metric derived from a given metric.) Now define d on $M \times M$ with $M := M_1 \sqcup M_2 \sqcup M_3$ by

$$\mathsf{d}(x,y) = \begin{cases} \mathsf{d}_{12}(x,y), & \text{if } x,y \in M_1 \sqcup M_2 \\ \mathsf{d}_{23}(x,y), & \text{if } x,y \in M_2 \sqcup M_3 \\ \inf_{z\in M_2}[\,\mathsf{d}_{12}(x,z) + \mathsf{d}_{23}(z,y)], & \text{if } x \in M_1, y \in M_3 \\ \inf_{z\in M_2}[\,\mathsf{d}_{23}(x,z) + \mathsf{d}_{12}(z,y)], & \text{if } x \in M_3, y \in M_1. \end{cases}$$

Obviously, d is a complete separable metric on M and, restricted to M_i it coincides with d_i (for each $i = 1,2,3$).
Then by the triangle inequality for d^W (Proposition 2.10 (i))

$$\begin{aligned} &\mathbb{D}((M_1, \mathsf{d}_1, m_1), (M_3, \mathsf{d}_3, m_3)) \\ &\leq \mathsf{d}^W(m_1, m_3) \\ &\leq \mathsf{d}^W(m_1, m_2) + \mathsf{d}^W(m_2, m_3) \\ &= \mathsf{d}_{12}^W(m_1, m_2) + \mathsf{d}_{23}^W(m_2, m_3) \\ &\leq \mathbb{D}((M_1, \mathsf{d}_1, m_1), (M_2, \mathsf{d}_2, m_2)) \\ &\qquad + \mathbb{D}((M_2, \mathsf{d}_2, m_2), (M_3, \mathsf{d}_3, m_3)) + 2\epsilon. \end{aligned}$$

This proves the claim.

(v) In order to prove completeness let $((M_n, \mathsf{d}_n, m_n))_{n\in\mathbb{N}}$ be a Cauchy sequence in $(\mathbb{X}_1, \mathbb{D})$. Let us choose a subsequence such that

$$\mathbb{D}((M_{n_k}, \mathsf{d}_{n_k}, m_{n_k}), (M_{n_{k+1}}, \mathsf{d}_{n_{k+1}}, m_{n_{k+1}})) \leq 2^{-k-1}$$

for all $k \in \mathbb{N}$. Then there exist a coupling $\hat{\mathsf{d}}_{k+1}$ of d_{n_k}, $\mathsf{d}_{n_{k+1}}$ and a coupling $\hat{q}_{k+1}$ of m_{n_k}, $m_{n_{k+1}}$ such that

$$\left(\int \hat{\mathsf{d}}_{k+1}^2(x,y)\, d\hat{q}_{k+1}(x,y)\right)^{1/2} \leq 2^{-k}.$$

Without restriction $\hat{\mathsf{d}}_{k+1}$ is a complete separable metric. Let us define recursively a sequence of complete separable metric spaces (M'_k, d'_k) as follows: $(M'_1, \mathsf{d}'_1) := (M_{n_1}, \mathsf{d}_{n_1})$ and $M'_{k+1} = M'_k \sqcup M_{n_{k+1}} / \sim$ with $x \sim y$ iff $\mathsf{d}'_{k+1}(x,y) = 0$ where

$$\mathsf{d}'_{k+1}(x,y) = \begin{cases} \mathsf{d}'_k(x,y), & \text{if } x,y \in M'_k \\ \hat{\mathsf{d}}_{k+1}(x,y) & \text{if } x,y \in M_{n_k} \sqcup M_{n_{k+1}} \\ \inf_{z \in M_{n_k}} [\, \mathsf{d}'_k(x,z) + \hat{\mathsf{d}}_{k+1}(z,y)], & \text{if } x \in M'_k, y \in M_{n_k} \sqcup M_{n_{k+1}}. \end{cases}$$

This way, (M'_k, d'_k) is a sequence of complete separable metric spaces with $M_{n_k} \subset M'_k$ and $M'_k \subset M'_{k+l}$ for all k, l. Hence, $M' = \cup_{k=1}^{\infty} M'_k$ is naturally equipped with a metric $\mathsf{d}' = \lim \mathsf{d}'_k$.
Let (M, d) be the completion of (M', d'). Then $(M_{n_k}, \mathsf{d}_{n_k})$ is isometrically embedded in (M, d) for each $k \in \mathbb{N}$ and the measure m_{n_k} on M_{n_k} defines a push forward measure $\bar{m}_{n_k}$ on M. By construction

$$\mathsf{d}_W(\bar{m}_{n_k}, \bar{m}_{n_{k+1}}) \leq \left(\int \hat{\mathsf{d}}^2_{k+1}(x,y)\, d\hat{q}_{k+1}(x,y) \right)^{1/2} \leq 2^{-k}$$

for all $k \in \mathbb{N}$. Hence, $(\bar{m}_{n_k})_{k \in \mathbb{N}}$ is a Cauchy sequence in $(\mathcal{P}_2(M), \mathsf{d}_W)$. According to Proposition 2.10 the latter is complete. That is, there exists a probability measure m on (M, d) such that

$$\mathbb{D}((M_{n_k}, \mathsf{d}_{n_k}, m_{n_k}), (M, \mathsf{d}, m)) \leq \mathsf{d}_W(\bar{m}_{n_k}, m) \to 0$$

as $k \to \infty$. This in turn implies

$$\mathbb{D}((M_n, \mathsf{d}_n, m_n), (M, \mathsf{d}, m)) \to 0$$

as $n \to \infty$ which proves the claim.

(vi) Nondegeneracy of $\mathbb{D}$ will follow from the corresponding property of Gromov's metric $\underline{\square}_1$ together with the following Lemma.

□

Lemma 3.7. *The metric $\mathbb{D}$ can be estimated from below in terms of Gromov's metric $\underline{\square}_1$ ([Gr99], $3\frac{1}{2}.12$) as follows:*

$$\frac{1}{2}\underline{\square}_1 \leq \mathbb{D}^{2/3}.$$

Proof. In order to prove this estimate, let normalized metric measure spaces (M, d, m) and (M', d', m') be given with $\mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m')) < \epsilon^{3/2}$ for some $\epsilon > 0$. Then for some metric $\hat{\mathsf{d}}$ on $M \sqcup M'$ extending d and d' and for some coupling q of m and m'

$$\int \hat{\mathsf{d}}^2(x,x')\, dq(x,x') < \epsilon^3.$$

Hence, $q(\{(x,x') \in M \times M' : \hat{\mathsf{d}}(x,x') \geq \epsilon\}) < \epsilon$. Therefore there exists a measurable map $\Phi : [0,1[\to M \times M'$ such that $\Phi_* \lambda = q$ where λ denotes the Lebesgue measure on $[0,1[$ ('parametrization of q') and there exists a measurable set $X_\epsilon \subset [0,1[$ with $\lambda(X_\epsilon) < \epsilon$ such that for all $x \in [0,1[\backslash X_\epsilon$:

$$\hat{\mathsf{d}}(\Phi(x)) < \epsilon. \tag{3.7}$$

If we write $\Phi(x) = (\varphi(x), \varphi'(x))$ with $\varphi : [0,1[\to M$, $\varphi' : [0,1[\to M'$ then $\varphi_* \lambda = m$ and $\varphi'_* \lambda = m'$. Moreover, for all $x, y \in [0,1[\backslash X_\epsilon$

$$\begin{aligned} & |\,\mathsf{d}(\varphi(x), \varphi(y)) - \mathsf{d}'(\varphi'(x), \varphi'(y))| \\ \leq \;& \hat{\mathsf{d}}(\varphi(x), \varphi'(x)) + \hat{\mathsf{d}}(\varphi(y), \varphi'(y)) \\ = \;& \hat{\mathsf{d}}(\Phi(x)) + \hat{\mathsf{d}}(\Phi(y)) < 2\epsilon \end{aligned}$$

according to (3.7). This proves

$$\sqsubseteq_1((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m')) < 2\epsilon.$$

□

3.2 Examples for $\mathbb{D}$-Convergence

Let us demonstrate the notion of $\mathbb{D}$-convergence with various examples.

Example 3.8. ('Products')
Let $(M_n,\, \mathsf{d}_n, m_n) \in \mathbb{X}_1$ for $n \in \mathbb{N}$. Then $\left(\bigotimes_{n=1}^{l}(M_n,\, \mathsf{d}_n, m_n)\right)_{l \in \mathbb{N}}$ is a $\mathbb{D}$-Cauchy sequence in $\mathbb{X}_1$ if (and only if)

$$\sum_{n=1}^{\infty} \mathbb{V}ar(M_n,\, \mathsf{d}_n, m_n) < \infty.$$

In this case, as $l \to \infty$

$$\bigotimes_{n=1}^{l}(M_n,\, \mathsf{d}_n, m_n) \;\xrightarrow{\mathbb{D}}\; \bigotimes_{n=1}^{\infty}(M_n,\, \mathsf{d}_n, m_n).$$

Proof. Obviously, for all k and l

$$\begin{aligned} & \mathbb{D}^2\left(\bigotimes_{n=1}^{l}(M_n,\, \mathsf{d}_n, m_n), \bigotimes_{n=1}^{l+k}(M_n,\, \mathsf{d}_n, m_n)\right) \\ & \leq \mathbb{D}^2\left((\{o\}, 0, \delta_o),\, \bigotimes_{n=l+1}^{l+k}(M_n,\, \mathsf{d}_n, m_n)\right) \\ & \leq \sum_{n=l+1}^{l+k} \mathbb{D}^2\left((\{o\}, 0, \delta_o), (M_n,\, \mathsf{d}_n, m_n)\right) = \sum_{n=l+1}^{l+k} \mathbb{V}ar(M_n,\, \mathsf{d}_n, m_n). \end{aligned}$$

(Actually, all these inequalities are equalities.) This proves the claim(s). □

Example 3.9. ('Dimension Increasing to Infinity')
Let $M_n = \mathbb{R}$ with Euclidean distance,

$$dm_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{x^2}{2\sigma_n^2}\right)$$

and $M^{(l)} = \mathbb{R}^l$ with Euclidean distance,

$$dm^{(l)}(x) = \frac{1}{(2\pi)^{l/2} \prod_{n=1}^{l} \sigma_n} \exp\left(-\frac{1}{2}\sum_{n=1}^{l}\left(\frac{x_n}{\sigma_n}\right)^2\right) dx$$

Then

$$(M^{(l)},\, \mathsf{d}^{(l)}, m^{(l)}) \xrightarrow{\mathbb{D}} (M^{(\infty)},\, \mathsf{d}^{(\infty)}, m^{(\infty)})$$

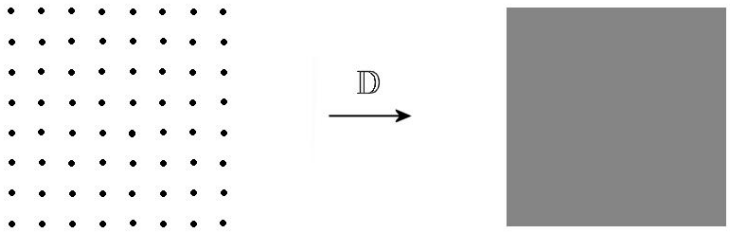
if and only if $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$.

Example 3.10. ('Increasing Finite Dimension')

(i) Let $M_n = (\frac{1}{n}\mathbb{Z} \cap [0,1])^k$ be the rescaled k-dimensional lattice, d be the Euclidean distance in $\mathbb{R}^k$ and m_n be the renormalized counting measure on M_n. Then as $n \to \infty$

$$(M_n, \mathsf{d}, m_n) \xrightarrow{\mathbb{D}} ([0,1]^k, \mathsf{d}, m)$$

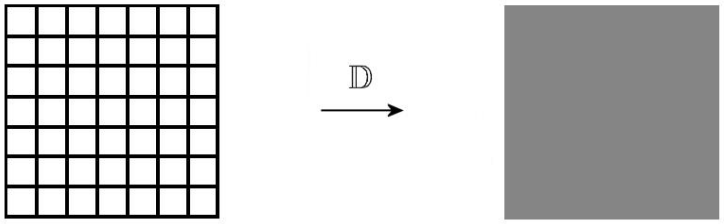
with m being the k-dimensional Lebesgue measure in $[0,1]^k$.



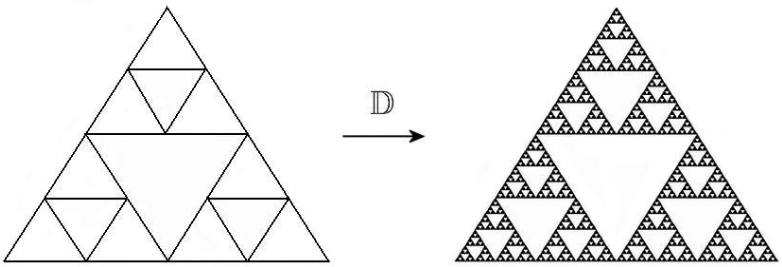
(ii) Similarly, if $\tilde{M}_n$ denotes the graph obtained from M_n with edges between next neighbors and $\tilde{m}_n$ being the 1-dimensional Lebesgue measure on the edges:

$$(\tilde{M}_n, \mathsf{d}, \tilde{m}_n) \xrightarrow{\mathbb{D}} ([0,1]^k, \mathsf{d}, m)$$

as $n \to \infty$.



Example 3.11. ('Increasing to Fractal Dimension')
Let $(M_n)_{n\in\mathbb{N}}$ be the usual approximation of the Sierpinski gasket $M \subset \mathbb{R}^2$ by graphs M_n with 3^n edges of sidelength 2^{1-n}, $n \in \mathbb{N}$. To be more specific, M_1 is the equilateral triangle with



sidelength 1 and for each $n \in \mathbb{N}$, the graph M_n is obtained from M_{n-1} by gluing together 3 copies and rescaling the whole by the factor $\frac{1}{2}$. Let d_n be the distance from the ambient two-dimensional Euclidean space (or *alternatively* the induced length distance on M_n) and let m_n be normalized one-dimensional Lebesgue measure on M_n. Then

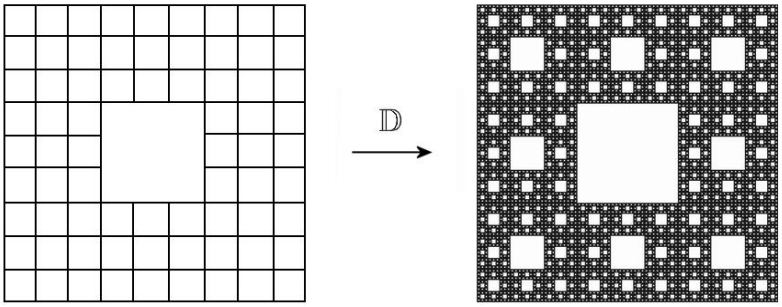
$$(M_n, \mathsf{d}_n, m_n) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$$

where M is the *Sierpinski gasket*, d is the two-dimensional Euclidean distance restricted to M (or the induced length distance on M, resp.) and m is the normalized $\log 3/\log 2$-dimensional Hausdorff measure on M.

Similarly, we can approximate the two-dimensional *Sierpinski carpet* $\tilde{M}$ (equipped with Euclidean distance $\tilde{\mathsf{d}}$ – or alternatively with the induced length distance – and with normalized $\log 8/\log 3$-Hausdorff measure $\tilde{m}$) by graphs $\tilde{M}_n$ with sidelength 3^{-n}. Here $\tilde{M}_1$ is the square with sidelength 1 and $\tilde{M}_n$ is obtained by gluing together 8 copies of $\tilde{M}_{n-1}$ and rescaling the whole by the factor $\frac{1}{3}$. Then

$$(\tilde{M}_n\, \tilde{\mathsf{d}}_n, \tilde{m}_n) \xrightarrow{\mathbb{D}} (\tilde{M},\, \tilde{\mathsf{d}}, \tilde{m}).$$

See for instance [Ki01].



