

Chapter 1

The Geometric Meaning of Curvature: Local and Nonlocal Aspects of Ricci Curvature

Frank Bauer, Bobo Hua, Jürgen Jost, Shiping Liu, and Guofang Wang

Abstract Curvature is a concept originally developed in differential and Riemannian geometry. There are various established notions of curvature, in particular sectional and Ricci curvature. An important theme in Riemannian geometry has been to explore the geometric and topological consequences of bounds on those curvatures, like divergence or convergence of geodesics, convexity properties of distance functions, growth of the volume of distance balls, transportation distance between such balls, vanishing theorems for Betti numbers, bounds for the eigenvalues of the Laplace operator or control of harmonic functions. Several of these geometric properties turn out to be equivalent to the corresponding curvature bounds in the context of Riemannian geometry. Since those properties often are also meaningful in the more general framework of metric geometry, in recent years, there have been several research projects that turned those properties into axiomatic definitions of curvature bounds in metric geometry. In this contribution, after developing the Riemannian geometric background, we explore some of these axiomatic approaches. In particular, we shall describe the insights in graph theory and network analysis following from the corresponding axiomatic curvature definitions.

F. Bauer

Department of Mathematics, Harvard University, Cambridge, MA, USA

Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

e-mail: bauerf80@gmail.com

B. Hua

School of Mathematical Sciences, Fudan University, Shanghai, China

e-mail: bobohua@fudan.edu.cn

J. Jost (✉)

Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

e-mail: jost@mis.mpg.de

S. Liu

School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, China

e-mail: spliu@ustc.edu.cn

G. Wang

Department of Mathematics, Freiburg University, Freiburg im Breisgau, Germany

e-mail: guofang.wang@math.uni-freiburg.de

1.1 The Origins of the Concept of Curvature

The concept of curvature was first introduced in mathematics to study curves in the plane or in space. The aim was to quantify the deviation of a curve from being a straight line, that is, how “curved” it is, in a way that was intrinsic, that is, did not depend on its parametrization. In fact, the shape of a curve in the plane is completely characterized by the—possibly varying—curvature at its points. (For a curve in three-dimensional space, an additional invariant, the torsion, is needed.) This having been relatively easily understood, mathematicians then wanted to proceed to surfaces in space. The obvious idea was to investigate the curvature of curves on the surface. In particular, through each point on such a surface, we have a one-dimensional family of orthogonal directions, and one can therefore look at suitable curves on the surface with those tangent directions. This was worked out by mathematicians like Monge and others, but in fact, this approach confused two different aspects. This was only clarified by Gauss [36] in 1827. The point is that while all curves are intrinsically alike and can only possibly differ by the way they sit in the plane or in space, surfaces possess their own intrinsic geometry. Different surfaces, regardless of how they sit in space, in general are not isometric, not even locally. That is, you cannot map a piece of one surface onto another one without stretching or squeezing it in some directions. The standard example is the distortion in maps of the surface of the earth where one projects a piece of a (roughly) spherical surface onto the flat plane and thereby necessarily distorts relative distances or angles, depending on the chosen projection scheme. On the other hand, one and the same piece of surface can sit differently in space. Here, the standard example is a sheet of paper that you can roll into a cylinder or (part of) a cone.

Gauss then disentangled these two aspects, the interior and exterior geometry of surfaces. His crucial discovery was that there exists a curvature measure, later called the Gauss curvature K , that solely depends on the interior geometry of a surface and is independent of how the surface sits in space. Another curvature measure, the mean curvature H , in contrast describes the exterior geometry, that is, how the surface is embedded or immersed in space. These curvature concepts, and what they mean for the geometry of curves and surfaces in space, are presented in [30, 47].

In this chapter, we shall only be concerned with the interior geometry of surfaces or other metric spaces. Therefore, our starting point is the Gauss curvature. It was Riemann [74] in his habilitation address in 1854 who conceived the grand picture of an intrinsic geometry of spaces of arbitrary dimension around the fundamental concept of curvature. This led to the development of Riemannian geometry. A reference is [47] which the reader is invited to consult for background, further developments and perspectives, and for proofs of the results from Riemannian geometry that we shall now discuss. For a historical commentary on the development of geometry, we refer to [74].

1.2 A Primer on Riemannian Geometry: Not Indispensable

1.2.1 Tangent Vectors and Riemannian Metrics

In this section, we work in the smooth category and assume that all objects possess all the differentiability properties that will be required for our computations. We consider a d -dimensional differentiable manifold M . Such a manifold can be locally described by coordinates taking their values in \mathbb{R}^d . These coordinates are more or less arbitrary, beyond some obvious requirements. The question then is how to switch from one coordinate system to another one. The convenient calculus for this purpose is the tensor calculus. This calculus employs some conventions:

- Einstein summation convention

$$a^i b_i := \sum_{i=1}^d a^i b_i \quad (1.1)$$

Thus, a summation sign is omitted when the same index occurs twice in a product, once as an upper and once as a lower index, with conventions about placing the indices to be described below. In particular:

- When $G = (g_{ij})_{i,j}$ is a metric tensor, the inverse metric tensor is written as $G^{-1} = (g^{ij})_{i,j}$, that is, by raising the indices. In particular

$$g^{ij} g_{jk} = \delta_k^i := \begin{cases} 1 & \text{when } i = k \\ 0 & \text{when } i \neq k. \end{cases} \quad (1.2)$$

- More generally,

$$v^i = g^{ij} v_j \text{ and } v_i = g_{ij} v^j. \quad (1.3)$$

In particular, this implies that for the Euclidean metric $g_{ij} = \delta_{ij}$, there is no difference between upper and lower indices.

A tangent vector for M at some point represented by x_0 in local coordinates x is an expression of the form

$$V = v^i \frac{\partial}{\partial x^i}; \quad (1.4)$$

this means that it operates on a function $\phi(x)$ in our local coordinates as

$$V(\phi)(x_0) = v^i \frac{\partial \phi}{\partial x^i} \Big|_{x=x_0}. \quad (1.5)$$

The tangent vectors at a point $p \in M$ form a d -dimensional vector space, called the tangent space $T_p M$ of M at p . Since we have written a tangent vector in local coordinates, the question then is how the same tangent vector is represented in different local coordinates y with $x = f(y)$. Applying here and in the sequel always the chain rule, we get

$$V = v^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}. \quad (1.6)$$

Thus, the coefficients of V in the y -coordinates are $v^i \frac{\partial y^j}{\partial x^i}$. With this transformation rule, the result of the operation of the tangent vector V on a function ϕ , $V(\phi)$, is independent of the choice of coordinates.

A vector field then is defined as $V(x) = v^i(x) \frac{\partial}{\partial x^i}$, that is, by having a tangent vector at each point of M . As indicated above, we assume here that the coefficients $v^i(x)$ are differentiable.

Returning to a single tangent vector, $V = v^i \frac{\partial}{\partial x^i}$ at some point x_0 , we consider a covector $\omega = \omega_i dx^i$ at this point as an object dual to V , with the rule

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_j^i \quad (1.7)$$

yielding

$$\omega_i dx^i \left(v^j \frac{\partial}{\partial x^j} \right) = \omega_i v^j \delta_j^i = \omega_i v^i. \quad (1.8)$$

We write this as $\omega(V)$, the application of the covector ω to the vector V , or as $V(\omega)$, the application of V to ω .

We have the transformation behavior

$$dx^i = \frac{\partial x^i}{\partial y^\alpha} dy^\alpha \quad (1.9)$$

required for the invariance of $\omega(V)$. Thus, the coefficients of ω in the y -coordinates are given by the identity

$$\omega_i dx^i = \omega_i \frac{\partial x^i}{\partial y^j} dy^j. \quad (1.10)$$

The transformation behavior of a tangent vector as in (1.6) is called contravariant, the opposite one of a covector as (1.10) covariant.

A 1-form then assigns a covector to every point in M , and thus, it is locally given as $\omega_i(x) dx^i$.

Having derived the transformation of vectors and covectors, we can then also determine the transformation rules for other tensors. A lower index always indicates

covariant, an upper one contravariant transformation. For example, the metric tensor, written as $g_{ij}dx^i \otimes dx^j$, with $g_{ij} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$ being the product of those two basis vectors, operates on pairs of tangent vectors. It therefore transforms doubly covariantly, that is, becomes

$$g_{ij}(f(y)) \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} dy^\alpha \otimes dy^\beta. \quad (1.11)$$

The metric tensor provides a Euclidean product of tangent vectors,

$$\langle V, W \rangle = g_{ij}v^i w^j \quad (1.12)$$

for $V = v^i \frac{\partial}{\partial x^i}$, $W = w^i \frac{\partial}{\partial x^i}$. In this formula, v^i and w^i transform contravariantly, while g_{ij} transforms doubly covariantly so that the product as a scalar quantity remains invariant under coordinate transformations.

Equipped with a Riemannian metric, one can introduce all the notions and carry out all the constructions that are familiar from Euclidean geometry. For instance, two vectors V, W are called orthogonal if $\langle V, W \rangle = 0$.

It is a basic property of tensors that computations can be carried out pointwise. Therefore, at a given point, we can choose our coordinates or our frame of tangent vectors conveniently. In particular, we can introduce Riemann normal coordinates at the point under considerations. In those coordinates, we have a basis $e_i = \frac{\partial}{\partial x^i}$ of tangent vectors that satisfy

$$g_{ij} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle = \delta_{ij}, \quad (1.13)$$

(with $\delta_{ij} = 1$ for $i = j$ and $= 0$ otherwise), and also

$$\Gamma_{jk}^i = 0 \text{ for all } i, j, k. \quad (1.14)$$

The Christoffel symbols Γ_{jk}^i will be defined below in (1.32).

Note, however, that (1.13) and (1.14) can only be achieved at a single point at a time, and not throughout a local neighborhood. In fact, the curvature tensor, which will be introduced below, provides the local obstruction for achieving these relations throughout some local neighborhood.

1.2.2 Differentials, Gradients, and the Laplace-Beltrami Operator

For a function ϕ , we have its differential

$$d\phi = \frac{\partial \phi}{\partial x^i} dx^i, \quad (1.15)$$

a 1-form; this depends on the differentiable structure, but not on the metric. The gradient of ϕ , however, involves the metric; it is defined as

$$\text{grad } \phi = g^{ij} \frac{\partial \phi}{\partial x^j} \frac{\partial}{\partial x^i}. \quad (1.16)$$

A characteristic property of the gradient of a function ϕ is that it is orthogonal to the level hypersurfaces $\phi \equiv c$. In fact, when $V \in T_p M$ is tangent to such a level hypersurface, it satisfies

$$V(\phi) = v^k \frac{\partial \phi}{\partial x^k} = 0. \quad (1.17)$$

When V then satisfies (1.17), we have

$$\langle \text{grad } \phi, V \rangle = g_{ik} g^{ij} \frac{\partial \phi}{\partial x^j} v^k = \frac{\partial \phi}{\partial x^k} v^k = 0, \quad (1.18)$$

that is, $\text{grad } \phi$ and V are orthogonal, indeed.

There also is a formula for the product of the gradients of two functions ϕ, ψ ,

$$\langle \text{grad } \phi, \text{grad } \psi \rangle = g_{ik} g^{ij} \frac{\partial \phi}{\partial x^j} g^{kl} \frac{\partial \psi}{\partial x^l} = g^{j\ell} \frac{\partial \phi}{\partial x^j} \frac{\partial \psi}{\partial x^\ell}. \quad (1.19)$$

Next, the divergence of a vector field $Z = Z^i \frac{\partial}{\partial x^i}$ is

$$\text{div } Z := \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^j} (\sqrt{g} Z^j) = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^j} \left(\sqrt{g} g^{ij} \left\langle Z, \frac{\partial}{\partial x^i} \right\rangle \right). \quad (1.20)$$

Geometrically, the divergence can be interpreted as the measure of the rate of change of the volume when flowing in the direction of the vector field Z .

This allows us to define the Laplace–Beltrami operator

$$\Delta f := \text{div grad } f = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^j} \left(\sqrt{g} g^{ij} \frac{\partial f}{\partial x^i} \right). \quad (1.21)$$

(Please note that the sign convention adopted here differs from that of Jost [47].)

A function $f : \Omega \rightarrow \mathbb{R}$ on some domain $\Omega \subset M$ is called *harmonic* if

$$\Delta f = 0. \quad (1.22)$$

Below, we shall introduce the volume form $d\text{vol}$ of a Riemannian metric, see (1.34) with the help of which we can then compute the L^2 -product of square integrable functions f, g as

$$(f, g) := \int f(x) g(x) d\text{vol}(x). \quad (1.23)$$

We then have

$$(\Delta f, g) = -(\text{grad} f, \text{grad} g) = (f, \Delta g) \quad (1.24)$$

for smooth functions f, g .

The eigenvalues of Δ on a compact Riemannian manifold M , that is, those λ for which there exists some nontrivial function f_λ , called an eigenfunction, with

$$\Delta f_\lambda + \lambda f_\lambda = 0, \quad (1.25)$$

contain important geometric invariants about the geometry of M . All eigenvalues are real because Δ is a symmetric operator by (1.24). The choice of sign in (1.25) is such that all eigenvalues are nonnegative. Of course, $\lambda_0 = 0$ always is an eigenvalue with a constant eigenfunction. Or putting it differently, a harmonic function is an eigenfunction for the eigenvalue 0, and on a compact M , all harmonic functions are constant by the maximum principle. When M is connected, all other eigenvalues are positive. (More generally, the multiplicity of the eigenvalue 0 equals the number of connected components of a Riemannian manifold.) The eigenvalues are usually numbered in increasing order, that is, when M is connected, they are

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \quad (1.26)$$

Of course, one may also study the spectrum of noncompact Riemannian manifolds, but in that case, the spectrum needs no longer be discrete, and we do not consider that here.

There is a more abstract and more general definition of the Laplace operator in Riemannian geometry. For a p -form $\omega = \phi(x) dx^{i_1} \wedge \dots \wedge dx^{i_p}$ with $1 \leq i_1 < \dots < i_p \leq d$, we have

$$d\omega = \sum_{j=1}^d \frac{\partial \phi(x)}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_p}. \quad (1.27)$$

We can then define the formal adjoint d^* of d w.r.t. the L^2 -product (1.23), that is,

$$(d^* f, g) = (f, dg) \quad (1.28)$$

for all functions f, g for which these expressions are well-defined, e.g., smooth with compact support. We can then define the Laplace operator on p -forms via

$$\Delta \omega = -(dd^* + d^*d)\omega. \quad (1.29)$$

A differential form ω is called *harmonic* if $\Delta \omega = 0$. On a compact Riemannian manifold, ω is harmonic if and only if

$$d\omega = 0 \text{ and } d^*\omega = 0. \quad (1.30)$$

For functions, that is, 0-forms, the definition (1.29) agrees with the earlier (1.21). On a compact Riemannian manifold, harmonic functions are constant, but in general, there exist nontrivial harmonic forms.

Details can be found in [47]. We should alert the reader to the fact that here we are using a different sign convention for Laplacians than in [47], in order to conform to usage in graph theory below.

1.2.3 Lengths and Distances

Equipped with a Riemannian metric, one can measure the length of curves. Let $[a, b]$ be a closed interval in \mathbb{R} , $\gamma : [a, b] \rightarrow M$ a (smooth) curve. The *length of γ* is defined as

$$L(\gamma) = \int_a^b \sqrt{g_{ij}(x(\gamma(t))) \dot{x}^i(t) \dot{x}^j(t)} dt. \quad (1.31)$$

$L(\gamma)$ does not depend on the parametrization of γ , that is if $\psi : [\alpha, \beta] \rightarrow [a, b]$ is a change of parameter, then

$$L(\gamma \circ \psi) = L(\gamma).$$

On a Riemannian manifold M , we can then define the *distance* between two points p, q via

$$d(p, q) := \inf\{L(\gamma) : \gamma : [a, b] \rightarrow M \text{ piecewise smooth curve with } \gamma(a) = p, \gamma(b) = q\}.$$

Any two points p, q in a *connected* Riemannian manifold can be connected by a piecewise smooth curve, and $d(p, q)$ therefore is always defined.

Shortest curves are called *geodesic*. When they are parametrized proportionally to arc length, that is, if

$$L(c_{|[t_1, t_2]}) = (t_2 - t_1)L(c) \text{ whenever } a \leq t_1 < t_2 \leq b,$$

they satisfy the following equations

$$\ddot{x}^i(t) + \Gamma_{jk}^i(x(t)) \dot{x}^j(t) \dot{x}^k(t) = 0, \quad i = 1, \dots, d \quad (1.32)$$

with the Christoffel symbols

$$\Gamma_{jk}^i = \frac{1}{2} g^{i\ell} (g_{j\ell, k} + g_{k\ell, j} - g_{jk, \ell}),$$

where

$$(g^{ij})_{i,j=1,\dots,d} = (g_{ij})^{-1} \quad (\text{i.e. } g^{i\ell} g_{\ell j} = \delta_{ij})$$

and

$$g_{j\ell,k} = \frac{\partial}{\partial x^k} g_{j\ell}.$$

In fact, even though not all solutions of (1.32) need to be globally length minimizing, they will nevertheless be called geodesics. Actually, any geodesic is locally minimizing, that is, it realizes the distances between any two sufficiently close points on it. As the example of the great circles on the sphere shows, which are geodesics, but no longer minimizing beyond a pair of antipodal points, geodesics need not globally minimize distance. (In fact, compact Riemannian manifolds like the sphere always possess closed geodesics, that is, geodesics that return to their starting point (with the same direction they were starting with) and are parametrized on the unit circle.)

We point out that a geodesic is not just a length minimizing curve, but also carries a particular parametrization, one that is proportional to arc length.

Equation (1.32) is a system of second order ODEs, and the Picard–Lindelöf theorem yields the local existence and uniqueness of a solution with prescribed initial values and derivatives, and this solution depends smoothly on the data.

1.2.4 Volumes

On a Riemannian manifold, we can not only measure lengths and distances, but also volumes. The issue of measurability can be checked in local coordinates, and so we need not address it here. When $\Omega \subset M$ then is measurable, we define its volume as

$$\text{Vol}(\Omega) := \int_{\Omega} d\text{vol}(x) := \int_{\Omega} \sqrt{g} dx \quad (1.33)$$

with the volume form

$$d\text{vol} := \sqrt{g} dx := \sqrt{\det(g_{ij})} dx. \quad (1.34)$$

Lengths of curves and volumes of sets, in particular of the distance balls

$$U(p, r) := \{q \in M : d(p, q) < r\} \text{ for } r > 0, \quad (1.35)$$

then are the relevant metric quantities. Their behavior characterizes the geometry of a Riemannian manifold. In order to control them, Riemann introduced the curvature tensor that yielded invariants with which one can control distances between points and volumes of balls.

