

A visual introduction to Riemannian curvatures and some discrete generalizations

Yann Ollivier

ABSTRACT. We try to provide a visual introduction to some objects used in Riemannian geometry: parallel transport, sectional curvature, Ricci curvature, Bianchi identities... We then explain some of the strategies used to define analogues of curvature in non-smooth or discrete spaces, beginning with Alexandrov curvature and δ -hyperbolic spaces, and insisting on various notions of generalized Ricci curvature, which we briefly compare.

The first part of this text covers in a hopefully intuitive and visual way some of the usual objects of Riemannian geometry: parallel transport, sectional curvature, Ricci curvature, the Riemann tensor, Bianchi identities... For each of those we try to provide one or several pictures that convey the meaning (or at least one possible interpretation) of the formal definition.

Next, we consider the problem of defining analogues of curvature for non-smooth or discrete objects. This is in the spirit of “synthetic” or “coarse” geometry, in which large-scale properties of metric spaces are investigated instead of fine small-scale properties; one of the aims being to be impervious to small perturbations of the underlying space. Motivation for these non-smooth or discrete extensions often comes from various other fields of mathematics, such as group theory or optimal transport. From this viewpoint is emerging a theory of *metric measure spaces*, see for instance [Gro99].

Thus we cover the definitions and mention some applications of δ -hyperbolic spaces and Alexandrov curvature (generalizing sectional curvature), and of two notions of generalized Ricci curvature: *displacement convexity of entropy* introduced by Sturm and Lott–Villani (following Renesse–Sturm and others), and *coarse Ricci curvature* as used by the author. We mention new theorems obtained for the original Riemannian case thanks to this more general viewpoint. We also briefly discuss the differences between these two approaches to Ricci curvature.

Note that scalar curvature is not covered, for lack of the author’s competence about recent developments and applications in this field.

1. Riemannian curvatures

The purpose of this section is not to give a formal course in Riemannian geometry. We will therefore not reproduce formal definitions of the basic objects, but try to convey some pictorial intuition behind them. For a more formal introduction we

refer, for instance, to [DC92], or to [Ber03] for a survey of numerous developments in Riemannian geometry.

Riemannian manifolds. The most basic example of a Riemannian manifold is a smooth surface embedded in three-dimensional Euclidean space. More generally, any smooth manifold can be seen as a set $X \subset \mathbb{R}^p$ for some integer p , such that, at any point $x \in X$, there exists an N -dimensional affine subspace of \mathbb{R}^p which coincides with X at first order around x . This subspace is the *tangent space* $T_x X$ at $x \in X$, and N is the *dimension* of X .

Note that if $t \mapsto c(t)$ is a smooth curve inside X , then its derivative $dc(t)/dt$ is a vector tangent to X at $c(t)$.

A *Riemannian* manifold is a manifold equipped with a *Riemannian metric*, that is, for each $x \in X$, a definite positive quadratic form defined on $T_x X$. For instance, if X is included in \mathbb{R}^p , such a quadratic form might be the restriction to $T_x X$ of the canonical Euclidean structure on \mathbb{R}^p . We will assume that the quadratic form depends smoothly on $x \in X$.

Such a quadratic form, applied to any tangent vector, allows to define the norm of this vector. By integration, one can then define the length of a curve in X . The distance (inside X) between two points of X is then defined as the infimum of the lengths of all curves between these two points. This turns X into a metric space.

We will always assume that X is connected and complete for this metric.

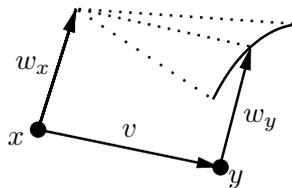
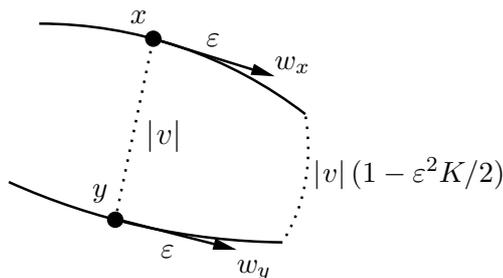
A *geodesic* is a curve γ in X such that, for any two close enough points on γ , the distance in X between those two points is obtained by travelling along γ . Locally, such curves always exist. Moreover, given a point $x \in X$ and a tangent vector $v \in T_x X$ at x , there exists exactly one geodesic starting at x such that its initial velocity is v and which has constant speed. This will be the *geodesic starting along* v .

The *endpoint* of v will be the point, denoted $\exp_x v$, obtained after following the geodesic starting along v for a unit time.

Parallel transport. Say we are given two very close points x and y in a Riemannian manifold. Is there a way to compare a tangent vector at x and a tangent vector at y , even though they live, a priori, in different vector spaces? This is done via *parallel transport*.

So let w_x be a tangent vector at x ; we are looking for a tangent vector w_y at y which would be “the same” as w_x . Since x and y are very close, we may assume that y is the endpoint of a small tangent vector v at x . For simplicity, we will assume that w_x is orthogonal to v , and that the norm of w_x is very small. Then, there exists a particular tangent vector w_y at y : it is the one whose endpoint is closest to the endpoint of w_x , given the restriction that w_y be orthogonal to v (if this orthogonality condition is lifted, then of course there is a tangent vector at y whose endpoint is exactly the endpoint of w_x , but this vector “turns back towards x ”). The vector w_y is the best candidate to be “the same” as w_x , translated to y (Fig. 1).

Parallel transport is the operation mapping w_x to w_y . (More exactly, since we have assumed that w_x is very small, we take the linear component of the map $w_x \mapsto w_y$ for small w_x and then extend by linearity to a map between the whole spaces $T_x X$ and $T_y X$. We have also assumed that w_x is orthogonal to v ; by definition, we

FIGURE 1. Parallel transport of w_x along v .FIGURE 2. The sectional curvature K .

define the parallel transport of v from x to y as the velocity at y of the geodesic starting along v .)

More generally, parallel transport of a vector w along any smooth curve starting at x can be defined by decomposing the curve into small intervals and performing successive parallel transports along these subintervals.

Sectional curvature and Ricci curvature. Let us now define various curvatures. The first one we consider is *sectional* curvature.

Let again x be a point in X , v a small tangent vector at x , y the endpoint of v , w_x a small tangent vector at x , and w_y the parallel transport of w_x from x to y along v . If, instead of a Riemannian manifold, we were working in ordinary Euclidean space, the endpoints x' and y' of w_x and w_y would constitute a rectangle with x and y . But in a manifold, generally these four points do not constitute a rectangle any more.

Indeed, because of curvature, the two geodesics starting along w_x and w_y may diverge from or converge towards each other. Thus, on a sphere (positive curvature), two meridians starting at two points on the equator have parallel initial velocities, yet they converge at the North (and South) pole. Since the initial velocities w_x and w_y are parallel to each other, this effect is at second order in the distance along the geodesics (Fig. 2).

Thus, let us consider the points lying at distance ε from x and y on the geodesics starting along w_x and w_y , respectively. In a Euclidean setting, the distance between those two points would be $|v|$, the same as the distance between x and y . The discrepancy from this Euclidean case is used as a definition of a curvature.

DEFINITION 1.1 (Sectional curvature). Let (X, d) be a Riemannian manifold. Let v and w_x be two unit-length tangent vectors at some point $x \in X$. Let $\varepsilon, \delta > 0$. Let y be the endpoint of δv and let w_y be obtained by parallel transport of w_x from

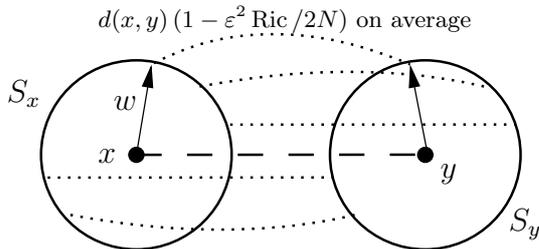


FIGURE 3. The Ricci curvature Ric.

x to y . Then

$$d(\exp_x \varepsilon w_x, \exp_y \varepsilon w_y) = \delta \left(1 - \frac{\varepsilon^2}{2} K(v, w) + O(\varepsilon^3 + \varepsilon^2 \delta) \right)$$

when $(\varepsilon, \delta) \rightarrow 0$. This defines a quantity $K(v, w)$, which is the *sectional curvature* at x in the directions (v, w) .

In contrast, *Ricci curvature* only depends on one tangent vector v ; it is obtained by averaging $K(v, w)$ over all the directions w (Fig. 3).

DEFINITION 1.2 (Ricci curvature). Let x be a point in an N -dimensional Riemannian manifold. Let v be a unit tangent vector at x . The *Ricci curvature along v* is the quantity $\text{Ric}(v)$ defined as N times the average of $K(v, w)$, where the average is taken over w running over the unit sphere in the tangent space $T_x X$.

The scaling factor N comes from the traditional definition of Ricci curvature as the trace of a linear map, which results in a sum over a basis rather than an average over the unit sphere. Actually $\text{Ric}(v)$ is a quadratic form in v and is usually denoted $\text{Ric}(v, v)$; the Ricci tensor is the corresponding bilinear form $\text{Ric}(v, v')$.

This definition can be rephrased as follows, due to the fact that we used parallel transport $w_x \mapsto w_y$ in the definition of curvature.