Example 3.12. ('Decreasing Dimension, Collapse')

(i) For each metric measure space $(M,\, \mathsf{d}, m)$ and each sequence $(M_n,\, \mathsf{d}_n, m_n), n \in \mathbb{N}$, with $\lim_{n\to\infty} \mathbb{V}ar(M_n,\, \mathsf{d}_n, m_n) = 0$ one has

$$(M \times M_n,\, \mathsf{d} \otimes \mathsf{d}_n, m \otimes m_n) \xrightarrow{\mathbb{D}} (M,\, \mathsf{d}, m)$$

as $n \to \infty$.

(ii) Let M be a finite graph, embedded in $\mathbb{R}^3$, let d be the graph distance and m be the 1-dimensional Lebesgue measure on M normalized to 1. Let

$$M_n := \left\{ x \in \mathbb{R}^3 : \mathsf{d}_{Euclid}(x, M) \le \frac{1}{n} \right\}$$

and

$$\tilde{M}_n := \left\{ x \in \mathbb{R}^3 : \mathsf{d}_{Euclid}(x, M) = \frac{1}{n} \right\}$$

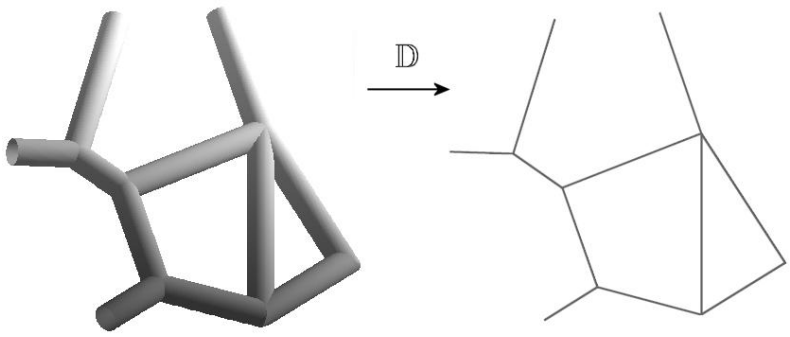
be the full (and surface, resp.) tubular neighborhood of M, let d_n (and $\tilde{\mathsf{d}}_n$) be the geodesic distance on M_n (or $\tilde{M}_n$, resp.) induced by the Euclidean distance d_{Euclid} on the ambient space $\mathbb{R}^3$, and let m_n (and $\tilde{m}_n$) be the 3- (or 2-, resp.) dimensional Lebesgue measure on M (or $\tilde{M}$, resp.), normalized to 1. Then

$$(M_n,\, \mathsf{d}_n, m_n) \xrightarrow{\mathbb{D}} (M,\, \mathsf{d}, m)$$

and

$$(\tilde{M}_n\, \tilde{\mathsf{d}}_n, \tilde{m}_n) \xrightarrow{\mathbb{D}} (M,\, \mathsf{d}, m)$$

as $n \to \infty$.



Being a length space is not preserved under isomorphisms of metric measure spaces. But the support $\mathrm{supp}[m]$ being a length space is preserved under isomorphisms. However, also this property is not preserved under $\mathbb{D}$-convergence.

Example 3.13. Let $M = \mathbb{R}$ with Euclidean distance d, $dm_n(x) = \varphi_n(x)dx$ with

$$\varphi_n(x) = \begin{cases} 1/(2n), & x \in]-1,1[, \\ 1/2, & x \in [-2+\frac{1}{n}, -1] \cup [1, 2-\frac{1}{n}], \\ 0, & \text{else} \end{cases}$$

and

$$dm(x) = \frac{1}{2}\left(1_{[-2,-1]}(x) + 1_{[1,2]}(x)\right) dx.$$

Then $(M, \mathsf{d}, m_n) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$ as $n \to \infty$ as well as $(\mathrm{supp}[m_n], \mathsf{d}, m_n) \xrightarrow{\mathbb{D}} (\mathrm{supp}[m], \mathsf{d}, m)$. However, $(\mathrm{supp}[m_n], \mathsf{d})$ as $n \to \infty$ does not converge w.r.t. D^{GH} towards $(\mathrm{supp}[m], \mathsf{d})$. It converges towards $([-2,2], \mathsf{d})$. On the other hand, Example 3.10 (i) demonstrates the opposite phenomenon: nonlength spaces converging to a length space.

3.3 Doubling Property under $\mathbb{D}$-Convergence

Definition 3.14. Given a number $C \in \mathbb{R}_+$, we say that a metric measure space (M, d, m) has the *restricted doubling property with doubling constant* C iff for all $x \in \mathrm{supp}[m]$ and all $r \in \mathbb{R}_+$

$$m(B_{2r}(x)) \le C \cdot m(B_r(x)).$$

A metric measure space (M, d, m) has the restricted doubling property if and only if for all $x \in \mathrm{supp}[m]$ and all $r, R \in \mathbb{R}_+$

$$m(B_R(x)) \le C^{\lfloor \frac{\log R/r}{\log 2} + 1 \rfloor} \cdot m(B_r(x))$$

where $\lfloor a \rfloor$ denotes the greatest integer $\le a$. It implies that for all x and r the sets $\overline{B}_r(x) \cap \mathrm{supp}[m]$ are compact.

Note that our definition differs from the usual definition of the doubling property: we only impose a condition on balls with center in the support of the measure. Some modification of this kind is necessary in order to obtain a property which is preserved under isomorphisms of metric measure spaces. The usual doubling property without this restriction implies that $\mathrm{supp}[m] = M$ whenever $m(M) \neq 0$.

For instance, let (M, d) be the two-dimensional Euclidean plane and let m be the one-dimensional Lebesgue measure on the x_1-axis. Then (M, d, m) has the restricted doubling property but not the doubling property in the usual sense.

There is a huge literature on metric measure spaces which have the doubling property, see e.g. [He01] and references cited therein.

Theorem 3.15. *The restricted doubling property is stable under $\mathbb{D}$-convergence.*
That is, if for all $n \in \mathbb{N}$ the normalized metric measure spaces (M_n, d_n, m_n) have the restricted doubling property with a common doubling constant C and if $(M_n, \mathsf{d}_n, m_n) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$ as $n \to \infty$ then also (M, d, m) has the restricted doubling property with the same constant C.

Proof. Assume that the normalized metric measure spaces (M_n, d_n, m_n), $n \in \mathbb{N}$, have the restricted doubling property with a common doubling constant C and that

$$\delta_n := \mathbb{D}((M_n, \mathsf{d}_n, m_n), (M, \mathsf{d}, m)) \to 0$$

as $n \to \infty$. Then for each $n \in \mathbb{N}$ the spaces $(\mathrm{supp}[m], \mathsf{d})$ and $(\mathrm{supp}[m_n], \mathsf{d}_n)$ can isometrically be embedded into some space $(\hat{M}, \hat{\mathsf{d}})$ such that

$$\hat{\mathsf{d}}_W(\hat{m}, \hat{m}_n) \leq 2\delta_n$$

where $\hat{m}$ and $\hat{m}_n$ denote the push forwards of the measures m and m_n, resp., under the embedding maps ψ and ψ_n, resp.
Let $x \in \mathrm{supp}[m]$, $r > 0$, $\epsilon > 0$ and $\alpha < 1$ be given (with $2r\alpha^2 - 6\epsilon > 0$ for simplicity). Our first observation is that

$$\hat{m}_n(\hat{B}_{2\epsilon}(\psi(x))) \quge \hat{m}(\hat{B}_\epsilon(\psi(x))) - \frac{1}{\epsilon^2}\,\hat{\mathsf{d}}_W^2(\hat{m}, \hat{m}_n) \tag{3.8}$$

since the mass which has to be transported from the interior of the small ball to the exterior of the large ball has to be moved by a distance of at least ϵ. Moreover, we know that $\hat{m}(\hat{B}_\epsilon(\psi(x))) = m(B_\epsilon(x)) > 0$ since by assumption $x \in \mathrm{supp}[m]$. Hence, for n large enough we conclude that $\hat{m}_n(\hat{B}_{2\epsilon}(\psi(x))) > 0$. Therefore, there exists a point $\hat{x}_n \in \mathrm{supp}[\hat{m}_n] \subset \hat{M}$ with $\hat{\mathsf{d}}(\hat{x}_n, \psi(x)) \leq 2\epsilon$. In particular, we may apply the restricted doubling property for balls centered at x_n. This yields

$$\begin{aligned}
m(B_{2\alpha^2 r - 6\epsilon}(x)) &= \hat{m}(\hat{B}_{2\alpha^2 r - 6\epsilon}(\psi(x))) \;\leq\; \hat{m}(\hat{B}_{2\alpha^2 r - 4\epsilon}(\hat{x}_n)) \\
&\leq \hat{m}_n(\hat{B}_{2\alpha r - 4\epsilon}(\hat{x}_n)) + \frac{1}{(2r\alpha - 2r\alpha^2)^2}\,\hat{\mathsf{d}}_W^2(\hat{m}, \hat{m}_n) \\
&\leq C \cdot \hat{m}_n(\hat{B}_{\alpha r - 2\epsilon}(\hat{x}_n)) + \frac{1}{(2r\alpha - 2r\alpha^2)^2}\,\hat{\mathsf{d}}_W^2(\hat{m}, \hat{m}_n) \\
&\leq C \cdot \hat{m}_n(\hat{B}_{\alpha r}(\psi(x))) + \frac{1}{(2r\alpha - 2r\alpha^2)^2}\,\hat{\mathsf{d}}_W^2(\hat{m}, \hat{m}_n) \\
&\leq C \cdot \hat{m}(\hat{B}_r(\psi(x))) + \left[\frac{1}{(2r\alpha - 2r\alpha^2)^2} + \frac{C}{(2r - 2r\alpha)^2}\right] \hat{\mathsf{d}}_W^2(\hat{m}, \hat{m}_n) \\
&\leq C \cdot m(B_r(x)) + \frac{1 + C\alpha^2}{[r\alpha(1-\alpha)]^2} \cdot \delta_n^2.
\end{aligned}$$

In the limit $n \to \infty$ we obtain

$$m(B_{2\alpha^2 r - 6\epsilon}(x)) \leq C \cdot m(B_r(x)).$$

Since this holds for any $\alpha < 1$ and any $\epsilon > 0$ we conclude

$$m(B_{2r}(x)) \leq C \cdot m(B_r(x)).$$

□

We close this chapter with an important result on compactness under $\mathbb{D}$-convergence.

Theorem 3.16. *('Compactness')*
For each pair $(C,R) \in \mathbb{R}_+ \times \mathbb{R}_+$ the family $\mathbb{X}_1(C,R)$ of all isomorphism classes of normalized metric measure spaces with ('restricted') doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence.

Proof. Let C and R be given and consider a space $(M,\,\mathsf{d},m) \in \mathbb{X}_1(C,R)$. Without restriction, we may assume that $\mathrm{supp}[m] = M$. Then each ball $B_\epsilon(x) \subset M$ has volume

$$m(B_\epsilon(x)) \geq \Big(\frac{\epsilon}{2R}\Big)^N$$

with 'doubling dimension' $N := \log C/\log 2$. Hence, each family of disjoint balls of radius ϵ in M contains at most $(2R/\epsilon)^N$ elements which in turn implies that M can be covered by $(2R/\epsilon)^N$ balls of radius 2ϵ. Firstly, this implies that M is compact. Secondly, since this covering property holds true uniformly in $(M,\,\mathsf{d},m) \in \mathbb{X}_1(C,R)$, according to Gromov's compactness theorem ([Gr81a] or [BH99], Theorem 5.41) it implies compactness of the family $\mathbb{X}_1(C,R)$ under Gromov-Hausdorff convergence. Due to the following Lemma 3.17 this in turn implies compactness under $\mathbb{D}$-convergence. □

Lemma 3.17. *Let $\{(M_i,\,\mathsf{d}_i,m_i), i \in I\}$ be an arbitrary family of normalized compact metric measure spaces which is closed under $\mathbb{D}$-convergence. If the family $\mathsf{X}' = \{(M_i,\,\mathsf{d}_i) : i \in I\}$ (of isometry classes) is compact w.r.t. Gromov-Hausdorff distance D^{GH} then the family $\mathbb{X}' = \{(M_i,\,\mathsf{d}_i,m_i) : i \in I\}$ (of isomorphism classes) is compact w.r.t. $\mathbb{D}$.*

Proof. Let a sequence $((M_n,\,\mathsf{d}_n,m_n))_{n\in\mathbb{N}}$ in $\mathbb{X}'$ be given. Then (by the assumption of the compactness of X') there exists a subsequence $((M_{n_k},\,\mathsf{d}_{n_k},m_{n_k}))_{k\in\mathbb{N}}$ and a compact metric space $(M,\,\mathsf{d})$ such that $(M_{n_k},\,\mathsf{d}_{n_k}) \to (M,\,\mathsf{d})$ in D^{GH}. Moreover, for an appropriately chosen subsequence one can achieve that M as well as all $M_{n_k}, k \in \mathbb{N}$, are isometrically embedded in some other compact metric space $(\hat{M},\,\hat{\mathsf{d}})$. Let us denote the embedded spaces and the associated push forward measures again by $(M_{n_k},\,\mathsf{d}_{n_k})$ and m_{n_k} (or $(M,\,\mathsf{d})$ and m, resp.). Then by compactness of $\hat{M}$ there exists a subsequence, again denoted by $(m_{n_k})_{k\in\mathbb{N}}$, converging to some $m \in \mathcal{P}_2(\hat{M})$. But then

$$\mathbb{D}((M_{n_k},\,\mathsf{d}_{n_k},m_{n_k}),(M,\,\mathsf{d},m)) \leq \hat{\mathsf{d}}_W(m_{n_k},m) \to 0$$

as $k \to \infty$. This proves the compactness of $\mathbb{X}'$. □

4 Curvature Bounds for Metric Measure Spaces

4.1 The Relative Entropy

Recall that a metric measure space always means a triple $(M,\,\mathsf{d},m)$ where $(M,\,\mathsf{d})$ is a complete separable metric space and m is a locally finite measure on M equipped with its Borel σ-algebra. To avoid pathologies, in the sequel we always exclude the case $m(M) = 0$.
Given a metric measure space $(M,\,\mathsf{d},m)$ we denote by $\mathcal{P}_2(M,\,\mathsf{d},m)$ the subspace of all $\nu \in \mathcal{P}_2(M,\,\mathsf{d})$ which are absolutely continuous w.r.t. m, that is, which can be written as $\nu = \rho \cdot m$ with Radon-Nikodym density ρ. In other words, $\mathcal{P}_2(M,\,\mathsf{d},m)$ can be identified with the set of all m-equivalence classes of nonnegative Borel-measurable functions $\rho : M \to \mathbb{R}$ satisfying $\int \rho(x)dm(x) = 1$ and $\int x^2\rho(x)dm(x) < \infty$.
For $\nu = \rho \cdot m \in \mathcal{P}_2(M,\,\mathsf{d},m)$ we define the *relative entropy* of ν w.r.t. m by

$$\mathrm{Ent}(\nu|m) := \lim_{\epsilon \searrow 0} \int\limits_{\{\rho>\epsilon\}} \rho \log \rho \, dm. \tag{4.1}$$