1.3 Curvature of Riemannian Manifolds

From our presentation of Riemannian geometry via tensor calculus, every quantity seemed to depend on the choice of local coordinates. In fact, it is not difficult to see that by a suitable choice of local coordinates, one can make the metric tensor become the unit matrix at any given point p ,

$$g_{ij}(p) = \delta_{ij}. \quad (1.36)$$

Moreover, Riemann discovered that in addition, one can make also all first derivatives vanish at that point by a suitable choice of local coordinates,

$$g_{ij,k} = 0 \text{ for all } i, j, k. \quad (1.37)$$

Coordinates satisfying (1.36) and (1.37) are called (Riemann) normal coordinates. However, these relations can in general only be achieved at a single point, that is, unless we are dealing with a Euclidean metric, we cannot have (1.36) or (1.37) simultaneously for all points in some open set.

Or to put it differently, we seek quantities that can distinguish between different metric structures, like the geometry on a sphere and that of Euclidean space. Ideally, in the spirit of Riemannian geometry which works with infinitesimal quantities, such invariants should be computable at any point. The preceding tells us that such quantities cannot be constructed from the metric tensor and its first derivatives at a given point. In contrast, in general the *second* derivatives of the metric cannot be made to vanish at a given point by a suitable choice of local coordinates. The obstruction will be given by the curvature tensor. And the curvature tensor then will furnish us a set of invariants that characterizes the local, and to a certain extent also the global geometry of a Riemannian manifold.

1.3.1 Covariant Derivatives

In order to have a more invariant scheme of computation, we shall work with the covariant derivative of Levi-Civita.

Definition 1.1 The covariant derivative ∇ satisfies

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k} \text{ for all } i, j \quad (1.38)$$

and is extended to all vector fields $V = v^i \frac{\partial}{\partial x^i}$ via the product rule

$$\nabla_{\frac{\partial}{\partial x^i}} v^j \frac{\partial}{\partial x^j} = \frac{\partial v^j}{\partial x^i} \frac{\partial}{\partial x^j} + v^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}. \quad (1.39)$$

Moreover,

$$\nabla_{w^i \frac{\partial}{\partial x^i}} v^j \frac{\partial}{\partial x^j} = w^i \nabla_{\frac{\partial}{\partial x^i}} v^j \frac{\partial}{\partial x^j}. \quad (1.40)$$

With this notation, the geodesic equation becomes

$$\nabla_{\frac{d}{dt}} \frac{dc}{dt} = 0 \quad (1.41)$$

where we use the transformation rule

$$\frac{d}{dt} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}. \quad (1.42)$$

The geometric meaning is that the tangent vector is covariantly constant along the curve, or in more intuitive terms, the tangent vectors at different points are parallel to each other. In that sense, a geodesic is the Riemannian analogue of a Euclidean straight line.

1.3.2 The Curvature Tensor; Sectional and Ricci Curvature

Since $\Gamma_{ij}^k = \Gamma_{ji}^k$, we have

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} \text{ for all } i, j. \quad (1.43)$$

Higher derivatives, however, in general do not commute, and we define

Definition 1.2 The curvature operator R is defined by

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \nabla_{\frac{\partial}{\partial x^i}} \nabla_{\frac{\partial}{\partial x^j}} - \nabla_{\frac{\partial}{\partial x^j}} \nabla_{\frac{\partial}{\partial x^i}}. \quad (1.44)$$

We shall see below that this operator contains the basic invariants of a Riemannian metric. But first, we want to express it in local coordinates and define some further quantities. In local coordinates, we write

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \frac{\partial}{\partial x^\ell} = R_{\ell ij}^k \frac{\partial}{\partial x^k}. \quad (1.45)$$

We put

$$R_{k\ell ij} := g_{km} R_{\ell ij}^m,$$

i.e.¹

$$R_{k\ell ij} = \left\langle R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \frac{\partial}{\partial x^\ell}, \frac{\partial}{\partial x^k} \right\rangle. \quad (1.46)$$

When we choose Riemann normal coordinates, this becomes

$$R_{k\ell ij} = \frac{1}{2}(g_{jk,\ell i} + g_{i\ell,kj} - g_{j\ell,ki} - g_{ik,\ell j}). \quad (1.47)$$

Definition 1.3 The *sectional curvature* of the plane spanned by the (linearly independent) tangent vectors $X = \xi^i \frac{\partial}{\partial x^i}, Y = \eta^j \frac{\partial}{\partial x^j} \in T_x M$ of the Riemannian manifold M is

$$\begin{aligned} K(X \wedge Y) &:= \frac{\langle R(X, Y)Y, X \rangle}{|X \wedge Y|^2} \\ &= \frac{R_{ijkl} \xi^i \eta^j \xi^k \eta^\ell}{g_{ik} g_{jl} (\xi^i \xi^k \eta^j \eta^\ell - \xi^i \xi^j \eta^k \eta^\ell)} \\ &= \frac{R_{ijkl} \xi^i \eta^j \xi^k \eta^\ell}{(g_{ik} g_{jl} - g_{ij} g_{kl}) \xi^i \eta^j \xi^k \eta^\ell} \end{aligned} \quad (1.48)$$

$$(|X \wedge Y|^2 = \langle X, X \rangle \langle Y, Y \rangle - \langle X, Y \rangle^2).$$

Definition 1.4 The *Ricci curvature* in the direction $X = \xi^i \frac{\partial}{\partial x^i} \in T_x M$ is

$$\text{Ric}(X, X) = g^{j\ell} \left\langle R\left(X, \frac{\partial}{\partial x^j}\right) \frac{\partial}{\partial x^\ell}, X \right\rangle. \quad (1.49)$$

In local coordinates, the Ricci tensor is

$$R_{ik} = g^{j\ell} R_{k\ell ij}. \quad (1.50)$$

The Ricci tensor is symmetric,

$$R_{ik} = R_{ki}. \quad (1.51)$$

Finally, the scalar curvature is

$$R = g^{ik} R_{ik}.$$

¹The indices k and l appear in different orders on the two sides of (1.46), according to the convention of Jost [47] that made an attempt to mediate between the different conventions in use in Riemannian geometry.

The preceding quantities are the basic invariants of a Riemannian metric. For a surface, at any point there only is a single tangent plane, and the corresponding sectional curvature is nothing but the Gauss curvature of that surface. The Ricci curvature for a vector X is (up to a normalization constant $\frac{1}{d-1}$ that will therefore occur repeatedly in subsequent formulae) the average of the sectional curvatures of all the different tangent planes containing the vector X . Again, for a surface, as there is only one tangent plane containing X , it is again the Gauss curvature. In higher dimensions, of course, the sectional curvatures provide more refined invariants than their averages, the Ricci curvatures. Nevertheless, the Ricci tensor does contain a lot of information about a Riemannian metric, and in the sequel we shall be concerned with analogues of the Ricci tensor. Finally, the scalar curvature is the average (again, up to a normalization factor $\frac{1}{d}$) of the Ricci curvatures of the tangent vectors at a given point.

So, what does the curvature of a Riemannian manifold tell us about its geometry?

1.3.3 The Geometric Meaning of Sectional Curvature

A curve is just a curve and nothing else, that is, an object without any interesting intrinsic geometry. Its internal structure is trivial, in the sense that any two curves have the same internal structure. And a geodesic simply is locally a shortest connection, and since any two points in a complete Riemannian manifold can be connected by a shortest geodesic, that fact does not carry any specific information. In order to probe the geometry of a Riemannian manifold, one needs to look at the relation between different geodesics. The simplest such setting is the collection of all geodesics emanating from one and the same point p . Let us consider two such geodesics, $c_1(t), c_2(t)$, with $c_1(0) = c_2(0) = p$, parametrized by arc length, that is, $d(c_i(t), p) = t$ for sufficiently small $t \geq 0, i = 1, 2$. The crucial quantity then is the distance between those geodesics as a function of t ,

$$f(t) := d(c_1(t), c_2(t)). \quad (1.52)$$

In the Euclidean plane, we have

$$f(t) = \gamma t, \quad (1.53)$$

with a constant γ that depends on the angle between c_1 and c_2 at the origin p . In contrast, on the unit sphere, we have

$$f(t) = \gamma \sin t, \quad (1.54)$$

whereas the relation in the hyperbolic plane is

$$f(t) = \gamma \sinh t. \quad (1.55)$$

This already is the typical behavior, in the sense that on spaces of positive sectional curvature, the distance between geodesics behaves like a trigonometric function, whereas in spaces of negative curvature, geodesics diverge at an exponential rate.

Of course, the curvature is a quantity that is defined pointwise, and on a general Riemannian manifold, it will therefore vary from point to point. But when it has a lower or an upper bound, geometric conclusions follow. That is, when it is $\leq \kappa$ or $\geq \lambda$, then geodesics locally diverge at least as fast or at most as slowly as on a space of constant curvature κ or λ , resp. That is, an upper/lower curvature bound controls the distance between geodesics from below/above. Here, a space of constant curvature K locally has the geometry of a scaled sphere when $K > 0$, of Euclidean space for $K = 0$, or of a scaled hyperbolic space for $K < 0$. The scaling factor is determined by the value of K . For instance, a space of constant curvature $K > 0$ is a sphere of radius $\frac{1}{\sqrt{K}}$. The smaller the radius, the larger the curvature, which of course agrees well with our intuition.

Thus, bounds on the sectional curvature control the distance function between geodesics, or conversely, when we have a control over the local divergence of geodesics, we can infer curvature bounds. This is a conceptually very useful result, as we can turn it around. Geodesics as locally distance minimizing curves exist in more general spaces than Riemannian manifolds, the so-called geodesic length spaces, see e.g. [46]. For instance, polytopes are not smooth, but one can easily define the lengths of curves and determine the distance minimizing ones. And on such spaces, one can then check for the divergence of geodesics. We can then simply declare such a geodesic length space to have curvature $\leq K$ or $\geq K$ when its geodesics locally exhibit the corresponding distance bounds. This was the approach taken by Busemann. A somewhat more restrictive curvature concept by Alexandrov also uses certain convexity properties of distances between geodesics. For a detailed treatment, we refer to [17, 46] and the references given there.

Putting it more abstractly, the idea simply is to identify a local property on a Riemannian manifold that is equivalent to a uniform infinitesimal curvature bound. And when this property then can be meaningfully defined on a class of spaces that is more general than Riemannian manifolds, we can then use that as a synthetic definition of a curvature bound. Of course, this then may be no longer compatible with a colloquial understanding of curvature as a deviation of *smooth* surfaces from being planar and thereby also abandon the aesthetic appeal of curvature, but such a state of affairs is not uncommon in mathematics.

1.3.4 The Geometric Meaning of Ricci Curvature

Since we have just found a local geometric characterization of sectional curvature bounds, we may now wish to ask whether something similar is possible for Ricci curvature as well. As it turns out there are two characteristic consequences of (lower) Ricci curvature bounds which we shall now describe.

1.3.4.1 Volume Growth

The first concerns the volume of balls. Whereas the sectional curvature contains information about the distance between geodesics, Ricci curvature yields estimates for the volume of distance balls $U(p, r) = \{q \in M : d(p, q) < r\}$. We have the Bishop–Gromov volume comparison theorem.

Theorem 1.1 *Let M be a d -dimensional Riemannian manifold with the lower Ricci curvature bound*

$$\text{Ric} \geq (d-1)\lambda, \quad (1.56)$$

which is an abbreviation for $\text{Ric}(X, X) \geq (d-1)\lambda\langle X, X \rangle$ for any tangent vector X , or in local coordinates $R_{ij}X^iX^j \geq (d-1)\lambda g_{ij}X^iX^j$. (The normalization here is such that the model space M_λ of dimension d and constant sectional curvature λ (i.e., a sphere, Euclidean or hyperbolic space depending on the sign of λ) has $\text{Ric} \equiv (d-1)\lambda$.) Let $V_\lambda(r)$ be the volume of a ball of radius r about any point in M_λ . Let $p \in M$.

Then

$$\frac{\text{Vol}(U(p, r))}{V_\lambda(r)} \text{ is monotonically decreasing in } r. \quad (1.57)$$

Of course, we have

$$\lim_{r \rightarrow 0} \frac{\text{Vol}(U(p, r))}{V_0(r)} = 1, \quad (1.58)$$

that is, infinitesimally, all volumes agree with the Euclidean one. The Ricci curvature then tells us about the local deviation from the Euclidean volume. However, only a lower Ricci curvature bound yields such a volume control. An upper Ricci curvature bound is not strong enough to make the quantity in (1.57) monotonically increasing. For that, one would rather need an upper *sectional* curvature bound; this is the Theorem of P. Günther.

1.3.4.2 The Weitzenböck and the Bochner Formula

The Weitzenböck–Bochner formula relates the Laplacian and the Ricci curvature. This is not surprising, as both involve the trace of a Hessian, that of a function in the case of the Laplacian and that of a metric in the case of Ricci.

Actually, the formulas are more general than that. Let e_1, \dots, e_d ($d = \dim M$) be a local orthonormal frame field (that is, $\langle e_i, e_j \rangle = \delta_{ij}$), with the dual coframe field e^1, \dots, e^d (that is, $e^i(e_j) = \delta_j^i$). Then the Laplace–Beltrami operator acting on

p -forms ($p = 0, 1, \dots, d$) is given by

$$\Delta = \nabla_{e_i e_i}^2 + e^i \wedge \iota(e_j) R(e_i, e_j). \quad (1.59)$$

Here, the second covariant derivative is defined as $\nabla_{XY}^2 = \nabla_X \nabla_Y - \nabla_{\nabla_X Y}$, and by letting the e_i form a Riemann normal frame, we can achieve $\nabla_{\nabla_{e_i} e_j} = 0$. ι stands for a contraction operator, and R denotes a certain curvature expression. The details are not so relevant here; they can be found in [47]. Here, we only want to emphasize the following general aspect. The Weitzenböck formula (1.59) expresses the Laplace-Beltrami operator on p -forms as the difference of a covariant second derivative operator $\nabla_{e_i e_i}^2$ and a curvature term. The second derivative operator is a negative operator. Therefore, when the curvature term is also negative, so then is Δ . This can be used as follows. Let ω be a harmonic p -form, that is, $\Delta\omega = 0$. We then compute $\Delta\langle\omega, \omega\rangle$ with the help of (1.59) and, assuming that the curvature term is positive, obtain an expression that is negative unless $\omega \equiv 0$. But on a compact manifold, $\int \Delta f = 0$. But by what we have just said, for $f = \langle\omega, \omega\rangle$, ω being a harmonic p -form, the integrand is pointwise negative unless $\omega \equiv 0$. Thus, ω must vanish identically, and there is no harmonic p -form. Here, we do not spell this out in detail, because in general, it is not so easy to interpret the curvature term in the Weitzenböck Formula (1.59) geometrically. For 1-forms, however, the curvature reduces to the negative Ricci curvature. This will allow us to derive interesting geometric consequences.

Theorem 1.2 (Bochner's Formula) *For a smooth function f on a Riemannian manifold M , we have*

$$\Delta\langle df, df \rangle = 2\langle \Delta df, df \rangle + 2|\nabla df|^2 + 2\text{Ric}(df, df). \quad (1.60)$$

Here, Δdf is the Laplacian of the 1-form df , as defined in (1.29). Note that since $d^2 = 0$, we have $\Delta df = -d\Delta^* df = d\Delta f$. Also, ∇df is the Riemannian version of the Hessian of the function f .

We can also desymmetrize (1.60) to obtain, for smooth functions f, g ,

$$\Delta\langle df, dg \rangle - \langle df, d\Delta g \rangle - \langle dg, d\Delta f \rangle = 2\langle \nabla df, \nabla dg \rangle + 2\text{Ric}(df, dg). \quad (1.61)$$

Proof Equation (1.60) can be derived from (1.59). Here, however, we want to provide a direct proof, which, in fact, is not very difficult. We use Riemann normal coordinates as in (1.13), (1.14). We have

$$\begin{aligned} \Delta df &= d\Delta f \\ &= d\left(g^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} - g^{ij} \Gamma_{ij}^k \frac{\partial f}{\partial x^k}\right) \\ &= \left(\frac{\partial^3 f}{\partial x^i \partial x^i \partial x^k} - \frac{1}{2}(g_{im,ik} + g_{im,ik} - g_{ii,km}) \frac{\partial f}{\partial x^m}\right) dx^k. \end{aligned}$$

We then compute

$$\begin{aligned}\Delta\langle df, df \rangle &= \frac{\partial^2}{(\partial x^k)^2} \left(g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} \right) \\ &= 2 \frac{\partial^2 f}{\partial x^k \partial x^i} \frac{\partial^2 f}{\partial x^k \partial x^i} + 2g^{ij} \frac{\partial^3 f}{(\partial x^k)^2 \partial x^i} \frac{\partial f}{\partial x^j} + \frac{\partial^2 g^{ij}}{(\partial x^k)^2} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j}.\end{aligned}$$

Using the preceding formula for $\frac{\partial^3 f}{(\partial x^k)^2 \partial x^i}$, we obtain

$$\begin{aligned}&= 2|\nabla df|^2 + 2\left(\langle \Delta df, df \rangle + \frac{1}{2}(g_{ki,kj} + g_{kj,ki} - g_{kk,ij}) \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j}\right) - g_{ij,kk} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} \\ &= 2\langle \Delta df, df \rangle + 2|\nabla df|^2 + (g_{ki,kj} + g_{kj,ki} - g_{kk,ij} - g_{ij,kk}) \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j}\end{aligned}$$

and with (1.47)

$$\begin{aligned}&= 2\langle \Delta df, df \rangle + 2|\nabla df|^2 + 2R_{ij} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} \\ &= 2\langle \Delta df, df \rangle + 2|\nabla df|^2 + 2 \operatorname{Ric}(df, df).\end{aligned}$$

Thus, we have

$$\Delta\langle df, df \rangle = 2\langle \Delta df, df \rangle + 2|\nabla df|^2 + 2 \operatorname{Ric}(df, df).$$

For a harmonic 1-form ω on a compact manifold M , which, according to (1.30) satisfies $d\omega = 0 = d^*\omega$, and which therefore, by the Poincaré Lemma, can locally be written as $\omega = df$ for some function f , (1.60) becomes

$$\Delta\langle \omega, \omega \rangle = 2|\nabla \omega|^2 + 2 \operatorname{Ric}(\omega, \omega). \quad (1.62)$$

As a consequence, we obtain Bochner's Theorem.

