Lecture Notes in Mathematics 2184

Laurent Najman
Pascal Romon *Editors*

Modern Approaches to Discrete Curvature

Springer

Lecture Notes in Mathematics 2184

More information about this series at http://www.springer.com/series/304

Laurent Najman • Pascal Romon
Editors

Modern Approaches to Discrete Curvature

Springer

Editors
Laurent Najman
Laboratoire d'Informatique Gaspard Monge
Université Paris-Est - ESIEE Paris
Noisy-le-Grand, France

Pascal Romon
Laboratoire d'Analyse et de Mathématiques Appliquées
Université Paris-Est Marne-la-Vallée
Marne-la-Vallée, France

ISSN 0075-8434 ISSN 1617-9692 (electronic)
Lecture Notes in Mathematics
ISBN 978-3-319-58001-2 ISBN 978-3-319-58002-9 (eBook)
DOI 10.1007/978-3-319-58002-9

Library of Congress Control Number: 2017949896

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Discrete curvature is one of the key concepts in modern discrete geometry. Discrete geometry itself is nothing new and boasts on the contrary an ancient and rich history, reaching back to Antiquity and the Platonic solids. The geometry of points, lines, polygons, and polyhedra predates differential geometry. The latter's progress in the last centuries though has made it a crucial field in mathematics and a tool of choice in many applications. And yet, discrete geometry has not ceased to stimulate researchers, as the works of Alexandrov and many others show. It has nevertheless taken a new twist at the end of the twentieth century, first with the onset of computer science and engineering, requiring a science based on discrete objects, bits and integers, molecules or bricks, graphs, and networks. This has greatly stimulated mathematic research and keeps doing so. But another impulse came from mathematics itself, from fields as varied as probabilities, integrable systems theory, or differential calculus; the aim is to construct a modern theory of discrete geometry akin to its smooth cousin while solving applied problems. This double motivation, a need for efficient tools and problems to solve, as well as a renewed approach on geometry in the discrete realm, has been a powerful incentive for research and has generated many advances in discrete geometry. And at the heart of this geometry lies the notion of discrete curvature, as it does in smooth geometry. Indeed, as we know now, curvature is often the key to understanding the metric and topological properties of geometric objects: genus (Gauss-Bonnet), diameter (Bonnet-Myers), functions and spectrum (Poincaré, Sobolev, Lichnerowicz), classification, flows and smoothing (Hamilton and Perelman), local description, etc. However, the challenge is to understand and use curvature in this non-smooth setting. Unlike the differential theory, where definitions have been cast more than a century ago by Riemann, discrete differential geometry still searches for the appropriate analog of the many curvatures (see [17]). It is not clear yet which notion is the most adapted to a given problem nor whether a single notion will (or should) arise. The reader will discover in this book many possible takes on curvature, with complex dependencies. This panoramic approach shows the wealth of contexts and ideas where curvature plays a role. And the material gathered here is but a sample—albeit a wide and representative one—of the dynamic publications currently happening in mathematics and computer science.

A word of warning might be necessary. Given the many possible approaches on discrete curvature, it is not possible to present the ideas with unified formalism and notations. Indeed, there is no unique definition of the various discrete settings, and thus the different notions in this book are different and apply to different objects. In this introduction, we try to provide some organization to this profusion and to present the different chapters in some logical way (see Fig. 1). But achieving such a task is probably wishful thinking. Nevertheless, we are deeply convinced that the interested reader will find her or his pain and efforts rewarded by the various gems that can be found in this monograph.

Discrete curvature has started its modern life with the approximation theory and finite differences, mimicking the smooth definition at a very basic level, derivatives being replaced with differences. This has met the needs of applied science for a