This coincides with

$$\int\limits_{\{\rho>0\}} \rho \log \rho \, dm$$

provided $\int\limits_{\{\rho>1\}} \rho \log \rho dm < \infty$. Otherwise $\mathrm{Ent}(\nu|m) := +\infty$. For $\nu \in \mathcal{P}_2(M, \mathsf{d}) \setminus \mathcal{P}_2(M, \mathsf{d}, m)$ we also define $\mathrm{Ent}(\nu|m) := +\infty$. Finally, we put

$$\mathcal{P}_2^*(M, \mathsf{d}, m) := \{\nu \in \mathcal{P}_2(M, \mathsf{d}) : \ \mathrm{Ent}(\nu|m) < \infty\}.$$

Lemma 4.1. *If m has finite mass, then the relative entropy $\mathrm{Ent}(\cdot\,|m)$ is lower semicontinuous and $\neq -\infty$ on $\mathcal{P}_2(M, \mathsf{d})$. More precisely, for all $\nu \in \mathcal{P}_2(M, \mathsf{d})$*

$$\mathrm{Ent}(\nu|m) \geq -\log m(M). \tag{4.2}$$

Proof. The lower estimate for the relative entropy is a simple application of Jensen's inequality. The lower semicontinuity is more subtle. For $N > 1$ define $U_N(r) = -Nr^{1-1/N}$. Then $r \to U_N(r)$ is convex on $\mathbb{R}_+$, hence,

$$U_N(r) \geq U_N(r_0) - (N-1) r_0^{-1/N} (r - r_0) \tag{4.3}$$

for all r, r_0. Moreover,

$$\lim_{N\to\infty} [Nr + U_N(r)] = \sup_N [Nr + U_N(r)] = r \log r.$$

Now consider $S_N : \mathcal{P}_2(M, \mathsf{d}) \to \mathbb{R}$ with

$$S_N(\nu) := \int U_N(\rho) dm + N$$

for $\nu = \nu_0 + \nu_* \in \mathcal{P}_2(M, \mathsf{d})$ with $\nu_* \perp m$ and $\nu_0 = \rho m$. Note that

$$N - N\nu_0(M)^{1-1/N} m(M)^{1/N} \leq S_N(\nu) \leq N$$

for all ν. Therefore, for all $\nu \in \mathcal{P}_2(M, \mathsf{d})$

$$\mathrm{Ent}(\nu|m) = \lim_{N\to\infty} S_N(\nu) = \sup_N S_N(\nu) \tag{4.4}$$

Hence, lower semicontinuity of $\mathrm{Ent}(\cdot|m)$ will follow from lower semicontinuity of S_N. In order to verify the latter, let $\nu_n = \rho_n m + \nu_n^*$ be any sequence in $\mathcal{P}_2(M, \mathsf{d})$ which converges to some $\nu = \rho m + \nu^* \in \mathcal{P}_2(M, \mathsf{d})$. It implies that $\nu_n \to \nu$ weakly in the sense of measures. Since ρ is a sub-probability density it lies in $L_{1-1/N}(M, m)$ and it can be approximated in the metric $D_{1-1/N}$ by nonnegative bounded, continuous $\rho^{(i)} \in L_{1-1/N}(M, m)$. Here $D_{1-1/N}(u, v) := \int |u - v|^{1-1/N} dm$. Put $\rho_n^{(i)} := |\rho_n - \rho + \rho^{(i)}|$. Then $|S_N(\rho^{(i)} m) - S_N(\rho m)| \leq N \cdot D_{1-1/N}(\rho^{(i)}, \rho) \to 0$ as well as $|S_N(\rho_n^{(i)} m) - S_N(\rho_n m)| \leq N \cdot D_{1-1/N}(\rho^{(i)}, \rho) \to 0$ for $i \to \infty$, uniformly in n. Moreover, since $\rho^{(i)}$ is continuous and bounded it follows from (4.3) that

$$S_N\left(\rho_n^{(i)} m\right) - S_N\left(\rho^{(i)} m\right) \geq -(N-1) \int \left[\rho^{(i)}\right]^{1-1/N} \left(\rho_n^{(i)} - \rho^{(i)}\right) dm \to 0$$

for $n \to \infty$. Summing up, we obtain

$$\liminf_{n\to\infty} S_N(\rho_n m) \geq S_N(\rho m)$$

and thus finally as $N \to \infty$

$$\liminf_{n\to\infty} \mathrm{Ent}(\rho_n m|m) \geq \mathrm{Ent}(\rho m|m).$$

□

Remark 4.2. The relative entropy can in the same manner also be defined for finite, non-normalized measures ν (and m) on M. Then for all $\alpha, \beta > 0$

$$\mathrm{Ent}(\alpha\nu|\beta m) = \alpha \mathrm{Ent}(\nu|m) + (\log\alpha - \log\beta)\alpha\nu(M). \tag{4.5}$$

Moreover, for all finite or countable sets I and all finite measures ν_i, $i \in I$,

$$\mathrm{Ent}(\sum_{i\in I} \nu_i|m) \geq \sum_{i\in I} \mathrm{Ent}(\nu_i|m) \tag{4.6}$$

with equality if and only if the ν_i, $i \in I$, are mutually singular, and

$$\mathrm{Ent}(\sum_{i\in I} \nu_i|m) \leq \sum_{i\in I} \mathrm{Ent}(\nu_i|m) - \sum_{i\in I} \nu_i(M)\log\nu_i(M). \tag{4.7}$$

For convex combinations of probability measures ν_i inequalities (4.6) and (4.7) read as follows

$$\sum \alpha_i \mathrm{Ent}(\nu_i|m) + \sum \alpha_i \log\alpha_i \leq \mathrm{Ent}(\sum \alpha_i\nu_i|m) \leq \sum \alpha_i \mathrm{Ent}(\nu_i|m). \tag{4.8}$$

Proof. Indeed,

$$\begin{aligned} \mathrm{Ent}(\sum_{i\in I} \nu_i|m) &= \sum_{i\in I} \int \rho_i \log\left(\sum_{k\in I} \rho_k\right) dm \\ &\overset{(*)}{\geq} \sum_{i\in I} \int \rho_i \log\rho_i dm \;=\; \sum_i \mathrm{Ent}(\nu_i|m) \end{aligned}$$

with equality in (*) if and only if $\rho_k\rho_i = 0$ m-a.e. on M for all $k \neq i$. On the other hand, according to Jensen's inequality (applied with the convex function $\varphi(r) = r\log r$):

$$\begin{aligned} \mathrm{Ent}(\sum \nu_i|m) &= \int \left(\sum \alpha_i\overline{\rho}_i\right) \log\left(\sum \alpha_i\overline{\rho}_i\right) dm \\ &\leq \int \sum \alpha_i\overline{\rho}_i \log\overline{\rho}_i dm \\ &= \sum \int \rho_i\log\rho_i dm - \sum \alpha_i\log\alpha_i \end{aligned}$$

with $\alpha_i = \nu_i(M)$, $\overline{\rho}_i = \frac{1}{\alpha_i}\rho_i$ and $\rho_i = \frac{d\nu_i}{dm}$. □

Remark 4.3. (i) If m has infinite mass then $\mathrm{Ent}(.|m)$ may exhibit strange behavior. In particular, it can attain the value $-\infty$ and also lower semicontinuity may fail. See the example below.
(ii) If m is finite on all balls and if $\mathrm{Ent}(\nu|m) < \infty$ then

$$\mathrm{Ent}(\nu|m) = \lim_{R\to\infty} \int_{B_R(o)} \rho\log\rho\, dm \tag{4.9}$$

for each $\nu = \rho m$ (with any $o \in M$). Indeed, due to the finiteness of m on $B_R(o)$ the integral on the RHS exists for all R and as $R \to \infty$ by monotone convergence

$$\int_{B_R(o)\cap\{\rho>1\}} \rho \log \rho dm \to \int_{\{\rho>1\}} \rho \log \rho dm < \infty$$

whereas

$$\int_{B_R(o)\cap\{\rho<1\}} \rho \log \rho dm \to \int_{\{\rho<1\}} \rho \log \rho dm \leq \infty.$$

Example 4.4. Let $M = \mathbb{R}$ with Euclidean distance d and $dm(x) = \exp(\exp(x^2))dx$, $d\mu_\alpha(x) = \frac{1}{2\alpha}\exp(\frac{-|x|}{\alpha})dx$.
(i) Then for all $\alpha > 0$, $\mu_\alpha \in \mathcal{P}_2(M)$ with $\mathsf{d}_W(\mu_\alpha, \delta_0) = \sqrt{2}\alpha$ and

$$\mathrm{Ent}(\mu_\alpha|m) = -\infty.$$

(ii) For each $\eta \in \mathcal{P}_2(M, d)$ with $\mathrm{Ent}(\eta|m) > -\infty$ the relative entropy $\mathrm{Ent}(\cdot|m)$ is *not* lower semicontinuous at η since

$$\eta_\alpha := (1-\alpha)\eta + \alpha\mu_\alpha \to \eta \quad \text{in} \quad (\mathcal{P}_2(M),\, \mathsf{d}_W).$$

as $\alpha \to 0$ and

$$-\infty = \lim_{\alpha\to 0} \mathrm{Ent}(\eta_\alpha|m) < \mathrm{Ent}(\eta|m)$$

(iii) Moreover, given any ν_0, $\nu_1 \in \mathcal{P}_2(M,\, \mathsf{d})$ there exists a midpoint η of them. If $\mathrm{Ent}(\eta|m) < \infty$ then for each $\epsilon > 0$ there exists an $\alpha > 0$ such that η_α (defined as before) is an ϵ-midpoint of ν_0 and ν_1 and

$$-\infty = \mathrm{Ent}(\eta_\alpha|m) \leq \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0, \nu_1)$$

for each $K \in \mathbb{R}$.

4.2 Curvature Bounds

Definition 4.5.
(i) We say that a metric measure space $(M,\, \mathsf{d}, m)$ has *curvature* $\geq K$ for some number $K \in \mathbb{R}$ iff the relative entropy $\mathrm{Ent}(\cdot|m)$ is weakly K-convex on $\mathcal{P}_2^*(M,\, \mathsf{d}, m)$ in the following sense: for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ there exists a geodesic $\Gamma : [0,1] \to \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ connecting ν_0 and ν_1 with

$$\mathrm{Ent}(\Gamma(t)|m) \leq (1-t)\mathrm{Ent}(\Gamma(0)|m) + t\mathrm{Ent}(\Gamma(1)|m) - \frac{K}{2}t(1-t)\,\mathsf{d}_W^2(\Gamma(0), \Gamma(1)) \tag{4.10}$$

for all $t \in [0,1]$. To be more specific, we say that in the previous case the metric measure space $(M,\, \mathsf{d}, m)$ has *globally curvature* $\geq K$. Moreover, we put

$$\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) := \sup\{K \in \mathbb{R} : (M,\, \mathsf{d}, m) \text{ has curvature } \geq K\}$$

(with $\sup\emptyset := -\infty$ as usual). Note that then $(M,\, \mathsf{d}, m)$ has curvature $\geq \underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m)$. Occasionally, we use slightly modified concepts:

(ii) We say that a metric measure space $(M,\, \mathsf{d}, m)$ has (globally) *curvature* $\geq K$ *in the lax sense* iff for each $\epsilon > 0$ and for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ there exists an ϵ-midpoint $\eta \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ of ν_0 and ν_1 with

$$\mathrm{Ent}(\eta|m) \leq \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon. \tag{4.11}$$

We denote the maximal K with this property by $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M, \mathsf{d}, m)$.

(iii) We say that a metric measure space (M, d, m) has *locally curvature* $\geq K$ if each point of M has a neighborhood M' such that (M', d, m) — with d and m also denoting the restrictions of d and m onto M' — has globally curvature $\geq K$. The maximal K with this property will be denoted by $\underline{\mathbb{C}\mathrm{urv}}_{loc}(M, \mathsf{d}, m)$.

Remark 4.6. Let (M, d, m) be a metric measure space of finite mass.

(i) Then $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ if and only if for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ there exists a midpoint $\eta \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ of ν_0 and ν_1 with

$$\mathrm{Ent}(\eta|m) \leq \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0, \nu_1). \tag{4.12}$$

(ii) Similarly, $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M, \mathsf{d}, m) \geq K$ if and only if for all $\epsilon > 0$ and all $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ there exists a curve $\Gamma : [0,1] \to \mathcal{P}_2^*(M, \mathsf{d}, m)$ connecting ν_0 and ν_1 with

$$\mathrm{Length}(\Gamma) \leq \mathsf{d}_W(\nu_0, \nu_1) + \epsilon \tag{4.13}$$

and

$$\mathrm{Ent}(\Gamma(t)|m) \leq (1-t)\mathrm{Ent}(\nu_0|m) + t\mathrm{Ent}(\nu_1|m) - \frac{K}{2}t(1-t)\,\mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon \tag{4.14}$$

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

(iii) $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M, \mathsf{d}, m) > -\infty$ implies that $\mathcal{P}_2^*(M, \mathsf{d}, m)$ is a length space (with metric d_W) and that $M_0 = \mathrm{supp}[m] \subset M$ is a length space (with metric d).

Proof. (i), (ii) We have to prove that the existence of (approximate) midpoints with property (4.12) (or (4.11)) implies the existence of (approximate) geodesics with property (4.10) (or (4.14), resp.).
Given $\epsilon = 0$ (or $\epsilon > 0$, resp.) define $\Gamma(\frac{1}{2})$ as ϵ-midpoint of $\Gamma(0) := \nu_0$ and $\Gamma(1) := \nu_1$ satisfying (4.11). Then define $\Gamma(\frac{1}{4})$ as $\epsilon/2$-midpoint of $\Gamma(0)$ and $\Gamma(\frac{1}{2})$ satisfying (4.11) with $\epsilon/2$ and define $\Gamma(\frac{3}{4})$ as $\epsilon/2$-midpoint of $\Gamma(\frac{1}{2})$ and $\Gamma(1)$ satisfying (4.11) with $\epsilon/2$. By iteration we obtain $\Gamma(t)$ for all dyadic $t \in [0,1]$. The continuous extension yields the required curve. [See for instance [St03], proof of Proposition 2.3, for a similar argument.] Lower semicontinuity of the relative entropy then proves the claim for all $t \in [0,1]$.