COROLLARY 1.3. Let v be a unit tangent vector at a point x in a Riemannian manifold. Let $\varepsilon, \delta > 0$ and let y be the endpoint of δv .

Let S_x be the set of endpoints of all tangent vectors at x with norm ε , and likewise let S_y be the set of endpoints of the sphere of radius ε in the tangent space at y . Then, if S_x is mapped to S_y using parallel transport, the average distance between a point of S_x and its image is

$$\delta \left(1 - \frac{\varepsilon^2}{2N} \text{Ric}(v, v) + O(\varepsilon^3 + \varepsilon^2 \delta) \right)$$

when $(\varepsilon, \delta) \rightarrow 0$.

If balls are used instead of spheres, the scaling factor is $\frac{\varepsilon^2}{2(N+2)}$ instead of $\frac{\varepsilon^2}{2N}$.

Thus, Ricci curvature is positive when “balls are closer than their centers are”. We will see that this property is well suited as a definition in more general, non-smooth spaces.

Let us mention another, more dynamical way to visualize Ricci curvature. Let again v be a unit tangent vector at some point x in the Riemannian manifold X . Let C be any small neighborhood of X , having an arbitrary shape. For each point

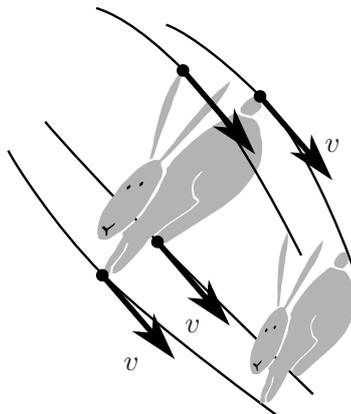


FIGURE 4. Curvature and volume change under geodesic flow.

z in C , let us consider the geodesic $t \mapsto z_t$ starting at z and whose initial velocity is v (we implicitly use parallel transport to identify v with a tangent vector at z). On average, the geodesics starting from various choices of z get closer or further apart depending on the sign of curvature. Now let us have C “slide along” these geodesics; more exactly, let C_t be the set $\{z_t, z \in C\}$. In particular $C_0 = C$ (Fig. 4).

Then we have

$$\text{vol } C_t = \text{vol } C \left(1 - \frac{t^2}{2} \text{Ric}(v) + \text{smaller terms} \right)$$

(note that the derivative of $\text{vol } C_t$ is 0 for $t = 0$ because we choose geodesics with parallel initial speeds). Thus Ricci curvature controls the evolution of volumes under the geodesic flow.

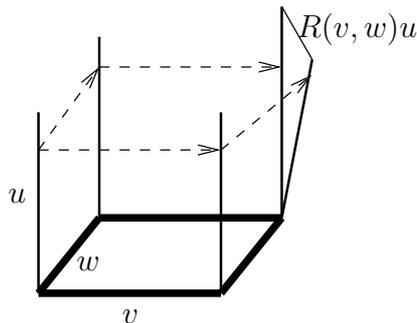
Signs of curvature. Positive curvature and negative curvature are very different worlds; we refer to [Gro91] for a thorough discussion of the “sign and geometric meaning of curvature”. It is often the case, in Riemannian geometry, that theorems in negative curvature assume that all sectional curvatures $K(v, w)$ are negative, whereas theorems in positive curvature assume that the Ricci curvatures $\text{Ric}(v, v)$ is positive for any v (which is weaker than having all sectional curvatures $K(v, w) \geq 0$).

We give here two basic theorems of Riemannian geometry illustrating each of these situations. When Ricci curvature is positive, the Bonnet–Myers theorem states that if a Riemannian manifold is more positively curved than a sphere, then its diameter is smaller.

THEOREM 1.4 (Bonnet–Myers). *Let X be an N -dimensional Riemannian manifold. Let $\inf \text{Ric}(X)$ be the infimum of the Ricci curvature $\text{Ric}(v, v)$ over all unit tangent vectors v .*

Let $S^N \subset \mathbb{R}^{N+1}$ be the unit sphere of the same dimension as X . Then, if $\inf \text{Ric}(X) \geq \inf \text{Ric}(S^N)$ then $\text{diam } X \leq \text{diam } S^N$.

Several other theorems about positive Ricci curvature follow this pattern of a comparison with the sphere of the same dimension.

FIGURE 5. The Riemann curvature $R(v, w)u$.

As for negative curvature, let us mention the Cartan–Hadamard theorem, which implies that all topological information about a negatively curved manifold is contained in its fundamental group. This is obviously false in positive curvature, since, for instance, a sphere is simply connected.

THEOREM 1.5 (Cartan–Hadamard). *Let X be an N -dimensional Riemannian manifold and assume that $K(v, w) \leq 0$ for any pair of tangent vectors v, w at the same point. Then the universal cover of X is homeomorphic to \mathbb{R}^N .*

The Riemann curvature tensor. The next object in our journey through curvatures is the *Riemann curvature tensor*. It depends on three tangent vectors u, v, w at a given point x , and its output is another tangent vector at x , denoted $R(v, w)u$. Before we define this tensor, we need to say more about parallel transport.

Given two tangent vectors v and w at point x , we may parallel-transport w along v ; this yields a tangent vector w' at the endpoint of v . Likewise, we may also parallel-transport v along w , getting a tangent vector v' at the endpoint of w . In Euclidean space, the vectors v, w, v', w' make up a parallelogram; in particular, the endpoints of v' and of w' coincide. But we could expect that, in a general Riemannian manifold, the endpoints of v' and w' may not coincide anymore. Actually things are not so bad: for small v and w , the endpoints of v' and w' do coincide at the first order at which this is non-trivially true (more precisely, up to $o(|v||w|)$). This property of parallel transport is called *torsion-freeness* and expresses the fact that “parallelograms close up”.

Actually this property was used implicitly before: when we defined curvature, we had to look at a phenomenon of order $O(|v||w|^2)$.

Let us now consider three vectors u, v, w instead of two. Let us keep the parallelogram built from u and v , and let us parallel-transport u along the path vw , or along the path wv . This time, we obtain two different results; their difference is the *Riemann curvature* $R(v, w)u$, which is again a tangent vector (Fig. 5).

This proves that, in contrast to parallelograms, cubes don’t close up in Riemannian geometry.

Let us take a further look at this. Given three directions u, v, w for the sides of a cube whose base corner is x , there are several ways to build the farthest corner of the cube: we saw that the paths vwu and wvu do not end up at the same point, and that the difference is $R(v, w)u$. (When we say “the path vwu ”, it means: follow the tangent vector v to its endpoint, parallel-transport w along this path, follow

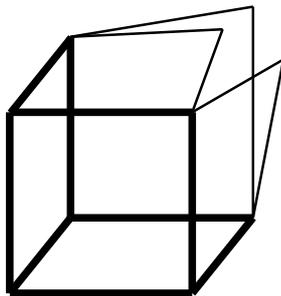


FIGURE 6. The first Bianchi identity.

the obtained vector to its endpoint, parallel-transport u along the whole path, and follow the obtained vector to its endpoint.)

All in all, given u , v and w , there are six possible paths to the farthest corner, hence six possible ways to define this farthest corner. However, these six paths define only three distinct endpoints: indeed, torsion-freeness guarantees that $vw = wv$, but torsion-freeness at the endpoint of u also guarantees that $uvw = uvw$. The latter equality just expresses that “lateral faces of the cube close up”.

These three distinct endpoints form a triangle. We saw that, by definition, one of the sides of this triangle is $R(v, w)u$; by symmetry the other sides are $R(w, u)v$ and $R(u, v)w$. Since the three sides of any triangle add up to 0 we get

$$R(u, v)w + R(v, w)u + R(w, u)v = 0$$

which is the *first Bianchi identity* and was discovered by Ricci.

With Bruno Sévenec (unpublished work), we turned this heuristic argument into a very short formal proof along the same lines, by interpreting expressions such as uvw in a suitable quotient of a jet space of curves in X starting at the base point x .

Riemannian volume measure. Another notion we will use from Riemannian geometry is that of *Riemannian volume measure*. This is a measure on an N -dimensional Riemannian manifolds which gives volume ε^N to a small cube made of N tangent vectors at the same point, each of length ε and orthogonal to each other.

This informal definition yields a volume n -form associated with a Riemannian manifold. This volume form can be denoted $\sqrt{\det g}$ where g is the bilinear form defining the metric. The reason is as follows: At a given point x , the metric g is a bilinear form on the tangent space $E = T_x X$. A bilinear form is an application from E to its dual E^* . Thus, from $g : E \rightarrow E^*$ we can apply the textbook definition of the determinant of a linear map, passing to the exterior product: $\det g = \bigwedge^N g$ is a map from $\bigwedge^N E$ to $\bigwedge^N E^* \simeq (\bigwedge^N E)^*$. That is to say, $\det g$ is a bilinear form on $\bigwedge^N E$ (which is a 1-dimensional space since $N = \dim E$). Moreover $\det g$ is definite positive whenever g is. The square root $\sqrt{\det g}$ is thus a *norm* on $\bigwedge^N E$, that is, a volume form.