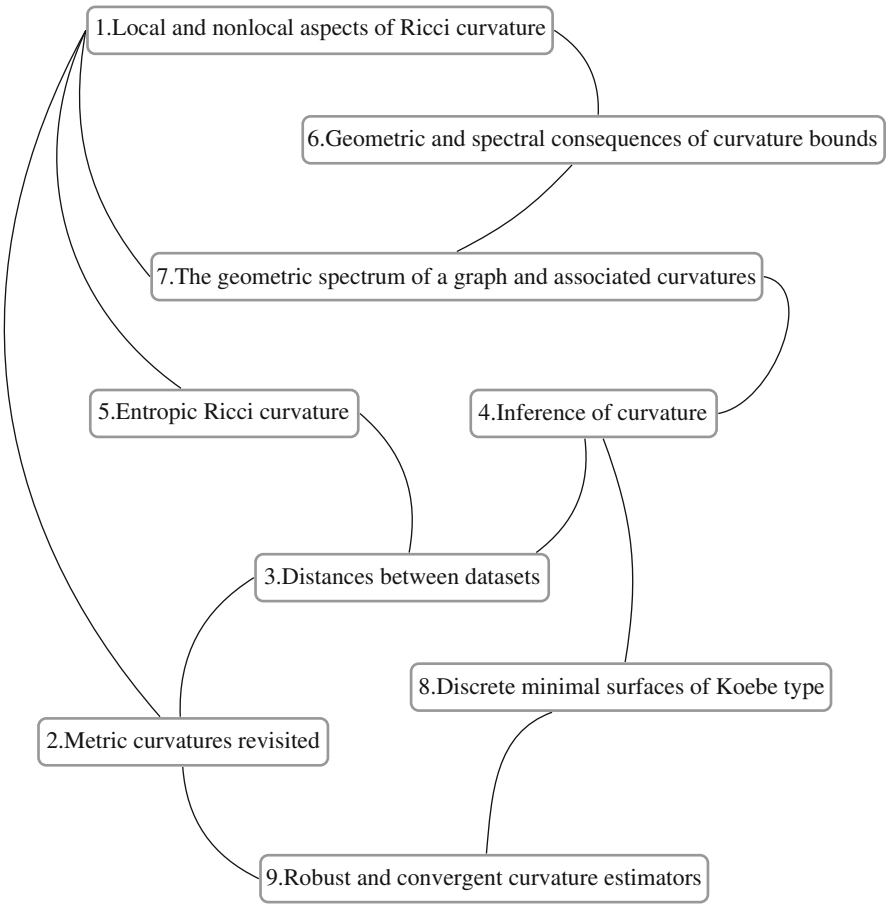


Fig. 1 A reading graph for this book, showing the strongest links between the chapters. The reader is thus invited to choose their own reading path

practical notion of curvature but has soon showed its limits. These were both on the theoretical side, since these approximations hardly ever obeyed the topological properties of their smooth ideal counterparts. But also on the practical side, since, (half-)surprisingly, such notions lack consistency, in general, the limit of such approximated curvatures either does not exist or does not converge to the smooth value. Such indeed was the (negative) answer of Xu et al. [26] to the question posed by Meek and Walton [18]: "Is there a closed-form expression based on local geometric data (lengths and angles) for the curvature of a discretized surface, that converges to the smooth one as the discretization is refined." This shortcoming challenges the mere idea of discretization. In reality, the lack of convergence of order two quantities is not so surprising to geometers, and counterexamples such as the Schwarz lantern have been known to mathematicians for a long time, especially in geometric measure theory. Yet, this does not preclude convergence results (see, for instance, Chaps. 2 and 9) under proper hypotheses, which are the scope of today's research. Typically, these rely either on stricter discretization (e.g., thick or fat triangulations (see Chap. 2), and also discretization along curvature lines as in [4]) or on adapting the type of convergence that is expected (e.g., Chaps. 2 and 3 on Gromov-Hausdorff convergence and Chap. 9) or more generally, as in the spirit of this book, on redefining the curvature itself, as a curvature measure (Chap. 4 and many others). A general overview of the lack of pointwise convergence and possible solutions can be found in [14].

Among the new approaches to discrete curvature, one finds a new take on local data (known to fail at pointwise convergence, as mentioned above), averaged as a measure rather than a vertex/edge-based quantity. In a way, the curvature ceases to be a Dirac measure to become more diffuse, by applying a (geometric) smoothing. A better and deeper understanding of this phenomenon can be found using normal cycles (in the Grassmannian) and geometric measure theory (see Chaps. 3 and 4 and the book of Morvan on generalized curvatures [7, 19], especially for extrinsic curvatures). As an advantage, this formalism gives a framework for managing the continuous, the discrete, and the semi-continuous cases, even point clouds. Very little regularity is required of the geometric object, a property which will also be encountered in the metric geometry approach. This is an asset when dealing with noisy real-world data; however this requires the choice of a characteristic length or convolution function to recover actual numeric values. Note that other geometric data such as sharp features can be obtained with these methods.