(iii) According to part (ii), it only remains to prove that M_0 is a length space. Given $x_0, x_1 \in M_0$ let ν_i for $i = 0, 1$ be the normalized volume in $B_\epsilon(x_i)$, i.e. $\nu_i = \frac{1}{m(B_\epsilon(x_i))} 1_{B_\epsilon(x_i)} m$, with $\epsilon > 0$ to be chosen later. (Note that $m(B_\epsilon(x_i)) > 0$ for all $\epsilon > 0$ since $x_i \in M_0$ and $m(B_\epsilon(x_i)) < \infty$ for all sufficiently small $\epsilon > 0$ since m is locally finite.) Then $\nu_i \in \mathcal{P}_2^*(M, \mathsf{d}, m)$. Hence, there exists $\eta \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ with

$$\mathsf{d}_W(\nu_i, \eta) \leq \frac{1}{2}\,\mathsf{d}_W(\nu_0, \nu_1) + \epsilon$$

for $i = 0, 1$. Therefore

$$\begin{aligned}
&\int \left[\,\mathsf{d}^2(x_0, y) + \mathsf{d}^2(x_1, y)\right] d\eta(y) \;=\; \mathsf{d}_W^2(\delta_{x_0}, \eta) + \mathsf{d}_W^2(\delta_{x_1}, \eta) \\
&\leq \left[\,\mathsf{d}_W(\nu_0, \eta) + \epsilon\right]^2 + \left[\,\mathsf{d}_W(\nu_1, \eta) + \epsilon\right]^2 \;\leq\; 2\left[\frac{1}{2}\,\mathsf{d}_W(\nu_0, \nu_1) + 2\epsilon\right]^2 \\
&\leq 2\left[\frac{1}{2}\,\mathsf{d}(x_0, x_1) + 3\epsilon\right]^2 \;=\; \frac{1}{2}\,\mathsf{d}^2(x_0, x_1) + \epsilon'
\end{aligned}$$

for arbitrarily small $\epsilon' > 0$. It implies that there exists a point $y \in \mathrm{supp}[\eta]$ with $\mathsf{d}^2(x_0, y) + \mathsf{d}^2(x_1, y) \le \frac{1}{2}\mathsf{d}^2(x_0, x_1) + \epsilon'$. In other words, y is an approximate midpoint and thus M_0 is a length space. □

Lemma 4.7. *If M is compact then curvature bounds in the usual sense and in the lax sense coincide:*

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = \underline{\mathbb{C}\mathrm{urv}}_{lax}(M, \mathsf{d}, m).$$

Proof. Given $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ let $\eta^{(i)}$ be a family of ϵ-midpoints of ν_0 and ν_1 satisfying (4.11) with $\epsilon = 1/i$. Let us consider the family of probability measures $Q := \{\eta^{(i)} : i \in \mathbb{N}\}$. This family is tight. Indeed, we may assume without restriction that M is a compact length space [otherwise, replace M by $M_0 = \mathrm{supp}[m]$, see Remark 4.6(iii)]. Hence, there exists a suitable subsequence $(\eta^{(i_j)})_{j\in\mathbb{N}}$ which converges to some $\eta \in \mathcal{P}_2(M, \mathsf{d})$. Continuity of the distance d_W and lower semicontinuity of the relative entropy $\mathrm{Ent}(\cdot|m)$ imply that η is a midpoint of ν_0 and ν_1 and (4.11) holds with $\epsilon = 0$. Iterating this procedure yields a geodesic connecting ν_0 and ν_1 and satisfying (4.10). □

The usual definition of K-convexity for the relative entropy would require that (4.10) holds for *each* geodesic connecting ν_0 and ν_1. This leads to the following definition which, however, will be not used in this paper.
We say that a metric measure space (M, d, m) has (globally) *curvature $\ge K$ in the restricted sense* iff $\mathcal{P}_2^*(M, \mathsf{d}, m)$ is a geodesic space and iff each geodesic Γ in $\mathcal{P}_2^*(M, \mathsf{d}, m)$ satisfies (4.10).

Remark 4.8. *Assume that M is a compact nonbranching geodesic space where each pair of points in M is connected by a unique geodesic which depends continuously on the endpoints. Then curvature bounds in the restricted sense and curvature bounds in the usual sense coincide.*

Proof. The (uniformly) continuous dependence of the geodesics on the endpoints implies (and actually is equivalent to the fact) that for each $\epsilon > 0$ there exists $\delta > 0$ such that the midpoint z' of $x' \in B_\delta(x)$ and $y' \in B_\delta(y)$ lies in $B_\epsilon(z)$ whenever z is the midpoint of x and y. Now let the probability measures q on $M \times M$ and η on M be given which are an optimal coupling and a midpoint, resp., of some ν_0 and ν_1. Decompose q into a sum $q = \sum_{i\in\mathbb{N}} q_i$ of mutually singular q_i, $i \in \mathbb{N}$, with $\mathrm{supp}[q_i] \subset B_\delta(x_i) \times B_\delta(y_i)$ for suitable $x_i, y_i \in M$, $i \in \mathbb{N}$. Let $\nu_{0,i}$ and $\nu_{1,i}$ denote the marginals of q_i. Assuming that (M, d, m) has curvature $\ge K$ *in the usual sense* then implies that for each $i \in \mathbb{N}$ there exists a midpoint $\tilde\eta_i$ of $\nu_{0,i}$ and $\nu_{1,i}$ satisfying

$$\mathrm{Ent}(\tilde\eta_i|m) \le \frac{1}{2}\mathrm{Ent}(\nu_{0,i}|m) + \frac{1}{2}\mathrm{Ent}(\nu_{1,i}|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_{0,i}, \nu_{1,i}).$$

The $\tilde\eta_i$ for $i \in \mathbb{N}$ are mutually singular since M is nonbranching and since the q_i are mutually singular (Lemma 2.11 (iii)). Hence, $\tilde\eta = \sum \tilde\eta_i$ satisfies

$$\mathrm{Ent}(\tilde\eta|m) \le \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0, \nu_1).$$

Moreover, $\mathsf{d}_W(\eta, \tilde\eta) \le 2\epsilon$ since for each $i \in \mathbb{N}$, $\mathrm{supp}[\eta_i] \subset B_\epsilon(z_i)$ as well as $\mathrm{supp}[\tilde\eta_i] \subset B_\epsilon(z_i)$ with z_i being the midpoint of x_i and y_i. By lower semicontinuity of $\mathrm{Ent}(\cdot|m)$ this implies

$$\mathrm{Ent}(\eta|m) \le \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0, \nu_1).$$

□

Theorem 4.9. *('Riemannian Spaces')*
Let M be a complete Riemannian manifold with Riemannian distance d and Riemannian volume m and put $m' = \exp(-V) \cdot m$ with a $\mathcal{C}^2$ function $V : M \to \mathbb{R}$. Then

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m') = \inf \{\mathrm{Ric}_M(\xi, \xi) + \mathrm{Hess}V(\xi, \xi) : \xi \in TM, |\xi| = 1\}. \tag{4.15}$$

In particular, (M, d, m) has curvature $\geq K$ if and only if the Ricci curvature of M is $\geq K$.

Note that in the above Riemannian setting for each pair of points ν_0, ν_1 in $\mathcal{P}_2(M, \mathsf{d}, m)$ there exists a *unique* geodesic connecting them. Hence, curvature bounds in the usual sense coincide with curvature bounds in the restricted sense. Moreover, note that in this setting, local curvature bounds always coincide with global curvature bounds.

Proof. Let us briefly sketch the main ideas of the proof, ignoring smoothness and regularity questions. For details, see [RS04] for the case $V = 0$ or [St04] for the general case.
Let $\nu_0 = \rho_0 m$ and $\nu_1 = \rho_1 m$ be given. According to Remark 2.12 (iii) there exists a function $\varphi : M \to \mathbb{R}$ such that

$$\nu_t = (F_t)_* \nu_0$$

with

$$F_t(x) = \exp_x(-t\nabla\varphi(x))$$

defines the unique geodesic $t \mapsto \nu_t$ in $\mathcal{P}_2(M, \mathsf{d})$ connecting ν_0 and ν_1. Change of variable formula then gives

$$\mathrm{Ent}(\nu_t | e^{-V} m) = \int \rho_0 \log \rho_0 dm - \int y_t \rho_0 dm + \int V(F_t)\rho_0 dm \tag{4.16}$$

with $y_t = \log \det dF_t$ being the logarithm of the determinant of the Jacobian of F_t (in some weak sense). Now for ν_0-a.e. $x \in M$ the function $t \mapsto y_t(x)$ satisfies the differential inequality

$$\ddot{y}_t(x) \leq -\frac{1}{n}(\dot{y}_t)^2(x) - \mathrm{Ric}(\dot{F}_t(x), \dot{F}_t(x)). \tag{4.17}$$

Together with (4.16) this yields

$$\begin{aligned} \frac{\partial^2}{\partial t^2}\mathrm{Ent}(\nu_t | e^{-V} m) &\geq \int \left[\mathrm{Ric}(\dot{F}_t, \dot{F}_t) + \mathrm{Hess}V(\dot{F}_t, \dot{F}_t)\right] \rho_0 dx \\ &\geq K \cdot \mathsf{d}_W^2(\nu_0, \nu_1) \end{aligned}$$

provided $\mathrm{Ric}(\xi, \xi) + \mathrm{Hess}V(\xi, \xi) \geq K|\xi|^2$ for all $\xi \in TM$. This 'proves' the K-convexity of $\mathrm{Ent}(\cdot | e^{-V} m)$. □

Some of the most simple examples are

Example 4.10. *(i) If M is a n-dimensional Riemannian manifold of constant sectional curvature κ then*

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = (n-1)\kappa.$$

(ii) If M is the Euclidean space $\mathbb{R}^n$ with the weighted measure $dm(x) = \exp(-K||x||^2/2)dx$ then

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = K.$$

(iii) If $\mathrm{supp}[m]$ *consists of one point then $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = +\infty$.*

Remark 4.11. *If m is finite on all balls then it suffices to verify (4.11) for all $\nu_i = \rho_i m$ with bounded density ρ_i and with bounded support* $\mathrm{supp}[\nu_i]$*, $i = 0, 1$.*

Proof. Let $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ be given, say $\nu_0 = \rho_0 m$ and $\nu_1 = \rho_1 m$. Fix $o \in M$ and define for $i = 0, 1$

$$\rho_{i,R} = \frac{1}{\alpha_{i,R}} 1_{B_R(o)}[\rho_i \wedge R] \qquad \text{with} \qquad \alpha_{i,R} = \int_{B_R(o)} (\rho_i \wedge R) dm.$$

Then according to Remark 4.3, $\alpha_{i,R} \to 1$ and $\mathrm{Ent}(\rho_{i,R} m | m) \to \mathrm{Ent}(\nu_i | m)$ as $R \to \infty$. Moreover, $\mathsf{d}_W(\rho_{i,R} m, \nu_i) \to 0$ and thus for sufficiently large R, each $\frac{\epsilon}{2}$-midpoint of $\rho_{0,R} m$ and $\rho_{1,R} m$ will be an ϵ-midpoint of ν_0 and ν_1. □

4.3 Basic Transformations

Proposition 4.12. *('Isomorphism')*
If (M, d, m) and (M', d', m') are isomorphic metric measure spaces then

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = \underline{\mathbb{C}\mathrm{urv}}(M', \mathsf{d}', m').$$

Thus $\underline{\mathbb{C}\mathrm{urv}}(\cdot)$ extends to a function on $\mathbb{X}$, the family of isomorphism classes of metric measure spaces.

Analogous statements hold true for $\underline{\mathbb{C}\mathrm{urv}}_{lax}(\cdot)$ and $\underline{\mathbb{C}\mathrm{urv}}_{loc}(\cdot)$.

Proof. Let $\Psi : M_0 \to M_0'$ be an isometry between $M_0 := \mathrm{supp}[m]$ and $M_0' := \mathrm{supp}[m']$ such that $\Psi_* m = m'$. Then for all $\nu = \rho m \in \mathcal{P}_2(M, \mathsf{d}, m)$ the push forward measure $\Psi_* \nu$ is absolutely continuous w.r.t. m' with density $\rho(\Psi^{-1})$. Indeed, for all bounded measurable $f : M \to \mathbb{R}$:

$$\int f d(\Psi_* \nu) = \int f(\Psi) \rho dm = \int \left[f \cdot \rho(\Psi^{-1}) \right] \circ \Psi dm = \int f \cdot \rho(\Psi^{-1}) dm'$$

and for all $\nu_0, \nu_1 \in \mathcal{P}_2(M, \mathsf{d})$:

$$\mathsf{d}_W'(\Psi_* \nu_0, \Psi_* \nu_1) = \mathsf{d}_W(\nu_0, \nu_1).$$

This is, Ψ induces an isometry $\nu \mapsto \Psi_* \nu$ between $\mathcal{P}_2(M, \mathsf{d})$ and $\mathcal{P}_2(M', \mathsf{d}')$ which maps $\mathcal{P}_2(M, \mathsf{d}, m)$ onto $\mathcal{P}_2(M', \mathsf{d}', m')$. Moreover,

$$\mathrm{Ent}(\Psi_* \nu | m') = \mathrm{Ent}(\nu | m)$$

since

$$\mathrm{Ent}(\Psi_* \nu | m') = \int \rho(\Psi^{-1}) \log(\rho(\Psi^{-1})) dm' \int \rho \log \rho dm = \mathrm{Ent}(\nu | m).$$

This proves the claim. □

Proposition 4.13. *('Scaled spaces')*
For each metric measure space (M, d, m) and all $\alpha, \beta > 0$:

$$\underline{\mathbb{C}\mathrm{urv}}(M, \alpha\, \mathsf{d}, \beta m) = \alpha^{-2}\, \underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m).$$

Analogous statements hold true for $\underline{\mathbb{C}\mathrm{urv}}_{lax}(\cdot)$ and $\underline{\mathbb{C}\mathrm{urv}}_{loc}(\cdot)$.

Proof. Obviously, $\mathrm{Ent}(\nu | \beta m) = \mathrm{Ent}(\nu | m) - \log \beta$ and $(\alpha \cdot \mathsf{d})_W(\nu_0, \nu_1) = \alpha \cdot \mathsf{d}_W(\nu_0, \nu_1)$. □

Proposition 4.14. *('Weighted Spaces')*
For each metric measure space (M, d, m) *and each lower bounded, measurable function* $V : M \to \mathbb{R}$

$$\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, e^{-V}m) \geq \underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) + \underline{\mathrm{Hess}}V$$

where $\underline{\mathrm{Hess}}V := \sup\{K \in \mathbb{R} : \ V \text{ is } K\text{-convex on } \mathrm{supp}[m]\}$. *If* V *is locally bounded from below then an analogous statement holds for* $\underline{\mathbb{C}\mathrm{urv}}_{loc}(\cdot)$.