Corollary 1.1 *If M is a compact Riemannian manifold of positive Ricci curvature, then all harmonic 1-forms vanish. Thus, the first cohomology group of M vanishes,*

$$H^1(M, \mathbb{R}) = 0. \quad (1.63)$$

Proof Integrating (1.62) and using $\int_M \Delta f d\operatorname{vol} = \int_M \operatorname{div} \operatorname{grad} f d\operatorname{vol} = 0$ for a function f on a compact manifold by the divergence theorem, yields

$$0 = \int_M \Delta\langle \omega, \omega \rangle d\operatorname{vol} = 2 \int_M (|\nabla \omega|^2 + \operatorname{Ric}(\omega, \omega)) d\operatorname{vol}. \quad (1.64)$$

When the Ricci curvature is positive, the integrand on the right-hand side is pointwise nonnegative. It therefore has to vanish identically, and since $\int \text{Ric}(\omega, \omega) d\text{vol} > 0$ unless $\omega \equiv 0$, the claim follows.

1.3.4.3 Eigenvalue Bounds

There is another important consequence of a lower Ricci curvature bound, namely a bound for the smallest nontrivial eigenvalue λ_1 (1.26) of the Laplace-Beltrami operator Δ from below. This is the estimate of Lichnerowicz. This also follows from Bochner's formula.

Theorem 1.3 *Let M be a compact d -dimensional Riemannian manifold with $\text{Ric} \geq (d-1)\rho$, with $\rho > 0$; this means that for every tangent vector X*

$$\text{Ric}(X, X) \geq (d-1)\rho \langle X, X \rangle, \quad (1.65)$$

or equivalently, in local coordinates

$$R_{ij}X^iX^j \geq (d-1)\rho g_{ij}X^iX^j.$$

Then the first eigenvalue of Δ satisfies

$$\lambda_1 \geq d\rho. \quad (1.66)$$

Proof The proof comes from Bochner's formula (1.60). As in the proof of Corollary 1.1, integrating this formula yields

$$0 = (\Delta df, df) + (\nabla df, \nabla df) + \int_M \text{Ric}(df, df). \quad (1.67)$$

We have

$$-(\Delta df, df) = (\Delta f, \Delta f) \leq d(\nabla df, \nabla df) \quad (1.68)$$

by the Schwarz inequality. Therefore, (1.67) and (1.65) yield

$$(\Delta f, \Delta f) \geq d \rho (df, df). \quad (1.69)$$

If now f is an eigenfunction of Δ for an eigenvalue λ , i.e.,

$$\Delta f + \lambda f = 0,$$

we obtain

$$\lambda(df, df) = -\lambda(\Delta f, f) = (\Delta f, \Delta f) \geq d \rho(df, df) \quad (1.70)$$

whence either $df = 0$, that is, f is constant, and hence $\lambda = 0$, or (1.66) holds.

By Obata's theorem, the estimate (1.66) is optimal, and when equality holds, M is a sphere of constant sectional curvature ρ .

Again, however, the situation is asymmetric in the sense that an upper Ricci bound does not imply an upper estimate for λ_1 .

1.3.5 Harmonic Functions

One may also use the Bochner formula (1.60) to derive local gradient estimates for harmonic functions f on domains $\Omega \subset M$ when the Ricci curvature of M is bounded from below. Here, we do not even need to assume that the bound be positive.

In order to develop our geometric intuition and to motivate some of the subsequent constructions, we now also want to relate harmonic functions to the other property following from a lower Ricci bound, the control of the volume growth. We start with the observation that a function h on a Euclidean domain $\Omega \subset \mathbb{R}^d$ is harmonic iff it satisfies the mean value property, that is,

$$h(x) = \frac{1}{\text{Vol}(U(x, r))} \int_{U(x, r)} h(y) dy \quad (1.71)$$

for all balls $U(x, r) \subset \Omega$. Again, in the spirit of the above local interpretation of curvature bounds, we can then *define* a generalized harmonic function h on a domain $\Omega \subset M$ by the requirement that

$$h(x) = \frac{1}{\text{Vol}(U(x, r))} \int_{U(x, r)} h(y) d\text{vol}(y), \quad (1.72)$$

where now, of course, the volume refers to the Riemannian metric on M . On a general Riemannian manifold, however, we can require (1.72) only for some fixed value $r > 0$, in contrast to the Euclidean mean volume property which holds for all $r > 0$. This approach was developed in [44] and [46]. With this, we can relate the regularity properties of such generalized harmonic functions to volumes of distance balls. In fact, when x, y satisfy $d(x, y) < 2r$, then $U(x, r) \cap U(y, r) \neq \emptyset$, and consequently

$$|h(x) - h(y)| \text{ is controlled by } \frac{1}{\text{Vol}(U(x, r))} \int_{(U(x, r) \cup U(y, r)) \setminus (U(x, r) \cap U(y, r))} |h(z)| d\text{vol}(z). \quad (1.73)$$

(There is the slight subtlety that the volumes of $U(x, r)$ and $U(y, r)$ need not agree, but this is not important for the geometric intuition we are trying to develop here.) The crucial observation is that the larger $\frac{\text{Vol}(U(x, r) \cap U(y, r))}{\text{Vol}(U(x, r))}$, the better the estimate (1.73) becomes. That is, for such an estimate, we should not only control the volumes of single balls, but rather the relative size of the volume of the intersection of two balls. Again, this can be controlled by a lower Ricci bound, as in Theorem 1.1. Even better, since

$$h(x) - h(y) = \frac{1}{\text{Vol}(U(x, r))} \int_{U(x, r)} h(z) d\text{vol}(z) - \frac{1}{\text{Vol}(U(y, r))} \int_{U(y, r)} h(z) d\text{vol}(z), \quad (1.74)$$

if we could somehow pair the points ξ of $U(x, r)$ with the points η of $U(y, r)$ in an optimal manner, that by some transfer $\eta = T(\xi)$ in such a manner that $h(\xi) - h(T(\xi))$ becomes small then we could even improve our estimate. Thus, we are naturally lead to the issue of the optimal transport of the points in one ball to those of another ball. The approach to the regularity of generalized harmonic maps by considering volumes of intersections of balls was first pursued in [45]. Anticipating some subsequent constructions, the connection between optimal transport and the regularity of generalized harmonic maps is developed in [18]. We should also mention the important result of Zhang [90] on the Lipschitz regularity of harmonic maps on Alexandrov spaces with lower Ricci bounds (a concept to be defined below).

1.4 A Nonlocal Approach to Geometry

Riemannian geometry constitutes an infinitesimal approach to geometry, in the sense that the crucial operators like tangent vectors operate by evaluating derivatives of smooth objects at a point. Nonlocal operations are derived operations in Riemannian geometry, insofar as they are obtained by processes of integration from infinitesimal ones. A nonlocal approach to geometry, in contrast, would take objects or operations that depend on two points as its basic ingredients. For instance, a vector field then is a function with two arguments, $p : M \times M \rightarrow \mathbb{R}$. Infinitesimal objects could then be obtained by limiting processes where the two points approach each other. However, there also exist spaces where such limiting processes do not make sense, and in those cases only a nonlocal approach to geometry is feasible.

Such nonlocal approaches have been much utilized in image processing. There, the basic model is that of Kindermann–Osher–Jones [53] and Gilboa–Osher [37]. Our approach to nonlocal geometry, however, will be different from that of Gilboa and Osher [37] and other papers in image processing, as well as from that of Bartholdi et al. [9]. It was developed in [41–43], and as we believe, is more systematic and natural.

We start with a function $\omega : M \times M \rightarrow \mathbb{R}$. $\omega(x, y)$ may express the similarity or vicinity between the points x and y . We shall usually assume that ω is nonnegative,

$$\omega(x, y) \geq 0, \text{ but } \omega \not\equiv 0, \quad (1.75)$$

and symmetric,

$$\omega(x, y) = \omega(y, x). \quad (1.76)$$

We also put

$$\bar{\omega}(x) := \int_M \omega(x, y) dy, \quad (1.77)$$

and assume, of course, that this is $< \infty$. We view $\bar{\omega}$ as the density of a metric.

We use $\bar{\omega}(x)$ and $\omega(x, y)$ to define the L^2 -norms for functions $u : M \rightarrow \mathbb{R}$ and vector fields $p : M \times M \rightarrow \mathbb{R}$. The scalar product for functions with respect $\bar{\omega}(x)$ is defined by

$$(u_1, u_2)_{L^2_{\bar{\omega}}} := \int u_1(x) u_2(x) \bar{\omega}(x) dx.$$

In particular, this yields the $L^2_{\bar{\omega}}$ -norm for functions,

$$\|u\|_{L^2_{\bar{\omega}}}^2 = (u, u)_{L^2_{\bar{\omega}}} = \int u^2(x) \bar{\omega}(x) dx.$$

For vector fields $p, q : M \times M \rightarrow \mathbb{R}$, we define their scalar product with respect to $\omega(x, y)$

$$(p, q)_{L^2_{\omega}} := \int p(x, y) q(x, y) \omega(x, y) dx dy.$$

The difference vector field of a function $u : M \rightarrow \mathbb{R}$ is defined by

$$Du(x, y) = u(y) - u(x). \quad (1.78)$$

The definition of the difference vector field does not depend on $\bar{\omega}$ or ω , but we shall now use the metric to define a divergence operator as an adjoint.

For a vector field $p : M \times M \rightarrow \mathbb{R}$, its *divergence operator* $\operatorname{div} p : M \rightarrow \mathbb{R}$ is now defined by

$$\operatorname{div} p(x) := \frac{1}{\bar{\omega}(x)} \int (p(x, y) - p(y, x)) \omega(x, y) dy. \quad (1.79)$$

The divergence operator satisfies for any $u : M \rightarrow \mathbb{R}$ and $p : M \times M \rightarrow \mathbb{R}$,

$$(Du, p)_{L^2_{\bar{\omega}}} = -(u, \operatorname{div} p)_{L^2_{\bar{\omega}}}. \quad (1.80)$$

Our nonlocal geometry becomes analogous to Riemannian geometry when we view $\bar{\omega}$ (and ω) as a (Riemannian) metric on Ω and Du as the differential of u which does not depend on the metric $\bar{\omega}$ and div as the gradient operator with respect to the metric $\bar{\omega}$ (and ω).

From our new nonlocal variational setting we also obtain a nonlocal Dirichlet functional

$$D(u) := \frac{1}{2}(Du, Du)_{L^2_{\bar{\omega}}} = \frac{1}{2} \int_{\Omega} \int_{\Omega} (u(x) - u(y))^2 \omega(x, y) dx dy.$$

(In image processing, the normalization factor $\frac{1}{4}$ is used instead, but here we work with $\frac{1}{2}$ for reasons of compatibility with the rest of this contribution.) Analogously to what we did above in Riemannian geometry (1.21), we define the Laplacian of a function as the divergence of the difference vector field of a function u .

$$\Delta_{\bar{\omega}} u(x) := \frac{1}{2} \operatorname{div}(Du) = \frac{1}{\bar{\omega}(x)} \int (u(y) - u(x)) \omega(x, y) dy = \bar{u}(x) - u(x). \quad (1.81)$$

The Laplace operator satisfies

$$(\Delta_{\bar{\omega}} u, v)_{L^2_{\bar{\omega}}} = -(Du, Dv)_{L^2_{\bar{\omega}}} = (u, \Delta_{\bar{\omega}} v)_{L^2_{\bar{\omega}}} \quad (1.82)$$

for all L^2 -functions u, v . We shall encounter this Laplace operator again below as the normalized Laplacian of graph theory. Equation (1.81) is also the Euler-Lagrange equation for the nonlocal Dirichlet functional.

When applied to image analysis (see e.g. [2, 76, 89] for background on PDE and variational methods in image processing), the nonlocal weight $\omega(x, y)$ should reflect the statistical dependencies between the pixels x and y in collections of images. (An alternative conceptualization would be that of Kimmel et al. [52] who work with a varying Riemannian metric obtained by pulling back a fixed metric in the feature space under consideration via a map u that represents the varying image during the denoising process.) We now sketch the application of the preceding to image denoising as developed in [42, 43]. The task of image denoising is to recover original images u from noise-corrupted versions f

$$f = u + v.$$

The variational or PDE based methods constitute an important class of image denoising strategies, see for [2, 25, 65, 75]. The variational methods for denoising images balance a fidelity term that measures the deviation of the resulting image u from the noisy input f and a regularizing or smoothing term that suppresses irregular

oscillations, which are supposed to stem from the noise, in the image u . The basic idea can be seen in the classical H^1 model

$$\int (|\operatorname{grad} u|^2(x) + \lambda(u-f)^2(x))dx, \quad (1.83)$$

where, in the absence of a specific Riemannian metric, $\operatorname{grad} u$ would be the Euclidean gradient. Of course, we can also assume that there is some other Riemannian metric in the background with respect to which gradient and integration can be defined. λ is a parameter that balances the relative weights of the two terms, and in some contexts, it also arises as a Lagrange multiplier. However, this scheme blurs images quickly, and hence is not directly useful in image denoising. The problems stem from the gradient term. Based on the nonlocal geometry just developed, we [42, 43] have then proposed the following H^1 model

$$F(u) = \frac{1}{2} \int \int (u(y) - u(x))^2 \omega(x, y) dy dx + \lambda \int (u-f)^2(x) \bar{\omega}(x) dx. \quad (1.84)$$

This is different from Gilboa-Osher's nonlocal H^1 model [37] (which simply works with $\lambda \int (u-f)^2(x) dx$ as the fidelity term) insofar as here we have derived the fidelity term from the same geometry as the regularization term. The Euler-Lagrange equation for (1.84) is

$$\Delta_{\bar{\omega}} u = \lambda(u-f), \quad (1.85)$$

or equivalently

$$u(x) - \bar{u}(x) = -\lambda(u-f)(x). \quad (1.86)$$

The variational problem can also be seen as a constrained problem:

$$u := \arg \min \{D(u) \mid \|u-f\|_{\omega}^2 = |M|_{\omega} \sigma^2\}, \quad (1.87)$$

where σ^2 is the variance of an additive noise added in a noisy image f and $|M|_{\omega}$ is the area of M with respect to ω , i.e. $|M|_{\omega} = \int_{\Omega} \bar{\omega}(x) dx$. Hence in this case λ is also a Lagrange multiplier of this constrained problem and can be computed by

$$\lambda = -\frac{1}{|M|_{\omega} \sigma^2} \int_M (u-f) \Delta_{\bar{\omega}} u \bar{\omega}. \quad (1.88)$$

Again, in order to understand the geometric properties of the solutions of (1.85), one should start with the corresponding harmonic functions, that is, the solutions of

$$\Delta_{\bar{\omega}} u = 0 \quad (1.89)$$

on some domain Ω , that is,

$$u(x) = \bar{u}(x) \quad (1.90)$$

for all $x \in \Omega$. Thus, we are again in the situation described at the end of Sect. 1.3.5. This time, we would need some control on intersections of weighted sets, that is, on

$$\bar{\omega}(x) - \bar{\omega}(y) \quad (1.91)$$

for $x, y \in M$. Again, as we shall argue below, some concept of generalized Ricci curvature is adequate here.

1.5 Generalized Ricci Curvature

Over the years, several notions of generalized curvature for metric spaces have been proposed and investigated. For a long time, research was concerned with sectional curvature. The reason is that, as explained above, sectional curvature inequalities can be characterized in terms of relations involving only distance functions. Thus, no further structure beyond the metric is needed in principle. In contrast, Ricci curvature in Riemannian geometry involves some averaging and contains information about volumes and eigenvalues of a Laplace operator which again we have introduced and discussed in terms of some local averaging. Such an operation, however, needs a measure. Therefore, in order to define some kind of generalized Ricci curvature, we need some measure on our space in addition to the metric, that is, the distance function. In recent years, this has become a rather active research topic, see for instance [64, 69, 81]. We shall not attempt to survey this here, but only mention that notions of generalized Ricci curvature for possibly discrete spaces have been introduced by Bonciocat–Sturm [22], Ollivier [70, 71] and Bakry and Emery [4]. Here, we shall discuss the latter two approaches, as they are more suited to those geometric topics that we are interested in, namely eigenvalue bounds and the regularity of harmonic functions. Another approach, also based on optimal transport and probability spaces, but using the gradient flow of the entropy, is presented in Chap. 5.

First, however, we shall develop another notion of generalized Ricci curvature for simplicial, or more generally, CW complexes, due to Forman [32], which takes Bochner’s Theorem 1.2 and its Corollary 1.1 as its starting point and does not require a measure.

1.5.1 Forman's Ricci Curvature

This notion of Ricci curvature is defined for CW complexes satisfying the following combinatorial condition: When, for two p -dimensional cells α_1, α_2 , it happens that $\beta \subset \bar{\alpha}_1 \cap \bar{\alpha}_2$ for some $(p-1)$ -cell β , then $\bar{\alpha}_1 \cap \bar{\alpha}_2 = \bar{\beta}$. This condition will henceforth be assumed. It is satisfied for simplicial, or more generally, polyhedral complexes.

The notation $\alpha < \beta$, or equivalently, $\beta > \alpha$ means that the cell α is contained in the boundary of the cell β . Two p -cells α_1, α_2 are called upward neighbors if there is a $(p+1)$ -cell γ with $\gamma > \alpha_i$ for both $i = 1, 2$, and they are called downward neighbors if there is $(p-1)$ -cell β with $\beta < \alpha_i$ for both $i = 1, 2$. They are called transverse neighbors if they are both up- and downward neighbors, and parallel neighbors if they are either up- and downward neighbors, but not both.

When we indicate the dimension p of a cell α , we shall also write α^p .

Definition 1.5 The curvature of the p -cell α is

$$F_p(\alpha) = \# \{ (p+1)\text{-cells } \gamma > \alpha \} + \# \{ (p-1)\text{-cells } \beta < \alpha \} - \# \{ \text{parallel neighbors of } \alpha \}. \quad (1.92)$$

More generally, when we have a weighted cell-complex with weights w_α , the curvature of a weighted p -cell α is

$$\begin{aligned} F_p(\alpha) = w_\alpha & \left(\sum_{\gamma^{p+1} > \alpha} \frac{w_\alpha}{w_\gamma} + \sum_{\beta^{p-1} < \alpha} \frac{w_\beta}{w_\alpha} \right. \\ & \left. - \sum_{\bar{\alpha}^p \neq \alpha} \left| \sum_{\gamma^{p+1} > \alpha, \gamma > \bar{\alpha}} \frac{\sqrt{w_\alpha w_{\bar{\alpha}}}}{w_\gamma} - \sum_{\beta^{p-1} < \alpha, \beta < \bar{\alpha}} \frac{w_\beta}{\sqrt{w_\alpha w_{\bar{\alpha}}}} \right| \right). \end{aligned} \quad (1.93)$$

With this notion of curvature, Forman [32] can derive an analogue of the Weitzenböck formula (1.59). In fact, by Eckmann [29], there is a natural analogy between the (co)homology of simplicial (or more general) complexes and the Hodge-de Rham cohomology of p -forms on Riemannian manifolds. For details of the following, see for instance [49]. Let C_p be the vector space of real p -chains, that is, formal linear combinations of the p -cells of our complex. We then have a boundary operator

$$\partial_p : C_p \rightarrow C_{p-1}, \quad (1.94)$$

which satisfies $\partial_{p-1} \circ \partial_p = 0$ and therefore defines a homology theory. We can also introduce a scalar product on C_p by letting different cells be orthogonal and setting

$$\langle \alpha, \alpha \rangle = w_\alpha \text{ for some } w_\alpha > 0. \quad (1.95)$$

(In the unweighted case, we should simply put $w_\alpha = 1$ for all cells.)