The starting point of these methods is the Steiner formula, which links the (extrinsic) curvatures to the area of the offsets of the surface under study, a method much used by Alexandrov. Interestingly enough, this approach turns out to be the starting point also of quadrangle/offset formalism for surfaces, developed in Chap. 8. This approach, derived from the integrable systems theory of geometric partial differential equations, is based on the following idea: A specific choice of parametrization (such as conformal, asymptotic, isothermic, etc.) transforms the study of a given (sub-)manifold into the study of a specific (simpler) PDE. In the discrete setup though, there is no such thing as a change of parametrization. On the contrary, it seems that one handles the surface itself instead of a parametrization,

as is the rule in differential geometry (although not in geometric measure theory, as the reader should notice). But even though changes of variables are hardly possible, special parametrizations do exist, e.g., conformal, isothermic, etc., inspired by Thurston's work in complex analysis (circle patterns). This leads to a beautiful theory of quadrangle-based discrete surfaces with special offsets, which allows to define Gaussian and mean curvature through offsets (à la Steiner) and minimal as well as constant mean or Gaussian curvature surfaces. The analogy with the smooth cases is justified by the convergence and the existence of similar structural properties (transformations of the space of solutions). In this approach, the geometric properties of the moduli space characterizes the discrete differential geometry. These special surfaces are—as expected—solution to an integrable system given by a Lax pair, analog to the smooth one, or described by conserved quantities [6]. In spite of their abstraction, these definitions often correspond to intuitive three-dimensional geometry; for instance, for K-nets, Steiner-defined curvature coincides with the product of orthogonal osculating circles. And logically but remarkably enough, this theoretical development has direct applications in architecture, in the construction of free form surfaces with constrained faces of fixed width and parallelism (see [24]). Surfaces made of panels of fixed width or with beams of fixed breadth are specific types of geometric quad-surfaces: Koebe and conical meshes. This leads to a new paradigm for discrete surfaces: instead of considering the vertices/edges/faces of one surface, one considers the surface together with its offset or equivalently the surface with its "Gauss map." By doing so, one avoids the tricky problem of defining the normals (thus guaranteeing convergence), but more than that, we have a consistent theory, even at a low discretization level [5]. Such indeed is the goal of discrete differential geometry.

Another and more classical view of curvature dates back to the early twentieth century. It is based on comparison between distances in a metric (actually a length) space and distances in a model space, usually of constant curvature, the Euclidean space, the round sphere, or the hyperbolic space with varying curvatures. By comparing triangles or quadrangles, Alexandrov, Toponogov, and Wald give bounds on the curvature, when it exists, and define a working curvature in most generalities. This is explored in Chaps. 2 and 3. Since it requires no smoothness, it applies well to discrete spaces (but not only to them). Convergence results are mentioned and we link to point cloud reconstruction as in Chap. 4. Moreover, this theory is deeply intertwined with image processing and other applications. There is also a tentative definition of what the Ricci flow might be for piecewise linear surfaces, a thriving topic in mathematics and computer science, with applications in smoothing (although Ricci curvature is intrinsic, while smoothing often relies on extrinsic curvature).

Adding a measure to a metric space, new ideas came from probability theory and optimal transportation, in the pioneering works of Bakry and Ledoux [3] and Ollivier [22]. The first authors, followed by many others, proposed a notion of curvature via Bochner-type functional inequalities on processes and functions (as explained in the discrete setting in Chap. 1). While coherent with Riemannian geometry, this extends easily to the discrete case (though not exclusively); see, for

instance [16], some recent estimates on graphs. This mixed two simple ideas: (1) work on the curvature via functional inequalities, often using the Laplace operator, and (2) separate the measure from the metric (which are linked in Riemannian geometry). The inspiration of Ollivier is quite different. Ricci curvature, as the rate of divergence of geodesics, is measured as a (logarithmic) difference between optimal transportation distance (Wasserstein distance; see also Chap. 3) and standard distance. This point of view applies very well to discrete spaces, in particular to graphs, and is the focus of Chap. 1 (which as the introductory chapter also recalls many basic definitions in classical geometry). It has been developed since by many authors, including algorithmically. These two approaches to intrinsic curvature are close but different: the former naturally yields lower bounds, while the latter gives an explicit—if difficult to compute—expression. Both cases though rely on a diffusion process (the Laplacian) which is known to be elusive in discrete geometry [25]. They also embody the same vision: By adding measures to a space, one reintroduces continuous objects that are lacking in a discrete space. Such is also the philosophy of Chap. 5, where curvature is defined through entropy. Continuity is realized by the geodesics in the probability space, over which functional inequalities are derived. This relies of course on metric ideas presented in Chaps. 2 and 3. Finally, let us note that optimal transportation appears also in shape reconstruction, to minimize the tension (or the Laplace energy); see [12].

Among discrete spaces, the graphs hold a particular position. It is certainly an old mathematical subject with strong research themes, stimulated by ever-growing applications and also, from our point of view, a geometrization taking place (after earlier works of Gromov). Graphs are interesting per se but may also be seen as a crude version of cell complexes or even discrete manifolds. Whatever works on the former can also be conceived for the latter, and we shall give finer results for those. In this book, Chaps. 1, 6, and 7 are devoted to graphs, with various geometric conditions attached. In these appears with the greatest clarity how much the curvature is linked with the Laplacian. Indeed bounds on the Ollivier-Ricci curvature yield Poincaré estimates, and so do other curvatures. Finer discrete geometric aspects, such as the frequency of triangles, the degree, etc., play a role in the curvature, or equivalently, in the Laplacian, and influence the global geometry, with Myers' and Cartan-Hadamard theorem. The reader will note with interest how different choices of the weights yield different results (combinatorial Laplacian vs. probabilistic, a.k.a. harmonic, Laplacian).

