Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 345 (2007) 643–646
Comptes Rendus Mathematique
http://france.elsevier.com/direct/CRASS1/

Differential Geometry/Probability Theory

Ricci curvature of metric spaces

Yann Ollivier

CNRS, UMPA, École normale supérieure de Lyon, 46, allée d'Italie, 69007 Lyon, France

Received 21 September 2007; accepted after revision 6 October 2007

Presented by Étienne Ghys

Abstract

We define a notion of Ricci curvature in metric spaces equipped with a measure or a random walk. For this we use a local contraction coefficient of the random walk acting on the space of probability measures equipped with a transportation distance. This notions allows to generalize several classical theorems associated with positive Ricci curvature, such as a spectral gap bound (Lichnerowicz theorem), Gaussian concentration of measure (Lévy–Gromov theorem), logarithmic Sobolev inequalities (a result of Bakry–Émery theory) or the Bonnet–Myers theorem. The definition is compatible with Bakry–Émery theory, and is robust and very easy to implement in concrete examples such as graphs. ***To cite this article: Y. Ollivier, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Résumé

Courbure de Ricci des espaces métriques. Nous définissons la courbure de Ricci d'un espace métrique muni d'une mesure ou d'une marche aléatoire. Notre outil est un coefficient de contraction local de la marche aléatoire agissant sur l'espace des mesures de probabilités muni d'une distance de transport. Nous pouvons ainsi généraliser des résultats classiques en courbure de Ricci minorée, comme la borne sur le trou spectral (théorème de Lichnerowicz), la concentration gaussienne de la mesure (théorème de Lévy–Gromov), l'inégalité de Sobolev logarithmique (conséquence de la théorie de Bakry–Émery) ou le théorème de Bonnet–Myers. Notre définition est compatible avec la théorie de Bakry–Émery, est robuste, et très simple à mettre en œuvre concrètement, par exemple sur un graphe. ***Pour citer cet article : Y. Ollivier, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

This Note presents some of the results from our preprint [14].

In Riemannian geometry, the natural framework for the study of spaces with positive curvature seems to be a lower bound on Ricci curvature (see e.g. the survey [9]). Positive Ricci curvature is also very relevant from a probabilistic or analytic point of view, as illustrated by the works of Gromov [6] and Bakry and Émery [2,3] on concentration of measure, contractivity properties of the heat equation and logarithmic Sobolev inequalities.

Some discrete spaces, such as the hypercube $\{0,1\}^N$ equipped with a ℓ^1 metric, seem to share several properties of the sphere S^N (the reference positively curved space), e.g. with respect to concentration of measure or logarithmic

E-mail address: yann.ollivier@umpa.ens-lyon.fr.

1631-073X/$ – see front matter
doi:10.1016/j.crma.2007.10.041

Sobolev inequalities [8]. Also keeping in mind the Dvoretzky theorem which compares convex bodies to a sphere, it is natural to wonder whether such spaces could be said to have positive Ricci curvature in some sense.

Several attempts have been made to generalize the Bakry–Émery approach, but to quote from [4], "*unfortunately, this technique seems useless in the finite setting*". Let us mention the convergent works of Sturm [16], Lott–Villani [10] and Ohta [12] (see also [15]), based on analysis along geodesics in the space of probability measures. However, their definition is difficult to check on examples, and, as it is infinitesimal, cannot be applied to discrete spaces such as graphs.

The definition presented here relies less on the infinitesimal structure of the space, and depends on the metric and measure in a more direct way, through the use of transportation distance between 'balls'. As it happens, somewhat related ideas have been used for a long time in the study of spin systems (Dobrushin criterion, see the review [11] or the very end of [5]) or similar product spaces [1], without the link with curvature. See also [7] and [13] for recent works developing ideas close to ours.

The definition. In Riemannian geometry, Ricci curvature can be apprehended as follows. Let x be a point in a Riemannian manifold, and let v be a unit tangent vector at x; follow v for a very small distance δ, to end up at a point y. Now let w be a second tangent vector at x, and translate it parallelly along v to get a tangent vector w_y at y. Now follow the vectors w and w_y for some small distance ε, to end up at points x' and y' respectively. Because of curvature, the distance between x' and y' is generally not the same as that from x to y: if curvature in the plane (v, w) is positive, it will be smaller, and larger if curvature is negative. Ricci curvature along v is this phenomenon, averaged on all directions w at x.