Recall that a function $V : M \to]-\infty, +\infty]$, defined on a geodesic space M, is called K-*convex* for some $K \in \mathbb{R}$ iff for each geodesic $\gamma : [0,1] \to M$ and for each $t \in [0,1]$

$$V(\gamma(t)) \leq (1-t)V(\gamma(0)) + tV(\gamma(1)) - \frac{K}{2}t(1-t)\,\mathsf{d}^2(\gamma(0),\gamma(1)). \tag{4.18}$$

Proof. A simple calculation yields

$$\mathrm{Ent}(\nu|e^{-V}m) = \mathrm{Ent}(\nu|m) + \int V d\nu.$$

Moreover, $\mathcal{P}_2^*(M, \mathsf{d}, e^{-V}m) \subset \mathcal{P}_2^*(M, \mathsf{d}, m)$ according to lower boundedness of V. Now put $K_0 := \underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m)$ and $K_1 := \underline{\mathrm{Hess}}V$. Given any geodesic Γ in $\mathcal{P}_2^*(M, \mathsf{d}, m)$ and any $t \in [0,1]$, choose an optimal coupling $\hat{q}$ on M^3 with marginals $\Gamma(0), \Gamma(t), \Gamma(1)$ in the sense of Lemma 2.11(ii). Then

$$\begin{aligned}
&\mathrm{Ent}(\Gamma(t)|e^{-V}m) - (1-t)\mathrm{Ent}(\Gamma(0)|e^{-V}m) - t\mathrm{Ent}(\Gamma(1)|e^{-V}m)\\
&= \mathrm{Ent}(\Gamma(t)|m) - (1-t)\mathrm{Ent}(\Gamma(0)|m) - t\mathrm{Ent}(\Gamma(1)|m)\\
&\quad + \int_M V d\Gamma(t) - (1-t)\int_M V d\Gamma(0) - t\int_M V d\Gamma(1)\\
&= \mathrm{Ent}(\Gamma(t)|m) - (1-t)\mathrm{Ent}(\Gamma(0)|m) - t\mathrm{Ent}(\Gamma(1)|m)\\
&\quad + \int_{M^3} [V(x_t) - (1-t)V(x_0) - tV(x_1)]\, d\hat{q}(x_0, x_t, x_1)\\
&\overset{(*)}{\leq} -\frac{K_0}{2}t(1-t)\,\mathsf{d}_W^2(\Gamma(0),\Gamma(1)) - \int_{M^3}\frac{K_1}{2}t(1-t)\,\mathsf{d}^2(x_0,x_1)d\hat{q}(x_0,x_t,x_1)\\
&= -\frac{K_0+K_1}{2}t(1-t)\,\mathsf{d}_W^2(\Gamma(0),\Gamma(1)).
\end{aligned}$$

The inequality (*) follows from K_1-convexity of V since for $\hat{q}$-almost every (x_0, x_t, x_1) the point x_t lies on a geodesic connecting x_0 and x_1 (Lemma 2.11(ii)). □

Proposition 4.15. *('Subsets')*
Let (M, d, m) *be a metric measure space and let* M' *be a convex subset of* M. *Then*

$$\underline{\mathbb{C}\mathrm{urv}}(M', \mathsf{d}, m) \geq \underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m).$$

Proof. Let ν_0, ν_1 be probability measures in $\mathcal{P}_2^*(M', \mathsf{d}, m)$. Regard them as probability measures on M. Let Γ be a geodesic in $\mathcal{P}_2^*(M, \mathsf{d}, m)$ connecting them and satisfying (4.10). It remains to prove that of each $t \in [0,1]$ the measures $\Gamma(t)$ is supported by M', i.e.

$$\Gamma(t)(M \setminus M') = 0.$$

According to Lemma 2.11(ii), there exists an optimal coupling $\hat{q}$ of $\Gamma(0) = \nu_0$, $\Gamma(t)$, and $\Gamma(1) = \nu_1$ such that for $\hat{q}$-a.e. $(x, z, y) \in M^3$ the point z lies on some geodesic connecting the points x and y. But then $\hat{q}$-almost surely z has to lie in M' since x and y lie in M' and the latter is assumed to be convex. This proves that $\Gamma(t)(M \setminus M') = 0$ and thus yields the claim for $\underline{\mathbb{C}\mathrm{urv}}(M', \mathsf{d}, m)$. □

Proposition 4.16. *('Products')*
Let $(M_i,\, \mathsf{d}_i, m_i)$ *for* $i = 1, \dots, l$ *be metric measure spaces and*

$$(M,\, \mathsf{d}, m) = \bigotimes_{i=1}^{l} (M_i,\, \mathsf{d}_i, m_i).$$

Assume that M *is nonbranching and compact. Then*

$$\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) = \inf_{i \in \{1,\dots,l\}} \underline{\mathbb{C}\mathrm{urv}}(M_i,\, \mathsf{d}_i, m_i). \tag{4.19}$$

Proof. (i) Let us first prove the inequality

$$\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) \le \inf_{i \in \{1,\dots,l\}} \underline{\mathbb{C}\mathrm{urv}}(M_i,\, \mathsf{d}_i, m_i).$$

Assume that this is not true. Then for some $K \in \mathbb{R}$ and $i \in \{1, \dots, l\}$

$$\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) \ge K > \underline{\mathbb{C}\mathrm{urv}}(M_i,\, \mathsf{d}_i, m_i). \tag{4.20}$$

Without restriction, we may assume $i = 1$. Then the last inequality implies that there exist $\nu_0^{(1)}, \nu_1^{(1)} \in \mathcal{P}_2^*(M_1,\, \mathsf{d}_1, m_1)$ such that for each midpoint $\eta^{(1)}$ in $\mathcal{P}_2^*(M_1,\, \mathsf{d}_1, m_1)$ between $\nu_0^{(1)}, \nu_1^{(1)}$ the inequality (4.12) is violated, i.e.

$$\mathrm{Ent}(\eta^{(1)}|m_1) > \frac{1}{2}\mathrm{Ent}(\nu_0^{(1)}|m_1) + \frac{1}{2}\mathrm{Ent}(\nu_1^{(1)}|m_1) - \frac{K}{8}\, \mathsf{d}_W^2(\nu_0^{(1)}, \nu_1^{(1)}). \tag{4.21}$$

Now for $j = 0, 1$ put $\nu_j := \nu_j^{(1)} \otimes \overline{m}_2 \otimes \ldots \otimes \overline{m}_l$ with normalized measures $\overline{m}_i = \frac{1}{m_i(M_i)} m_i$ for $i = 2, \dots, l$. Then obviously

$$\mathrm{Ent}(\nu_j|m) = \mathrm{Ent}(\nu_j^{(1)}) - \sum_{i=2}^{l} \log m_i(M_i). \tag{4.22}$$

and $\nu_j \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$. Moreover

$$\mathsf{d}^W(\nu_0, \nu_1) = \mathsf{d}_1^W(\nu_0^{(1)}, \nu_1^{(1)})$$

(where for typographical reasons, we replace the lower index 'W' by an upper index, again denoting the L_2-Wasserstein distances derived from d or d_1, resp.).
Now the first inequality in (4.20) implies that there exists a midpoint η of ν_0 and ν_1 satisfying (4.12). According to Remark 2.2 it implies

$$\mathsf{d}^W(\nu_0, \eta)^2 + \mathsf{d}^W(\eta, \nu_1)^2 \le \frac{1}{2}\, \mathsf{d}^W(\nu_0, \nu_1)^2$$

which in turn implies

$$\mathsf{d}_1^W(\nu_0^{(1)}, \eta^{(1)})^2 + \mathsf{d}_1^W(\eta^{(1)}, \nu_1^{(1)})^2 \le \frac{1}{2}\, \mathsf{d}_1^W(\nu_0^{(1)}, \nu_1^{(1)})^2.$$

Again according to Remark 2.2 this yields that $\eta^{(1)}$ is a midpoint of $\nu_0^{(1)}$ and $\nu_1^{(1)}$. But (4.21) and (4.22) imply that (4.12) is violated, – contradicting our previous assertion. Thus $\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) < K$ which proves our first claim.

(ii) To prove the reverse implication, we first treat the particular case $\nu_0 = \nu_0^{(1)} \otimes \ldots \otimes \nu_0^{(l)}$ and $\nu_1 = \nu_1^{(1)} \otimes \ldots \otimes \nu_1^{(l)}$. Assume that $\underline{\mathbb{C}\mathrm{urv}}(M_i, d_i, m_i) \geq K$ for each $i = 1, \ldots, l$. Then for each i there exists a midpoint $\eta^{(i)}$ of $\nu_0^{(i)}$ and $\nu_1^{(i)}$ with

$$\mathrm{Ent}(\eta^{(i)}|m_i) \leq \frac{1}{2}\mathrm{Ent}(\nu_0^{(i)}|m) + \frac{1}{2}\mathrm{Ent}(\nu_1^{(i)}|m) - \frac{K}{8} \cdot \mathsf{d}_i^W(\nu_0^{(i)}, \nu_1^{(i)})^2.$$

Put $\eta := \eta^{(1)} \otimes \ldots \otimes \eta^{(l)}$. Then η is a midpoint of ν_0 and ν_1 since

$$\mathsf{d}^W(\eta, \nu_0)^2 = \sum_{i=1}^{l} \mathsf{d}_i^W(\eta^{(i)}, \nu_0^{(i)})^2 \leq \sum_{i=1}^{l} \left[\frac{1}{2}\,\mathsf{d}_i^W(\nu_0^{(i)}, \nu_1^{(i)})\right]^2 = \left[\frac{1}{2}\,\mathsf{d}^W(\nu_0, \nu_1)\right]^2.$$

Moreover,

$$\mathrm{Ent}(\nu_0|m) = \sum_{i=1}^{l} \mathrm{Ent}(\nu_0^{(i)}, m_i), \quad \mathrm{Ent}(\nu_1|m) = \sum_{i=1}^{l} \mathrm{Ent}(\nu_1^{(i)}, m_i)$$

and

$$\mathrm{Ent}(\eta|m) = \sum_{i=1}^{l} \mathrm{Ent}(\eta^{(i)}, m_i).$$

Hence,

$$\begin{aligned} \mathrm{Ent}(\eta|m) &\leq \sum_{i=1}^{l} \left[\frac{1}{2}\mathrm{Ent}(\nu_0^{(i)}|m) + \frac{1}{2}\mathrm{Ent}(\nu_1^{(i)}|m) - \frac{K}{8} \cdot \mathsf{d}_i^W(\nu_0^{(i)}, \nu_1^{(i)})^2\right] \\ &\leq \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}^W(\nu_0, \nu_1)^2. \end{aligned}$$

This proves the claim in the particular case.

(iii) Now let arbitrary $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ and $\epsilon > 0$ be given. Then there exist

$$\tilde{\nu}_0 = \frac{1}{n}\sum_{j=1}^{n} \nu_{0,j}$$

with mutually singular product measures $\nu_{0,j}$, $j = 1, \ldots, n$ and

$$\tilde{\nu}_1 = \frac{1}{n}\sum_{j=1}^{n} \nu_{1,j}$$

with mutually singular product measures $\nu_{1,j}$, $j = 1, \ldots, n$ such that

$$\mathrm{Ent}(\tilde{\nu}_0|m) \leq \mathrm{Ent}(\nu_0|m) + \epsilon, \quad \mathrm{Ent}(\tilde{\nu}_1|m) \leq \mathrm{Ent}(\nu_1|m) + \epsilon,$$

$$\mathsf{d}^W(\nu_0, \tilde{\nu}_0) \leq \epsilon, \quad \mathsf{d}^W(\nu_1, \tilde{\nu}_1) \leq \epsilon$$

and

$$\mathsf{d}^W(\tilde{\nu}_0, \tilde{\nu}_1) \geq \left[\frac{1}{n}\sum_{j=1}^{n} \mathsf{d}^W(\nu_{0,j}, \nu_{1,j})^2\right]^{1/2} - \epsilon.$$

Furthermore, since ν_0 is the sum of mutually singular $\nu_{0,j}$

$$\mathrm{Ent}(\tilde{\nu}_0|m) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{Ent}(\nu_{0,j}|m) - \log n$$

and similarly,

$$\mathrm{Ent}(\tilde{\nu}_1|m) = \frac{1}{n}\sum_{j=1}^{n}\mathrm{Ent}(\nu_{1,j}|m) - \log n.$$

According to part (ii) for each $j = 1, \dots, n$ there exists a midpoint η_j of $\nu_{0,j}$ and $\nu_{1,j}$ satisfying

$$\mathrm{Ent}(\eta_j|m) \le \frac{1}{2}\mathrm{Ent}(\nu_{0,j}|m) + \frac{1}{2}\mathrm{Ent}(\nu_{1,j}|m) - \frac{K}{8}\cdot \mathsf{d}^W(\nu_{0,j},\nu_{1,j})^2.$$

According to Lemma 2.11(iii), since M is nonbranching and since the $\nu_{0,j}$ for $j = 1, \dots, n$ are mutually singular, also the η_j for $j = 1, \dots, n$ must be mutually singular. Hence,

$$\eta := \frac{1}{n}\sum_{j=1}^{n}\eta_j$$

satisfies

$$\mathrm{Ent}(\eta|m) = \frac{1}{n}\sum_{j=1}^{n}\mathrm{Ent}(\eta_j|m) - \log n$$

and thus

$$\begin{aligned}\mathrm{Ent}(\eta|m) &\le \frac{1}{2}\mathrm{Ent}(\tilde{\nu}_0|m) + \frac{1}{2}\mathrm{Ent}(\tilde{\nu}_1|m) - \frac{K}{8}\frac{1}{n}\sum_{j=1}^{n}\mathsf{d}^W(\nu_{0,j},\nu_{1,j})^2 \\ &\le \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}[\mathsf{d}^W(\nu_0,\nu_1) \mp 3\epsilon]^2 + \epsilon\end{aligned}$$

(with $\mp$ to be chosen according to the sign of K). Moreover, η is an approximate midpoint of ν_0 and ν_1:

$$\begin{aligned}2\,\mathsf{d}^W(\nu_0,\eta) &\le 2\left[\frac{1}{n}\sum_j \mathsf{d}^W(\nu_{0,j},\eta_j)^2\right]^{1/2} \le \left[\frac{1}{n}\sum_j \mathsf{d}^W(\nu_{0,j},\nu_{1,j})^2\right]^{1/2} \\ &\le \mathsf{d}^W(\tilde{\nu}_0,\tilde{\nu}_1) + \epsilon \le \mathsf{d}^W(\nu_0,\nu_1) + 3\epsilon\end{aligned}$$

and similarly for $\mathsf{d}^W(\nu_1,\eta)$. This proves that $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M,\mathsf{d},m) \ge K$. Together with compactness of M this finally yields the claim. □

4.4 From Local to Global

A crucial implication of our definition of lower curvature bounds for metric measure spaces is the following Globalization Theorem which states that local curvature bounds imply global curvature bounds. This is in the spirit of the Globalization Theorem of Topogonov for lower curvature bounds (in the sense of Alexandrov) for metric spaces.

Theorem 4.17. *('Globalization')*
Let (M,d,m) be a compact, nonbranching metric measure space and assume that $\mathcal{P}_2^*(M,\mathsf{d},m)$ is a geodesic space. Then

$$\underline{\mathbb{C}\mathrm{urv}}(M,\mathsf{d},m) = \underline{\mathbb{C}\mathrm{urv}}_{loc}(M,\mathsf{d},m). \tag{4.23}$$

Proof. Put $K = \underline{\mathbb{C}\mathrm{urv}}_{loc}(M,\mathsf{d},m)$ and consider for each number $k \in \mathbb{N}\cup\{0\}$ the following property:

C(k): For each geodesic $\Gamma : [0,1] \to \mathcal{P}_2^*(M, \mathsf{d}, m)$ and for each pair $s, t \in [0,1]$ with $0 \le t - s \le 2^{-k}$ there exists a midpoint $\eta(s,t)$ of $\Gamma(s)$ and $\Gamma(t)$ such that

$$\mathrm{Ent}(\eta(s,t)|m) \le \frac{1}{2}\mathrm{Ent}(\Gamma(s)|m) + \frac{1}{2}\mathrm{Ent}(\Gamma(t)|m) - \frac{K}{8}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t)). \tag{4.24}$$

Our first claim is that

- For each $k \in \mathbb{N}$: C(k) implies C(k-1).