2. Discrete curvatures

Let us now consider some generalizations of these notions to more general spaces than smooth manifolds. This is motivated by the consideration of “synthetic” geometry: a geometry that focuses on “large-scale” instead of “small-scale” properties, and uses mainly metric comparisons, avoiding local properties such as differentiability, with the goal of being invariant under small perturbations of the underlying space. We refer for instance to Gromov’s influential book [Gro99].

Two types of spaces can be considered. *Geodesic spaces* are continuous spaces in which the distance between points is always realized by the length of a continuous curve; important examples are piecewise smooth objects, or objects obtained as limits of smooth manifolds, which can be thought of as manifolds with singularities.

Just as important are truly *discrete spaces*; typical examples are discrete groups equipped with a Cayley graph metric, or spaces of configurations of discrete statistical physics systems, for instance.

We will briefly mention sectional curvature, the generalizations of which are now well understood especially in negative curvature, with deep connections for instance with group theory. We will focus on generalizations of Ricci curvature, which were developed more recently. In passing we will mention new results obtained thanks to a discrete viewpoint on Ricci curvature, both in the original setting of Riemannian manifolds and in a surprisingly diverse array of discrete settings.

2.1. Non-smooth and discrete sectional curvature: Alexandrov curvature and Gromov-hyperbolicity. There are at least two notions of metric spaces having bounds on their sectional curvature. The first one is restricted to geodesic spaces and is that of *spaces with curvature bounds in the sense of Alexandrov* (also termed *locally CAT(k) spaces* for curvature bounded above, *locally CAT⁺(k) spaces* for curvature bounded below). The second is that of *δ -hyperbolic spaces* (or *Gromov-hyperbolic spaces*) and applies to any metric space, but only for negative curvature “at large scale”.

Both notions have been extremely successful, in particular in negative curvature, as δ -hyperbolic spaces and CAT(0) spaces have deep connections to geometric group theory, for instance. We cannot list all the relevant work and we refer the reader to [BH99] or [BBI01] for a more extensive treatment.

Curvature bounds in the sense of Alexandrov go as follows. Let X be a geodesic space. A *triangle* in X is a triplet of points $(a, b, c) \in X^3$, together with three curves from a to b , b to c and c to a , respectively (the *sides* of the triangle), such that the lengths of these curves realize the distances $d(a, b)$, $d(b, c)$ and $d(c, a)$, respectively.

The curvature criterion of Alexandrov (also called CAT criterion, for Cartan–Alexandrov–Topogonov), states that triangles become “thinner” in negative curvature and “fatter” in positive curvature. This will be measured by the distance between a vertex of the triangle, and a point on the opposite side, compared to that expected in the Euclidean situation (Fig. 7).

For any triangle in X , there exists a *comparison triangle* (A, B, C) in the Euclidean plane whose sides have the same lengths. More generally, let S_k is the standard 2-dimensional space of constant curvature k (Euclidean plane if $k = 0$, Euclidean sphere of radius $1/\sqrt{k}$ with its intrinsic metric when $k > 0$, and hyperbolic plane with curvature k if $k < 0$), and let r_k be the diameter of S_k (which is infinite if $k \leq 0$). Then, for any triangle in X with perimeter at most $2r_k$, there

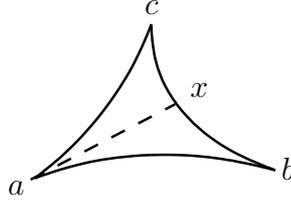


FIGURE 7. Comparison triangles and Alexandrov curvature.

exists a comparison triangle (A_k, B_k, C_k) in S_k , whose sides have the same lengths as (a, b, c) . Such a comparison triangle is unique up to isometry, and will be called *the comparison triangle of (a, b, c) in curvature k* .

DEFINITION 2.1. Let X be a geodesic space and k a real number. X is a *space of curvature $\leq k$ in the sense of Alexandrov* if, for any small enough triangle (a, b, c) in X , and for any point x on the side bc of this triangle, the following holds:

$$d(a, x) \leq d_{S_k}(A_k, X_k)$$

where (A_k, B_k, C_k) is the comparison triangle of (a, b, c) in curvature k , and X_k is the point on the side B_kC_k corresponding to x , i.e. such that $d(X_k, B_k) = d(x, b)$. Likewise, X is a *space of curvature $\geq k$ in the sense of Alexandrov* if, in the same situation, the reverse inequality

$$d(a, x) \geq d_{S_k}(A_k, X_k)$$

holds.

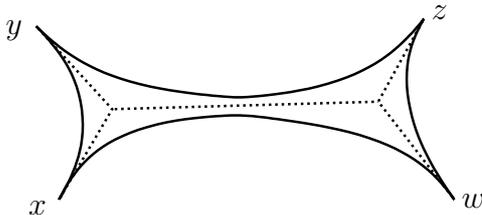
This definition is compatible with the Riemannian notion of sectional curvature. For instance, we have the following ([BH99], Theorem II.1A.6).

THEOREM 2.2. *Let X be a complete smooth Riemannian manifold and $k \in \mathbb{R}$. Then X has curvature at most k in the sense of Alexandrov if and only if for any two tangent vectors v, w at the same point, with unit length and orthogonal to each other, one has $K(v, w) \leq k$.*

The terminology *CAT(k) space*, for curvature $\leq k$, applies to the case when *all* triangles, not only small ones, satisfy the comparison criterion. For complete smooth Riemannian manifolds this is equivalent to negative sectional curvature and being simply connected. On the other hand, for spaces of curvature bounded *below*, imposing the condition on small triangles implies the same for all triangles (Topogonov's theorem, Theorem 10.5 in [BBI01]).

Spaces with curvature bounds in the sense of Alexandrov keep a number of properties of Riemannian manifolds with controlled sectional curvature. For instance, the Cartan–Hadamard theorem (Theorem 1.5) in this context states that any CAT(0) space is contractible ([BH99], Corollary II.1.5).

The Bonnet–Myers theorem (Theorem 1.4) applies as well: a space with curvature at least $k > 0$ in the sense of Alexandrov has diameter at most π/\sqrt{k} ([BBI01], Theorem 10.4.1); otherwise, triangles would get too big with respect to their comparison triangle in the sphere. Note that here, contrary to Theorem 1.4, we do not work with a Ricci curvature bound but with a (stronger) sectional curvature bound instead.

FIGURE 8. A δ -hyperbolic quadrilateral.

Spaces with curvature bounds in the sense of Alexandrov do not stray too far from smooth manifolds. For instance ([BBI01], Theorem 10.8.3), an Alexandrov space with Hausdorff dimension n contains an open dense subset which is an n -dimensional manifold.

For *discrete* spaces, there is a very successful notion of negative sectional curvature, namely δ -hyperbolicity (also called Gromov-hyperbolicity). This notion is usually attributed to Rips and was thoroughly developed by Gromov [Gro87], especially in the context of geometric group theory, where it led to the notion of hyperbolic groups.

δ -hyperbolicity states that, at large scales, a space behaves metrically almost like a tree. It disregards any phenomena occurring at small scales; in particular, a bounded metric space X is always δ -hyperbolic with $\delta = \text{diam } X$.

The simplest definition of δ -hyperbolicity involves triangles, and thus assumes that the space is geodesic. For non-necessarily geodesic metric spaces, the condition is expressed in terms of quadrilaterals, as follows.

The underlying idea is that large n -gons in the hyperbolic plane look almost like trees, their sides clinging to each other. Let us pass to the ideal limit case and consider four points w, x, y, z in a tree. Assume the general situation where no branch of the subtree joining these points is degenerate: this subtree is made of five segments, one of which lies in the middle.

Let us consider the three pairwise sums of distances $d(w, x) + d(y, z)$, $d(w, y) + d(x, z)$ and $d(w, z) + d(x, y)$. Of these three sums, only one (the smallest one) does not involve the length of the middle segment of the tree; the other two are equal. This can be summed up by the ultrametric-like inequality

$$d(w, x) + d(y, z) \leq \max(d(w, y) + d(x, z), d(w, z) + d(x, y))$$

for all points w, x, y, z in a metric tree. By permuting the roles of the points, we get that, among the three sums involved, the largest two must be equal (because the largest is bounded by the max of the smallest and second-largest).

Now we just perturb this inequality by an additive term. This means that, in a hyperbolic space, the metric relations between any four points are the same as in a tree up to a small error (Fig. 8).

DEFINITION 2.3. Let $\delta \geq 0$. A metric space X is δ -hyperbolic if

$$d(w, x) + d(y, z) \leq \max(d(w, y) + d(x, z), d(w, z) + d(x, y)) + 2\delta$$

for any points w, x, y, z in X .

This does not imply any local topological property when $\delta > 0$: indeed, anything can happen at scales smaller than δ . One key feature of this definition is its invariance under *quasi-isometries* [BH99], which makes it very robust.

This notion is an extension of negative sectional curvature “at large scales” in the sense that any Riemannian manifold with negative sectional curvature and which is simply connected, is δ -hyperbolic. Moreover $\text{CAT}(-k)$ spaces for $k > 0$ are δ -hyperbolic, so that δ -hyperbolicity is compatible with global negative curvature in the sense of Alexandrov. Finally, δ -hyperbolic groups (groups whose Cayley graph is δ -hyperbolic) are important objects and keep many of the features of fundamental groups of negatively curved manifolds. Interestingly, these groups turn out to be, in some sense, “generic” among discrete groups [Gro87, Oll05].