Chapter 7 tackles the "extrinsic" part (which is also present in Chap. 4); curvature appears again via function spaces, i.e., via the geometric spectrum of the Laplacian on graphs. Similar and yet different realizations of the eigenfunctions, as local or global embeddings in space, are presented, the goal being a reconstruction of the geometry. The definitions in Chap. 7 show nevertheless the strong dependence on intrinsic data.

Finally, this work also tackles the topic of digital spaces (i.e., the geometry of $\mathbb{Z}^n$), another clear application of discrete geometry, with obvious applications to imaging. This is typically a setup where a naive approximation of smooth geometry may lead to vast errors, such as metrication artifacts. Chapter 9 presents

smarter and geometrically meaningful integral invariants as curvature estimators. The issue of approximation is handled through *multigrid convergence* to their continuous counterpart when the digital shape is a digitization of a sufficiently smooth Euclidean shape. Furthermore a good noise robustness is shown.

The crucial problem of connectedness in digital shapes is a key issue in Chap. 9, inducing definitions of curvature, determined by their consistency (the Gauss-Bonnet theorem). The different local configurations that can arise on a digital shape are studied, and curvature indices and normals are introduced, leading to an important transform called digital curvature flow, which can be seen as a digitized version of the classical curvature flow. Several applications of this flow are illustrated, such as thinning and snakes.

There are of course many aspects of discrete curvature that could not fit in this book. A wider overview of the topic may already be found in the 2014 CIRM Proceedings of the Meeting on Discrete Curvature [21]. But, of all the missing aspects, discrete exterior calculus (DEC) is in our opinion one of the most important. Luckily, many fine articles have been written on this topic recently, and we will happily direct the reader to them (see references below).

Let us just make a brief presentation of the topic and recall that discrete curvature, while not the main subject of DEC, is present through the ubiquitous Laplace operator, which is central in discrete calculus. DEC starts by defining the basic objects: discrete differential forms and vector fields. Simple operators are then constructed: the discrete exterior derivative d, the Hodge star $\star$ and the codifferential $\mathrm{d}^\star$ operating on forms, the discrete wedge product for combining forms, the discrete flat and sharp operators to interchange vector fields and one forms, the discrete interior product operator for combining forms and vector fields, and the Lie derivative. Using these, more sophisticated operators can be constructed, such as the gradient, divergence, and curl operators. Because the basic operators were defined by their algebraic properties, we obtain easily *exact* Green, Gauss, and Stokes theorems (for instance, the differential must commute with the boundary operator: $\int_c \mathrm{d}\omega = \int_{\partial c} \omega$).

There is no unique or unequivocally better discretization of differential calculus, and the construction depends on the choice of the building bricks: discrete differential forms and vector fields. Often, these must be thought of as integrals of continuous forms over discrete elements (simplices or cells). For instance, Polthier and Preuß [23] presents a calculus based on finite elements on triangles exclusively. Another approach, followed by Desbrun et al., rests upon the (discrete) dual, which must be endowed with metric as well. For instance, dual vertices, which are primal faces, can be identified with a point in the face, e.g., the center of mass, or the circumcenter when the faces are circular, thus defining dual edge lengths. However, each choice of the dual geometry will generate a different DEC (see [2] for Laplacians on general polygonal meshes). Once again, the problem determines the approach. DEC applications are manifold, including finding geodesics, solving PDEs, remeshing, and optimal transportation again. For further reading, in addition to the works already quoted, we recommend the following papers and monographs: [8, 10, 11, 13, 15] and, for further reading also, [1, 9, 20].

Acknowledgments

We wish to thank first of all the authors, who have made this book possible, but also all the supportive members of the scientific community, who have encouraged us and given us their kind and useful advice, in particular Jérôme Bertrand, François Fillastre, Yukiko Kenmochi, Benoît Kloeckner, Christian Léonard, Yong Lin, Pooran Memari, Thomas Richard, and Cyril Roberto, as well as Springer's anonymous referees.