In order to prove this claim, let $k \in \mathbb{N}$ be given with property C(k). Moreover, let a geodesic Γ and numbers $s, t \in [0,1]$ be given with $0 \le t - s \le 2^{1-k}$. Define iteratively a sequence $(\Gamma^{(i)})_{i\in\mathbb{N}}$ of geodesics in $\mathcal{P}_2^*(M, \mathsf{d}, m)$ which coincide with Γ on $[0,s] \cup [t,1]$ as follows:
start with $\Gamma^{(0)} := \Gamma$; assuming that $\Gamma^{(2i)}$ is already given, let $\Gamma^{(2i+1)} : [0,1] \to \mathcal{P}_2^*(M, \mathsf{d}, m)$ be any geodesic which coincides with Γ on $[0,s] \cup [t,1]$ and for which $\Gamma^{(2i+1)}(s + \frac{t-s}{4})$ is a midpoint of $\Gamma(s) = \Gamma^{(2i)}(s)$ and $\Gamma^{(2i)}(s + \frac{t-s}{2})$ and for which $\Gamma^{(2i+1)}(s + 3\frac{t-s}{4})$ is a midpoint of $\Gamma^{(2i)}(s + \frac{t-s}{2})$ and $\Gamma(t) = \Gamma^{(2i)}(t)$ satisfying

$$\begin{aligned}\mathrm{Ent}(\Gamma^{(2i+1)}(s + \frac{t-s}{4})|m) &\le \frac{1}{2}\mathrm{Ent}(\Gamma(s)|m) \\ &+ \frac{1}{2}\mathrm{Ent}(\Gamma^{(2i)}(s + \frac{t-s}{2})|m) - \frac{K}{32}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t))\end{aligned} \tag{4.25}$$

and

$$\begin{aligned}\mathrm{Ent}(\Gamma^{(2i+1)}(s + 3\frac{t-s}{4})|m) &\le \frac{1}{2}\mathrm{Ent}(\Gamma(t)|m) \\ &+ \frac{1}{2}\mathrm{Ent}(\Gamma^{(2i)}(s + \frac{t-s}{2})|m) - \frac{K}{32}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t)).\end{aligned} \tag{4.26}$$

Such midpoints exist according to assumption C(k).
Then let $\Gamma^{(2i+2)} : [0,1] \to \mathcal{P}_2^*(M, \mathsf{d}, m)$ be any geodesic which coincides with Γ on $[0,s] \cup [t,1]$ and for which $\Gamma^{(2i+2))}(s + \frac{t-s}{2})$ is a midpoint of $\Gamma^{(2i+1)}(s + \frac{t-s}{4})$ and $\Gamma^{(2i+1)}(s + 3\frac{t-s}{4})$ satisfying

$$\begin{aligned}\mathrm{Ent}(\Gamma^{(2i+2)}(s + \frac{t-s}{2})|m) &\le \frac{1}{2}\mathrm{Ent}(\Gamma^{(2i+1))}(s + \frac{t-s}{4})|m) \\ &+ \frac{1}{2}\mathrm{Ent}(\Gamma^{(2i+1)}(s + 3\frac{t-s}{4})|m) - \frac{K}{32}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t)).\end{aligned} \tag{4.27}$$

Again, such a midpoint exists according to assumption C(k). This yields a sequence of geodesics $\Gamma^{(i)}$, $i \in \mathbb{N}$.
Combining (4.25) – (4.27) gives

$$\begin{aligned}\mathrm{Ent}(\Gamma^{(2i+2)}(s + \frac{t-s}{2})|m) &\le \frac{1}{2}\mathrm{Ent}(\Gamma^{(2i))}(s + \frac{t-s}{2})|m) \\ &+ \frac{1}{4}\mathrm{Ent}(\Gamma(s)|m) + \frac{1}{4}\mathrm{Ent}(\Gamma(t)|m) - \frac{K}{16}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t)).\end{aligned}$$

By iteration, it yields

$$\begin{aligned}\mathrm{Ent}(\Gamma^{(2i)}(s + \frac{t-s}{2})|m) &\le 2^{-i}\mathrm{Ent}(\Gamma(s + \frac{t-s}{2})|m) \\ &+ (1 - 2^{-i})\left[\frac{1}{2}\mathrm{Ent}(\Gamma(s)|m) + \frac{1}{2}\mathrm{Ent}(\Gamma(t)|m) - \frac{K}{8}\,\mathsf{d}_W^2(\Gamma(s),\Gamma(t))\right].\end{aligned}$$

Compactness of $\mathcal{P}_2(M, \mathsf{d})$ now implies that there exists a subsequence of $\left(\Gamma^{(2i)}(s + \frac{t-s}{2})\right)_{i\in\mathbb{N}}$ converging to some $\eta \in \mathcal{P}_2(M, \mathsf{d})$. Continuity of the distance implies that η is a midpoint of

$\Gamma(s)$ and $\Gamma(t)$ (since each of the $\Gamma^{(2i)}(s + \frac{t-s}{2})$ is a midpoint) and lower semicontinuity of the relative entropy implies

$$\mathrm{Ent}(\eta|m) \le \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}^2(\nu_0,\nu_1).$$

This proves property C(k-1).

Now according to our curvature assumption, each point $x \in M$ has a neighborhood $M(x)$ such that $(M(x),\, \mathsf{d}, m)$ has globally curvature $\ge K$ in the usual sense. By compactness of M, there exist $\lambda > 0$, finitely many disjoint sets $L_1, \dots, L_n$ which cover M, and closed sets $M_j \supset B_\lambda(L_j)$ for $j = 1, \dots, n$ such that $(M_j,\, \mathsf{d}, m)$ has globally curvature $\ge K$ in the usual sense. Choose $k' \in \mathbb{N}$ such that

$$2^{-k'} \cdot \mathrm{diam}(M,\, \mathsf{d}, m) \le \lambda. \tag{4.28}$$

Our next claim is that

- Property C(k') is satisfied.

In order to prove this claim, fix Γ as well as s, t and let $\hat{q}$ be a coupling of $\Gamma(0), \Gamma(s), \Gamma(t), \Gamma(1)$ on M^4. Then according to Lemma 2.11, for $\hat{q}$-almost every $(x_0, x_s, x_t, x_1) \in M^4$ the points x_s, x_t lie on some geodesic connecting x_0 and x_1 with

$$d(x_s, x_t) = |t-s| \cdot d(x_0, x_1) \le 2^{-k'} \cdot \mathrm{diam}(M,\, \mathsf{d}, m) \le \lambda \tag{4.29}$$

according to (4.28). Define probability measures $\Gamma_j(s)$ and $\Gamma_j(t)$ for $j = 1, \dots, n$ by

$$\Gamma_j(s)(A) := \frac{1}{\alpha_j}\Gamma(s)(A \cap L_j) = \frac{1}{\alpha_j}\hat{q}(M \times (A \cap L_j) \times M \times M)$$

and

$$\Gamma_j(t)(A) := \frac{1}{\alpha_j}\hat{q}(M \times L_j \times A \times M)$$

provided $\alpha_j := \Gamma_s(L_j) \ne 0$. [Otherwise, define $\Gamma_j(s)$ and $\Gamma_j(t)$ arbitrarily.] Then $\mathrm{supp}[\Gamma_j(s)] \subset \overline{L}_j$ which together with (4.29) implies

$$\mathrm{supp}[\Gamma_j(s)] \cup \mathrm{supp}[\Gamma_j(t)] \subset \overline{B}_\lambda(L_j) \subset M_j. \tag{4.30}$$

Therefore, for each $j \in \{1, \dots, n\}$ the assumption $\underline{\mathbb{C}\mathrm{urv}}(M_j,\, \mathsf{d}, m) \ge K$ can be applied to the probability measures $\Gamma_j(s)$ and $\Gamma_j(t) \in \mathcal{P}_2^*(M_j,\, \mathsf{d}, m)$. It yields the existence of a midpoint $\eta_j(s,t)$ of them with the property

$$\mathrm{Ent}(\eta_j(s,t)|m) \le \frac{1}{2}\mathrm{Ent}(\Gamma_j(s)|m) + \frac{1}{2}\mathrm{Ent}(\Gamma_j(t)|m) - \frac{K}{8}\,\mathsf{d}_W^2(\Gamma_j(s), \Gamma_j(t)). \tag{4.31}$$

Define

$$\eta(s,t) := \sum_{j=1}^n \alpha_j \eta_j(s,t).$$

Then $\eta(s,t)$ is a midpoint of $\Gamma(s) = \sum_{j=1}^n \alpha_j \Gamma_j(s)$ and $\Gamma(t) = \sum_{j=1}^n \alpha_j \Gamma_j(t)$. Moreover, since the $\Gamma_j(s)$ for $j = 1, \dots, n$ are mutually singular and since M is nonbranching, also the $\eta_j(s,t)$ for $j = 1, \dots, n$ are mutually singular, Lemma 2.11(iii). Hence, by (4.6)

$$\mathrm{Ent}(\eta(s,t)|m) = \sum_j \alpha_j \mathrm{Ent}(\eta_j(s,t)|m) + \sum_j \alpha_j \log \alpha_j \tag{4.32}$$

and

$$\mathrm{Ent}(\Gamma(s)|m) = \sum_j \alpha_j \mathrm{Ent}(\Gamma_j(s)|m) + \sum_j \alpha_j \log \alpha_j \tag{4.33}$$

whereas

$$\mathrm{Ent}(\Gamma(t)|m) \geq \sum_j \alpha_j \mathrm{Ent}(\Gamma_j(t)|m) + \sum_j \alpha_j \log \alpha_j \tag{4.34}$$

since the $\Gamma_j(t)$ for $j = 1, \ldots, n$ are not necessarily mutually singular. Summing up (4.31) for $j = 1, \ldots, n$ and using (4.32) – (4.34) yields (4.24). This proves property C(k').

In order to finish the proof of the Theorem, let two probability measures $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ be given. By assumption, there exists a geodesic Γ in $\mathcal{P}_2^*(M, \mathsf{d}, m)$ connecting them. According to our second claim, property C(k') is satisfied and according to our first claim, this implies property C(k) for all $k = k' - 1, k' - 2, \ldots, 0$. Property C(0) finally states that there exists a midpoint η of $\Gamma(0)$ and $\Gamma(1)$ with

$$\mathrm{Ent}(\eta|m) \leq \frac{1}{2}\mathrm{Ent}(\Gamma(0)|m) + \frac{1}{2}\mathrm{Ent}(\Gamma(1)|m) - \frac{K}{8}\,\mathsf{d}_W^2(\Gamma(0), \Gamma(1)). \tag{4.35}$$

This proves the Theorem. □

Remark 4.18. Let M be a compact space.

(i) The condition on $\mathcal{P}_2^*(M, \mathsf{d}, m)$ to be a geodesic space is always satisfied if $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) > -\infty$ (Remark 4.6(iii)).

(ii) If $\mathcal{P}_2^*(M, \mathsf{d}, m)$ is a geodesic space then $\mathrm{supp}[m]$ is a geodesic space. The converse is not true in general. Conjecture: it is true under the additional assumption $\underline{\mathbb{C}\mathrm{urv}}_{loc}(M, \mathsf{d}, m) > -\infty$.

(iii) If $M_0 := \mathrm{supp}[m]$ is a geodesic space then $\mathcal{P}_2(M_0, \mathsf{d})$ is a geodesic space . Moreover, the space $\mathcal{P}_2^*(M_0, \mathsf{d}, m)$ is dense in $\mathcal{P}_2(M_0, \mathsf{d})$. Indeed, given any $\mu \in \mathcal{P}_2(M_0, \mathsf{d})$ and any $\epsilon > 0$ there exist $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in M_0$ such that $\mathsf{d}_W(\mu, \mu') \leq \epsilon$ where $\mu' := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Moreover, $\mathsf{d}_W(\mu', \mu'') \leq \epsilon$ where $\mu'' := \frac{1}{n}\sum_{i=1}^n \frac{1}{m(B_\epsilon(x_i))} \cdot 1_{B_\epsilon(x_i)}\, m$ and

$$\mathrm{Ent}(\mu''|m) \leq \sup_{x \in M_0} \log \left[\frac{1}{n}\sum_{i=1}^n \frac{1}{m(B_\epsilon(x_i))} \cdot 1_{B_\epsilon(x_i)}(x)\right] \leq -\inf_{i=1,\ldots,n} \log\, m(B_\epsilon(x_i)) < \infty.$$

That is, $\mu'' \in \mathcal{P}_2^*(M_0, \mathsf{d}, m)$ and $\mathsf{d}_W(\mu, \mu'') \leq 2\epsilon$ which proves the density.

4.5 Stability under Convergence

One of the most important results in this paper is that our curvature bounds for metric measure spaces are stable under convergence. The key to this result is the fact that we are able to construct a transformation Q' from the L_2-Wasserstein space over one metric measure space (M, d, m) to the L_2-Wasserstein space over any other metric measure space (M', d', m') which reduces the relative entropy and which is almost an isometry between these spaces, provided the underlying spaces are close in the metric $\mathbb{D}$.
Actually, this is quite easy in the particular case where m is the push forward of m' under a map $\psi' : M' \to M$. In this case we can define (similarly to the construction in the proof of Proposition 4.12)

$$Q' : \mathcal{P}_2(M, \mathsf{d}, m) \to \mathcal{P}_2(M', \mathsf{d}', m'), \qquad \rho \cdot m \mapsto (\rho \circ \psi') \cdot m'.$$

The general case is more subtle since we may not restrict ourselves to transformations derived from push forward maps. For instance, if m' is the Riemannian measure of a collapsed space then the Riemannian measure m of the initial space cannot be represented as a push forward measure.

Given two normalized metric measure spaces $(M,\, \mathsf{d}, m)$ and $(M',\, \mathsf{d}', m')$ we will define a canonical map

$$Q' : \mathcal{P}_2(M,\, \mathsf{d}, m) \to \mathcal{P}_2(M',\, \mathsf{d}', m')$$

as follows: Let q be a coupling of m and m' and $\hat{\mathsf{d}}$ be a coupling of d and d' such that

$$\int \hat{\mathsf{d}}^2(x, x')\, dq(x, x') \leq 2\mathbb{D}^2((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m')).$$

Let Q' and Q be the disintegrations of q w.r.t. m' and m, resp., that is,

$$dq(x, x') = Q'(x', dx)dm'(x') = Q(x, dx')dm(x)$$

and let $\hat{\Delta}$ denote the m-essential supremum of the map

$$x \mapsto \left[\int_{M'} \hat{\mathsf{d}}^2(x, x')\, Q(x, dx')\right]^{1/2}.$$

In general, of course, $\hat{\Delta}$ may attain the value ∞. However, if both metric measure spaces have finite diameter we have

$$\hat{\Delta} \leq \operatorname{diam}(M,\, \mathsf{d}, m) + \operatorname{diam}(M',\, \mathsf{d}', m') < \infty.$$

For $\nu = \rho m \in \mathcal{P}_2(M,\, \mathsf{d}, m)$ define $Q'(\nu) \in \mathcal{P}_2(M',\, \mathsf{d}', m')$ by $Q'(\nu) := \rho' m'$ where

$$\rho'(x') := \int_M \rho(x) Q'(x', dx). \tag{4.36}$$

In other words, for all measurable $A \subset M'$

$$\begin{aligned} Q'(\nu)(A) &= \int_{M'} 1_A(x')\rho'(x')dm'(x') \\ &= \int_{M'}\int_M 1_A(x')\rho(x)Q'(x', dx)dm'(x') \\ &= \int_{M\times M'} 1_A(x')\rho(x)dq(x, x'). \end{aligned}$$

Lemma 4.19. *The map Q' defined as above satisfies $Q'(m) = m'$ and for all $\nu = \rho\, m$:*

$$\operatorname{Ent}(Q'(\nu)|m') \leq \operatorname{Ent}(\nu|m) \tag{4.37}$$

and

$$\mathsf{d}_W^2(\nu, Q'(\nu)) \leq \frac{2 + \hat{\Delta}^2 \cdot \operatorname{Ent}(\nu|m)}{-\log \mathbb{D}((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m'))}. \tag{4.38}$$

provided $\mathbb{D}((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m')) < 1$.