2.2. Discrete Ricci curvature: the coarse Ricci curvature approach.

Let us now turn to generalizations of Ricci curvature. The survey [Lot07] contains a discussion of the geometric interest of bounds on Ricci curvature, and the need for a generalized notion of positive Ricci curvature for metric measure spaces.

This time, let us begin with the discrete case, and the notion of *coarse Ricci curvature* used by the author [Oll07, Oll09]. Another approach introduced by Sturm [Stu06] and Lott–Villani [LV09], based on *displacement convexity of entropy*, is also presented below. Coarse Ricci curvature keeps fewer properties from Riemannian manifolds, but is somewhat simpler to present and has a wider range of examples.

Consider the following motivation: From a statistical physics viewpoint, at a macroscopic level there should not be much difference between a system of n particles whose total energy is exactly E , and a system of n particles whose energies are chosen independently with average E/n , when n is large. For free particles, the energy E is (proportional to) the sum of square velocities; so in the first situation the configuration space is a sphere of radius \sqrt{E} in speed space \mathbb{R}^{3n} , while for the case of independent particles, the configuration space for speeds is the whole of \mathbb{R}^{3n} , but equipped with a Gaussian measure (the Gaussian measure being the Maxwell–Boltzmann distribution $\exp(-E/kT)$ for this energy E). If we pick a large number of points according to this latter Gaussian distribution, a set will emerge which is, in some sense, close to a sphere. A third model for our particles of average energy E would be, for each particle, to choose at random an energy either 0 or $2E/3n$ along each axis; in this third model, the configuration space would be the discrete hypercube $\{0, 2E/3n\}^{3n}$.

One can wonder whether these three macroscopically similar viewpoints have anything in common from a geometric point of view. Since the sphere (in the first example) is the archetypical space of positive curvature, one can wonder whether the last two examples (Euclidean space with a Gaussian measure, and discrete hypercube) somehow have positive curvature. For instance, Gromov argues in [Gro99] that the discrete hypercube shares a lot of geometric properties with a sphere.

A successful approach in that direction, with a rich progeny over the years, is Bakry–Émery theory [BE84, BE85]. It allows to define a notion of Ricci curvature for a manifold equipped with a measure (more exactly, for a manifold equipped with a diffusion process whose invariant measure is the measure considered). This way, the Euclidean space endowed with a Gaussian measure acquires positive Ricci curvature. However, the use of diffusions make its adaptation to a discrete setting

delicate, and this theory is not stable if we replace the space considered with a “close enough” space.

Thus it would be nice to have at hand a notion of Ricci curvature, which would be valid “at a certain scale” and thus robust and insensitive to the local structure, discrete or continuous. Ideally, this should be compatible with the Bakry–Émery approach, be easy to check on examples, capture some common features for our three examples above, and allow to generalize some of the properties of Riemannian manifolds with positive Ricci curvature.

Coarse Ricci curvature directly transposes Definition 1.2, or more precisely the Corollary 1.3. The point is to compare the distances between two small balls, with the distance between their centers. Ricci curvature will be positive if *small balls are closer than their centers are*.

For this we need to define what the distance between balls is. It will be practical to think of a ball as a measure of mass 1 around some point (in line with Definition 1.2 which involves an average over the sphere). Then we can use the well-known *transportation distance* (also called Wasserstein distance, Monge–Kantorovich–Rubinstein distance, or earth-mover distance), defined as follows [Vil03].

DEFINITION 2.4 (Transportation distance). Let μ_1, μ_2 be two measures with mass 1 in a metric space (X, d) . A *transference plan* (or *coupling*) from μ_1 to μ_2 is a measure ξ on $X \times X$ such that $\int_y d\xi(x, y) = d\mu_1(x)$ and $\int_x d\xi(x, y) = d\mu_2(y)$. (Thus $d\xi(x, y)$ represents the amount of mass travelling from x to y .)

The L^1 *transportation distance* between μ_1 and μ_2 , denoted $W_1(\mu_1, \mu_2)$, is the best average travelling distance that can be achieved:

$$W_1(\mu_1, \mu_2) = \inf_{\xi \in \Pi(\mu_1, \mu_2)} \iint d(x, y) d\xi(x, y)$$

where $\Pi(\mu_1, \mu_2)$ is the set of transference plans from μ_1 to μ_2 .

The transportation distance between small balls can then be used to define a Ricci curvature at some scale. The relevant notion of a “small ball” depends on the situation at hand: for instance, in a graph it is natural to use balls of radius 1, whereas in manifolds, arbitrarily small balls are used.

Let us assume that, for each point x in space X , we have chosen a measure b_x to play the role of “small ball around x ”. The basic example is when X itself comes with a global measure μ (such as the Riemannian volume measure on a manifold, or the counting measure on a graph) and we have chosen a scale parameter $\varepsilon > 0$; then we can define b_x by restricting the measure to the ε -ball around x and scaling:

$$b_x^{\varepsilon, \mu} = \frac{\mu|_{B(x, \varepsilon)}}{\mu(B(x, \varepsilon))}$$

where $B(x, \varepsilon)$ is the closed ball of radius ε around x . This is a very common choice for b_x but is not the only one; thus, we will keep full generality in the definition.

We are now ready to transpose Corollary 1.3.

DEFINITION 2.5 (Coarse Ricci curvature). Let (X, d) be a metric space. We assume that, for each $x \in X$, a probability measure b_x on X is given. Let x and y be two distinct points in X . The *coarse Ricci curvature* along xy is the quantity $\kappa(x, y)$ defined by the relation

$$W_1(b_x, b_y) = (1 - \kappa(x, y)) d(x, y)$$

(see Fig. 9).

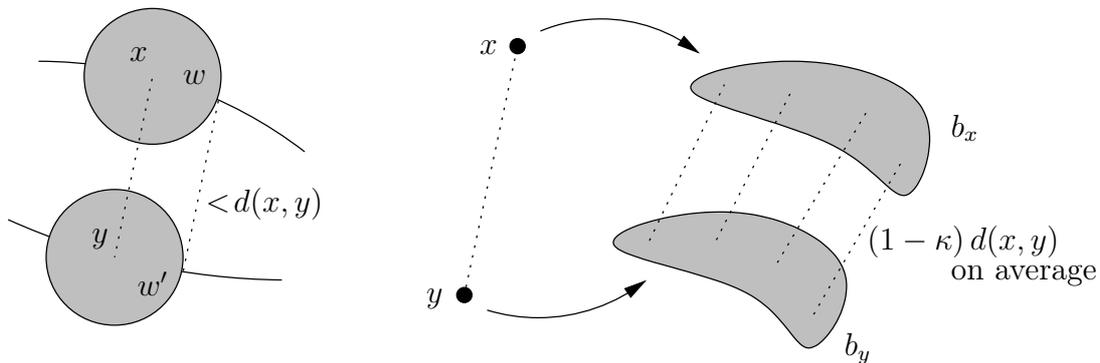


FIGURE 9. In positive curvature, balls are closer than their centers are.

Compared to the Riemannian case, note that we chose not to apply the $\frac{\varepsilon^2}{2N}$ scaling factor. Also, in the Riemannian case Ricci curvature was defined along a tangent vector. Here, in a metric setting, the best we can do for tangent vectors is to have a pair of close points. A simple but key feature of coarse Ricci curvature is that it is enough to compute it for pairs of points which are close enough [Oll07].

The technical assumptions implicit in this definition are basically the ones needed for good behavior of the Wasserstein distance: the space X should be Polish (metric, separable, complete), and the measures b_x should all have a finite first moment [Vil03]. Moreover the map $x \mapsto b_x$ should be measurable.

This kind of condition appears in [RS05] as one among several ways to characterize a lower Ricci bound on manifolds. It was proposed in [Oll07, Oll09] (see also [Jou07]) as a possible definition of Ricci curvature extending to metric spaces several results known for positively curved manifolds.

The data $(b_x)_{x \in X}$ exactly define a *Markov chain* on X : the point x is sent to a point chosen at random according to the measure b_x . The condition $W_1(b_x, b_y) \leq (1 - \kappa)d(x, y)$ actually has quite a long history in the Markov chain literature, under various names: *Dobrushin–Shlosman criterion* [Dob70, DS85, Dob96], or [CW94], or *path coupling* [BD97]; though, without the link to curvature. It appears in some recent Markov chain textbooks [Che04, LPW09]. It is most often used to prove fast convergence of the measure of the Markov chain to its invariant distribution. It is operational in a number of settings: to name just a few, the Ising model of ferromagnetism [DS85], graph colouring [BD97], fast generation of random orthogonal matrices [Oli09], sampling of paths with constraints [Ger10] (see also [Wil04]), simulations of waiting queues [Jou09, JO10]...

The definition also works well if one replaces the Markov chain b_x with a continuous-time Markov process (e.g., [CW94]). In this setting, a lower bound K on Ricci curvature for a Markov process (P^t) states that $W_1(P_x^t, P_y^t) \leq \exp^{-Kt} d(x, y)$, and it is enough to check this for close enough points x, y and small enough time t (see [Vey-b] for details). Note an interesting connection: in the manifold case, if the distance d happens to change with time according to the backwards Ricci flow, then the transportation distance between P_x^t and P_y^t (measured using the distance at time t) will stay bounded, because the two effects compensate [MT10]. This could lead to notions of Ricci flow in metric spaces [MT10, Oll10].