Noisy-le-Grand, France — Laurent Najman
Marne-la-Vallée, France — Pascal Romon
January 2017

References

1. Alexa, M., Cohen-Or, D., Levin, D.: As-rigid-as-possible shape interpolation. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, pp. 157–164. ACM Press/Addison-Wesley (2000)
2. Alexa, M., Wardetzky, M.: Discrete laplacians on general polygonal meshes. ACM Trans. Graph. **30**(4), 102:1–102:10 (2011). doi:10.1145/2010324.1964997. http://doi.acm.org/10.1145/2010324.1964997
3. Bakry, D., Ledoux, M.: Levy-Gromov's isoperimetric inequality for an infinite dimensional diffusion generator. Invent. Math. **123**(1), 259–281 (1996)
4. Bauer, U., Polthier, K., Wardetzky, M.: Uniform convergence of discrete curvatures from nets of curvature lines. Discrete Comput. Geom. **43**(4), 798–823 (2010)
5. Bobenko, A.I., Suris, Y.B.: Discrete differential geometry. Integrable structure, *Graduate Studies in Mathematics*, vol. 98. AMS, Providence (2008). http://dx.doi.org/10.1090/gsm/098
6. Burstall, F., Hertrich-Jeromin, U., Rossman, W., Santos, S.: Discrete special isothermic surfaces. Geom. Dedicata **174**(1), 1–11 (2015)
7. Cohen-Steiner, D., Morvan, J.M.: Restricted Delaunay triangulations and normal cycle. In: Proceedings of the Nineteenth Annual Symposium on Computational Geometry, SCG'03, pp. 312–321. ACM, New York (2003). http://doi.acm.org/10.1145/777792.777839

8. Crane, K., de Goes, F., Desbrun, M., Schröder, P.: Digital geometry processing with discrete exterior calculus. In: ACM SIGGRAPH 2013 Courses, p. 7. ACM, New York (2013)
9. Crane, K., Weischedel, C., Wardetzky, M.: Geodesics in heat: a new approach to computing distance based on heat flow. ACM Trans. Graph. **32**(5), 152 (2013)
10. Desbrun, M., Hirani, A.N., Leok, M., Marsden, J.E.: Discrete exterior calculus. arXiv preprint math/0508341 (2005)
11. Desbrun, M., Kanso, E., Tong, Y.: Discrete differential forms for computational modeling. In: Discrete Differential Geometry, pp. 287–324. Springer, Berlin (2008)
12. de Goes, F., Cohen-Steiner, D., Alliez, P., Desbrun, M.: An optimal transport approach to robust reconstruction and simplification of 2d shapes. Comput. Graphics Forum **30**(5), 1593–1602 (2011). doi:10.1111/j.1467-8659.2011.02033.x. http://dx.doi.org/10.1111/j.1467-8659.2011.02033.x
13. Grady, L.J., Polimeni, J.: Discrete Calculus: Applied Analysis on Graphs for Computational Science. Springer Science + Business Media, Berlin (2010)
14. Hildebrandt, K., Polthier, K., Wardetzky, M.: On the convergence of metric and geometric properties of polyhedral surfaces. Geom. Dedicata **123**(1), 89–112 (2006)
15. Hirani, A.N.: Discrete exterior calculus. Ph.D. thesis, California Institute of Technology (2003)
16. Klartag, B., Kozma, G., Ralli, P., Tetali, P.: Discrete curvature and abelian groups. arXiv preprint arXiv:1501.00516 (2015)
17. Magid, E., Soldea, O., Rivlin, E.: A comparison of Gaussian and mean curvature estimation methods on triangular meshes of range image data. Comput. Vis. Image Underst. **107**(3), 139–159 (2007)
18. Meek, D.S., Walton, D.J.: On surface normal and Gaussian curvature approximations given data sampled from a smooth surface. Comput. Aided Geom. Des. **17**(6), 521–543 (2000)
19. Morvan, J.M.: Generalized Curvatures. Springer, Berlin (2008)
20. Mullen, P., Memari, P., de Goes, F., Desbrun, M.: Hot: Hodge-optimized triangulations. ACM Trans. Graph. **30**(4), 103 (2011)
21. Najman, L., Romon, P. (eds.): Discrete Curvature: Theory and Applications. No. 1 in Actes des rencontres du CIRM. CEDRAM, France (2013). https://hal.archives-ouvertes.fr/hal-01090755
22. Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256**(3), 810–864 (2009). http://dx.doi.org/10.1016/j.jfa.2008.11.001
23. Polthier, K., Preuß, E.: Identifying vector field singularities using a discrete hodge decomposition. In: Visualization and Mathematics III, pp. 113–134. Springer, Berlin (2003)
24. Wallner, J., Pottmann, H.: Geometric computing for freeform architecture. J. Math. Ind. **1**(1), 4 (2011)
25. Wardetzky, M., Mathur, S., Kälberer, F., Grinspun, E.: Discrete laplace operators: no free lunch. In: Symposium on Geometry processing, pp. 33–37 (2007)
26. Xu, Z., Xu, G., Sun, J.G.: Convergence analysis of discrete differential geometry operators over surfaces. In: Mathematics of Surfaces XI, pp. 448–457. Springer, Berlin (2005)