In other words, consider the sphere (or ball) of radius ε around x, and use parallel transport to move it onto the sphere of radius ε around y; we have just seen that if Ricci curvature along v is positive, then on average, points of the sphere travel less than the distance from x to y, and the difference gives the value of Ricci curvature. This is made precise below through the use of transportation distances. We refer to [17] for the definition and intuition behind transportation distances.

In our general setting, we will use an arbitrary probability measure around x as a replacement for the ε-ball centered at x, hence the following definition:

Definition 1. Let (X, d) be a Polish metric space, equipped with its Borel σ-field. Let $(m_x)_{x\in X}$ be a family of probability measures on X, such that (i) the measure m_x depends measurably on $x \in X$ and (ii) for every $x \in X$, the first moment $\int d(x, y)\,\mathrm{d}m_x(y)$ is finite.

Let $x, y \in X$, $x \neq y$. The coarse Ricci curvature $\kappa(x, y)$ of $(X, d, (m_x))$ along (xy) is defined by the relation $\frac{\mathcal{T}_1(m_x,m_y)}{d(x,y)} =: 1 - \kappa(x, y)$ where $\mathcal{T}_1(m_x, m_y)$ is the L^1 transportation distance from m_x to m_y.

If the measures m_x are ε-balls centered at x in an N-dimensional Riemannian manifold, we saw above that when y tends to x along a unit tangent vector v, $\kappa(x, y)$ gives back the usual Ricci curvature $\mathrm{Ric}(v, v)$ (actually up to a multiplicative factor $\frac{\varepsilon^2}{N+2}$). More generally, if the data consists of a metric space (X, d) equipped with a locally finite measure μ, one can for example take m_x to be the measure μ restricted to the ε-ball around x (rescaled to mass 1); this will define the 'Ricci curvature of (X, d, μ) at scale ε'.

The probability measures m_x, taken as the law of a jump from x, define a random walk on X. This is consistent with the spirit of Bakry–Émery theory, in which emphasis is on the process rather than the invariant measure. For example, if on a Riemannian manifold we take m_x to be a discrete-time approximation of the law of the random process associated with some second-order elliptic operator, then $\kappa(x, y)$ gives back the Ricci–Bakry–Émery curvature of this operator, and actually in a more visual way.

It is expected for a notion of curvature to be locally testable, hence the following, which applies, e.g., to a graph with $a = 1$, or with a Riemannian manifold with any value of a:

Exercise 2. Suppose that the space (X, d) is a-geodesic, i.e., for any $x, y \in X$, there exists a sequence $x = x_0, x_1, \ldots, x_k = y$ of points in X with $d(x, y) = \sum d(x_i, x_{i+1})$ and $d(x_i, x_{i+1}) \leqslant a$. Let $\kappa \in \mathbb{R}$. If $\kappa(x, y) \geqslant \kappa$ for any $x, y \in X$ with $d(x, y) \leqslant a$, then $\kappa(x, y) \geqslant \kappa$ for any $x, y \in X$.

Examples. Examples of positively curved spaces in this sense include: Riemannian manifolds with positive Ricci curvature; discrete-time approximations of processes on manifolds with positive Ricci curvature in the sense of Bakry and Émery (e.g. the Ornstein–Uhlenbeck process associated with the Gaussian measure on $\mathbb{R}^N$); the discrete cube $\{0,1\}^N$ with its graph metric, using the counting measure on the 1-ball around x for m_x (Exercise); the discrete cube again, equipped with the random walk naturally associated with any binomial distribution; the Glauber dynamics for spin systems in high temperature (actually positive Ricci curvature is exactly equivalent to the Dobrushin criterion, or to the Dobrushin–Shlosman version if block dynamics are used); the $M/M/\infty$ queue and other waiting queues (e.g. with heterogeneous server rates), in a suitable continuous-time limit; Kac's random walk on orthogonal matrices (this is the main result of Imbuzeiro Oliveira in [13]). More examples are presented in the preprint [14].