With this product, we can define the adjoint $\partial_p^* : C_{p-1} \rightarrow C_p$ of ∂_p via

$$\langle \partial_p^* \beta^{p-1}, \alpha^p \rangle = \langle \beta^{p-1}, \partial_p \alpha^p \rangle. \quad (1.96)$$

As in (1.29), we can then define a Laplace operator via

$$\square_p = -\partial_p^* \partial_p - \partial_p \partial_p^* \quad \text{on } p\text{-cells}. \quad (1.97)$$

Importantly, Forman found a decomposition that is analogous to (1.59),

$$\square_p = B_p - F_p \quad (1.98)$$

where B_p is a negative operator analogous to ∇^2 in (1.59), and F_p is given by (1.92) in the unweighted and by (1.93) in the weighted case.

In particular, one can then derive an analogue of Bochner's vanishing Theorem 1.1,

Corollary 1.2 *Let M be a finite regular CW-complex satisfying the above combinatorial condition. If $F_p(\alpha) > 0$ for all p -cells α , then*

$$H_p(M, \mathbb{R}) = 0. \quad (1.99)$$

We notice that the preceding corollary holds for all p , and not just for $p = 1$. There are, of course, also corresponding versions of Corollary 1.1, but we don't want to enter the underlying algebraic aspects of the curvatures appearing in those versions, which all come, of course, from (1.59). Rather, we specialize Forman's curvature to dimension 1, to obtain the Forman-Ricci curvature for 1-dimensional cells, that is, edges e . Then (1.92) becomes

$$\text{Ric}(e) := F_1(e) = \sharp\{2\text{-cells } f > e\} + 2 - \sharp\{\text{parallel neighbors of } e\}. \quad (1.100)$$

In particular, when M is a graph, that is, there are only vertices (0-cells) and edges (1-cells), then for an edge $e = (v_1, v_2)$

$$\text{Ric}(e) = 4 - \deg v_1 - \deg v_2, \quad (1.101)$$

where the degree $\deg v$ of a vertex is defined as the number of its neighbors, that is, other vertices connected to v by an edge; see Sect. 1.6.1. Of course, in that case, the condition of Corollary 1.2 is only satisfied in the trivial case where the graph consists of a single edge. Actually, the conclusion of that Corollary continues to hold on a connected graph when $\text{Ric}(e) \geq 0$ for all e and $\text{Ric}(e_0) > 0$ for at least one e_0 . That is satisfied when the graph is a path (see Sect. 1.6.1 for the definition). The Corollary becomes more powerful on complexes that also contain 2-dimensional cells, because in that case, we may get a positive contribution $\sharp\{2\text{-cells } f > e\}$ in (1.100).

In contrast to the other generalized Ricci curvature notions that we shall present below, Forman's version is of a purely combinatorial nature. This is in line with the approach of Eckmann [29] who, as already noted, developed the analogy between the theory of the Hodge-Laplace operator (1.29) operating on p -forms on Riemannian manifolds and the combinatorial properties of the discrete Laplace operator operating on p -simplices in simplicial complexes. For the issue of consistent choices of weights w_α across the different dimensions p , we refer to [39]. In this context, we should also mention the work of Garland [35] who derived and used a combinatorial Bochner formula on Bruhat-Tits buildings (for a geometric interpretation, see [51]).

1.5.2 Ollivier's Ricci Curvature

We now present Ollivier's definition. We first need to introduce the L^1 -Wasserstein distance W_1 . For more general on Wasserstein distances (also called Wasserstein metrics), we refer to Chap. 5.

Definition 1.6 Let (X, d) be a metric space equipped with its Borel sigma algebra,² and let m_1, m_2 be (Radon) probability measures on X . The L^1 -Wasserstein or transportation distance between the probability measures m_1 and m_2 is

$$W_1(m_1, m_2) = \inf_{\xi \in \Pi(m_1, m_2)} \int_{(x,y) \in X \times X} d(x, y) d\xi(x, y), \quad (1.102)$$

where $\Pi(m_1, m_2)$ is the set of probability measures ξ that satisfy

$$\int_{y \in X} d\xi(x, y) = m_1(x), \quad \int_{x \in X} d\xi(x, y) = m_2(y). \quad (1.103)$$

Of course, on a discrete space, like a graph, the integrals are replaced by sums.

The conditions (1.103) mean that we start with the measure m_1 and end up with m_2 , or in stochastic terminology, that the marginales of ξ be m_1 and m_2 . When we consider the distance $d(x, y)$ as the transportation cost from x to y , then $W_1(m_1, m_2)$ is the minimal cost to transport the mass of m_1 to that of m_2 . ξ is considered as a transfer plan between m_1 and m_2 , or a coupling of the two random walks governed by m_1 and m_2 , respectively. Those ξ which attain the infimum in (1.102) are called optimal couplings. Optimal coupling exist under rather general conditions, but they

²The Borel sigma algebra is the set of all subsets of X that are obtained from the open balls by taking complements, finite intersections and countable unions. For the sets in the Borel sigma, one can then define their volumes w.r.t. to a Radon probability measure. The technical details are not so important for understanding the essence of the subsequent constructions.

need not be unique. A comprehensive reference for the theory is [83]. A shorter introduction is [31].

The transportation distance $W_1(m_1, m_2)$ can also be expressed by the Kantorovich duality formula,

$$W_1(m_1, m_2) = \sup_{f: \text{Lip}(f) \leq 1} \left[\int_{x \in X} f(x) dm_1(x) - \int_{y \in X} f(y) dm_2(y) \right], \quad (1.104)$$

where $\text{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}$ is the Lipschitz seminorm of f .

Definition 1.7 Let (X, d) be a complete and separable metric space equipped with its Borel sigma algebra and a family of probability measures $m_x, x \in M$ which depend measurably on x and which have finite first moments, i.e., $\int_M d(x, y) dm_x(y) < \infty$. For any two distinct points $x, y \in X$, the (Ollivier-) Ricci curvature of (X, d, m) then is defined as

$$\kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)}. \quad (1.105)$$

The probability measures m_x could also be interpreted as the probability densities associated to a random walk, as we shall elaborate upon below when we discuss graphs.

A positive lower bound for $\kappa(x, y)$ has many geometric consequences. For instance Ollivier [71] observed the following Bonnet-Myers type result.

Theorem 1.4 *Suppose that $\kappa(x, y) \geq \kappa > 0$ for all $x, y \in X$. Then for any $x, y \in X$ one has*

$$d(x, y) \leq \frac{W_1(\delta_x, m_x) + W_1(\delta_y, m_y)}{\kappa(x, y)}, \quad (1.106)$$

and hence

$$\text{diam}(X) \leq \frac{2 \sup_x W_1(\delta_x, m_x)}{\kappa}. \quad (1.107)$$

Proof $d(x, y) = W_1(\delta_x, \delta_y) \leq W_1(\delta_x, m_x) + W_1(m_x, m_y) + W_1(\delta_y, m_y) \leq W_1(\delta_x, m_x) + (1 - \kappa)d(x, y) + W_1(\delta_y, m_y)$.

On a Riemannian manifold, however, this result is weaker than the usual Myers Theorem [68], which scales differently with the Ricci bound. For other Myers type theorems for tessellations, we refer to Keller's contribution to this volume (Chap. 6, Sect. 6.2.3).

1.5.3 Curvature Dimension Inequality

We now present Bakry and Emery's approach to a generalized lower bound for the Ricci curvature [4]. The theory is systematically developed in [8]. As the approach of Forman, it starts from a Weitzenböck-Bochner identity, but proceeds very differently, by abstracting the algebraic aspects of the Bochner formula (1.61) which states that on Riemannian manifolds

$$\frac{1}{2}\Delta\langle df, dg \rangle - \frac{1}{2}\langle df, d\Delta g \rangle - \frac{1}{2}\langle dg, d\Delta f \rangle = \langle \nabla df, \nabla dg \rangle + \text{Ric}(df, dg). \quad (1.108)$$

The symmetric version of this formula, (1.60), is

$$\frac{1}{2}\Delta|df|^2 = \langle df, d\Delta f \rangle + \|\nabla df\|_2^2 + \text{Ric}(df, df). \quad (1.109)$$

As a motivation for the algebra, we recall the product formula

$$\frac{1}{2}\Delta(fg) = \frac{1}{2}f\Delta g + \frac{1}{2}g\Delta f + \langle df, dg \rangle. \quad (1.110)$$

Bochner's formula establishes an important connection between geometric and analytic properties of a manifold. Many analytical consequences of a lower Ricci curvature bound are established through it, see for instance the proof of the Lichnerowicz estimate Theorem 1.3. However on more general spaces than Riemannian manifolds, it is not clear how to define the Hessian ∇df and the Ricci tensor. But using the simple inequality

$$\|\nabla df\|_2^2 \geq \frac{(\Delta f)^2}{n},$$

an immediate consequence of the Bochner identity is that on an n -dimensional manifold whose Ricci curvature is bounded from below by K one has

$$\frac{1}{2}\Delta|df|^2 \geq \langle df, d\Delta f \rangle + \frac{1}{n}(\Delta f)^2 + K|df|^2. \quad (1.111)$$

We have used this already in the proof of Theorem 1.3. The advantage of this inequality over the Bochner identity (1.109) is that now all the objects in (1.111) can easily be defined on metric measure spaces.

It was the important insight by Bakry and Emery [4] that one can use the inequality (1.111) as a substitute for the lower Ricci curvature bound on spaces where a direct generalization of Ricci curvature is not possible. Indeed, Bakry and Emery take Eq. (1.111) as the starting point of their approach. We will briefly outline their approach now.

For the sake of generality, we state the following definitions for a general differential operator L instead of restricting ourselves to the Laplace-Beltrami operator Δ on a Riemannian manifold. However, it might be helpful to keep the Laplace-Beltrami operator as one particular example in mind.

Definition 1.8 For a differential operator L we define the gradient form Γ by

$$2\Gamma(f, g)(x) = (L(f \cdot g) - f \cdot L(g) - L(f) \cdot g)(x) \quad (1.112)$$

and the iterated gradient form Γ_2 by

$$2\Gamma_2(f, g) = L\Gamma(f, g) - \Gamma(f, Lg) - \Gamma(Lf, g). \quad (1.113)$$

Here, (1.112) should be seen as an abstract version of (1.110), and (1.113) then of (1.108).

Definition 1.9 We say that an operator L satisfies the curvature dimension inequality $CD(n, K)$ (CD-inequality for short) if, for any function f

$$\Gamma_2(f) \geq \frac{1}{n}(Lf)^2 + K\Gamma(f).$$

Note that for $L = \Delta$, this definition is nothing but the inequality (1.111) written in the Γ notation.

The curvature dimension inequality has proven to be useful in various situations and many results (including the Lichnerowicz estimate Theorem 1.3 and Myers theorem), that require a lower bound on the Ricci curvature, could be generalized to metric measure spaces, see [4, 5, 7]. Another important result that could be proved in the curvature dimension inequality formalism was a generalization of the Li-Yau gradient estimates. In the special case of an n -dimensional compact manifold with non-negative Ricci curvature, the Li-Yau gradient estimates [58] for positive solutions u of the heat equation $\mathcal{L}u := (\Delta - \partial_t)u = 0$ read

$$\frac{|\nabla u|^2}{u^2} - \frac{\partial_t u}{u} \leq \frac{n}{2t}. \quad (1.114)$$

Bakry and Ledoux [6] generalized Li and Yau's result and could show that under the assumption of $CD(n, 0)$ the gradient estimate (1.114) is satisfied for diffusion semigroups (for a definition see below) generated by an operator L .

Definition 1.10 Given an operator L , the semigroup $P_t = e^{tL}$ is said to be a *diffusion semigroup* if the following identities are satisfied for any smooth function $\Phi : \mathbb{R} \rightarrow \mathbb{R}$:

$$\Gamma(f, gh) = g\Gamma(f, h) + h\Gamma(f, g) \quad (1.115)$$

$$\Gamma(\Phi \circ f, g) = \Phi'(f)\Gamma(f, g) \quad (1.116)$$

$$L(\Phi \circ f) = \Phi'(f)L(f) + \Phi''(f)\Gamma(f). \quad (1.117)$$

Gradient estimates are very powerful tools in geometric analysis. In particular they imply Harnack inequalities, heat kernel and eigenvalue estimates, see [47, 57] for an overview. We shall discuss in Sects. 1.6.3–1.6.5 the curvature dimension inequality and the Li-Yau gradient estimates for graphs.

The curvature dimension inequality formalism is also very useful for infinite dimensional analysis. In particular, we mention that $CD(\infty, K)$ implies a dimension-free version of the gradient estimate, the Bakry-Emery gradient estimate (see Theorem 1.13 below).

1.6 Ricci Curvature and the Geometry of Graphs

We now apply both the geometric intuition developed in the previous sections and Ollivier’s concept of generalized Ricci curvature (Definition 1.7), Bakry and Emery’s curvature dimension inequality (Definition 1.9) to the special case where the underlying metric space is a graph.

1.6.1 Basic Notions from Graph Theory

In order to prepare for the discussion about the relation between Ricci curvature and the geometry, we introduce some basic definitions and constructions from graph theory, including the (normalized) graph Laplacian. For more details, see [48] and the references given there.

We first consider a locally finite unweighted graph $G = (V, E)$. V is the vertex and E the edge set. We say that $x, y \in V$ are neighbors, and write $x \sim y$, when they are connected by an edge. The degree d_x of a vertex x is defined as the number of its neighbors. “Locally finite” then means that every vertex has only finitely many neighbors, or equivalently, that d_x is finite for every $x \in V$.

While for the moment, we might wish to exclude self-loops, that is, edges connecting a vertex with itself, subsequently, in Sect. 1.6.2.2, we shall have to allow for their possibility. We also assume that G is connected, that is, for every pair of distinct vertices $x, y \in V$, there exists a path between them, that is, a sequence $x = x_0, x_1, \dots, x_m = y$ of distinct vertices such that $x_{v-1} \sim x_v$ for $v = 1, \dots, m$. Since we can decompose graphs that are not connected into their connected components, the connectivity assumption is no serious restriction. A cycle in G is a closed path $x_0, x_1, \dots, x_m = x_0$ for which all the vertices x_1, \dots, x_m are distinct. For $m = 3, 4, 5, \dots$, we speak of a triangle, quadrangle, pentagon, ... A graph without cycles is called a tree. A graph is called bipartite if its vertex set can be decomposed into two disjoint components V_1, V_2 such that whenever $x \sim y$, then x and y are in different components. Any tree is bipartite. More generally, a graph is bipartite iff it has no cycles of odd length. In particular, it has no triangles.

Triangles will play a crucial role in our discussion of Ricci curvature on graphs. Therefore, we now introduce some corresponding notation. For two vertices x, y , we let N_{xy} be the set of all vertices z that are neighbors of both x and y . Equivalently, this is the set of all vertices z for which x, y, z constitute a triangle. We let $\sharp(x, y)$ then be the number of vertices in N_{xy} , that is, the number of joint neighbors of x and y , or equivalently, the number of triangles containing x and y .

We have an obvious metric d on the vertex set V . For neighbors x, y , $d(x, y) = 1$. For arbitrary vertices x, y , $d(x, y)$ is the length of the shortest path connecting x and y , i.e. the minimal number of edges that needs to be traversed to get from x to y .

We next introduce the (normalized) graph Laplacian operating on L^2 -functions on the vertex set V . Here, we use the scalar product

$$(v, u) := \sum_{x \in V} d_x v(x) u(x) \quad (1.118)$$

to define $L^2(G)$. We then put

$$\begin{aligned} \Delta : L^2(G) &\rightarrow L^2(G) \\ \Delta v(x) &:= \frac{1}{d_x} \left(\sum_{y, y \sim x} v(y) - d_x v(x) \right) = \frac{1}{d_x} \sum_{y, y \sim x} v(y) - v(x). \end{aligned} \quad (1.119)$$

When we attach to each vertex $x \in V$ the measure

$$m_x(y) = \begin{cases} \frac{1}{d_x} & \text{if } y \sim x; \\ 0 & \text{else,} \end{cases} \quad (1.120)$$

we see that this is the discrete version of (1.81). We point out that the definition of the Laplacian utilized here is equivalent to that used in [26], but different from the algebraic graph Laplacian often considered in graph theory; the latter would not have the factor $\frac{1}{d_x}$. We can also consider, for neighbors $x \sim y$, the discrete differential

$$Du(x, y) := u(y) - u(x), \quad (1.121)$$

the analogue of (1.78). D can be considered as a map from functions on the vertices of G to functions on the edges of G . In order to make the latter space also an L^2 -space, we introduce the product

$$(Du, Dv) := \sum_{e=(x,y)} (u(y) - u(x))(v(y) - v(x)). \quad (1.122)$$

Note that we are summing here over edges, and not over vertices. If we did the latter, we would need to put in a factor $1/2$ because each edge would then be counted twice. We then have

$$(\Delta u, v) = -(Du, Dv) \quad (1.123)$$

for all $u, v \in L^2(G)$, as in (1.82).

We now list some basic properties of Δ .

1. Δ is selfadjoint w.r.t. (\cdot, \cdot) :

$$(u, \Delta v) = (\Delta u, v) \quad (1.124)$$

for all $u, v \in L^2(G)$. This is the analogue of (1.82). Of course, it follows from (1.123).

2. Δ is nonpositive:

$$(\Delta u, u) \leq 0 \quad (1.125)$$

for all u . This follows from the Cauchy-Schwarz inequality.

3. $\Delta u = 0$ iff u is constant. In fact, when $\Delta u = 0$, there can neither be a vertex x with $u(x) \geq u(y)$ for all $y \sim x$ with strict inequality for at least one such y , since $\Delta u(x) = 0$ means that the value $u(x)$ is the average of the values at the neighbors of x . Since G is assumed to be connected, u then has to be a constant (if G were not connected, a solution of $\Delta u = 0$ would have to be constant on every connected component of G .) Of course, this is a discrete version of the standard maximum principle argument.