Contents

Contributors

Paul Baird Laboratoire de Mathématiques de Bretagne Atlantique, Brest Cedex 3, France

Frank Bauer Department of Mathematics, Harvard University, Cambridge, MA, USA
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Alexander I. Bobenko Technische Universität Berlin, Institut für Mathematik, Berlin, Germany

Ulrike Bücking Technische Universität Berlin, Institut für Mathematik, Berlin, Germany

Frédéric Chazal Inria Saclay, Palaiseau, France

David Coeurjolly Université de Lyon, CNRS, INSA-Lyon, LIRIS, UMR 5205, Lyon, France

David Cohen-Steiner Inria Sophia-Antipolis, Sophia-Antipolis, France

Bobo Hua School of Mathematical Sciences, Fudan University, Shanghai, China

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

Matthias Keller Universität Potsdam, Institut für Mathematik, Potsdam, Germany

Jacques-Olivier Lachaud Université de Savoie, CNRS, LAMA, UMR 5127, Chambéry, France
Université Grenoble-Alpes, CNRS, LJK, UMR 5224, Saint-Martin-d'Héres, France

Jérémy Levallois Université de Lyon, CNRS, INSA-Lyon, LIRIS, UMR 5205, Lyon, France
Université de Savoie, CNRS, LAMA, UMR 5127, Chambéry, France

André Lieutier Dassault Systèmes, Aix-en-Provence, France

Shiping Liu School of Mathematical Sciences, University of Science and Technology of China, Hefei, China

Jan Maas Institute of Science and Technology Austria (IST Austria), Klosterneuburg, Austria

Facundo Mémoli Department of Mathematics, The Ohio State University, Columbus, OH, USA

Quentin Mérigot Laboratoire de Mathématiques d'Orsay, University of Paris-Sud, CNRS, Université Paris-Saclay, Orsay, France

Emil Saucan Electrical Engineering and Mathematics Departments, Technion, Haifa, Israel

Stefan Sechelmann Technische Universität Berlin, Institut für Mathematik, Berlin, Germany

Boris Thibert Laboratoire Jean Kuntzmann, Université Grenoble-Alpes, CNRS, Saint-Martin-d'Hères, France

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

List of Figures

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

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_1

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 lead 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

$$a^i b_i := \sum_{i=1}^{d} a^i b_i \tag{1.1}$$

  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

$$g^{ij} g_{jk} = \delta^i_k := \begin{cases} 1 & \text{when } i = k \\ 0 & \text{when } i \neq k. \end{cases} \tag{1.2}$$

- More generally,

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

  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

$$V = v^i \frac{\partial}{\partial x^i}; \tag{1.4}$$

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

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

The tangent vectors at a point $p \in M$ form a d-dimensional vector space, called the tangent space T_pM 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

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

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

$$dx^i(\frac{\partial}{\partial x^j}) = \delta^i_j \tag{1.7}$$

yielding

$$\omega_i dx^i(v^j \frac{\partial}{\partial x^j}) = \omega_i v^j \delta^i_j = \omega_i v^i. \tag{1.8}$$

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

$$dx^i = \frac{\partial x^i}{\partial y_\alpha} dy^\alpha \tag{1.9}$$

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

$$\omega_i dx^i = \omega_i \frac{\partial x^i}{\partial y_j} dy^j. \tag{1.10}$$

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

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

The metric tensor provides a Euclidean product of tangent vectors,

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

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

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

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

$$\Gamma^i_{jk} = 0 \text{ for all } i, j, k. \tag{1.14}$$

The Christoffel symbols Γ^i_{jk} 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

$$d\phi = \frac{\partial \phi}{\partial x^i}dx^i, \tag{1.15}$$

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

$$\operatorname{grad}\phi = g^{ij}\frac{\partial\phi}{\partial x^j}\frac{\partial}{\partial x^i}. \tag{1.16}$$

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 $V \in T_pM$ is tangent to such a level hypersurface, it satisfies

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

When V then satisfies (1.17), we have

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

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

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

$$\langle \operatorname{grad}\phi, \operatorname{grad}\psi\rangle = g_{ik}g^{ij}\frac{\partial\phi}{\partial x^j}g^{k\ell}\frac{\partial\psi}{\partial x^\ell} = g^{j\ell}\frac{\partial\phi}{\partial x^j}\frac{\partial\psi}{\partial x^\ell}. \tag{1.19}$$