Proof. Inequality (4.37) is a consequence of Jensen's inequality, applied to the convex function $r \mapsto r \log r$, as follows

$$\begin{aligned}
\mathrm{Ent}(Q'(\nu)|m') &= \int \rho' \log \rho' dm' \\
&= \int \left[\int \rho(x) Q'(x', dx) \right] \cdot \log \left[\int \rho(x) Q'(x', dx) \right] dm'(x') \\
&\leq \int \int \rho(x) \log \rho(x) Q'(x', dx) dm'(x') \\
&= \int \rho(x) \log \rho(x) dm(x) \\
&= \mathrm{Ent}(\nu|m).
\end{aligned}$$

Inequality (4.38) follows from the fact that the measure $\rho(x)dq(x,x') = \rho(x)Q(x,dx')dm(x)$ is a coupling of $\rho(x)dm(x)$ and $\rho'(x')dm'(x') = \int_M \rho(x)Q'(x',dx)dm'(x') = \int_M \rho(x)dq(x,x')$ and thus

$$\mathsf{d}_W^2(\rho m, \rho' m') \leq \int \int \hat{\mathsf{d}}^2(x,x') Q(x,dx') \rho(x) dm(x) =: \Phi(\rho).$$

Now again with Jensen's inequality applied to the convex function $\psi(r) := r \log r$,

$$\psi\left(\frac{\Phi(\rho)}{\Phi(1)}\right) \leq \frac{1}{\Phi(1)} \int \int \hat{\mathsf{d}}^2(x,x') Q(x,dx') \psi(\rho(x))\, dm(x) \leq \frac{\hat{\Delta}^2}{\Phi(1)} \mathrm{Ent}(\nu|m).$$

Hence, since by assumption $\Phi(1) \leq 2\mathbb{D}((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m')) < 2$

$$\begin{aligned}
\mathsf{d}_W^2(\rho m, \rho' m') &\leq \Phi(\rho) \leq \Phi(1) \cdot \psi^{-1}\left(\hat{\Delta}^2 \mathrm{Ent}(\nu|m)/\Phi(1) \right) \\
&\leq \Phi(1) \cdot \psi^{-1}\left([2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)]/\Phi(1) \right) \\
&= \frac{2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)}{\log \psi^{-1}\left([2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)]/\Phi(1) \right)} \\
&\leq 2 \frac{2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)}{\log\left([2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)]/\Phi(1) \right)} \\
&\leq 2 \frac{2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)}{\log\left([2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)]/2\mathbb{D}^2 \right)} \\
&\leq \frac{2 + \hat{\Delta}^2 \mathrm{Ent}(\nu|m)}{-\log \mathbb{D}}
\end{aligned}$$

where we have used the abbreviation $\mathbb{D} := \mathbb{D}((M,\, \mathsf{d}, m), (M',\, \mathsf{d}', m'))$. □

Theorem 4.20. *('Convergence')*
Let $((M_n,\, \mathsf{d}_n, m_n))_{n\in\mathbb{N}}$ be a sequence of normalized metric measure spaces with uniformly bounded diameter. If

$$(M_n,\, \mathsf{d}_n, m_n) \xrightarrow{\mathbb{D}} (M,\, \mathsf{d}, m)$$

as $n \to \infty$ then

$$\limsup_{n\to\infty} \underline{\mathbb{C}\mathrm{urv}}_{lax}(M_n,\, \mathsf{d}_n, m_n) \leq \underline{\mathbb{C}\mathrm{urv}}_{lax}(M,\, \mathsf{d}, m).$$

In particular, for each $K \in \mathbb{R}$ and $\Delta \in \mathbb{R}_+$ the family $\mathbb{X}_1(K, \Delta)$ of isomorphism classes of normalized metric measure spaces with curvature $\geq K$ and diameter $\leq \Delta$ is closed w.r.t $\mathbb{D}$.

Proof. Let $((M_n,\, \mathsf{d}_n, m_n))_{n\in\mathbb{N}}$ be a sequence in $(\mathbb{X}_1, \mathbb{D})$ with $(M_n,\, \mathsf{d}_n, m_n) \to (M,\, \mathsf{d}, m)$ and assume that $\operatorname{diam}(M,\, \mathsf{d}, m) \le \Delta$ and $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M_n,\, \mathsf{d}_n, m_n) \ge K$ for some $\Delta, K \in \mathbb{R}$ and all sufficiently large $n \in \mathbb{N}$. Now let $\epsilon > 0$ and $\nu_0 = \rho_0 m$, $\nu_1 = \rho_1 m \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ be given. Choose R with

$$\sup_{i=0,1} \operatorname{Ent}(\nu_i|m) + \frac{|K|}{8}[\,\mathsf{d}_W(\nu_0,\nu_1) + 2\epsilon]^2 + \epsilon \le R. \tag{4.39}$$

We have to deduce the existence of an ϵ-midpoint η which satisfies inequality (4.11). Choose $n \in \mathbb{N}$ with

$$\mathbb{D}((M_n,\, \mathsf{d}_n, m_n), (M,\, \mathsf{d}, m)) \le \exp\left(-\frac{2 + (\Delta+\Delta')^2 K}{\epsilon^2}\right). \tag{4.40}$$

Define the map $Q_n' : \mathcal{P}_2(M,\, \mathsf{d}, m) \to \mathcal{P}_2(M_n,\, \mathsf{d}_n, m_n)$ as in the previous lemma, now with m_n in the place of m', and analogously the map $Q_n : \mathcal{P}_2(M_n,\, \mathsf{d}_n, m_n) \to \mathcal{P}_2(M,\, \mathsf{d}, m)$. Put

$$\nu_{i,n} := Q_n'(\nu_i) = \rho_{i,n} m_n$$

with $\rho_{i,n}(y) = \int \rho_i(x) Q_n'(y, dx)$ for $i = 0, 1$ and let η_n be an ϵ-midpoint of $\nu_{0,n}$ and $\nu_{1,n}$ with

$$\operatorname{Ent}(\eta_n|m_n) \le \frac{1}{2}\operatorname{Ent}(\nu_{0,n}|m_n) + \frac{1}{2}\operatorname{Ent}(\nu_{1,n}|m_n) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_{0,n},\nu_{1,n}) + \epsilon. \tag{4.41}$$

From (4.38) – (4.40) we conclude

$$\begin{aligned}
\mathsf{d}_W^2(\nu_0,\nu_{0,n}) &\le \frac{2 + \hat{\Delta}^2 \cdot \operatorname{Ent}(\nu_0|m)}{-2\log \mathbb{D}((M,\, \mathsf{d}, m), (M_n,\, \mathsf{d}_n, m_n))} \\
&\le \frac{2 + (\Delta+\Delta')^2 \cdot R}{-2\log \mathbb{D}((M,\, \mathsf{d}, m), (M_n,\, \mathsf{d}_n, m_n))} \le \epsilon^2
\end{aligned}$$

and analogously $\mathsf{d}_W^2(\nu_1,\nu_{1,n}) \le \epsilon^2$. Moreover, (4.37) and (4.41) imply

$$\begin{aligned}
\operatorname{Ent}(\eta_n|m_n) &\le \frac{1}{2}\operatorname{Ent}(\nu_{0,n}|m_n) + \frac{1}{2}\operatorname{Ent}(\nu_{1,n}|m_n) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_{0,n},\nu_{1,n}) + \epsilon \\
&\le \frac{1}{2}\operatorname{Ent}(\nu_0|m) + \frac{1}{2}\operatorname{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0,\nu_1) + \epsilon'
\end{aligned}$$

with $\epsilon' = [1 + \frac{|K|}{2}(\,\mathsf{d}_W(\nu_0,\nu_1) + \epsilon)] \cdot \epsilon$. Finally, put

$$\eta = Q_n(\eta_n).$$

Then again by (4.38) – (4.40) and by the previous estimate for $\operatorname{Ent}(\eta_n|m)$

$$\begin{aligned}
\mathsf{d}_W^2(\eta_n,\eta) &\le \frac{2 + \hat{\Delta}^2 \cdot \operatorname{Ent}(\eta_n|m)}{-2\log \mathbb{D}((M,\, \mathsf{d}, m), (M_n,\, \mathsf{d}_n, m_n))} \\
&\le \frac{2 + (\Delta+\Delta')^2 \cdot R}{-2\log \mathbb{D}((M,\, \mathsf{d}, m), (M_n,\, \mathsf{d}_n, m_n))} \le \epsilon^2.
\end{aligned}$$

Hence,

$$\sup_{i=0,1} \mathsf{d}_W(\eta,\nu_i) \le \frac{1}{2}\,\mathsf{d}_W(\nu_0,\nu_1) + 4\epsilon,$$

i.e. η is a (4ϵ)-midpoint of ν_0 and ν_1. Furthermore, by (4.37)

$$\begin{aligned}
\operatorname{Ent}(\eta|m) &\le \operatorname{Ent}(\eta_n|m_n) \\
&\le \frac{1}{2}\operatorname{Ent}(\nu_0|m) + \frac{1}{2}\operatorname{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0,\nu_1) + \epsilon'
\end{aligned}$$

with ϵ' as above. This proves that $\underline{\mathbb{C}\mathrm{urv}}_{lax}(M,\, \mathsf{d}, m) \ge K$. □

As an immediate consequence of the previous theorem together with Proposition 4.16 we obtain:

Corollary 4.21. *('Infinite Products')*
Let $(M,\, \mathsf{d}, m) = \bigotimes_{n\in\mathbb{N}}(M_n,\, \mathsf{d}_n, m_n)$ *where* $(M_n,\, \mathsf{d}_n, m_n)$ *for* $n \in \mathbb{N}$ *are normalized metric measure spaces with compact nonbranching* M_n*. Assume* $\sum_{n\in\mathbb{N}} \mathbb{V}ar(M_n,\, \mathsf{d}_n, m_n) < \infty$*. Then*

$$\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m) = \inf_{n\in\mathbb{N}} \underline{\mathbb{C}\mathrm{urv}}(M_n,\, \mathsf{d}_n, m_n).$$

Important infinite dimensional examples are given by abstract Wiener spaces. Let (M, H, m) be an abstract Wiener space, that is, M is a separable Banach space, m is a Gaussian measure on M, and H is a separable Hilbert space that is continuously and densely embedded in M such that

$$\int_M \exp(i\langle x, y\rangle) dm(x) = \exp\left(-\frac{1}{2}\|y\|_H^2\right)$$

for any $y \in M^* \subset H$ (where we identify H with its dual). For the classical Wiener space, $M = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ is the path space of one-dimensional Brownian motion,

$$H = \left\{u \in M :\ u \text{ is abs. cont. with } \int_{\mathbb{R}_+} |\dot{u}(t)|^2 dt < \infty\right\}$$

is the Cameron-Martin space, and m is the Wiener measure.
Given any abstract Wiener space (M, H, m) define a pseudo metric on M by

$$\mathsf{d}(x, y) := \|x - y\|_H$$

if $x - y \in H$ and $\mathsf{d}(x, y) := \infty$ else and consider the 'pseudo metric measure space' $(M,\, \mathsf{d}, m)$. Of course, formally this does not fit in our framework. Nevertheless, the definition of the L_2-Wasserstein distance d_W derived from this pseudo metric d perfectly makes sense. It is a pseudo metric on the space of probability measures on M and a metric on the subspace of all those probability measures which have finite d_W-distance from m. Also the relative entropy $\mathrm{Ent}(.|m)$ and the curvature bound $\underline{\mathbb{C}\mathrm{urv}}(M,\, \mathsf{d}, m)$ are well-defined. Particular attention, however, has to be paid to the fact that d_W is not continuous w.r.t. the weak topology of measures on M; it behaves very singular. For a detailed analysis, we refer to [FÜ04a, FÜ04b]. Here we restrict ourselves to the following result.

Proposition 4.22. *('Wiener Space')*

$$\underline{\mathbb{C}\mathrm{urv}}_{lax}(M,\, \mathsf{d}, m) \geq 1.$$

Proof. Let $\nu_0, \nu_1 \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ and $\epsilon > 0$ be given. Choose a increasing ('total') sequence $(M_n)_n$ of regular finite dimensional subspaces with $\bigcup_n M_n$ being dense in W and H. Let m_n, $\nu_{0,n}$ and $\nu_{1,n}$ be the images measures of m, ν_0 and ν_1, resp., under the projections $\pi_n : M \to M_n$. Then

$$\mathsf{d}_W(\nu_{0,n}, \nu_0) \to 0, \quad \mathsf{d}_W(\nu_{1,n}, \nu_1) \to 0$$

and

$$\mathrm{Ent}(\nu_{0,n}|m_n) \to \mathrm{Ent}(\nu_0|m), \quad \mathrm{Ent}(\nu_{1,n}|m_n) \to \mathrm{Ent}(\nu_1|m)$$

as $n \to \infty$. The space (M_n, M_n, m_n) is a finite dimensional abstract Wiener space. Therefore, it is isomorphic to

$$(\mathbb{R}^N, \mathbb{R}^N, \exp(-\|x\|^2/2)\, dx)$$

for some $N = N(n) \in \mathbb{N}$ (where dx denotes the Lebesgue measure in $\mathbb{R}^N$). Hence, according to Theorem 4.9

$$\underline{\mathbb{C}\mathrm{urv}}(M_n,\, \mathsf{d}, m_n) = 1.$$

Thus for each $n \in \mathbb{N}$ there exists a midpoint η_n of $\nu_{0,n}$ and $\nu_{1,n}$ with

$$\begin{aligned}\mathrm{Ent}(\eta_n|m) &= \mathrm{Ent}(\eta_n|m_n)\\ &\leq \frac{1}{2}\mathrm{Ent}(\nu_{0,n}|m_n) + \frac{1}{2}\mathrm{Ent}(\nu_{1,n}|m_n) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_{0,n},\nu_{1,n})\\ &\leq \frac{1}{2}\mathrm{Ent}(\nu_0|m) + \frac{1}{2}\mathrm{Ent}(\nu_1|m) - \frac{K}{8}\,\mathsf{d}_W^2(\nu_0,\nu_1) + \epsilon\end{aligned}$$

for n large enough. This proves the claim since η_n is an ϵ-midpoint of ν_0 and ν_1 (again, for large enough n). □

4.6 Volume Growth Estimates

In the Riemannian setting, it is well-known that lower bounds for the Ricci curvature of the underlying space imply upper bounds for the growth

$$R \mapsto m(\overline{B}_R(x))$$

of the volume of concentric balls. In particular, this growth is at most exponentially in R. This is the content of the famous Bishop-Gromov volume comparison theorem.
Also for general metric measure spaces, lower bounds for the curvature will imply upper estimates for the volume growth of concentric balls. These estimates, however, have to take into account that in the general case (without any dimensional restriction) the volume can grow much faster than exponentially. For instance, already in the following standard example we observe squared exponential volume growth.

Example 4.23. Let (M, d) be the one-dimensional Euclidean space equipped with the measure $dm(x) = \exp(-\frac{K}{2}x^2)\,dx$ for some $K \in \mathbb{R}$. Then $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) = K$ and, if $K < 0$

$$m(\overline{B}_R(x)) \geq \exp\left(\frac{|K|}{2}\left(R - \frac{1}{2}\right)^2\right)$$

for each $x \in M$ and $R \geq \frac{1}{2}$.

Theorem 4.24. *Let (M, d, m) be an arbitrary metric measure space with $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ for some $K \leq 0$. For fixed $x \in \mathrm{supp}[m] \subset M$ consider the volume growth*

$$v_R := m(\overline{B}_R(x))$$

of closed balls centered at x. Then for all $R \geq 2\epsilon > 0$

$$v_R \leq v_{2\epsilon} \cdot \left(\frac{v_{2\epsilon}}{v_\epsilon}\right)^{R/\epsilon} \cdot \exp\left(\frac{|K|}{2}\left(R + \frac{\epsilon}{2}\right)^2\right). \tag{4.42}$$

In particular, each ball in M has finite volume.