Selected results. We now present some results from [14] that generalize, respectively, the Lichnerowicz spectral gap theorem, the Gaussian concentration of measure for Lipschitz functions (one aspect of the Lévy–Gromov theorem), and the logarithmic Sobolev inequality obtained by Bakry and Émery. When X is a Riemannian manifold and we use ε-balls to define the measures m_x, letting $\varepsilon \to 0$ allows to recover these theorems from the results below, although with some loss in the numerical constants.

First, it not difficult to see that the condition $\kappa(x,y) \geqslant \kappa > 0$ for all x, y, is equivalent to the random walk operator being a $(1-\kappa)$-contraction in the space of probability measures on X equipped with the transportation distance $\mathcal{T}_1$. Consequently, positive Ricci curvature implies the existence of a unique probability distribution invariant under the random walk, which we henceforth denote ν.

Definition 3. Let $x \in X$. The spread at x is defined as $\sigma(x) := (\frac{1}{2}\iint d(y,z)^2\,\mathrm{d}m_x(y)\,\mathrm{d}m_x(z))^{1/2}$. The *average spread* is $\sigma := \|\sigma(x)\|_{L^2(X,\nu)}$. The local dimension at x is

$$n_x := \frac{\sigma(x)^2}{\sup\{\mathrm{Var}_{m_x} f,\ f\ 1\text{-}\textit{Lipschitz}\}}.$$

For m_x the ε-ball in an N-dimensional Riemannian manifold, we have $\sigma(x) \approx \varepsilon$ and $n_x \approx N$.

The following is easy once one realizes that $\kappa(x,y) \geqslant \kappa$ is equivalent to the random walk operator sending 1-Lipschitz functions to $(1-\kappa)$-Lipschitz functions:

Proposition 4. *Suppose that $\kappa(x,y) \geqslant \kappa > 0$ for any $x, y \in X$. Suppose that ν is reversible and $\sigma < \infty$. Then the smallest eigenvalue of the discrete Laplacian of the random walk, acting on $L^2(X,\nu)/\{const\}$, is at least κ.*

The following theorem expresses that positive curvature entails Gaussian-then-exponential concentration of measure, with variance ('observable diameter') $D^2 := \mathbb{E}_\nu \frac{\sigma(x)^2}{n_x\kappa}$. Some very simple examples show that the Gaussian-exponential transition is genuine. For a Riemannian manifold, D^2 behaves like the inverse of the lower bound on Ricci curvature as expected from the Lévy–Gromov theorem, and $t_{\max} \to \infty$ if small enough balls for m_x are taken, so that the exponential regime disappears. For the discrete cube, it yields the optimal result up to numerical constants.

Theorem 5. *Suppose that $\kappa(x,y) \geqslant \kappa > 0$ for any $x, y \in X$. Let $D_x^2 := \frac{\sigma(x)^2}{n_x\kappa}$ and $D^2 := \mathbb{E}_\nu D_x^2$. Suppose that the function $x \mapsto D_x^2$ is C-Lipschitz and set $t_{\max} := \frac{D^2}{\max(\varepsilon, 2C/3)}$ where $\varepsilon := \sup_{x\in X} \operatorname{diam} \operatorname{Supp} m_x$.*

Then for any 1-Lipschitz function f, we have concentration of measure as follows:

$$\nu\big(\{x,\ f(x) \geqslant t + \mathbb{E}_\nu f\}\big) \leqslant \begin{cases} \exp\big(-t^2/6D^2\big) & \text{for } t \leqslant t_{\max}, \\ \exp\big(-t_{\max}^2/6D^2 - (t-t_{\max})/\max(3\varepsilon, 2C)\big) & \text{for } t \geqslant t_{\max}. \end{cases}$$

We now turn to the logarithmic Sobolev inequality. For this, we will need a notion of norm of the gradient valid in any metric space, which is as follows: Choose some $\lambda > 0$ and, for a function $f : X \to \mathbb{R}$ and $x \in X$, set $|\nabla_\lambda f|(x) := \sup_{y,y'\in X} \frac{|f(y)-f(y')|}{d(y,y')} \mathrm{e}^{-\lambda d(x,y) - \lambda d(y,y')}$, which is a kind of 'blurred' Lipschitz constant of f at x. For example, for a graph we will be able to take $\lambda \approx 1$, and arbitrarily large λ for a Riemannian manifold, hence recovering the usual gradient.