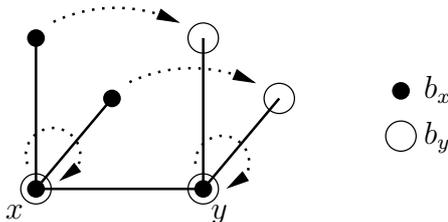


FIGURE 10. Coarse Ricci curvature of the cube.

As examples, let us illustrate how coarse Ricci curvature applies, first, to the original case of Riemannian manifolds, and, second, to the discrete hypercube $\{0, 1\}^N$ mentioned above.

EXAMPLE 2.6 (Riemannian manifolds). Let (X, d) be a smooth Riemannian manifold and let vol be the Riemannian volume measure. Let $\varepsilon > 0$ small enough and consider the ball of radius ε around each point x , i.e., set $b_x = b_x^{\varepsilon, \text{vol}}$ as above. Let $x, y \in X$ be two close enough points. Let v be the unit tangent vector at x directed towards y . Then the coarse Ricci curvature along xy is

$$\kappa(x, y) = \frac{\varepsilon^2 \text{Ric}(v, v)}{2(N + 2)} + O(\varepsilon^3 + \varepsilon^2 d(x, y))$$

This is basically a reformulation of Corollary 1.3. (Actually the latter only yields an inequality, as there might be better transference plans than parallel transport. But parallel transport gives the exact value of the Wasserstein distance at this order [Oll09].)

The next example is the discrete hypercube.

EXAMPLE 2.7 (Discrete hypercube). Let $X = \{0, 1\}^N$ be the N -dimensional discrete hypercube, equipped with its graph (L^1) metric, and let μ be the uniform measure on X . Let $b_x = b_x^{1, \mu}$ be the uniform probability measure on the 1-ball around x (i.e., on x and its N neighbors). Then the coarse Ricci curvature along two adjacent points x, y is $\kappa(x, y) = \frac{2}{N+1}$.

This is quite clear on Figure 10: indeed, to move b_x to b_y , out of the $N + 1$ points in b_x , $N - 1$ of them have to move by one edge, and 2 of them do not have to move at all.

Note that we have computed $\kappa(x, y)$ only for adjacent points x and y ; but, as mentioned before, this is enough to imply the same lower bound on curvature for *all* pairs of points, by the triangle inequality for W_1 [Oll07, Oll09].

The inclusion of x in b_x , i.e., the use of balls rather than spheres, is identical to the well-known need to use *lazy* random walks on bipartite graphs: otherwise, points located at an odd distance never meet each other. In general metric spaces, spheres may be empty and balls are preferable.

Other very simple examples include the grid \mathbb{Z}^n in \mathbb{R}^n with its graph metric, or \mathbb{R}^n itself equipped with any arbitrary Euclidean or non-Euclidean norm: these have 0 curvature as expected. We mentioned above numerous more sophisticated examples from Markov chains or statistical physics; see also the examples in [Oll09].

Since Ricci curvature is an average of sectional curvature, bounds on sectional curvature imply the same bounds on Ricci curvature, at least for ordinary curvature. One can wonder whether the same holds for the generalized notions used here.

This is indeed the case. For δ -hyperbolic groups, for instance, with b_x the uniform measure on a ball of radius $\gg \delta$ around x , one can check [Oll09] that coarse Ricci curvature is indeed negative (and actually takes the smallest possible value). For spaces with curvature bounded below in the sense of Alexandrov, equipped with their (continuous-time) canonical diffusion semigroup (see e.g. [GKO]), a result of Gigli, Kuwada and Ohta [GKO, Oht09a] (combined with [Pet, ZZ10]) implies that if Alexandrov curvature is bounded below, then coarse Ricci curvature is bounded below by the corresponding value.

Let us review some of the theorems obtained thanks to this approach. They are of two kinds: extensions of known theorems about Ricci curvature to a more general setting, and, more interestingly, new theorems proven thanks to the insight provided by the coarse viewpoint. We will say a few words about the former and insist on the latter.

Coarse Ricci curvature does not keep too many topological or metric properties from the manifold case: this can be seen from the discrete examples listed above. For instance, the Bonnet–Myers theorem is true only in a very weak sense [Oll09]. The kind of properties kept from manifolds relate mostly to functional inequalities or *concentration of measure* inequalities, i.e. various inequalities involving measures.

A metric space satisfies a concentration of measure property if every real-valued 1-Lipschitz function on the space is “almost constant”, that is to say, if there exists a value such that the function is close to that value on most of the space. The basic example is that of the function “proportion of tails” on the space of length- N sequences of heads and tails: it is very close to $1/2$ on most sequences, with fluctuations of order $1/\sqrt{N}$. Concentration of measure asserts that this is not a property of this particular function, but of the underlying space. In this case the underlying space is the hypercube $\{H, T\}^N$, and the Lipschitz functions are those for which the influence of any single symbol in the sequence on the value of the function is at most $1/N$; any such function will be almost constant. We refer to [Led01] for more on this topic, or to [Gro99] for a geometric viewpoint.

Thus the hypercube exhibits this concentration phenomenon. A theorem of Gromov [Gro86] implies that manifolds with positive Ricci curvature also share this property, using a method very different from that used for the hypercube. It is shown in [Oll09] that concentration follows from positive coarse Ricci curvature. This allows to put both examples in a common framework with a unified and simple proof (though with some small loss in the numerical constants).

Last, let us give examples of new theorems proven using the technique or intuition of coarse Ricci curvature. Some deal with non-Riemannian settings such as waiting queues or statistical physics [JO10], but it is interesting that new results have been obtained even in the original Riemannian case, as follows.

First, let us give a result for the empirical average of a function over a trajectory of the Brownian motion. When time grows, this empirical average converges to the average of the function over the manifold, since the invariant law of Brownian motion is uniform. The following is a concentration result for the variations of this quantity around its mean.

THEOREM 2.8 ([JO10]). *Let $(Y_t)_{t \geq 0}$ be a trajectory of the Brownian motion starting at a fixed point x_0 on a smooth, compact N -dimensional Riemannian manifold X with Ricci curvature at least $K > 0$. Let*

$$\bar{f}_t = \frac{1}{t} \int_{s=0}^t f(Y_s) ds$$

be the empirical mean of the 1-Lipschitz function $f : X \rightarrow \mathbb{R}$ under the diffusion process (Y_t) starting at some point $x \in X$. Then, for any fixed x_0 one has

$$\Pr(|\bar{f}_t - \mathbb{E}\bar{f}_t| > r) \leq 2e^{-\frac{tK^2r^2}{32}}.$$

[GLWY09] contains related results and a beautiful characterization of this kind of property in terms of transportation-information inequalities.

Our second example deals with the spectral gap of a manifold. Let Δ be the Laplace(–Beltrami) operator on the compact Riemannian manifold X . It is a negative operator on L^2 ; its spectral gap λ_1 is the smallest non-zero eigenvalue of $-\Delta$. Alternatively, λ_1 can be defined as the best constant in either one of those inequalities: the norm of the Poisson equation operator $\|\Delta^{-1}f\|_{L^2} \leq \frac{1}{\lambda_1} \|f\|_{L^2}$, the rate of smoothing by the heat equation $\|f_t\|_{L^2} \leq C_f e^{-\lambda_1 t} \|f_0\|_{L^2}$ with f_t solving $\partial f_t / \partial t = \Delta f_t$, or the Poincaré(–Sobolev) inequality $\|f\|_{L^2}^2 \leq \frac{1}{\lambda_1} \|\nabla f\|_{L^2}^2$, all these inequalities applying for functions in L^2 whose integral over X is 0. Thus, for instance, $1/\lambda_1$ is the characteristic decay time for the heat equation.

A classical theorem of Lichnerowicz (see e.g. [Ber03]) asserts that if X is an N -dimensional manifold with positive Ricci curvature, then $1/\lambda_1$ is controlled by the largest radius of curvature:

$$\frac{1}{\lambda_1} \leq \frac{N-1}{N} \sup_{x \in X} \frac{1}{\text{Ric}(x)}$$

where $\text{Ric}(x)$ is the infimum of $\text{Ric}(v, v)$ for v a unit tangent vector at $x \in X$. This was improved by Veysseire as follows.

THEOREM 2.9 ([Vey10, Vey-a]). *Let X be a smooth, compact N -dimensional manifold with positive Ricci curvature and let λ_1 be the spectral gap of the Laplace operator on X . Then*

$$\frac{1}{\lambda_1} \leq \int_{x \in X} \frac{1}{\text{Ric}(x)} \frac{d \text{vol}(x)}{\text{vol}(X)}$$

Thus we get a harmonic mean of Ricci curvature over the manifold, instead of the worst value. A slightly more complicated formula in [Vey10, Vey-a] allows to recover the Lichnerowicz theorem with the $(N-1)/N$ factor.