We are again interested in the eigenvalues of the Laplacian, that is, in those λ with

$$\Delta u + \lambda u = 0 \quad (1.126)$$

for some nontrivial function $u \in L^2(G)$, called an eigenfunction for λ . From the properties of Δ just listed, we can infer some immediate consequences for the eigenvalues.

- All eigenvalues are real, because Δ is selfadjoint.
- All eigenvalues are nonnegative, because Δ is a nonpositive operator.
- On a finite graph, the smallest eigenvalue is $\lambda_0 = 0$, with a constant eigenfunction (when the graph is not finite, a constant function is no longer in L^2). Since we assume that Γ is connected, this eigenvalue is simple. In other words,

$$\lambda_k > 0 \quad (1.127)$$

for $k > 0$ where we order the eigenvalues as

$$\lambda_0 = 0 < \lambda_1 \leq \dots \leq \lambda_K$$

and put $K := N - 1$.

- The largest eigenvalue λ_{N-1} is 2 iff G is bipartite and is < 2 else.

The eigenfunctions v_i, v_j for different eigenvalues λ_i, λ_j are orthogonal to each other,

$$(v_i, v_j) = 0. \quad (1.128)$$

In particular, since the constants are the eigenfunctions for the eigenvalue $\lambda_0 = 0$, for all $i > 0$, we then have

$$\sum_x d_x v_i(x) = 0. \quad (1.129)$$

We do not want to go into the more detailed properties of the eigenfunctions here, but only mention the fact that when G is bipartite, then an eigenfunction for the largest eigenvalue λ_{N-1} equals a constant on one of the two classes and a different constant on the other classes, where these two constants need to be such that (1.129) is satisfied. For a non-bipartite graph, we do not have such a simple highest eigenfunction, and in some sense, this is the reason why $\lambda_{N-1} < 2$ in that case. We refer to [11] for details and a systematic analysis of the highest eigenvalue.

The eigenvalues can be obtained from a variational principle, the Courant-Fischer-Weyl min-max principle,

$$\lambda_k = \min_{\substack{u_0, \dots, u_k \neq 0 \\ (u_i, u_j) = 0, \forall i \neq j}} \max_{\substack{u \in \text{span}\{u_0, \dots, u_k\} \\ u \neq 0}} \frac{(Du, Du)}{(u, u)}. \quad (1.130)$$

In fact, the min-max is obtained for a corresponding eigenfunction. The above facts about λ_0 and λ_K can also be obtained from this formula and (1.123).

The normalized graph Laplacian that we have introduced here and whose properties we shall also investigate below is also called Tutte's Laplacian or the harmonic Laplacian (though with the opposite sign convention) in graph theory, and it should be distinguished from the algebraic or combinatorial Laplacian which is more commonly used in graph theory and investigated in Keller's contribution to this volume (Chap. 6, Sect. 1.3) where it is called the uniform Laplacian. That Laplacian is defined as

$$\bar{\Delta}v(x) := \sum_{y, y \sim x} v(y) - d_x v(x) = \sum_{y, y \sim x} (v(y) - v(x)), \quad (1.131)$$

that is, without the normalization factor $\frac{1}{d_x}$. That combinatorial Laplacian also encodes many important properties of graphs, and in particular, it leads to a trace formula. Here, however, we work with the normalized instead of the combinatorial Laplacian, because the former is the operator underlying random walks and diffusion processes and therefore also seems to be better adapted for our approach to discrete Ricci curvature. In particular, the reader should note that the Laplacian used by Keller is different from that employed here, the spectral bounds of Chap. 6, Sect. 1.3 are not directly comparable with those presented here. On the other hand, the Laplacian discussed in Baird's contribution (Chap. 7) is the same as ours.

1.6.2 Ricci Curvature and Clustering

In this section, we essentially describe the results of Jost and Liu [50]. As explained, in order to define Ricci curvature, we not only need a metric, but also a measure. Therefore, we recall the probability measures from (1.120)

$$m_x(y) = \begin{cases} \frac{1}{d_x} & \text{if } y \sim x; \\ 0 & \text{otherwise.} \end{cases} \quad (1.132)$$

We can interpret this in terms of a random walker that sits at x at time $t \in \mathbb{N}$ and then selects a neighbor of x with equal probability $\frac{1}{d_x}$ as the target of his walk at time $t + 1$.

Theorem 1.5 *On a locally finite graph $G = (V, E)$, we have for any pair of neighboring vertices x, y ,*

$$\kappa(x, y) \geq - \left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \wedge d_y} \right)_+ - \left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \vee d_y} \right)_+ + \frac{\sharp(x, y)}{d_x \vee d_y},$$

where we have put

$$d_x \wedge d_y := \min\{d_x, d_y\}, \quad d_x \vee d_y := \max\{d_x, d_y\}.$$

Remark For the case where $\sharp(x, y) = 0$, this result was obtained in [59]. For our purposes, however, the key point is to understand how the presence of triangles in a graph improves the lower Ricci bound.

Proof (Sketch of the Proof of Theorem 1.5) We first establish some notation. A vertex z is called a common neighbor of x and y if $z \sim x$ and $z \sim y$. It is called an exclusive neighbor of x if $z \sim x, z \not\sim y, z \neq y$.

We suppose w.l.o.g.,

$$d_x = d_x \vee d_y, \quad d_y = d_x \wedge d_y.$$

In order to estimate $\kappa(x, y)$ from below, we need a good transfer plan that moves m_x to m_y ; here is the idea.

1. Move the mass of $\frac{1}{d_x}$ from y to y 's exclusive neighbors;
2. Move a mass of $\frac{1}{d_y}$ from x 's exclusive neighbors to x ;
3. Fill gaps using the mass at x 's exclusive neighbors. Filling the gaps at common neighbors costs 2 and the one at y 's exclusive neighbors costs 3.

The question then is whether (1) and (2) can be realized. For (1), this means that the share of mass that y 's exclusive neighbors should receive, $1 - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_y}$ (the total mass minus what has to go to x or to the common neighbors of x and y) is at least what is originally at y , i.e.,

$$1 - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_y} \geq \frac{1}{d_x}, \quad \text{or} \quad 1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \wedge d_y} \geq 0, \quad (1.133)$$

recalling that we assumed $d_x \geq d_y$. In the situation depicted in Figs. 1.1 and 1.2 this is possible. But if, for instance, y had no exclusive neighbors, this would not be possible.

Fig. 1.1 Starting configuration; mass 0 at all vertices without number attached

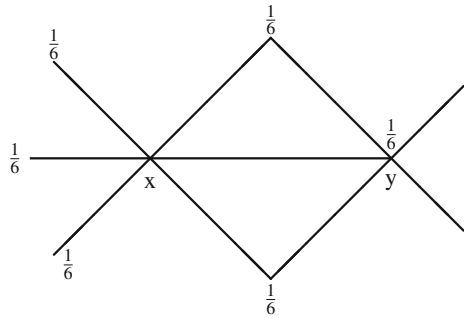


Fig. 1.2 Target configuration

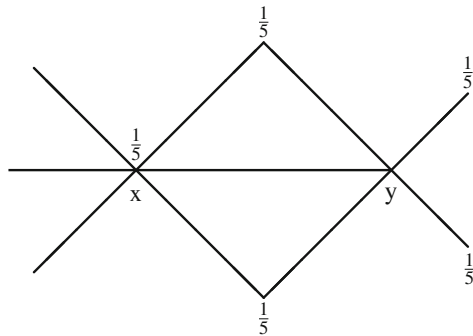
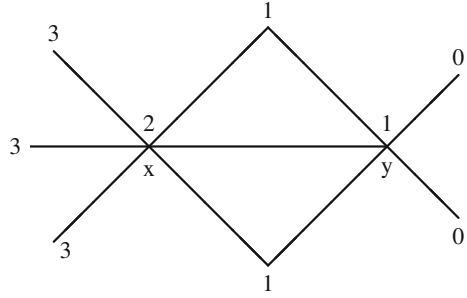


Fig. 1.3 Mass moved from vertices with larger value to those with smaller ones



Similarly, for step (2) we would need

$$1 - \frac{1}{d_x} - \frac{\sharp(x, y)}{d_x} \geq \frac{1}{d_y}, \text{ or } 1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \vee d_y} \geq 0. \quad (1.134)$$

The construction of the actual transport plan then needs to consider 3 cases according to whether the first two steps can be realized or not. For the details, we refer to [50]; suffice it here to consider the following 1-Lipschitz function as depicted in Fig. 1.3 and recall the duality formula (1.104).

That is, we put

$$f(z) = \begin{cases} 0, & \text{at } y\text{'s exclusive neighbors;} \\ 1, & \text{at } y \text{ or common neighbors;} \\ 2, & \text{at } x; \\ 3, & \text{at } x\text{'s exclusive neighbors,} \end{cases}$$

If there are no paths of length 1 between common neighbors and x 's exclusive neighbors, nor paths of length 1 or 2 between the exclusive neighbors of x and y , we have by Kantorovich duality,

$$\begin{aligned} W_1(m_x, m_y) &\geq \frac{1}{d_x} [f(y) + 3(d_x - 1 - \sharp(x, y)) + \sharp(x, y)] - \frac{1}{d_y} (f(x) + \sharp(x, y)) \\ &= 3 - \frac{2}{d_x} - \frac{2}{d_y} - \frac{\sharp(x, y)}{d_y} - \frac{2\sharp(x, y)}{d_x}. \end{aligned}$$

That is, in this case, the estimate in our theorem should be an equality. In other cases, f is not optimal for Kantorovich duality, and the estimate can be further improved. In other words, paths of length 1 or 2 between neighbors of x and y also affect the curvature. Thus, not only triangles, but also quadrangles and pentagons (but not polygons with more edges) influence Ricci curvature. This aspect has been investigated in detail in [21].

In conclusion,

$$\kappa(x, y) \geq -2 + \frac{2}{d_x} + \frac{2}{d_y} + \frac{\sharp(x, y)}{d_x \wedge d_y} + \frac{2\sharp(x, y)}{d_x \vee d_y}.$$

So much for a sketch of the proof.

Looking at examples, the simplest one is the lattice \mathbb{Z}^d with edges between vertices of Euclidean distance 1, that is, the Cayley graph for the abelian group \mathbb{Z}^d with the obvious set of generators. Here, when we look at the neighborhoods of two lattice points x, y of distance 1, we get the optimal transport plan by moving each vertex in the neighborhood of x by a distance of 1 to the nearest one among the neighbors of y . Therefore, $W_1(m_x, m_y) = 1$, and consequently, for this graph, the Ricci curvature vanishes.

The lower bound of Theorem 1.5 is sharp both for complete graphs and for trees, as we shall now explain. On a complete graph \mathcal{K}_n ($n \geq 2$) with n vertices, $\sharp(x, y) = n - 2$ for any x, y . Hence the inequality

$$\kappa(x, y) \geq \frac{n-2}{n-1}$$

is sharp.

For some other graphs, the lower bound of Theorem 1.5 is not sharp, however. For instance, for polyhedral surfaces, recently Loisel and Romon [63] obtained more precise results.

We shall now show that trees also attain the lower bound of Theorem 1.5. This coincides with the geometric intuition of Ricci curvature developed in Sect. 1.3.4.1. Since trees have the fastest volume growth rate, they should have the lowest Ricci curvature.

Proposition 1.1 *On a tree $T = (V, E)$, for any neighboring x, y ,*

$$\kappa(x, y) = -2 \left(1 - \frac{1}{d_x} - \frac{1}{d_y} \right)_+. \quad (1.135)$$

Proof We shall prove that $1 + 2 \left(1 - \frac{1}{d_x} - \frac{1}{d_y} \right)_+$ is also a lower bound of W_1 . If x or y has degree 1, say $d_x = 1$, so that y is its only neighbor, then obviously $W_1(m_x, m_y) = 1$. So only the case $1 - \frac{1}{d_x} - \frac{1}{d_y} \geq 0$ remains.

We consider the 1-Lipschitz function

$$f(z) = \begin{cases} 0, & \text{if } z \sim y, z \neq x; \\ 1, & \text{if } z = y; \\ 2, & \text{if } z = x; \\ 3, & \text{if } z \sim x, z \neq y. \end{cases} \quad (1.136)$$

Since on a tree, there is only one path joining two vertices, there is no further path between neighbors of x and y . So f can be easily extended to a 1-Lipschitz function on the whole graph. Then by Kantorovich duality, we have

$$\begin{aligned} W_1(m_x, m_y) &\geq \frac{1}{d_x}(3(d_x - 1) + 1) - \frac{1}{d_y} \cdot 2 \\ &= 3 - \frac{2}{d_x} - \frac{2}{d_y}. \end{aligned} \quad (1.137)$$

We can also relate this to the above heuristic discussion of the relation between Ricci curvature and the relative volume of the intersection of balls. In fact, $\sharp(x, y)/d_x \vee d_y$ is $m_x \wedge m_y(G) := m_x(G) - (m_x - m_y)_+(G)$, i.e. the intersection measure of m_x and m_y . The vertices x_1 that satisfy $x_1 \sim x$, $x_1 \sim y$ constitute the intersection of the unit metric spheres centered at x and y , resp.

We also have an easy upper bound for the Ricci curvature of a graph.

Theorem 1.6 *On a locally finite graph $G = (V, E)$, for any neighboring x, y , we have*

$$\kappa(x, y) \leq \frac{\sharp(x, y)}{d_x \vee d_y}. \quad (1.138)$$

Proof All masses, except those at common neighbors, have to be moved at least a distance 1. Hence

$$W_1(m_x, m_y) \geq \left(1 - \frac{\sharp(x, y)}{d_x \vee d_y}\right) \times 1,$$

and the conclusion follows from the definition of $\kappa(x, y)$.

We now return to graphs that may contain triangles. Watts-Strogatz [85] have introduced the local clustering coefficient

$$c(x) := \frac{1}{d_x(d_x - 1)} \sum_{y, y \sim x} \sharp(x, y) \quad (1.139)$$

in order to measure the extent to which neighbors of x are directly connected. Expressed in words,

$$c(x) = \frac{\text{number of realized edges between neighbors of } x}{\text{number of possible edges between neighbors of } x}. \quad (1.140)$$

This clustering coefficient is an important quantity in network analysis. For instance, in social networks where the vertices represent individuals and the edges friendship relations, the question addressed by the clustering coefficient is “How many of the friends of my friends are also my friends?”.

We may also consider this local clustering coefficient as an average over the $\sharp(x, y)$ for the neighbors of x . As such an average, we should also try to compare it to averaged Ricci curvature. In other words, we should consider the discrete version of scalar curvature,

$$\kappa(x) := \frac{1}{d_x} \sum_{y, y \sim x} \kappa(x, y). \quad (1.141)$$

This scalar curvature $\kappa(x)$ and the local clustering coefficient $c(x)$ then control each other.

Corollary 1.3 *With $D(x) := \max_{y, y \sim x} d_y$, we have*

$$\frac{d_x - 1}{d_x} c(x) \geq \kappa(x) \geq -2 + \frac{d_x - 1}{d_x \vee D(x)} c(x).$$

Proof From Theorems 1.5 and 1.6.

1.6.2.1 Stochastic Processes on Graphs

As a preparation, we consider a graph with a lower Ricci bound

$$\kappa(x, y) \geq k \text{ for all } x \sim y, \quad (1.142)$$

or equivalently,

$$W_1(m_x, m_y) \leq (1 - k)d(x, y) = 1 - k \text{ for all } x \sim y. \quad (1.143)$$

We shall now interpret this in probabilistic terms as a path coupling criterion for random walks. This translates a lower bound of the Ollivier-Ricci curvature into a control on the expectation value of the distance between two coupled random walks. The general tool is the Bubley-Dyer Theorem which tells us that when the contraction property (1.143) holds for the measures m_x , then it also holds for any other pair of measures (see [23] or [56, 72]).

Theorem 1.7 *For a probability measure μ , we put*

$$\mu P(\cdot) := \sum_x \mu(x) m_x(\cdot). \quad (1.144)$$

If (1.143) holds for each pair $x \sim y \in V$, then also for any probability measures μ and ν on V

$$W_1(\mu P, \nu P) \leq (1 - k)W_1(\mu, \nu). \quad (1.145)$$

The important consequence for us is that we can iterate (1.143) during a random walk. Initially, two walkers are starting at x and y , with transition probabilities m_x and m_y . With δ_x the Dirac measure at x , we have after the first step $\delta_x P^1(\cdot) := \delta_x P(\cdot) = m_x(\cdot)$. By iteration the distribution of a t -step random walk starting from x with a transition probability m_x becomes

$$\delta_x P^t(\cdot) = \sum_{x_1, \dots, x_{t-1}} m_x(x_1) m_{x_1}(x_2) \cdots m_{x_{t-1}}(\cdot) \quad (1.146)$$

for $t > 1$.

Theorem 1.7 therefore implies that when (1.142) and hence (1.143) holds, then for any t and any \bar{x}, \bar{y} , not necessarily neighbors,

$$W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t) \leq (1 - k)^t d(\bar{x}, \bar{y}). \quad (1.147)$$

In order to link this to Ricci curvature, we now consider two random walks (\bar{X}_t, \bar{Y}_t) with distributions $\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t$ that are coupled in the sense that the joint probabilities satisfy

$$p(\bar{X}_t = \bar{x}', \bar{Y}_t = \bar{y}') = \xi_t^{\bar{x}, \bar{y}}(\bar{x}', \bar{y}'),$$

where $\xi_t^{\bar{x}, \bar{y}}(\cdot, \cdot)$ is the optimal coupling of $\delta_{\bar{x}} P^t$ and $\delta_{\bar{y}} P^t$ as in the definition of the Wasserstein distance W_1 . The term $W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t)$ then becomes the expectation value of the distance $\mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t)$ between the coupled random walks \bar{X}_t and \bar{Y}_t .

Corollary 1.4 *If (1.142) holds, then for any $\bar{x}, \bar{y} \in V$,*

$$\mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t) = W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t) \leq (1 - k)^t d(\bar{x}, \bar{y}). \quad (1.148)$$

1.6.2.2 Weighted and Neighborhood Graphs

Following [11], we now translate the properties of random walks into geometric structures, the neighborhood graphs. In Sect. 1.6.2.3, we shall then use this construct to derive eigenvalue bounds in terms of lower Ricci curvature bounds on graphs.