Theorem 6. *Suppose that $\kappa(x,y) \geqslant \kappa > 0$ for any $x, y \in X$. Then for some (explicit) $\lambda > 0$, any positive function $f : X \to \mathbb{R}$ with $|\nabla_\lambda f| < \infty$ satisfies*

$$\mathrm{Ent}_\nu f := \int f \log f \, \mathrm{d}\nu \leqslant \left(\sup_x \frac{4\sigma(x)^2}{\kappa n_x} \right) \int \frac{|\nabla_\lambda f|^2}{f} \, \mathrm{d}\nu$$

and moreover, as in Bakry–Émery theory, we have the contraction property $|\nabla_\lambda(Mf)| \leqslant (1 - \kappa/2) M(|\nabla_\lambda f|)$ where M is the discrete heat equation operator associated with the random walk.

Acknowledgements

I would like to thank Vincent Beffara, Fabrice Debbasch, Alessio Figalli, Pierre Pansu, Bruno Sévennec, Romain Tessera and Cédric Villani for inspiring conversations about coarse geometry and Ricci curvature, and Étienne Ghys for presenting this Note. Special thanks to Pierre Py for the two points x and y.

References

[1] R. Bubley, M.E. Dyer, Path Coupling: A technique for proving rapid mixing in Markov chains, in: FOCS 1997, pp. 223–231.
[2] D. Bakry, M. Émery, Hypercontractivité de semi-groupes de diffusion, C. R. Acad. Sci. Paris Sér. I Math. 299 (15) (1984) 775–778.
[3] D. Bakry, M. Émery, Diffusions hypercontractives, in: Séminaire de probabilités, XIX, 1983/84, in: Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206.
[4] P. Diaconis, L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Probab. 6 (3) (1996) 695–750.
[5] R. Dobrushin, Perturbation methods of the theory of Gibbsian fields, in: R. Dobrushin, P. Groeneboom, M. Ledoux (Eds.), Lectures on Probability Theory and Statistics Lectures from the 24th Saint-Flour Summer School held July 7–23, 1994, in: P. Bernard (Ed.), Lecture Notes in Mathematics, vol. 1648, Springer, Berlin, 1996, pp. 1–66.
[6] M. Gromov, V. Milman, G. Schechtman, Asymptotic Theory of Finite Dimensional Normed Spaces, Lecture Notes in Mathematics, vol. 1200, Springer, Berlin, 1986.
[7] A. Joulin, Poisson-type deviation inequalities for curved continuous time Markov chains, preprint.
[8] M. Ledoux, The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs, vol. 89, Amer. Math. Soc., 2001.
[9] J. Lott, Optimal transport and Ricci curvature for metric-measure spaces, expository manuscript.
[10] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, preprint.
[11] F. Martinelli, Relaxation times of Markov chains in statistical mechanics and combinatorial structures, in: H. Kesten (Ed.), Probability on Discrete Structures, in: Encyclopaedia of Mathematical Sciences, vol. 110, Springer, Berlin, 2004, pp. 175–262.
[12] S.-i. Ohta, On the measure contraction property of metric measure spaces, preprint.
[13] R. Imbuzeiro Oliveira, On the convergence to equilibrium of Kac's random walk on matrices, preprint, arXiv: 0705.2253.
[14] Y. Ollivier, Ricci curvature of Markov chains on metric spaces, preprint, arXiv: math/0701886.
[15] M.-K. von Renesse, K.-T. Sturm, Transport inequalities, gradient estimates, and Ricci curvature, Comm. Pure Appl. Math. 68 (2005) 923–940.
[16] K.-T. Sturm, On the geometry of metric measure spaces, Acta Math. 196 (1) (2006) 65–177.
[17] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, vol. 58, Amer. Math. Soc., 2003.