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

$$\operatorname{div} Z := \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^j}(\sqrt{g}Z^j) = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^j}\Big(\sqrt{g}g^{ij}\Big\langle Z, \frac{\partial}{\partial x^i}\Big\rangle\Big). \tag{1.20}$$

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

$$\Delta f := \operatorname{div}\operatorname{grad} f = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^j}\Big(\sqrt{g}g^{ij}\frac{\partial f}{\partial x^i}\Big). \tag{1.21}$$

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

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

$$\Delta f = 0. \tag{1.22}$$

Below, we shall introduce the volume form $dvol$ 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

$$(f, g) := \int f(x)g(x)dvol(x). \tag{1.23}$$

We then have

$$(\Delta f, g) = -(\operatorname{grad} f, \operatorname{grad} g) = (f, \Delta g) \tag{1.24}$$

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

$$\Delta f_\lambda + \lambda f_\lambda = 0, \tag{1.25}$$

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

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots. \tag{1.26}$$

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

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

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

$$(d^* f, g) = (f, dg) \tag{1.28}$$

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

$$\Delta\omega = -(dd^* + d^* d)\omega. \tag{1.29}$$

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

$$d\omega = 0 \text{ and } d^*\omega = 0. \tag{1.30}$$

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 $\mathbb{R}$, $\gamma : [a, b] \to M$ a (smooth) curve. The *length of* γ is defined as

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

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

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

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

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

they satisfy the following equations

$$\ddot{x}^i(t) + \Gamma^i_{jk}(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0, \quad i = 1, \ldots, d \tag{1.32}$$

with the Christoffel symbols

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

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

$$\mathrm{Vol}(\Omega) := \int_\Omega dvol(x) := \int_\Omega \sqrt{g} dx \tag{1.33}$$

with the volume form

$$dvol := \sqrt{g} dx := \sqrt{det(g_{ij})} dx. \tag{1.34}$$

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

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

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,

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

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

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

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-Cività.

Definition 1.1 The covariant derivative ∇ satisfies

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

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

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

Moreover,

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

With this notation, the geodesic equation becomes

$$\nabla_{\frac{d}{dt}} \frac{dc}{dt} = 0 \tag{1.41}$$

where we use the transformation rule

$$\frac{d}{dt} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}. \tag{1.42}$$

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 $\Gamma^k_{ij} = \Gamma^k_{ji}$, we have

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

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

Definition 1.2 The curvature operator R is defined by

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

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

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

We put

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

i.e.[1]

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

When we choose Riemann normal coordinates, this becomes

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

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

$$\begin{aligned} K(X \wedge Y) :&= \langle R(X,Y)Y, X\rangle \frac{1}{|X \wedge Y|^2} \\ &= \frac{R_{ijk\ell}\xi^i\eta^j\xi^k\eta^\ell}{g_{ik}g_{j\ell}(\xi^i\xi^k\eta^j\eta^\ell - \xi^i\xi^j\eta^k\eta^\ell)} \\ &= \frac{R_{ijk\ell}\xi^i\eta^j\xi^k\eta^\ell}{(g_{ik}g_{j\ell} - g_{ij}g_{k\ell})\xi^i\eta^j\xi^k\eta^\ell} \end{aligned} \tag{1.48}$$

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

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

$$\mathrm{Ric}(X, X) = g^{j\ell}\Big\langle R\Big(X, \frac{\partial}{\partial x^j}\Big)\frac{\partial}{\partial x^\ell}, X\Big\rangle. \tag{1.49}$$

In local coordinates, the Ricci tensor is

$$R_{ik} = g^{j\ell}R_{k\ell ij}. \tag{1.50}$$

The Ricci tensor is symmetric,

$$R_{ik} = R_{ki}. \tag{1.51}$$

Finally, the scalar curvature is

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

[1]The 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,

$$f(t) := d(c_1(t), c_2(t)). \tag{1.52}$$

In the Euclidean plane, we have

$$f(t) = \gamma t, \tag{1.53}$$

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

$$f(t) = \gamma \sin t, \tag{1.54}$$

whereas the relation in the hyperbolic plane is

$$f(t) = \gamma \sinh t. \tag{1.55}$$

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*

$$\mathrm{Ric} \geq (d-1)\lambda, \tag{1.56}$$

which is an abbreviation for $\mathrm{Ric}(X,X) \geq (d-1)\lambda\langle X,X\rangle$ *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* $\mathrm{Ric} \equiv (d-1)\lambda$.*) Let $V_\lambda(r)$ be the volume of a ball of radius r about any point in M_λ. Let $p \in M$.*

Then

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

Of course, we have

$$\lim_{r\to 0} \frac{\mathrm{Vol}(U(p,r))}{V_0(r)} = 1, \tag{1.58}$$

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, $e^i(e_j) = \delta^i_j$). Then the Laplace–Beltrami operator acting on