For this purpose, we shall need to work with a somewhat more general class of graphs than before. More precisely, we shall need to consider weighted graphs, and also allow for the possibility of self-loops. That is, for any $x, y \in V$, not necessarily different, we have a symmetric, nonnegative connection weight

$$w_{xy} = w_{yx} \geq 0. \quad (1.149)$$

We can then declare x and y to be neighbors, $x \sim y$, iff $w_{xy} > 0$. Of course, the unweighted graphs that we have considered before constitute the special cases where $w_{xy} = 1$ iff $x \sim y$ and $w_{xy} = 0$ else. As mentioned, here, we also allow

for the possibility of self-loops, that is, vertices x with $w_{xx} > 0$. A weighted graph is connected if for every $x, y \in V$, there exists a path $x_0 = x, x_1, \dots, x_n = y$ with $w_{x_{i-1}x_i} > 0$ for $i = 1, \dots, n$.

Remark Of course, one could also allow for non-symmetric or negative weights. The spectrum of non-symmetric graphs was systematically investigated in [10], and some results on graphs with possibly negative connection weights can be found, for instance, in [12, 13]. For our present purposes, however, the class of weighted graphs satisfying (1.149) suffices.

We now need to adapt some of the preceding constructions and results to weighted graphs. First of all, we now define the measure m_x by

$$m_x(y) := \frac{w_{xy}}{d_x}, \text{ where now } d_x := \sum_y w_{xy}. \quad (1.150)$$

Of course, all this and what follows reduces to our previous definitions for unweighted graphs. We can again consider $m_x(y)$ as the probability that a random walker starting at x moves to y in one time step. Since now possibly $m_x(x) > 0$, because there might be a self-loop at x , the random walker might now be lazy and simply stay at x .

Again, the L^2 -product is given by

$$(u, v) = \sum_x d_x u(x) v(x). \quad (1.151)$$

The Laplacian now is

$$\Delta v(x) = \frac{1}{d_x} \sum_y w_{xy} v(y) - v(x) = \sum_y m_x(y) v(y) - v(x). \quad (1.152)$$

As before, the Laplacian is self-adjoint and nonpositive so that, with the same conventions as before, the eigenvalues are nonnegative real numbers. We also have a version of Theorem 1.5 for weighted graphs, taken from [14].

Theorem 1.8 *On a weighted graph, we have for neighbors x, y*

$$\begin{aligned} \kappa(x, y) \geq & - \left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \vee \frac{w_{x_1 y}}{d_y} \right)_+ \\ & - \left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \right)_+ \\ & + \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} + \frac{w_{xx}}{d_x} + \frac{w_{yy}}{d_y}. \end{aligned}$$

Again, this inequality is sharp.

With the notation (1.144), i.e.,

$$\mu P(\cdot) = \sum_x \mu(x) m_x(\cdot),$$

the Dirac measure δ_x at x and $\delta_x P^1(\cdot) = \delta_x P(\cdot) = m_x(\cdot)$, the distribution of a t -step random walk starting at x with transition probability m_x becomes

$$\delta_x P^t(\cdot) = \sum_{x_1, \dots, x_{t-1}} m_x(x_1) m_{x_1}(x_2) \cdots m_{x_{t-1}}(\cdot) \quad (1.153)$$

for $t > 1$. The probability that the random walker moves from x to y in t steps then is

$$\delta_x P^t(y) = \begin{cases} \sum_{x_1, \dots, x_{t-1}} \frac{w_{xx_1}}{d_x} \frac{w_{x_1 x_2}}{d_{x_1}} \cdots \frac{w_{x_{t-1} y}}{d_{x_{t-1}}}, & \text{if } t > 1; \\ \frac{w_{xy}}{d_x}, & \text{if } t = 1. \end{cases} \quad (1.154)$$

We now define a family of graphs $G[t]$ for $t \geq 1$ whose weights equal the transition probabilities of the t -step random walks on the graph G .

Definition 1.11 The neighborhood graph $G[t] = (V, E[t])$ of the graph $G = (V, E)$ of order $t \geq 1$ is the weighted graph with vertex set V and edge weights

$$w_{xy}[t] := \delta_x P^t(y) d_x \quad (1.155)$$

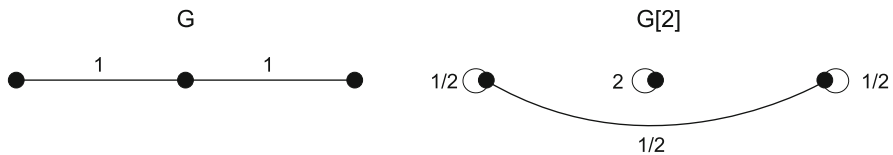
from (1.154).

Obviously, $G = G[1]$. Also, $w_{xy}[t] > 0$ if and only if there exists a path of length t between x and y in G . We also remark here, without exploring this further, that the discrete heat kernel $p_t(x, y)$ on G is

$$p_t(x, y) = \frac{w_{xy}[t]}{d_x d_y},$$

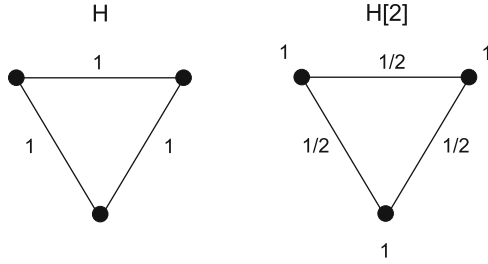
see for instance Grigor'yan [38].

Example 1.1 We consider the following two examples.



Note that the neighborhood graph $G[2]$ is disconnected. In fact the next lemma shows that this is the case because G is bipartite. Note furthermore that

$$E(G) \not\subseteq E(G[2]).$$



For this example we have $E(H) \subseteq E(H[2])$.

We now list some elementary properties of the neighborhood graph $G[t]$, its Laplacian $\Delta[t]$ and the latter's eigenvalues $\lambda_i[t]$, taken from [11, 14].

Lemma 1.1

- (i) t even: $G[t]$ is connected if G is not bipartite, but disconnected if G is bipartite. And $G[t]$ is not bipartite.
- (ii) t odd: $G[t]$ is always connected (since G is assumed to be connected) and $G[t]$ is bipartite iff G is bipartite.
- (iii) $d_x[t] = d_x$ for all $x \in V$. Hence the inner product (1.151) is the same on all the $G[t]$.
- (iv) The Laplacian on $G[t]$ is

$$\Delta[t] = -\text{id} + (\text{id} + \Delta)^t. \quad (1.156)$$

- (v) Therefore, for even t , the eigenvalues of $\Delta[t]$ satisfy

$$0 = \lambda_0[t] \leq \lambda_1[t] \leq \dots \leq \lambda_{N-1}[t] \leq 1. \quad (1.157)$$

- (vi) Let $d[t](x, y)$ be the distance on $G[t]$ defined as the smallest number of edges needed for a path connecting x and y (this is independent of the weights, except that vertices ξ and η are connected by an edge iff $w_{\xi\eta} > 0$). Then

$$\frac{1}{t}d(x, y) \leq d[t](x, y), \quad (1.158)$$

with the convention $d[t](x, y) = \infty$ if $G[t]$ is not connected and x and y are in different components. Conversely, if $E \subseteq E[t]$, then

$$d[t](x, y) \leq d(x, y). \quad (1.159)$$

Note that at the end of Sect. 1.6.1, we had observed that the largest eigenvalue is 2 for a bipartite graph. In (1.157), in contrast, that eigenvalues is bounded by 1

for even t . This discrepancy stems from the fact that for even t , the graph $G[t]$ has self-loops which were not permitted for the graph $G = G[1]$ in Sect. 1.6.1.

In [11], an important relation between the eigenvalues of the original graph G and those of its neighborhood graphs was found.

Proposition 1.2

(i) If $\lambda_1[t] \geq \mathcal{A}[t]$, then

$$1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - \mathcal{A}[t])^{\frac{1}{t}} \quad (1.160)$$

if t is even and

$$1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1 \quad (1.161)$$

if t is odd.

(ii) If $\lambda_{N-1}[t] \leq \mathcal{B}[t]$, then all eigenvalues of Δ are contained in

$$\left[0, 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}\right] \cup \left[1 + (1 - \mathcal{B}[t])^{\frac{1}{t}}, 2\right]$$

for even t , whereas

$$\lambda_{N-1} \leq 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}$$

for odd t .

That is, eigenvalues bounds on $G[t]$ translate into eigenvalue bounds on the original graph G . This is a powerful principle for estimating the eigenvalues of G as we shall see. As the neighborhood graphs constitute a geometric representation of the random walk on G , this can be seen as a scheme for translating properties of the random walk into eigenvalue bounds. The scheme itself is not new, but here we can offer an intuitive and easy to apply geometric version of it. Just read on to the next section.

1.6.2.3 Ricci Curvature and Eigenvalues of Graphs

In this section, we assume that the graph G is finite, that is, it has finitely many, say N , vertices, and then also finitely many edges. Here, we present the theory developed in [14] which partially builds upon the neighborhood graph concept of Bauer and Jost [11].

We now come to the estimates of the eigenvalues in terms of the Ricci curvature. We can build here upon well established relations between the coupling of stochastic processes and eigenvalue estimates, see [23, 28]. In this connection, Ollivier [70] showed

Theorem 1.9 *When we have a lower Ricci curvature bound*

$$\kappa(x, y) \geq k, \quad (1.162)$$

(in fact, it suffices to have this for all $x \sim y$), then

$$k \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 2 - k. \quad (1.163)$$

A problem with this estimate is that for most graphs, $k \leq 0$ in (1.162), so that (1.163) only yields a trivial estimate. We shall subsequently present the estimate of Bauer et al. [14] which is nontrivial for all connected finite graphs that are not bipartite. Nevertheless, in order to understand the relation between Ricci curvature and eigenvalues, let us derive (1.163) here.

Proof We consider the transition probability operator

$$\begin{aligned} P : L^2(G) &\rightarrow L^2(G) \\ Pf(x) &:= \sum_y f(y) m_x(y) = \sum_y f(y) \delta_x P(y). \end{aligned} \quad (1.164)$$

Then

$$P^t f(x) = \sum_y f(y) \delta_x P^t(y). \quad (1.165)$$

We construct a discrete time heat equation,

$$f(x, t+1) - f(x, t) = \Delta f(x, t), \quad (1.166)$$

where the initial state $f(x, 0) = f_1(x)$ satisfies $\Delta f_1(x) = -\lambda f_1(x) = Pf_1(x) - f_1(x)$ for $\lambda \neq 0$. By iteration, the solution of (1.166) is

$$f(x, t) = P^t f_1(x) = (1 - \lambda)^t f_1(x). \quad (1.167)$$

Then we have for any $\bar{x}, \bar{y} \in V$

$$\begin{aligned} |1 - \lambda|^t |f_1(\bar{x}) - f_1(\bar{y})| &= |f(\bar{x}, t) - f(\bar{y}, t)| \\ &= |P^t f_1(\bar{x}) - P^t f_1(\bar{y})| \\ &\leq \sum_{\bar{x}', \bar{y}'} |f(\bar{x}') - f(\bar{y}')| \xi_t^{\bar{x}, \bar{y}}(\bar{x}', \bar{y}') \\ &\leq \text{Lip}(f_1) \mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t) \\ &\leq \text{Lip}(f_1) (1 - k)^t d(\bar{x}, \bar{y}). \end{aligned}$$

Here, $\text{Lip}(f)$ is always finite since the underlying space V is a finite set. In the last inequality we used Corollary 1.4.

Since the eigenfunction f_1 for the eigenvalue λ is orthogonal to the constant function, i.e. $(f_1, \mathbf{1}) = 0$, we can always find $x_0, y_0 \in V$ such that $|f_1(x_0) - f_1(y_0)| > 0$. It follows that

$$\varepsilon < \left(\frac{1-k}{|1-\lambda|} \right)^t \text{Lip}(f_1) d(x_0, y_0)$$

for some positive ε . This, however, leads to a contradiction when $t \rightarrow \infty$, unless

$$\frac{1-k}{|1-\lambda|} \geq 1. \quad (1.168)$$

(1.168) easily implies (1.163).

Of course, the preceding proof is analogous to methods familiar in Riemannian geometry. More precisely, the solution of the heat equation

$$\frac{\partial f(x, t)}{\partial t} = \Delta f(x, t) \quad (1.169)$$

on a Riemannian manifold with the eigenfunction f_1 as the initial value is $f(x, t) = f_1(x)e^{-\lambda t}$, containing information about both the eigenvalue λ and the eigenfunction $f_1(x)$. The calculation in the preceding proof then is the discrete analogue of the gradient estimate for the solution of the heat equation in Riemannian geometry.

This discrete gradient estimate is provided in

Theorem 1.10 *On a graph G , the following are equivalent:*

- (i) *The Ricci curvature is bounded from below by k , i.e. $\kappa(x, y) \geq k$, for all $x \sim y$;*
- (ii) *$|P^t f(x) - P^t f(y)| \leq (1-k)^t \text{Lip}(f)$ holds for any function f , $x \sim y$ and $t \in \mathbb{Z}^+$.*

Proof We have already shown that (i) implies (ii). The reverse direction follows from the Kantorovich duality (1.104).

We shall now show how the Ricci geometry of neighborhood graphs can improve the estimate of Theorem 1.9 and in fact obtain an estimate that is nontrivial for any graph that is not bipartite, following [14].

Lemma 1.2 *Let k be a lower bound of κ on G . If $E \subseteq E[t]$, then the curvature $\kappa[t]$ of the neighborhood graph $G[t]$ satisfies*

$$\kappa[t](x, y) \geq 1 - t(1-k)^t, \quad \forall x, y \in V. \quad (1.170)$$

Proof From Lemma 1.1 (vi) and Corollary 1.4 and using that the transportation distance (1.102) is linear in the graph distance $d(\cdot, \cdot)$, we obtain

$$\begin{aligned} W_1^{d[t]}(\delta_x P^t, \delta_y P^t) &\leq W_1^d(\delta_x P^t, \delta_y P^t) \\ &\leq (1-k)^t d(x, y) \\ &\leq t(1-k)^t d[t](x, y). \end{aligned}$$

By the definition of the Ricci curvature, we obtain (1.170).

We can now see the upper bound of the largest eigenvalue in Theorem 1.9. W.l.o.g. $k > 0$, in which case $E \subset E[t]$. From Lemma 1.2 and $\lambda_1 \geq k$, we know on $G[t]$,

$$\lambda_1[t] \geq 1 - t(1-k)^t.$$

Then with Proposition 1.2 (i), for even t ,

$$\lambda_{N-1} \leq 1 + t^{\frac{1}{t}}(1-k).$$

Letting $t \rightarrow +\infty$ yields $\lambda_{N-1} \leq 2 - k$, indeed.

The neighborhood graph technique then leads to the following generalization of Theorem 1.9, the main result of Bauer [14].

Theorem 1.11 *Let $k[t]$ be a lower bound of Ollivier-Ricci curvature of the neighborhood graph $G[t]$. Then for all $t \geq 1$ the eigenvalues of Δ on G satisfy*

$$1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \dots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}. \quad (1.171)$$

If G is not bipartite, then for all sufficiently large t , $k[t] > 0$, and hence (1.171) is nontrivial in the sense that the lower bound is positive and the upper bound is < 2 .

1.6.3 Curvature Dimension Inequality and Eigenvalue Ratios

In Sect. 1.5.3 we introduced Bakry and Emery's curvature dimension inequality in a general setting. Here we will discuss the curvature dimension inequality for graphs. Apparently, the first paper on this subject is [77]. We also mention the recent contribution [1] where a concept of coarse Ricci curvature from the Bakry-Emery perspective is developed.

For simplicity, we will restrict ourselves to the graph Laplace operator Δ . We recall the definition of the curvature dimension inequality from Sect. 1.5.3:

Definition 1.12 We say that a graph G satisfies the curvature dimension inequality $CD(n, K)$ if, for any function f

$$\Gamma_2(f) \geq \frac{1}{n}(\Delta f)^2 + K\Gamma(f).$$

Note that for graphs Γ is given by

$$\Gamma(f, g)(x) = \frac{1}{2d_x} \sum_{y \sim x} w_{xy} (f(y) - f(x))(g(y) - g(x))$$

and Δ is given by Eq. (1.119). To ease notations, we assume that the graph G is finite and unweighted in the remaining part of this section. In the graph setting, the curvature dimension inequality was studied in [50]. We proved the following theorem, thereby generalizing an earlier result of Lin and Yau [59].

Theorem 1.12 *Every graph satisfies $CD(2, t(x) - 1)$, where*

$$t(x) = \frac{1}{2} \min_{y, y \sim x} \left(\frac{4}{d_x} + \frac{\sharp(x, y)}{D(x)} \right).$$

Again, the presence of triangles play a crucial role for the lower curvature bound.

Bakry-Emery gradient estimates for a solution u of the continuous time heat equation $(\Delta - \partial_t)u = 0$ still hold for a graph G satisfying $CD(n, K)$. Actually, if we ignore the role of dimension in Definition 1.12, that is, taking $n = \infty$, the CD -inequality is characterized by such kind of gradient estimates. Let us denote by $u(x, t) = P_t f(x)$ a solution of the heat equation with $u(\cdot, 0) = f(\cdot)$. (Recall that we used $P^t f$ for solutions of discrete time heat equations in Sect. 1.6.2.3.)

Theorem 1.13 *On a graph G , the following are equivalent:*

- (i) $CD(\infty, K)$ holds;
- (ii) $\Gamma(P_t f) \leq e^{-2Kt} P_t(\Gamma(f))$ holds for all $t > 0$ and all functions f .

Proof (i) \Rightarrow (ii): For $0 \leq s \leq t$, define $F(s) := e^{-2Ks} P_s(\Gamma(P_{t-s} f))$ and calculate

$$\frac{d}{ds} F(s) = 2e^{-2Ks} P_s(\Gamma_2(P_{t-s} f) - K\Gamma(P_{t-s} f)) \geq 0.$$

Hence $F(s) \leq F(t)$.

(ii) \Rightarrow (i): Employ the fact that $P_t f = f + t\Delta f + o(t)$ to look at the gradient estimate at $t = 0$.

For more details of the proof, see Proposition 3.3 in Bakry [3] or [61]. Note that the proof does not require the diffusion property (Definition 1.10).

We should compare this result with Theorem 1.10; the latter can thus be seen as a time-discrete version of the Bakry-Emery gradient estimate.

Partially building upon this characterization of the CD-inequality, in [61] the following eigenvalue ratio estimates were obtained.

Theorem 1.14 *There exists an absolute constant C , such that for any graph G satisfying $CD(\infty, 0)$ and any natural number k ,*

$$\lambda_k \leq Cd_G k^2 \lambda_1, \quad (1.172)$$

where $d_G := \max_{x \in V} d_x$.