Proof. Apply the following Lemma with $r = \epsilon$. □

Lemma 4.25. *Let (M, d, m), K and x as in the above theorem. Then for all $\epsilon, R > 0$ and all $t \in]0, 1]$*

$$\log v_R \leq \frac{1}{t}\log v_{\epsilon+t(R+\epsilon)} + (1 - \frac{1}{t})\log v_\epsilon + \frac{|K|}{2}(1-t)(R+\epsilon)^2. \tag{4.43}$$

In other words, for all $\epsilon, r > 0$ and all $R > \epsilon + r$

$$v_R \leq v_\epsilon \cdot \left(\frac{v_{\epsilon+r}}{v_\epsilon}\right)^{\frac{R+\epsilon}{r}} \cdot \exp\left(\frac{|K|}{2}(R + \epsilon - r)(R+\epsilon)\right). \tag{4.44}$$

Proof. Fix $x \in \mathrm{supp}[m]$ and $\epsilon, R > 0$. Let ν_0 and ν_1 denote the uniform distributions in $\overline{B}_\epsilon(x)$ and $\overline{B}_R(x)$, resp. That is,

$$d\nu_0(x) = \frac{1}{v_\epsilon} \cdot 1_{\overline{B}_{\epsilon(x)}}\, dm(x), \quad d\nu_1(x) = \frac{1}{v_R} \cdot 1_{\overline{B}_{R(x)}}\, dm(x).$$

Then obviously $\nu_0, \nu_1 \in \mathcal{P}_2^*(M,\, \mathsf{d}, m)$ with

$$\mathrm{Ent}(\nu_0|m) = -\log v_\epsilon, \quad \mathrm{Ent}(\nu_1|m) = -\log v_R.$$

Let $\nu_t, t \in [0,1]$, be a geodesic in $\mathcal{P}_2^*(M,\, \mathsf{d}, m)$ connecting ν_0, ν_1 such that

$$\mathrm{Ent}(\nu_t|m) \le (1-t)\mathrm{Ent}(\nu_0|m) + t\mathrm{Ent}(\nu_1|m) - \frac{K}{2}t(1-t)\,\mathsf{d}_W^2(\nu_0,\nu_1).$$

Such a geodesic exists according to our curvature assumption. Since $\mathsf{d}_W(\nu_0, \delta_x) \le \epsilon$ and $\mathsf{d}_W(\nu_1, \delta_x) \le R$ it follows

$$\mathsf{d}_W(\nu_0,\nu_1) \le R + \epsilon. \tag{4.45}$$

Moreover, if $\hat{q}$ is an optimal coupling of ν_0, ν_t, ν_1 then for $\hat{q}$-almost every (y_0, y_t, y_1) the point y_t lies on a geodesic connecting y_0, y_1 with $\mathsf{d}(y_0, y_t) = t \cdot \mathsf{d}(y_0, y_1)$. Together with (4.45), the latter implies

$$\mathrm{supp}[\nu_t] \subset \overline{B}_{\epsilon + t(R+\epsilon)}(x). \tag{4.46}$$

Now according to Jensen's inequality, for all ν_t satisfying (4.46)

$$\mathrm{Ent}(\nu_t|m) \ge \mathrm{Ent}(m_t|m)$$

where $m_t := \frac{1}{v_{\epsilon+t(R+\epsilon)}} 1_{\overline{B}_{\epsilon+t(R+\epsilon)}(x)} m$ denotes uniform distribution in the closed ball $\overline{B}_{\epsilon+t(R+\epsilon)}(x)$. Hence,

$$\begin{aligned} -\log v_{\epsilon+t(R+\epsilon)} &= \mathrm{Ent}(m_t|m) \le \mathrm{Ent}(\nu_t|m) \\ &\le (1-t)\mathrm{Ent}(\nu_0|m) + t\mathrm{Ent}(\nu_1|m) - \frac{K}{2}t(1-t)\,\mathsf{d}_W^2(\nu_0,\nu_1) \\ &\le -(1-t)\log v_\epsilon - t\log v_R + \frac{|K|}{2}t(1-t)(R+\epsilon)^2. \end{aligned}$$

This proves the first claim. For the second claim, choose $t = \frac{r}{R+\epsilon}$ and apply the first claim. □

Slightly modifying the previous arguments also yields estimates for the volume of spherical shells

$$v_{R,\delta} := m(\overline{B}_R(x) \setminus B_{R-\delta}(x)).$$

Let ν_1 denote uniform distribution in the shell $\overline{B}_R(x) \setminus B_{R-\delta}(x)$ and let ν_0 (as before) be uniform distribution in $\overline{B}_\epsilon(x)$). Then we now obtain

$$R - \epsilon - \delta \le \mathsf{d}_W(\nu_0,\nu_1) \le R + \epsilon$$

and

$$\mathrm{supp}[\nu_t] \subset \overline{B}_{\epsilon+t(R+\epsilon)}(x) \setminus B_{R-\delta-(1-t)(R+\epsilon)} \tag{4.47}$$

for the probability measures on the geodesic connecting ν_0 and ν_1. Hence, arguing similarly as before, we deduce

Theorem 4.26. *Let (M, d, m) be an arbitrary metric measure space with $\underline{\mathbb{C}\mathrm{urv}}(M, \mathsf{d}, m) \geq K$ for some $K \in \mathbb{R}$. For fixed $x \in \mathrm{supp}[m]$ consider $v_{R,\delta} := m(\overline{B}_R(x) \setminus B_{R-\delta}(x))$. Then for all $\epsilon, \delta, r > 0$ and all $R > r > 2\epsilon + \delta$*

$$v_{R,\delta} \leq v_\epsilon \cdot \left(\frac{v_{\epsilon+r,2\epsilon+\delta}}{v_\epsilon}\right)^{\frac{R+\epsilon}{r}} \cdot \exp\left(-\frac{K}{2}\left(1 - \frac{r}{R+\epsilon}\right) \cdot \left(R - \frac{\delta}{2} \pm \frac{2\epsilon+\delta}{2}\right)^2\right) \tag{4.48}$$

where $\pm$ has to be chosen as $+$ if $K \leq 0$ and as $-$ if $K > 0$.
Choosing $\epsilon = \delta = r/2$ this yields in the case $K \geq 0$ for all $R \geq 3\epsilon > 0$

$$v_{R,\epsilon} \leq v_{3\epsilon} \cdot \left(\frac{v_{3\epsilon}}{v_\epsilon}\right)^{R/2\epsilon} \cdot \exp\left(-\frac{K}{2}\left[(R-3\epsilon)^2 - \epsilon^2\right]\right). \tag{4.49}$$

In particular, $K > 0$ implies that m has finite mass and finite variance.

In general, estimating the volume of concentric balls in terms of squared exponential growing functions is best possible, as demonstrated in the previous Example. In a forthcoming paper [St05], we discuss metric measure spaces satisfying a so-called curvature-dimension condition (K, N) (replacing the condition that the curvature is $\geq K$) with some additional number $N \in \mathbb{R}_+$, playing the role of a dimension. We will prove that under this condition the volume of balls grows at most exponentially.

References

[Al51] A. D. Alexandrov (1951): A theorem on triangles in a metric space and some applications. *Trudy Math. Inst. Steklov* **38**, 5-23. (Russian; translated into German and combined with more material in [Al57])

[Al57] A. D. Alexandrov (1957): Über eine Verallgemeinerung der Riemannschen Geometrie. *Schr. Forschungsinst. Math. Berlin* **1**, 33-84.

[AT04] L. Ambrosio, P. Tilli (2004): *Topics on analysis in metric spaces.* Oxford Lecture Series in Mathematics and its Applications **25**, Oxford University Press, Oxford.

[BE85] D. Bakry, M. Émery (1985): Diffusions hypercontractives. *Séminaire de probabilités XIX*, Lecture Notes in Math. **1123**, Springer, Berlin, 177–206.

[BG99] S. G. Bobkov, F. Götze (1999): Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163**, no. 1, 1–28.

[BH99] M. R. Bridson, A. Haefliger (1999): *Metric spaces of non-positive curvature.* Grundlehren der Mathematischen Wissenschaften, **319**, Springer-Verlag, Berlin.

[Br91] Y. Brenier (2001): Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44**, no. 4, 375–417.

[BBI01] D. Burago, Y. Burago, S. Ivanov (2001): *A course in metric geometry.* Graduate Studies in Mathematics **33**. American Mathematical Society, Providence, RI.

[BGP92] Y. Burago, M. Gromov, G. Perelman (1992): A. D. Aleksandrov spaces with curvatures bounded below. (Russian) *Uspekhi Mat. Nauk* **47**, no. 2(284), 3–51, 222; translation in *Russian Math. Surveys* **47** (1992), no. 2, 1–58.

[Ch93] I. Chavel (1993): *Riemannian geometry—a modern introduction.* Cambridge Tracts in Mathematics **108**, Cambridge University Press, Cambridge.

[Ch70] J. Cheeger (1970): Finiteness theorems for Riemannian manifolds. *Amer. J. Math.* **92**, 61–74.

[Ch99] J. Cheeger (1999): Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.* **9**, no. 3, 428–517.

[CC97/00] J. Cheeger, T. H. Colding (1997,2000,2000): On the structure of spaces with Ricci curvature bounded below I, II, III. *J. Differential Geom.* **46**, no. 3, 406–480; *J. Differential Geom.* **54**, no. 1, 13–35; *J. Differential Geom.* **54**, no. 1, 37–74.

[CMS01] D. Cordero-Erausquin, R. McCann, M. Schmuckenschläger (2001): A Riemannian interpolation inequality à la Borell, Brascamb and Lieb. *Invent. Math.* **146**, 219-257.

[Da89] E. B. Davies (1989): *Heat kernels and spectral theory.* Cambridge Tracts in Mathematics **92**, Cambridge University Press, Cambridge.

[DD02] M. Del Pino, J. Dolbeault (2002): Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. *J. Math. Pures Appl.* **81**, no. 9, 847–875.

[Du89] R. M. Dudley (1989): *Real analysis and probability.* The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

[FÜ04a] D. Feyel, A. S. Üstünel (2004): Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theory Related Fields* **128**, no. 3, 347–385.

[FÜ04b] D. Feyel, A. S. Üstünel (2004): The strong solution of the Monge-Ampère equation on the Wiener space for log-concave densities. *C. R. Math. Acad. Sci. Paris* **339**, no. 1, 49–53.

[Gr81a] M. Gromov (1981): Groups of polynomial growth and expanding maps. inst. Hautes Études Sci. Publ. Math. **53**, 53-78

[Gr81] M. Gromov (1981): *Structures Métriques des Variétes Riemanniennes.* Edited by J. LaFontaine and P. Pansu. Textes Mathématiques, 1. CEDIC, Paris, 1981.

[Gr99] M. Gromov (1999): *Metric structures for Riemannian and non-Riemannian spaces.* Based on [Gr81] (English translation by Sean Michael Bates). With appendices by M. Katz, P. Pansu and S. Semmes. Birkhäuser Boston Inc., Boston, MA.

[GP97] K. Grove, P. Petersen (1997): *Comparison geometry.* Papers from the Special Year in Differential Geometry held in Berkeley, CA, 1993-94. Research Institute Publications, **30**, Cambridge University Press, Cambridge.

[He01] J. Heinonen (2001): *Lectures on analysis on metric spaces.* Universitext, Springer-Verlag, New York.

[HK95] P. Hajłasz, P. Koskela (1995): Sobolev meets Poincaré. *C. R. Acad. Sci. Paris Sr. I Math.* **320**, no. 10, 1211–1215.

[HK00] P. Hajłasz, P. Koskela (2000): Sobolev met Poincaré. *Mem. Amer. Math. Soc.* **145**, no. 688.

[JKO98] R. Jordan, D. Kinderlehrer, F. Otto (1998): The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**, no. 1, 1–17 (electronic).

[Ka42] L. V. Kantorovich (1942): On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **37**, 199–201.

[KR57] L. V. Kantorovich, G. S. Rubinstein (1957): On a functional space and certain extremum problems (Russian). *Dokl. Akad. Nauk SSSR (N.S.)* **115**, 1058–1061.

[KK94] A. Kasue, H. Kumura (1994): Spectral convergence of Riemannian manifolds. *Tohoku Math. J.* **46**, no. 2, 147–179.

[Ka04] A. Kasue (2004): Convergence of Riemannian manifolds and Laplace operators II, Preprint 2004.

[Ki01] J. Kigami (2001): *Analysis on fractals*. Cambridge Tracts in Mathematics **143**. Cambridge University Press, Cambridge.

[Ko00] P. Koskela (2000): Upper gradients and Poincaré inequalities. *Lecture notes on analysis in metric spaces (Trento, 1999)*, 55–69, Appunti Corsi Tenuti Docenti Sc., Scuola Norm. Sup., Pisa.

[KS03] K. Kuwae, T. Shioya (2003): Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. *Comm. Anal. Geom.* **11**, no. 4, 599–673.

[Le01] M. Ledoux (2001): *The concentration of measure phenomenon*. Mathematical Surveys and Monographs **89**, American Mathematical Society, Providence, RI, 2001.

[LY86] P. Li, S.-T. Yau (1986): On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156**, no. 3-4, 153–201.

[Mc95] R. J. McCann (1995): Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80**, no. 2, 309–323.

[Mc97] R. J. McCann (1997): A convexity principle for interacting gases. *Adv. Math.* **128**, no. 1, 153–179.

[M1781] G. Monge (1871): *Mémoire sur la théorie des déblais et des remblais*. Histoire de l'Académie Royale des Sciences, Paris.

[OV00] F. Otto, C. Villani (2000): Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173**, no. 2, 361–400.

[Ot01] F. Otto (2001): The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26**, no. 1-2, 101–174.

[Pe02] G. Perelman (2002): *The entropy formula for the Ricci flow and its geometric applications*, Preprint 2002.

[Pl02] C. Plaut (2002): Metric spaces of curvature $\geq k$, Chapter 19 in *Handbook of Geometric Topology*, Elsevier Science.

[RR98] S. T. Rachev, L. Rüschendorf (1998): *Mass transportation problems. Vol. I. Theory.* Probability and its Applications. Springer, New York.

[RS04] M.-K. von Renesse, K. T. Sturm (2003): Transport inequalities, gradient estimates, entropy and Ricci curvature. To appear in Comm. Pure Appl. Math.

[SC02] L. Saloff-Coste (2002): *Aspects of Sobolev-type inequalities.* London Mathematical Society Lecture Note Series **289**, Cambridge University Press, Cambridge.

[St98] K. T. Sturm (1998): Diffusion processes and heat kernels on metric spaces. *Ann. Probab.* **26**, no. 1, 1–55.

[St99] K. T. Sturm (1999): Metric spaces of lower bounded curvature. *Exposition. Math.* **17** no. 1, 35–47.

[St03] K. T. Sturm (2003): Probability Measures on Metric Spaces of Nonpositive Curvature. In *"Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces"* (edited by P. Auscher, T. Coulhon, A. Grigor'yan) pp. 357-390. Contemporary Mathematics **338**. AMS 2003.

[St04] K. T. Sturm (2004): Convex functionals of probability measures and nonlinear diffusions on manifolds, to appear in J. Math. Pures Appl.

[St04a] K. T. Sturm (2004): Generalized Ricci Bounds and Convergence of Metric Measure Spaces, to appear in C. R. Acad. Sci. Paris, Ser. I.

[St05] K. T. Sturm (2005): A curvature-dimension condition for metric measure spaces. In preparation.

[Ta95] M. Talagrand (1995): Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81**, 73–205.

[Vi03] C. Villani (2003): *Topics in Mass Transportation.* Graduate Studies in Mathematics, American Mathematical Society.

[Wa69] L. N. Wasserstein [or Vasershtein] (1969): Markov processes over denumerable products of spaces describing large system of automata. *Problems of Information Transmission* **5** no. 3, 47–52 (translated from *Problemy Peredači Informacii* **5** (1969), no. 3, 64–72 (Russian))

Institut für Angewandte Mathematik
Universität Bonn
Wegelerstrasse 6
53115 Bonn
Germany

sturm@uni-bonn.de