The proof given in [Vey10] is very short and purely computational; it was found afterwards. The original intuition [Vey-a] was to couple two Brownian motions starting at very close points and using curvature, an approach used in [Ken86] (see [ACT11] for a new elegant viewpoint). If Brownian motion is seen as a large number of small steps, using parallel transport of the trajectory (x_t) starting at x provides a Brownian trajectory (y_t) starting at y close to x ; since, at each step, the distance between x_t and y_t decreases by a factor depending on Ricci curvature, the distance after time t decreases exponentially fast: Ricci curvature is a kind of “Lyapunov exponent along Brownian trajectories”. This is turned into a spectral gap estimate using the correspondance between Brownian motion and the heat kernel. The new idea here is that the exponential rate at which the trajectories

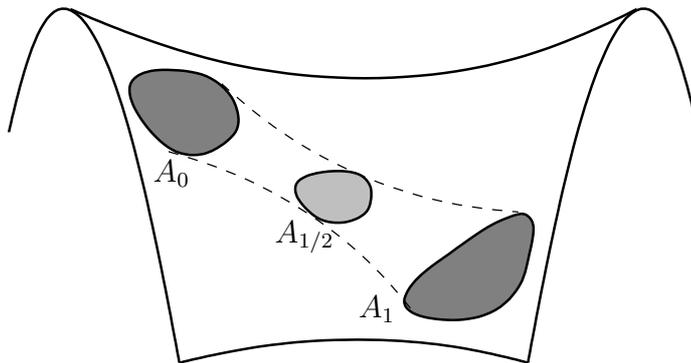


FIGURE 11. In negative curvature, midpoints contract.

(x_t) and (y_t) get closer only depends on curvature *at the current point* (x_t) , and that, since x_t covers the whole manifold with uniform measure, an averaging effect over the value of curvature should occur, instead of using only the worst curvature. This proof is longer than the short computational proof in [Vey10], but may be more adapted to generalization, e.g. for discrete settings.

Finally, let us mention that [Oll10] lists a number of questions about coarse Ricci curvature. Some have been solved (e.g., the relationship with Alexandrov curvature), but most remain open.

2.3. Non-smooth Ricci curvature: the displacement convexity of entropy approach. The other main approach to Ricci curvature for non-smooth spaces is *displacement convexity of entropy*. It is introduced in [Stu06] and [LV09], following a characterization of Ricci curvature for manifolds presented in [RS05] (elaborating on [OV00, JKO98, McC97]). It is also based on the use of Wasserstein distances.

This approach allows to define a property $CD(K, N)$ which, for Riemannian manifolds, is equivalent to having Ricci curvature at least K and dimension at most N . We will only present here the case $N = \infty$; the case $N < \infty$ involves very precise metric comparisons with N -dimensional manifolds. The property $CD(K, N)$ in this sense will be an extension of Bakry–Émery’s $CD(K, N)$ property [BE84, BE85], though from a different viewpoint.

The idea is that, in positive Ricci curvature, “midpoints spread out”: if we take two sets or measures in the manifold X , and consider the set of points that lie “halfway” between the two sets (in a sense to be made precise below), then the set of midpoints is wider than expected from the Euclidean case. For instance, on a sphere, for any reasonable notion of what a midpoint is, the set of midpoints of the two poles will be the whole equator. The reverse is true in negative Ricci curvature. This intuition, illustrated in Figures 11 and 12, was formalized in various ways in [CMS01, RS05, CMS06].

Once more, we will use probability measures instead of sets. So let A_0 and A_1 be two probability measures in the geodesic space X . The extent to which they are spread will be evaluated using the *relative entropy* (or *Kullback–Leibler divergence*) with respect to the uniform measure μ on X (μ can actually be any

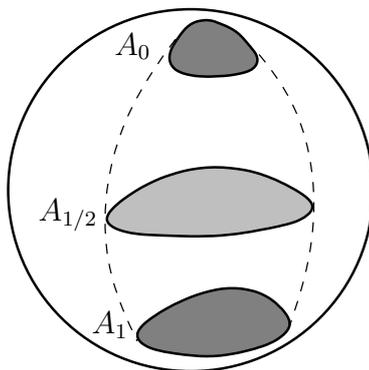


FIGURE 12. In positive curvature, midpoints spread out.

reference measure of particular interest; in general there is no notion of “uniform measure” on arbitrary spaces). Namely, for a probability measure A define

$$H(A|\mu) = \int A(dx) \log \frac{A(dx)}{\mu(dx)}$$

assuming the density $A(dx)/\mu(dx)$ exists and the integral makes sense (and $+\infty$ otherwise). This quantity is well-known in information theory [CT91]; the lower it is, the more “uniform” or “spread out” the measure A is (w.r.t. the reference measure μ).

Next, given two probability measures A_0 and A_1 , we need to define the “measure of midpoints” $A_{1/2}$ of A_0 and A_1 . We will define $A_{1/2}$ as a measure that lies midway between A_0 and A_1 in L^2 Wasserstein distance. The latter is defined like the L^1 Wasserstein distance introduced above [Vil03].

DEFINITION 2.10. Let μ_1, μ_2 be two measures with mass 1 in a metric space (X, d) . The L^2 transportation distance (or L^2 Wasserstein distance) between μ_1 and μ_2 , denoted $W_2(\mu_1, \mu_2)$, is the best average travelling distance that can be achieved:

$$W_2(\mu_1, \mu_2) = \inf_{\xi \in \Pi(\mu_1, \mu_2)} \sqrt{\iint d(x, y)^2 d\xi(x, y)}$$

where $\Pi(\mu_1, \mu_2)$ is the set of transference plans from μ_1 to μ_2 , as introduced in Definition 2.4.

Assume for simplicity that the geodesic space X is compact. Then, given any two probability measures A_0 and A_1 , there exists a (not necessarily unique) probability measure $A_{1/2}$ such that

$$W_2(A_0, A_{1/2}) = W_2(A_{1/2}, A_1) = \frac{1}{2} W_2(A_0, A_1)$$

where W_2 is the L^2 Wasserstein distance. We will call such an $A_{1/2}$ a *midpoint measure* of A_0 and A_1 .

Curvature will be positive when the relative entropy of a midpoint measure $A_{1/2}$ is smaller than the average of the relative entropies of A_0 and A_1 . The difference is used to define (a lower bound on) curvature.

DEFINITION 2.11. Let (X, d) be a metric space and μ be a reference measure on X , with $\mu(X)$ finite. Then (X, d, μ) satisfies the $\text{CD}(K, \infty)$ if, for any two probability measures A_0 and A_1 with finite entropy with respect to μ , there exists a midpoint measure $A_{1/2}$ of A_0 and A_1 such that

$$H(A_{1/2}|\mu) \leq \frac{1}{2} (H(A_0|\mu) + H(A_1|\mu)) - \frac{K}{8} W_2(A_0, A_1)^2$$

The reason for the name *displacement convexity of entropy* [McC97] is clear: the inequality above states that entropy is a convex function on the space of probability measures, where the convex interpolation between two measures is obtained by transportation. (Actually, we used convexity only for *one* midpoint measure, not for all of them, hence the qualification of *weak* convexity sometimes encountered.)

The restriction that $\mu(X)$ is finite is not essential; it only makes for a slightly less technical definition. The $K/8$ scaling factor reproduces what happens in Riemannian manifolds with Ricci curvature equal to K .

This defines only a global lower bound for curvature; but this definition can be localized by taking A_0 and A_1 in the neighborhood of some point. Then, under various technical assumptions, the global lower bound on curvature is the infimum of the local lower bounds (Theorem I.4.17 in [Stu06], extended in [BS10]). This is analogous to the property mentioned above, that it is enough to check coarse Ricci curvature $\kappa(x, y)$ for close pairs of points x and y .

Let us describe examples of spaces known to have Ricci curvature bounds in this sense.

First, this definition is a consistent extension of Ricci curvature: For Riemannian manifolds with μ the Riemannian volume measure, the $\text{CD}(K, \infty)$ condition is satisfied exactly when Ricci curvature is at least K ([Stu06], Theorem I.4.9), and likewise for the refined $\text{CD}(K, N)$ curvature-dimension condition.

Second, spaces with a lower bound on sectional curvature in the sense of Alexandrov (Definition 2.1) have a corresponding lower bound on Ricci curvature in the present sense [Pet, ZZ10].

Another source of examples are Gromov–Hausdorff limits [Gro99, BBI01] of Riemannian manifolds with positive Ricci curvature. These limits can have singularities or exhibit collapsing (reduction of dimension) of some parts. One of the main results here is that $\text{CD}(K, \infty)$ and $\text{CD}(K, N)$ lower bounds on curvature are preserved when passing to these limits [Stu06, LV09]. Actually, the study of the phenomena arising when taking Gromov–Hausdorff limits of manifolds was a strong motivation for the search of generalized notions of Ricci curvature [Lot07], especially in view of Gromov’s precompactness theorem (see e.g. [Gro99]) which states that the set of all Riemannian manifolds of a given dimension, with diameter bounded above and Ricci curvature bounded below, is precompact in Gromov–Hausdorff topology. This line of thought led to sustained research, for example a series of results by Cheeger and Colding [CC]. An elegant way to incorporate the measured aspect in these limits is provided by a “Gromov–Wasserstein” distance [Stu06].