We remark that this estimate does not depend on the size of G . Such dimension-free eigenvalue ratio estimates also hold in the continuous category, see [60], improving the previous results of Funano and Shioya [33, 34]. Examples in [61] show that in (1.172) the order of k is optimal and the dependence on d_G is necessary and optimal.

We shall not elaborate on the proof for Theorem 1.14 here, but only discuss an interesting application for the analysis of spectral clustering algorithms. Spectral clustering algorithms are very powerful tools for data mining, see e.g. [54, 55, 82]. Such algorithms typically consist of two steps. In the first step, the first $k + 1$ eigenfunctions of Δ are used to provide coordinates for the vertices of a graph G , thereby embedding G into \mathbb{R}^{k+1} . The second step consists in partitioning the vertex set V into small groups via the Euclidean metric (or the spherical metric after normalization). The output will be a partition S_0, S_2, \dots, S_k of V , such that each S_i has small expansion. Here, the expansion $\phi(S_i)$ of S_i is defined as

$$\phi(S_i) := \frac{|E(S_i, V \setminus S_i)|}{\text{vol}(S_i)}, \quad (1.173)$$

where $|E(S_i, V \setminus S_i)| := \sum_{x \in S_i, y \in V \setminus S_i} 1$, and $\text{vol}(S_i) := \sum_{x \in S_i} d_x$. In fact, the underlying mathematical problem is to find the $(k + 1)$ -partition of V that attains

$$h_k := \min_{S_0, \dots, S_k} \max_{1 \leq i \leq k} \phi(S_i), \quad (1.174)$$

where the minimum is taken over all collections of $k + 1$ non-empty, mutually disjoint subsets $\{S_i\}_{i=0}^k$ with $\bigcup_{i=0}^k S_i = V$. Roughly speaking, in the algorithm, one tries to use the solutions of the optimization problem in (1.130), i.e. the eigenfunctions of Δ , to approximate the solution of the optimization problem (1.174). Therefore, the efficiency of the clustering algorithm depends on the relation between λ_k and h_k .

Solving a conjecture of Miclo [66], Lee et al. [55] proved the following so-called higher order Cheeger inequalities. There exists an absolute constant C such that for any graph G and all natural numbers k ,

$$\frac{\lambda_k}{2} \leq h_k \leq Ck^2 \sqrt{\lambda_k}. \quad (1.175)$$

Roughly speaking, h_k lies between λ_k and $\sqrt{\lambda_k}$. It is an interesting question for which kind of graphs, h_k is equivalent to λ_k^α up to constants, for some $\alpha \in [1/2, 1]$. This would indicate the efficiency of the spectral clustering algorithm when applied to different kinds of graphs.

In [61], the following higher order Buser type inequalities were derived from Theorem 1.14.

Theorem 1.15 *There exists an absolute constant C , such that for any graph G satisfying $CD(\infty, 0)$ and any natural number k ,*

$$C(d_G k)^{-1} \sqrt{\lambda_k} \leq h_2 \leq h_k. \quad (1.176)$$

This implies that h_k is equivalent to $\sqrt{\lambda_k}$ up to constants for graphs satisfying $CD(\infty, 0)$. Therefore, the curvature condition helps to identify a class of graphs on which the algorithm performs poorly. A deeper understanding of the structure of non-negatively curved graphs would provide further insight into the spectral clustering algorithm. In this respect, it is shown that the Cartesian product of two regular graphs satisfying $CD(\infty, 0)$ satisfies again $CD(\infty, 0)$ in [61]. (If we consider the non-normalized Laplacian in the CD-inequality, the regularity constraints are not needed.)

1.6.4 Exponential Curvature Dimension Inequality on Graphs

We will now discuss a modification of the curvature dimension inequality that was introduced in [15].

Definition 1.13 We say that a graph G satisfies the *exponential curvature dimension inequality* at the point $x \in V$, $CDE(x, n, K)$ if for any positive function $f : V \rightarrow \mathbb{R}$ such that $(\Delta f)(x) < 0$ we have

$$\Gamma_2(f)(x) - \Gamma\left(f, \frac{\Gamma(f)}{f}\right)(x) \geq \frac{1}{n}(\Delta f)(x)^2 + K\Gamma(f)(x).$$

We say that $CDE(n, K)$ is satisfied if $CDE(x, n, K)$ is satisfied for all $x \in V$. It is useful to note that

$$\Gamma_2(f) - \Gamma\left(f, \frac{\Gamma(f)}{f}\right) = \frac{1}{2}\Delta\Gamma(f) - \Gamma\left(f, \frac{\Delta(f^2)}{2f}\right), \quad (1.177)$$

and we define

$$\tilde{\Gamma}_2(f) = \Gamma_2(f) - \Gamma\left(f, \frac{\Gamma(f)}{f}\right).$$

This definition might seem to be rather artificial and not well motivated. However the exponential curvature dimension inequality is quite natural in several respects. In [15] it was shown that for diffusion semigroups defined in Definition 1.10 (and thus in particular of the Laplace-Beltrami operator), the exponential curvature dimension inequality is in fact weaker than the original curvature dimension inequality.

Theorem 1.16 *If the semigroup generated by L is a diffusion semigroup, then the condition $CD(n, K)$ implies $CDE(n, K)$.*

An advantage of the exponential curvature dimension inequality over other curvature notions on graphs is that the curvature can be arbitrarily negative. In contrast, for Ollivier's Ricci curvature and the classical curvature dimension inequality the curvature is always bounded from below, see Theorems 1.5 and 1.12. There are other properties of the exponential curvature dimension inequality that make it a useful curvature notion. Here however, we only mention that it is the right curvature notion for Li-Yau gradient estimates on graphs. We will discuss this issue in the next section.

1.6.5 Li-Yau Gradient Estimate on Graphs and Its Applications

In this section we discuss the gradient estimates obtained in [15]. Bakry and Ledoux's general result on gradient estimates [6], discussed in Sect. 1.5.3, cannot be applied to graphs. The reason is that the graph Laplace operator does not generate a diffusion semigroup. However in [15] it was observed that, on graphs, for the choice of $\Phi(f) = \sqrt{f}$ a key formula similar to a combination of (1.116) and (1.117) still holds:

$$2\sqrt{f}\Delta\sqrt{f} = \Delta f - 2\Gamma(\sqrt{f}) \quad (1.178)$$

In fact in the proof of the gradient estimate this simple equality will take over the role of the key identity

$$\Delta \log u = \frac{\Delta u}{u} - |\nabla \log u|^2 \quad (1.179)$$

in the proofs in the continuous setting. In the special case of a finite graph with non-negative curvature we have the following gradient estimate:

Theorem 1.17 *Let G be a finite graph satisfying $CDE(n, 0)$, and let u be a positive solution to the heat equation on G . Then for all $t > 0$*

$$\frac{\Gamma(\sqrt{u})}{u} - \frac{\partial_t(\sqrt{u})}{\sqrt{u}} \leq \frac{n}{2t}. \quad (1.180)$$

After having established the right notion of curvature and having identified the key identity (1.179), we can now give a rather simple proof of this theorem. But first we state a simple lemma from [15].

Lemma 1.3 *Let $G(V, E)$ be a (finite or infinite) graph, and let $g, F : V \times [0, T] \rightarrow \mathbb{R}$ be functions. Suppose that $g \geq 0$, and F has a local maximum at $(x^*, t^*) \in V \times]0, T]$. Then*

$$\mathcal{L}(gF)(x^*, t^*) \leq (\mathcal{L}g)F(x^*, t^*),$$

where $\mathcal{L} := \Delta - \partial_t$.

Proof (Proof of the Theorem) Let

$$F = t \left(\frac{2\Gamma(\sqrt{u})}{u} - \frac{2\partial_t(\sqrt{u})}{\sqrt{u}} \right). \quad (1.181)$$

Fix an arbitrary $T > 0$. Our goal is to show that $F(x, T) \leq n$ for every $x \in V$. Let (x^*, t^*) be a maximum point of F in $V \times [0, T]$. We may assume $F(x^*, t^*) > 0$. Hence $t^* > 0$. Moreover, by identity (1.178) which is true both in the continuous and the discrete setting, we know that

$$F = t \left(\frac{2\Gamma(\sqrt{u})}{u} - \frac{\Delta u}{u} \right) = t \cdot \frac{-2\Delta\sqrt{u}}{\sqrt{u}}, \quad (1.182)$$

where we used the fact that $\mathcal{L}u = 0$ (recall that $\mathcal{L} = \Delta - \partial_t$) which implies

$$2 \frac{\partial_t \sqrt{u}}{\sqrt{u}} = \frac{\partial_t u}{u} = \frac{\Delta u}{u}. \quad (1.183)$$

We conclude from (1.182) that

$$(\Delta\sqrt{u})(x^*, t^*) < 0. \quad (1.184)$$

In what follows all computations are understood to take place at the point (x^*, t^*) . We apply Lemma 1.3 with the choice $g = u$. This gives

$$\begin{aligned} \mathcal{L}(u) \cdot F &\geq \mathcal{L}(u \cdot F) = \mathcal{L}(t^* \cdot (2\Gamma(\sqrt{u}) - \Delta u)) \\ &= t^* \cdot \mathcal{L}(2\Gamma(\sqrt{u}) - \Delta u) - (2\Gamma(\sqrt{u}) - \Delta u), \end{aligned}$$

where we used (1.182) and the definition of \mathcal{L} . We know that $\mathcal{L}(u) = 0$. Also, since Δ and \mathcal{L} commute, $\mathcal{L}(\Delta u) = 0$. So we are left with

$$\begin{aligned} \frac{uF}{t^*} &= 2\Gamma(\sqrt{u}) - \Delta u \geq t^* \cdot \mathcal{L}(2\Gamma(\sqrt{u})) \\ &= t^* \cdot (2\Delta\Gamma(\sqrt{u}) - 4\Gamma(\sqrt{u}, \partial_t \sqrt{u})) = 4t^* \cdot \widetilde{F}_2(\sqrt{u}). \end{aligned} \quad (1.185)$$

The last equality is true by (1.177) and (1.183). By (1.184) and the $CDE(n, 0)$ -inequality applied to $\sqrt{u}(\cdot, t^*)$ we get

$$\frac{uF}{t^*} \geq \frac{4t^*}{n} (\Delta(\sqrt{u}))^2 \stackrel{(1.182)}{=} \frac{t^*}{n} \left(-\frac{\sqrt{u}F}{t^*} \right)^2 = \frac{u}{nt^*} F^2.$$

Thus $F \leq n$ at (x^*, t^*) as desired.

Let us briefly discuss the differences between the gradient estimates (1.114) on Riemannian manifolds and (1.180) on graphs. There exists a one parameter family of gradient estimates \mathcal{G}_p for $p > 0$,

$$\frac{|\nabla u^p|^2}{u^{2p}} - \frac{\partial_t u}{u} \leq \frac{n}{2t}.$$

Note that the larger p , the stronger \mathcal{G}_p is. In the Riemannian case, the original Li-Yau inequality (1.114) corresponds to \mathcal{G}_1 . It is known that the Li-Yau gradient estimate is sharp, that is $p = 1$ is optimal on Riemannian manifolds. In the discrete setting it was shown in [15] that \mathcal{G}_p cannot hold for any graph with $p > 0.5$. Thus the gradient estimate (1.180) which corresponds to $p = 0.5$ is in this sense, although weaker than its continuous counterpart (1.114), optimal.

For simplicity of exposition, we only present the most simplest case of the Li-Yau gradient estimate on graphs here. In [15] local and global gradient estimates were also obtained for graphs with general lower curvature bounds and more general operators than the Laplacian. For possible applications and further generalizations of the Li-Yau gradient estimates on graphs, including heat kernel estimates and Harnack inequalities we refer the reader to [15, 16, 73]. See also [67] for related work.

1.6.6 Applications to Network Analysis

Some of the preceding tools are quite useful for the analysis of empirical networks. Depending on the data, such networks can be represented by unweighted or weighted and possibly also directed graphs. For instance, one can then study the eigenvalue spectrum. More in line with the present contribution, one can also look at the distribution of their Ricci curvatures. The Forman-Ricci curvature is computationally easiest, and there are systematic correlations with the Ollivier-Ricci curvature. This is an ongoing research project with Samal, Saucan, Sreejith, Mohanraj, and Weber, see [78–80, 86–88].

1.6.7 Other Curvature Notions for Graphs

At the end, we briefly mention some curvature notions for graphs other than Ricci curvature.

For the notion of combinatorial curvature, we need to fill in faces into the graph. We therefore assume that the possibly infinite graph G is embedded into a 2-manifold $S(G)$ such that each face is homeomorphic to a closed disk with finite edges as the boundary. For instance, G could be a planar graph, that is, a graph embedded into the plane. Therefore, we call such a $G = (V, E, F)$ that can be embedded into a 2-manifold as described, with its sets of vertices V , edges E , and faces F , a *sempianar graph*. For each vertex $x \in V$, the combinatorial curvature at x is defined as

$$\Phi(x) = 1 - \frac{d_x}{2} + \sum_{f \ni x} \frac{1}{|f|}, \quad (1.186)$$

where, as before, d_x is the degree of the vertex x , whereas $|f|$ is the degree of the face f . The sum is taken over all faces incident to x (i.e. $x \in f$). For more details on the combinatorial curvature, see the contribution of Keller in this volume (Chap. 6).

When we replace each face of G with a regular polygon of side lengths one and glue them along the common edges and equip the polygonal surface $S(G)$ with the resulting metric structure, then (1.186) simply measures the difference of 2π and the total angle Σ_x at the vertex x ,

$$2\pi \Phi(x) = 2\pi - \Sigma_x. \quad (1.187)$$

Let $\chi(S(G))$ denote the Euler characteristic of the surface $S(G)$. We then have the Gauss-Bonnet formula of G of DeVos and Mohar [27],

$$\sum_{x \in G} \Phi(x) \leq \chi(S(G)), \quad (1.188)$$

whenever $\sum_{x \in G: \Phi(x) < 0} \Phi(x)$ converges. Thus, the combinatorial curvature captures a topological property of sempianar graphs.

We can also compare the combinatorial curvature with another version of curvature naturally obtained from the surface $S(G)$, its generalized sectional (Gaussian) curvature. It turns out that the sempianar graph G has nonnegative combinatorial curvature precisely if the polygonal surface $S(G)$ is an Alexandrov space with nonnegative sectional curvature, i.e. $\text{Sec } S(G) \geq 0$ (or $\text{Sec}(G) \geq 0$ for short). This principle is systematically explored in [40].

Here, a metric space (X, d) is called an Alexandrov space if it is a geodesic space (i.e. each pair of points in X can be joined by a shortest path called a geodesic) and locally satisfies the Toponogov triangle comparison. Essentially, nonnegative curvature in the present context means that the total angles of geodesic triangles

are at least 2π . Upper curvature bounds, like nonpositive sectional curvature can be interpreted as convexity properties for the distance function. The basic geometric setting for Alexandrov curvature type bounds is this. One starts with a geodesic triangle, that is, three points $a_1, a_2, a_3 \in X$ with mutual distances $d(a_i, a_j)$ that satisfy the triangle inequality, as it befits a metric space. And since X is a geodesic space, they can be pairwise joined by shortest geodesics. Such a configuration is called a *triangle*. Such a triangle, however, does not yet possess nontrivial geometric content, as we can find a comparison triangle in any surface C_K of constant curvature K (with the only restriction that for positive K , there is some restriction on the size of our triangle so that it fits into a hemisphere of C_K) with the same side lengths. (Thus, for positive K , C_K is a sphere of curvature K , for $K = 0$, it is the Euclidean plane, and for $K < 0$, it is a scaled version of the hyperbolic plane with curvature K .) That is, we choose points $\bar{a}_1, \bar{a}_2, \bar{a}_3 \in C_K$ with

$$d(a_i, a_j) = d_K(\bar{a}_i, \bar{a}_j), \quad \text{for } i, j = 1, 2, 3,$$

where d_K is the distance in C_K . In order to get at specific properties of (X, d) , we need to consider a fourth point. Alexandrov takes the midpoint a_4 of a_1 and a_2 , that is,

$$d(a_4, a_1) = d(a_4, a_2) = \frac{1}{2}d(a_1, a_2).$$

In particular, a_4 sits on a shortest geodesic from a_1 to a_2 . One then compares $d(a_3, a_4)$ with $d_K(a_3, a_4)$. When the former is smaller (larger) than the latter for every such triangle, one says that (X, d) has curvature smaller (larger) than K . In particular, an upper curvature bound implies uniqueness of the geodesic from a_1 to a_2 , as one readily observes. (Of course, we have to keep in mind here that for positive K , we had to restrict the size of our triangle, so that it could be realized inside a hemisphere of C_K .) Monographs on Alexandrov spaces are [20, 24]. In fact, there was an earlier notion of curvature bounds for metric spaces, by Wald [84], which looked at general configurations of four points $a_1, \dots, a_4 \in X$ with their mutual distances and checked into which constant curvature spaces such a quadrilateral can be isometrically embedded. This works nicely for surfaces (which was Wald's purpose, as the title of his paper [84] already clarifies), because a quadrilateral on a surface can be isometrically embedded into some constant curvature surface, and one can use the latter's curvature to assign a curvature to the original quadrilateral. One then gets curvature bounds in the sense of Wald when every quadrilateral satisfies a corresponding bound. For higher dimensional spaces, the requirement is perhaps somewhat too general. For instance, a tetrahedron in Euclidean 3-space, that is, a configuration of four points with all non-zero distances being equal can be isometrically embedded into some two-dimensional sphere, but not into the Euclidean plane. The notion of Wald curvature and its relation to Alexandrov curvature is discussed in more detail in Saucan's contribution to this volume (Chap. 2, Sects. 2.3 and 2.5).

Importantly, the preceding notions of Alexandrov and others refer to sectional, and not to Ricci curvature. As is already clear from the classical setting of Riemannian geometry, the content of the two notions is different. Ricci curvature is an average of the sectional curvatures containing a fixed tangent vectors, and as such, it is naturally a coarser notion than the latter. In closing this article, we would like to elucidate this aspect from the perspective gained from the preceding considerations. Ricci curvature, as we have seen, is essentially about the relation between two distance balls in a metric space. Ricci curvature is about such quantities as the relative size of their overlap as a function of their radii and the distances of their centers, or more precisely, how easy or difficult it is to transport the mass of one of them to the other. Thus, it is natural to speculate that we should get more refined invariants when we look at the overlap patterns of three (or perhaps more?) instead of two balls. In fact, it was found in [19] that this can be used to define sectional curvature bounds in general metric spaces. Those spaces, in contrast to the situation covered by Alexandrov's approach, need not be continuous, but could well be discrete. Let us now describe this concept.

Again, in the metric space (X, d) , we consider a triangle, that is, a triple of points (a_1, a_2, a_3) in X , and the comparison triangle in \mathbb{R}^2 (for simplicity of exposition) with the same side lengths. That is, we choose points $\bar{a}_1, \bar{a}_2, \bar{a}_3 \in \mathbb{R}^2$ with

$$d(a_i, a_j) = \|\bar{a}_i - \bar{a}_j\|, \quad \text{for } i, j = 1, 2, 3,$$

where $\|\cdot\|$ is the Euclidean norm. The idea is now to look at the smallest radius $r > 0$ such that the three closed balls around the a_i with radius r have a nonempty intersection,

$$B(a_1, r) \cap B(a_2, r) \cap B(a_3, r) \neq \emptyset, \quad (1.189)$$

(of course, $B(a, r) = \{p \in X : d(p, a) \leq r\}$) and to compare this with the corresponding radius \bar{r} for the Euclidean comparison triangle. We then say that (X, d) has nonpositive sectional curvature if

$$r \leq \bar{r}. \quad (1.190)$$

In more detail, we define the functions

$$\rho_{(a_1, a_2, a_3)}(x) = \max_{i=1,2,3} d(x, a_i), \quad x \in X,$$

and,

$$\rho_{(\bar{a}_1, \bar{a}_2, \bar{a}_3)}(x) = \max_{i=1,2,3} \|x - \bar{a}_i\|, \quad x \in \mathbb{R}^2$$

and call

$$r(a_1, a_2, a_3) \alpha \inf_{x \in X} \rho_{(a_1, a_2, a_3)}(x) \quad \text{and} \quad r(\bar{a}_1, \bar{a}_2, \bar{a}_3) \alpha \min_{x \in \mathbb{R}} \rho_{(\bar{a}_1, \bar{a}_2, \bar{a}_3)}(x)$$

circumradii of the respective triangles. The definition then is

Definition 1.14 (Nonpositive Curvature) We say that the metric space (X, d) has generalized nonpositive sectional curvature if, for each triangle (a_1, a_2, a_3) in X , we have

$$r(a_1, a_2, a_3) \leq r(\bar{a}_1, \bar{a}_2, \bar{a}_3), \quad (1.191)$$

where \bar{a}_i with $i = 1, 2, 3$ are the vertices of an associated comparison triangle. Of course, one checks that when (M, g) is a Riemannian manifold, then it possesses generalized nonpositive sectional curvature in the sense of Definition 1.14 iff it has nonpositive sectional curvature in the sense of Riemannian geometry, see Sect. 1.3.2. And of course, the construction can be naturally extended to define other upper sectional curvature bounds, by taking appropriate 2-dimensional spheres or hyperbolic spaces instead of the Euclidean plane as comparison spaces. Also, by reversing the inequality in (1.191), one may also define lower curvature bounds, as in Alexandrov's approach described above.

This leads us to a final remark. As just observed, we can as well define upper as lower sectional curvature bounds, and either of them has nontrivial geometric content. In contrast, in our discussion of Ricci curvature bounds, we have exclusively discussed lower bounds. The reason for this restriction appears already in the classical context of Riemannian geometry. In fact, an important theorem of Lohkamp [62] says that every differentiable manifold can be equipped with a metric of negative Ricci curvature. Therefore, carrying a metric of negative Ricci curvature imposes no topological restriction whatsoever on a manifold. By way of contrast, nonnegative Ricci curvature, or more generally, a lower Ricci curvature bound, is a contentful condition that implies many further geometric properties, as we have discussed and explored in this article.

Acknowledgements The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Advanced Investigator Grant Agreement no. 267087. Frank Bauer was partially supported by the Alexander von Humboldt foundation and partially supported by the NSF Grant DMS-0804454 Differential Equations in Geometry. Shiping Liu was partially supported by the EPSRC Grant EP/K016687/1 Topology, Geometry and Laplacians of Simplicial Complexes. Bobo Hua was supported by the NSFC 11401106.

References

1. Ache, A., Warren, M.: Coarse Ricci curvature as a function on $M \times M$ (2015). arXiv preprint 1505.04461v1
2. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations. Springer, Berlin (2006)
3. Bakry, D.: Functional inequalities for Markov semigroups. In: Probability Measures on Groups: Recent Directions and Trends, pp. 91–147. Tata Institute of Fundamental Research, Mumbai (2006)
4. Bakry, D., Émery, M.: Diffusions hypercontractives. In: Séminaire de probabilités, XIX, 1983/84. Lecture Notes in Mathematics, vol. 1123, pp. 177–206. Springer, Berlin (1985). <http://dx.doi.org/10.1007/BFb0075847>
5. Bakry, D., Ledoux, M.: Sobolev inequalities and Myers's diameter theorem for an abstract Markov generator. *Duke Math. J.* **85**(1), 253–270 (1996). doi:10.1215/S0012-7094-96-08511-7. <http://dx.doi.org/10.1215/S0012-7094-96-08511-7>
6. Bakry, D., Ledoux, M.: A logarithmic Sobolev form of the Li-Yau parabolic inequality. *Rev. Mat. Iberoamericana* **22**, 683–702 (2006)
7. Bakry, D., Qian, Z.: Some new results on eigenvectors via dimension, diameter, and Ricci curvature. *Adv. Math.* **155**(1), 98–153 (2000). doi:<http://dx.doi.org/10.1006/aima.2000.1932>
8. Bakry, D., Gentil, I., Ledoux, M.: Analysis and Geometry of Markov Diffusion Operators. Springer, Berlin (2014)
9. Bartholdi, L., Schick, T., Smale, N., Smale, S.: Hodge theory on metric spaces. *Found. Comput. Math.* **12**, 1–48 (2012)
10. Bauer, F.: Normalized graph Laplacians for directed graphs. *Linear Algebra Appl.* **436**, 4193–4222 (2012)
11. Bauer, F., Jost, J.: Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian. *Commun. Anal. Geom.* **21**, 787–845 (2013)
12. Bauer, F., Atay, F., Jost, J.: Synchronization in discrete-time networks with general pairwise coupling. *Nonlinearity* **22**, 2333–2351 (2009)
13. Bauer, F., Atay, F., Jost, J.: Synchronized chaos in networks of simple units. *Europhys. Lett.* **89**, 20002–p1–p6 (2010)
14. Bauer, F., Jost, J., Liu, S.: Ollivier-Ricci curvature and the spectrum of the normalized graph Laplace operator. *Math. Res. Lett.* **19**, 1185–1205 (2012)
15. Bauer, F., Horn, P., Lin, Y., Lippner, G., Mangoubi, D., Yau, S.: Li-Yau inequality on graphs. *J. Differ. Geom.* **99**, 359–405 (2015)
16. Bauer, F., Hua, B., Yau, S.T.: Davies-Gaffney-Grigor'yan lemma on graphs. *Commun. Anal. Geom.* **23**, 1031–1068 (2015)
17. Bačák, M.: Convex Analysis and Optimization in Hadamard Spaces. De Gruyter, Berlin (2014)
18. Bačák, M., Hua, B., Jost, J., Kell, M.: (in preparation)
19. Bačák, M., Hua, B., Jost, J., Kell, M., Schikorra, A.: A notion of nonpositive curvature for general metric spaces. *Differ. Geom. Appl.* **38**, 22–32 (2015)
20. Berestovskij, V., Nikolaev, I.: Multidimensional generalized Riemannian spaces. In: Reshetnyak, Y.G. (ed.) *Geometry IV. Encyclopedia of Mathematical Sciences*, vol. 70, pp. 165–243. Springer, Berlin (1993)
21. Bhattacharya, B., Mukherjee, S.: Exact and asymptotic results on coarse Ricci curvature of graphs. *Discret. Math.* **338**(1), 23–42 (2015)
22. Bonciocat, A.I., Sturm, K.T.: Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.* **256**(9), 2944–2966 (2009). doi:10.1016/j.jfa.2009.01.029. <http://dx.doi.org/10.1016/j.jfa.2009.01.029>
23. Bubley, R., Dyer, M.E.: Path coupling: a technique for proving rapid mixing in Markov chains. In: 38th Annual Symposium on Foundations of Computer Science (FOCS '97), pp. 223–231 (1997)

24. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence, RI (2001)
25. Chan, T., Shen, J.: Image Processing and Analysis. Variational, PDE, Wavelet, and Stochastic Methods. SIAM, Philadelphia, PA (2005)
26. Chung, F.: Spectral Graph Theory. American Mathematical Society, Providence, RI (1997)
27. DeVos, M., Mohar, B.: An analogue of the Descartes-Euler formula for infinite graphs and Higuchi's conjecture. *Trans. Am. Math. Soc.* **359**(7), 3287–3300 (electronic) (2007). doi:<http://dx.doi.org/10.1090/S0002-9947-07-04125-6>
28. Dobrushin, R.L.: Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.* **15**, 458–486 (1970)
29. Eckmann, B.: Harmonische Funktionen und Randwertaufgaben in einem Komplex. *Comment. Math. Helv.* **17**(1), 240–255 (1944)
30. Eschenburg, J., Jost, J.: Differentialgeometrie und Minimalflächen. Springer, Berlin (2013)
31. Evans, L.: Partial differential equations and Monge-Kantorovich mass transfer. *Curr. Dev. Math.* **1997**, 65–126 (1999)
32. Forman, R.: Bochner's method for cell complexes and combinatorial Ricci curvature. *Discret. Comput. Geom.* **29**(3), 323–374 (2003)
33. Funano, K.: Eigenvalues of Laplacian and multi-way isoperimetric constants on weighted Riemannian manifolds (2013). <http://arxiv.org/abs/1307.3919>
34. Funano, K., Shioya, T.: Concentration, Ricci curvature, and eigenvalues of Laplacian. *Geom. Funct. Anal.* **23**(3), 888–936 (2013)
35. Garland, H.: p -adic curvature and the cohomology of discrete subgroups of p -adic groups. *Ann. Math.* **97**, 375–423 (1973)
36. Gauss, C.: Disquisitiones generales circa superficies curvas. In: Dombrowski, P. (ed.) 150 years after Gauss' "Disquisitiones generales circa superficies curvas". Société Mathématique de France, Paris (1979)
37. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* **7**(3), 1005–1028 (2008)
38. Grigoryan, A.: Analysis on graphs. Technical Report, University of Bielefeld (2009). <https://www.math.uni-bielefeld.de/~grigor/aglect.pdf>
39. Horak, D., Jost, J.: Spectra of combinatorial laplace operators on simplicial complexes. *Adv. Math.* **244**, 303–336 (2013)
40. Hua, B., Jost, J., Liu, S.: Geometric analysis aspects of infinite semiplanar graphs with nonnegative curvature. *J. Reine Angew. Math.* **700**, 1–36 (2015). <http://dx.doi.org/10.1515/crelle-2013-0015>
41. Jin, Y., Jost, J., Wang, G.: A nonlocal version of the Osher-Sole-Vese model. *J. Math. Imaging Vision* **44**(2), 99–113 (2012)
42. Jin, Y., Jost, J., Wang, G.: A new nonlocal H^1 model for image denoising. *J. Math. Imaging Vision* **48**(1), 93–105 (2014)
43. Jin, Y., Jost, J., Wang, G.: A new nonlocal variational setting for image processing. *Inverse Prob. Imaging* **9**, 415–430 (2015)
44. Jost, J.: Equilibrium maps between metric spaces. *Calc. Var.* **2**, 173–204 (1994)
45. Jost, J.: Generalized harmonic maps between metric spaces. In: Jost, J. (ed.) *Geometric Analysis and the Calculus of Variations for Stefan Hildebrandt*, pp. 143–174. International Press, Boston (1996)
46. Jost, J.: Nonpositive Curvature: Geometric and Analytic Aspects. Birkhäuser, Basel (1997)
47. Jost, J.: Riemannian Geometry and Geometric Analysis. Springer, Berlin (2011)
48. Jost, J.: Mathematical Methods in Biology and Neurobiology. Springer, Berlin (2014)
49. Jost, J.: Mathematical Concepts. Springer, Berlin (2015)
50. Jost, J., Liu, S.: Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs. *Discrete Comput. Geom.* **51**, 300–322 (2014)
51. Jost, J., Yau, S.: Harmonic maps and superrigidity. *Proc. Symp. Pure Math.* **54**(I), 245–280 (1993)

52. Kimmel, R., Malladi, R., Sochen, N.: Images as embedding maps and minimal surfaces: movies, color, texture, and volumetric medical images. *Int. J. Comput. Vis.* **39**(2), 111–129 (2000)
53. Kindermann, S., Osher, S., Jones, P.W.: Deblurring and denoising of images by nonlocal functionals. *Multiscale Model. Simul.* **4**(4), 1091–1115 (2005)
54. Kwok, T.C., Lau, L.C., Lee, Y.T., Oveis Gharan, S., Trevisan, L.: Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap. In: *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, STOC '13*, pp. 11–20. ACM, New York (2013). doi:<http://doi.acm.org/10.1145/2488608.2488611>
55. Lee, J.R., Oveis Gharan, S., Trevisan, L.: Multi-way spectral partitioning and higher-order Cheeger inequalities. In: *STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing*, pp. 1117–1130. ACM, New York (2012). <http://dx.doi.org/10.1145/2213977.2214078>
56. Levin, D.A., Peres, Y., Wilmer, E.L.: *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI (2009). With a chapter by James G. Propp and David B. Wilson
57. Li, P.: *Geometric Analysis*. Cambridge Studies in Advanced Mathematics, vol. 134. Cambridge University Press, Cambridge (2012)
58. Li, P., Yau, S.T.: On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156**(1), 153–201 (1986). doi:10.1007/BF02399203. <http://dx.doi.org/10.1007/BF02399203>
59. Lin, Y., Yau, S.T.: Ricci curvature and eigenvalue estimate on locally finite graphs. *Math. Res. Lett.* **17**, 343–356 (2010)
60. Liu, S.: An optimal dimension-free upper bound for eigenvalue ratios (2014). <http://arxiv.org/abs/1405.2213>
61. Liu, S., Peyerimhoff, N.: Eigenvalue ratios of nonnegatively curved graphs (2014). <http://arxiv.org/abs/1406.6617>
62. Lohkamp, J.: Metrics of negative Ricci curvature. *Ann. Math.* **140**, 655–683 (1994)
63. Loisel, B., Romon, P.: Ricci curvature on polyhedral surfaces via optimal transportation. *Axioms* **3**(1), 119–139 (2014). <https://hal.archives-ouvertes.fr/hal-00941486v2>
64. Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math. (2)* **169**(3), 903–991 (2009). doi:10.4007/annals.2009.169.903. <http://dx.doi.org/10.4007/annals.2009.169.903>
65. Meyer, Y.: *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. American Mathematical Society, Providence, RI (2001)
66. Miclo, L.: On eigenfunctions of Markov processes on trees. *Probab. Theory Relat. Fields* **142**, 561–594 (2008)
67. Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions (2014). arXiv:1412.3340
68. Myers, S.B.: Riemannian manifolds with positive mean curvature. *Duke Math. J.* **8**, 401–404 (1941)
69. Ohta, S.I.: On the measure contraction property of metric measure spaces. *Comment. Math. Helv.* **82**, 805–828 (2007)
70. Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **256**(3), 810–864 (2009). doi:<http://dx.doi.org/10.1016/j.jfa.2008.11.001>
71. Ollivier, Y.: A survey of Ricci curvature for metric spaces and Markov chains. In: Kotani, M., Hino, M., Kumagai, T. (eds.) *Probabilistic Approach to Geometry*. Advanced Studies in Pure Mathematics, vol. 57, pp. 343–381. Mathematical Society of Japan, Tokyo (2010)
72. Peres, Y.: *Mixing for Markov chains and spin systems*. Lecture Notes (2005). <http://www.stat.berkeley.edu/~peres/ubc.pdf>
73. Qian, B.: Remarks on Li-Yau inequality on graphs (2013). arXiv:1311.3367
74. Riemann, B.: *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*. Springer, Berlin (2013). Edited with a commentary by J. Jost
75. Sapiro, G.: *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, Cambridge (2006)

76. Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., Lenzen, F.: *Variational Methods in Imaging*. Springer, Berlin (2009)
77. Schmuckenschläger, M.: *Curvature of Nonlocal Markov Generators*. *Convex Geometric Analysis*, vol. 34, pp. 189–197. MSRI Publications, Berkeley, CA (1998)
78. Sreejith, R., Jost, J., Saucan, E., Samal, A.: Forman curvature for directed networks (2016). arXiv preprint arXiv:1605.04662
79. Sreejith, R., Mohanraj, K., Jost, J., Saucan, E., Samal, A.: Forman curvature for complex networks. *J. Stat. Mech: Theory Exp.* **2016**(6), 063206 (2016)
80. Sreejith, R., Jost, J., Saucan, E., Samal, A.: Systematic evaluation of a new combinatorial curvature for complex networks (2016). arXiv preprint arXiv:1610.01507
81. Sturm, K.T.: On the geometry of metric measure spaces. I and II. *Acta Math.* **196**(1), 65–177 (2006). doi:10.1007/s11511-006-0003-7. <http://dx.doi.org/10.1007/s11511-006-0003-7>
82. von Luxburg, U.: A tutorial on spectral clustering. *Stat. Comput.* **17**, 395–416 (2007)
83. Villani, C.: *Optimal Transport, Old and New*. Grundlehren der Mathematischen Wissenschaften, vol. 338. Springer, Berlin (2009). doi:10.1007/978-3-540-71050-9. <http://dx.doi.org/10.1007/978-3-540-71050-9>
84. Wald, A.: Begründung einer koordinatenlosen Differentialgeometrie der Flächen. *Ergebnisse eines Math. Kolloquiums*, 1. Reihe **7**, 24–46 (1936)
85. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998)
86. Weber, M., Saucan, E., Jost, J.: Can one see the shape of a network? (2016). arXiv preprint arXiv:1608.07838
87. Weber, M., Saucan, E., Jost, J.: Characterizing complex networks with Forman-Ricci curvature and associated geometric flows. arXiv preprint arXiv:1607.08654 (2016)
88. Weber, M., Jost, J., Saucan, E.: Forman-Ricci flow for change detection in large dynamic data sets. *Axioms* **5**(4), 26 (2016)
89. Weickert, J.: *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart (1998)
90. Zhang, H., Zhu, X.: Lipschitz continuity of harmonic maps between Alexandrov spaces (2013). arXiv:1311.1331

Modern Approaches to Discrete Curvature

Najman, L.; Romon, P. (Eds.)

2017, XXVI, 353 p. 80 illus., 35 illus. in color., Softcover

ISBN: 978-3-319-58001-2