Displacement convexity of entropy has also been successfully applied to manifolds equipped with non-Riemannian metrics (where an arbitrary non-Euclidean norm is used on each tangent space), i.e. Finsler manifolds [Oht09b, OS09, OS],

allowing to prove new results such as Poincaré-like inequalities on some of these spaces.

This approach has also been conducted on the Heisenberg group [Jui09]: no Ricci curvature lower bound $\text{CD}(K, N)$ with $K > -\infty$ holds in this space. On the other hand, in [Jui09] it is shown that a related but weaker property holds, the *measure contraction property* [Stu06, Oht07].

Let us now quickly mention some of the theorems from Riemannian manifolds that carry out to spaces with positive Ricci curvature in the present sense. We refer to [Stu06, LV09, Vil08] for more details.

The condition $\text{CD}(K, \infty)$ entails several powerful functional inequalities, in particular the logarithmic Sobolev inequality, which in turn implies other properties such as Gaussian concentration of measure (as discussed above) or spectral gap (Poincaré) inequalities.

The condition $\text{CD}(K, N)$ is stronger and closer to the Riemannian situation. It implies, for instance, the Bishop–Gromov volume growth inequality, which is essentially a restatement of the definition of $\text{CD}(K, N)$ with the sets A_0 and A_1 taken to be concentric balls. The Bonnet–Myers theorem also holds (keep in mind, though, that the definition of $\text{CD}(K, N)$ is precisely modeled after the Riemannian situation, and explicitly uses functions such as $\sin(d(x, y)\sqrt{K/(N-1)})$). The Brunn–Minkowski inequality (similar to the pictures above, but with a larger set of midpoints $A_{1/2}$) is another consequence. This illustrates the power implicit in the $\text{CD}(K, N)$ condition, for various choices of A_0 and A_1 ; the difficult part, of course, being to prove that the condition holds.

The constants obtained in these various inequalities are often sharp compared to the Riemannian case; for comparison, using coarse Ricci curvature leads to a numerical loss in the constants (even if the order of magnitude stays the same).

It is impossible to mention the displacement convexity approach without hinting at the deep context from which it arose, especially, the important relationship between the heat equation and entropy on metric spaces. The fundamental idea is that “the heat equation is the best way to increase entropy while moving the particles as least as possible” [JKO98]. More precisely, if a measure μ_t on the base space X evolves according to the heat equation $\partial\mu_t/\partial t = \Delta\mu_t$, this evolution can be thought of in the space of all measures on X in the following way: start with μ_0 , consider the entropy function around μ_0 , and follow the gradient of this entropy function. Here, to define gradients, the (infinite-dimensional) space of measures is equipped with a formally Riemannian metric for which the distance between two very close measures is defined as their W_2 transportation distance [JKO98, OV00]. This approach can be used to *define* a heat equation in metric measured spaces [AGS05, Gig10] and has been very successful.

Let us finally discuss displacement convexity of entropy in relation with coarse Ricci curvature.

First, there is no analogue to $\text{CD}(K, N)$ with $N < \infty$ for coarse Ricci curvature; the latter appears more comparable to $\text{CD}(K, \infty)$.

Second, the CD condition was not designed for discrete spaces; moreover good technical behavior of the CD condition requires that the geodesic space be *non-branching*, i.e. that two geodesics cannot coincide for some time then part away, as

typically happens in graphs. On the other hand, coarse Ricci curvature bypasses any local topological structure. An easy way to discretize displacement convexity of entropy is to use *approximate* midpoints for the set $A_{1/2}$, instead of exact midpoints. This approach yields partial results: in [BS10] it is applied to homogeneous planar graphs, and in [OV] to the discrete hypercube. The case of the discrete hypercube is not fully solved: only a weaker property, the Brunn–Minkowski inequality, is obtained in [OV] (though intriguing possible connections with combinatorics appear).

Another recent and promising approach to displacement convexity of entropy in discrete spaces has been proposed independently in [Maa11, CHLZ, EM, Mie]: these authors introduce a new metric on the space of probability measures on a discrete space, such that the [JKO98] interpretation of the heat equation holds, namely, such that a given Markov kernel on the underlying space coincides with the gradient of the entropy functional on the space of probability measures. Such an approach gives positive curvature to the discrete hypercube [EM], as expected. The functional inequalities proven in [EM] under discrete positive curvature in this sense are comparable to those obtained from $\text{CD}(K, \infty)$: logarithmic Sobolev inequalities and Gaussian concentration of measure, for instance, with optimal numerical constants recovered for the hypercube.

Coarse Ricci curvature, on the other hand, is based on a simpler criterion and may apply in a wider range of situations. For instance, no reversibility assumption is needed. More, it allows for non-Gaussian concentration: discrete examples in significant applications exhibit, e.g., Poissonian behavior. This is not compatible with $\text{CD}(K, \infty)$.

Another situation in which $\text{CD}(K, \infty)$ and coarse Ricci curvature differ is Finsler spaces. For instance, consider \mathbb{R}^n equipped with a non-Euclidean norm, e.g. the ℓ^1 norm. It is immediate to check that such a space has vanishing coarse Ricci curvature, by translation invariance. These spaces are also known to satisfy the $\text{CD}(0, \infty)$ property; yet the natural heat equation defined by the [JKO98] interpretation is nonlinear (hence non-Markovian, so that the setting of coarse Ricci curvature does not apply to it), and does not satisfy the contraction in transportation distance appearing in the definition of coarse Ricci curvature [OS09, OS].

Finally, let us mention three variants around the $\text{CD}(K, \infty)$ condition, each with independent interest. The nineteenth-century-old *Brunn–Minkowski inequality* (see e.g. [Gar02] for an introduction) can be seen in geodesic spaces as a weaker version of $\text{CD}(K, N)$, using a larger set of midpoints $A_{1/2}$; the equivalence with Ricci curvature in the Riemannian case is proven in [CMS06, CMS01]. As a general property of geodesic spaces, it is weaker than $\text{CD}(K, N)$ but already has a number of the same geometrical consequences [Stu06]. It has been studied in, e.g., Finsler spaces [Oht09b], discretizes well [Bon09] and is known to hold in the discrete hypercube [OV]. See also [Bac10, Hil] for a functional inequality version. The property $\text{RCD}(K, \infty)$, introduced in [AGS], is a reinforcement both of $\text{CD}(K, \infty)$ and of coarse Ricci curvature. It allows for an elegant control of fine properties of the heat equation and related problems. The *measure contraction property* [Stu06, Oht07] is weaker than the CD condition and can be thought of as the particular case when the measure A_0 in the pictures above is a Dirac mass and A_1 is the uniform measure on the whole space. Any of those properties is yet another reasonable contender for the title of generalized Ricci curvature.

This closes this tour of some possible generalizations of sectional and Ricci curvature, and their connections to various other areas of mathematics. The field is still very active.

References

- [ACT11] M. Arnaudon, K. A. Coulibaly, A. Thalmaier, *Horizontal diffusion in C^1 path space*, Séminaire de Probabilités XLIII, Lecture Notes in Math. **2006** (2011), 73–94.
- [AGS05] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.
- [AGS] L. Ambrosio, N. Gigli, G. Savaré, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, preprint, [arXiv:1109.0222](https://arxiv.org/abs/1109.0222)
- [Bac10] K. Bacher, *On Borell-Brascamp-Lieb inequalities on metric measure spaces*, Potential Anal. **33** (2010), n° 1, 1–15.
- [BBI01] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics **33**, AMS (2001).
- [BD97] R. Bubley, M. E. Dyer, *Path coupling: a technique for proving rapid mixing in Markov chains*, FOCS 1997, 223–231.
- [Ber03] M. Berger, *A panoramic view of Riemannian geometry*, Springer, Berlin (2003).
- [BE84] D. Bakry, M. Émery, *Hypercontractivité de semi-groupes de diffusion*, C. R. Acad. Sci. Paris Sér. I Math. **299** (1984), n° 15, 775–778.
- [BE85] D. Bakry, M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84. Lecture Notes in Math. **1123**, Springer, Berlin (1985), 177–206.
- [BH99] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der mathematischen Wissenschaften **319**, Springer (1999).
- [Bon09] M. Bonnefont, *A discrete version of the Brunn–Minkowski inequality and its stability*, Ann. Math. Blaise Pascal **16** (2009), no. 2, 245–257.
- [BS09] A.-I. Bonciocat, K.-T. Sturm, *Mass Transportation and rough curvature bounds for discrete spaces*, J. Funct. Anal. **256** (2009), n° 9, 2944–2966.
- [BS10] K. Bacher, K.-T. Sturm, *Localization and tensorization properties of the curvature-dimension condition for metric measure spaces*, J. Funct. Anal. **259** (2010), n° 1, 28–56.
- [CC] J. Cheeger, T. H. Colding, *On the structure of spaces with Ricci curvature bounded below, I, II, III*, J. Differential Geom. **46** (1997), 37–74; *ibid.* **54** (2000), 13–35; *ibid.* **54** (2000), 37–74.
- [Che04] M.-F. Chen, *From Markov chains to non-equilibrium particle systems*, second edition, World Scientific (2004).
- [CHLZ] S.-N. Chow, W. Huang, Y. Li, H. Zhou, *Fokker-Planck equations for a free energy functional or Markov process on a graph*, preprint.
- [CMS01] D. Cordero-Erausquin, R.J. McCann, M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math. **146** (2001), 219–257.
- [CMS06] D. Cordero-Erausquin, R.J. McCann, M. Schmuckenschläger, *Prékopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields and optimal transport*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), no. 4, 613–635.
- [CT91] T. M. Cover, J. A. Thomas, *Elements of information theory*, Wiley (1991).
- [CW94] M.-F. Chen, F.-Y. Wang, *Application of coupling method to the first eigenvalue on manifold*, Sci. China Ser. A **37** (1994), n° 1, 1–14.
- [DC92] M. P. do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser (1992).
- [DGW04] H. Djellout, A. Guillin, L. Wu, *Transportation cost–information inequalities and applications to random dynamical systems and diffusions*, Ann. Prob. **32** (2004), n° 3B, 2702–2732.
- [Dob70] R. L. Dobrušin, *Definition of a system of random variables by means of conditional distributions* (Russian), Teor. Veroyatnost. i Primenen. **15** (1970), 469–497. English translation: *Prescribing a system of random variables by conditional expectations*, Theory of Probability and its Applications **15** (1970) n° 3, 458–486.

- [Dob96] R. Dobrushin, *Perturbation methods of the theory of Gibbsian fields*, in R. Dobrushin, P. Groeneboom, M. Ledoux, *Lectures on probability theory and statistics*, Lectures from the 24th Saint-Flour Summer School held July 7–23, 1994, edited by P. Bernard, Lecture Notes in Mathematics **1648**, Springer, Berlin (1996), 1–66.
- [DS85] R. L. Dobrushin, S. B. Shlosman, *Constructive criterion for the uniqueness of Gibbs field*, in J. Fritz, A. Jaffe and D. Szász (eds), *Statistical physics and dynamical systems*, papers from the second colloquium and workshop on random fields: rigorous results in statistical mechanics, held in Kőszeg, August 26–September 1, 1984, Progress in Physics **10**, Birkhäuser, Boston (1985), 347–370.
- [EM] M. Erbar, J. Maas, *Ricci curvature of finite Markov chains via convexity of the entropy*, preprint, [arXiv:1111.2687](https://arxiv.org/abs/1111.2687)
- [FSS10] K.-T. Sturm, S. Fang, J. Shao, *Wasserstein spaces over Wiener spaces*, Probab. Theory Related Fields **146** (2010), n° 3–4, 535–565.
- [Gar02] R. J. Gardner, *The Brunn–Minkowski inequality*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 3, 355–405.
- [Ger10] L. Gerin, *Random sampling of lattice paths with constraints, via transportation*, Proc. AofA’10, DMTCS Proceedings, vol.AM (2010), 317–328.
- [Gig10] N. Gigli, *On the heat flow on metric measure spaces: existence, uniqueness and stability*, Calc. Var. Partial Differential Equations **39** (2010), 101–120.
- [GKO] N. Gigli, K. Kuwada, S.-i. Ohta, *Heat flow on Alexandrov spaces*, preprint, [arXiv:1008.1319](https://arxiv.org/abs/1008.1319)
- [GLWY09] A. Guillin, C. Léonard, L. Wu, N. Yao, *Transportation-information inequalities for Markov processes*, Probab. Theory Related Fields **144** (2009), n° 3–4, 669–695.
- [Gro86] M. Gromov, *Isoperimetric inequalities in Riemannian manifolds*, in V. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Mathematics **1200**, Springer, Berlin (1986), 114–129.
- [Gro87] M. Gromov, *Hyperbolic groups*, in *Essays in group theory*, ed. S. M. Gersten, Springer (1987), 75–265.
- [Gro91] M. Gromov, *Sign and geometric meaning of curvature*, Rend. Sem. Mat. Fis. Milano **61** (1991), 9–123 (1994).
- [Gro99] M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Progress in Math. **152**, Birkhäuser (1999).
- [Hil] E. Hillion, *On Prekopa-Leindler inequalities on metric-measure spaces*, preprint, [arXiv:0912.3593](https://arxiv.org/abs/0912.3593)
- [JKO98] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal. **29** (1998), n° 1, 1–17.
- [Jou07] A. Joulin, *Poisson-type deviation inequalities for curved continuous time Markov chains*, Bernoulli **13** (2007), n°3, 782–798.
- [Jou09] A. Joulin, *A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature*, Bernoulli **15** (2009), n° 2, 532–549.
- [JO10] A. Joulin, Y. Ollivier, *Curvature, concentration, and error estimates for Markov chain Monte Carlo*, Ann. Probab. **38** (2010), n° 6, 2418–2442.
- [Jui09] N. Juillet, *Geometric inequalities and generalized Ricci bounds in the Heisenberg group*, Int. Math. Res. Not. IMRN (2009), n° 13, 2347–2373.
- [Ken86] W. Kendall, *Nonnegative Ricci curvature and the Brownian coupling property*, Stochastics **19** (1986), n° 1–2, 111–129.
- [Led01] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89**, AMS (2001).
- [LPW09] D. Levin, Y. Peres, E. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence (2009).
- [Lot07] J. Lott, *Optimal transport and Ricci curvature for metric-measure spaces*, Surv. Differ. Geom. **11** (2007), 229–257.
- [LV09] J. Lott, C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), n° 3, 903–991.
- [Maa11] J. Maas, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal. **261** (2011), n° 8, 2250–2292.
- [McC97] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), 153–179.

- [Mie] A. Mielke, *Geodesic convexity of the relative entropy in reversible Markov chains*, preprint.
- [MT10] R. McCann, P. Topping, *Ricci flow, entropy and optimal transportation*, Amer. J. Math. **132** (2010), 711–730.
- [Oht07] S.-i. Ohta, *On the measure contraction property of metric measure spaces*, Comment. Math. Helv. **82** (2007), 805–828.
- [Oht09a] S.-i. Ohta, *Gradient flows on Wasserstein spaces over compact Alexandrov spaces*, Amer. J. Math. **131** (2009), 475–516.
- [Oht09b] S.-i. Ohta, *Finsler interpolation inequalities*, Calc. Var. Partial Differential Equations **36** (2009), 211–249.
- [Oli09] R. Imbuzeiro Oliveira, *On the convergence to equilibrium of Kac’s random walk on matrices*, Ann. Appl. Probab. **19** (2009), n° 3, 1200–1231.
- [Oll05] Y. Ollivier, *A January 2005 invitation to random groups*, Ensaios Matemáticos **10**, Sociedade Brasileira de Matemática, Rio de Janeiro (2005).
- [Oll07] Y. Ollivier, *Ricci curvature of metric spaces*, C. R. Math. Acad. Sci. Paris **345** (2007), n° 11, 643–646.
- [Oll09] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), n° 3, 810–864.
- [Oll10] Y. Ollivier, *A survey of Ricci curvature for metric spaces and Markov chains*, in *Probabilistic approach to geometry*, Adv. Stud. Pure Math. **57**, Math. Soc. Japan (2010), 343–381.
- [OS09] S.-i. Ohta, K.-T. Sturm, *Heat flow on Finsler manifolds*, Comm. Pure. Appl. Math. **62** (2009), 1386–1433.
- [OS] S.-i. Ohta, K.-T. Sturm, *Non-contraction of heat flow on Minkowski spaces*, preprint, [arXiv:1008.1319](https://arxiv.org/abs/1008.1319)
- [OV00] F. Otto, C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), 361–400.
- [OV] Y. Ollivier, C. Villani, *A curved Brunn–Minkowski inequality on the discrete hypercube*, preprint, [arXiv:1011.4779](https://arxiv.org/abs/1011.4779)
- [Pet] A. Petrunin, *Alexandrov meets Lott–Villani–Sturm*, preprint, [arXiv:1003.5948](https://arxiv.org/abs/1003.5948)
- [RS05] M.-K. von Renesse, K.-T. Sturm, *Transport inequalities, gradient estimates, and Ricci curvature*, Comm. Pure Appl. Math. **68** (2005), 923–940.
- [Stu06] K.-T. Sturm, *On the geometry of metric measure spaces*, Acta Math. **196** (2006), n° 1, 65–177.
- [Vey10] L. Veysseire, *A harmonic mean bound for the spectral gap of the Laplacian on Riemannian manifolds*, C. R. Acad. Sci. Paris, Ser. I **348** (2010), 1319–1322.
- [Vey-a] L. Veysseire, *Improved spectral gap bounds on positively curved manifolds*, preprint, [arXiv:1105.6080](https://arxiv.org/abs/1105.6080)
- [Vey-b] L. Veysseire, *Coarse Ricci curvature for continuous-time Markov processes*, preprint, [arXiv:1202.0420](https://arxiv.org/abs/1202.0420)
- [Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics **58**, American Mathematical Society, Providence (2003).
- [Vil08] C. Villani, *Optimal transport, old and new*, Grundlehren der mathematischen Wissenschaften **338**, Springer (2008).
- [Wil04] D. B. Wilson, *Mixing times of Lozenge tiling and card shuffling Markov chains*, Ann. Appl. Probab. **14** (2004), n° 1, 274–325.
- [ZZ10] H.-C. Zhang, X.-P. Zhu, *Ricci curvature on Alexandrov spaces and rigidity theorems*, Comm. Anal. Geom. **18** (2010), n° 3, 503–553.