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

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

Here, the second covariant derivative is defined as $\nabla^2_{XY} = \nabla_X \nabla_Y - \nabla_{\nabla_X Y}$, 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 $\nabla^2_{e_i e_i}$ 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*

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

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

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

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

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

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

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

and with (1.47)

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

Thus, we have

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

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

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

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,*

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

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

$$0 = \int_M \Delta\langle \omega, \omega\rangle dvol = 2\int_M (|\nabla\omega|^2 + \mathrm{Ric}(\omega, \omega))dvol. \tag{1.64}$$

When the Ricci curvature is positive, the integrand on the right-hand side is pointwise nonnegative. It therefore has to vanish identically, and since $\int \mathrm{Ric}(\omega,\omega))dvol > 0$ 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* $\mathrm{Ric} \geq (d-1)\rho$*, with $\rho > 0$; this means that for every tangent vector X*

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

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

$$\lambda_1 \geq d\rho. \tag{1.66}$$

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

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

We have

$$-(\Delta df, df) = (\Delta f, \Delta f) \leq d\,(\nabla df, \nabla df) \tag{1.68}$$

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

$$(\Delta f, \Delta f) \geq d\,\rho(df, df). \tag{1.69}$$

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

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

we obtain

$$\lambda(df, df) = -\lambda(\Delta f, f) = (\Delta f, \Delta f) \geq d\ \rho(df, df) \tag{1.70}$$

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,

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

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

$$h(x) = \frac{1}{\mathrm{Vol}(U(x,r))} \int_{U(x,r)} h(y)dvol(y), \tag{1.72}$$

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

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

(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 $\frac{\mathrm{Vol}(U(x,r)\cap U(y,r))}{\mathrm{Vol}(U(x,r))}$, 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

$$h(x)-h(y)=\frac{1}{\mathrm{Vol}(U(x,r))}\int_{U(x,r)}h(z)dvol(z)-\frac{1}{\mathrm{Vol}(U(y,r))}\int_{U(y,r)}h(z)dvol(z), \tag{1.74}$$

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

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

and symmetric,

$$\omega(x, y) = \omega(y, x). \tag{1.76}$$

We also put

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

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 $u : M \to \mathbb{R}$ and vector fields $p : M \times M \to \mathbb{R}$. 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,

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

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

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

The difference vector field of a function $u : M \to R$ is defined by

$$Du(x, y) = u(y) - u(x). \tag{1.78}$$

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 $p : M \times M \to \mathbb{R}$, its *divergence operator* $\operatorname{div} p : M \to \mathbb{R}$ is now defined by

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

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

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

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.

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

The Laplace operator satisfies

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

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

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

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

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

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

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

or equivalently

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

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

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

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

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

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

$$\Delta_{\bar{\omega}}u = 0 \tag{1.89}$$

on some domain Ω, that is,

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

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

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

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

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

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

$$F_p(\alpha) = w_\alpha\Big(\sum_{\gamma^{p+1}>\alpha} \frac{w_\alpha}{w_\gamma} + \sum_{\beta^{p-1}<\alpha} \frac{w_\beta}{w_\alpha} - \sum_{\tilde{\alpha}^p \neq \alpha} \Big| \sum_{\gamma^{p+1}>\alpha, \gamma>\tilde{\alpha}} \frac{\sqrt{w_\alpha w_{\tilde{\alpha}}}}{w_\gamma} - \sum_{\beta^{p-1}<\alpha, \beta<\tilde{\alpha}} \frac{w_\beta}{\sqrt{w_\alpha w_{\tilde{\alpha}}}} \Big|\Big). \tag{1.93}$$

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

$$\partial_p : C_p \to C_{p-1}, \tag{1.94}$$

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

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

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

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

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

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

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

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

$$\square_p = B_p - F_p \tag{1.98}$$

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*

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

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

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

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

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

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 $\mathrm{Ric}(e) \geq 0$ for all e and $\mathrm{Ric}(e_0) > 0$ 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

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

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

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

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

[2]The 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,

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

where $\mathrm{Lip}(f) := \sup_{x\neq y} \frac{|f(x)-f(y)|}{d(x,y)}$ 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

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

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*

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

and hence

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

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

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

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

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

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

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

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

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

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

$$2\Gamma(f,g)(x) = \big(L(f\cdot g) - f\cdot L(g) - L(f)\cdot g\big)(x) \tag{1.112}$$

and the iterated gradient form Γ_2 by

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

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 $\mathscr{L}u := (\Delta - \partial_t)u = 0$ read

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

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 $\Phi : \mathbb{R} \to \mathbb{R}$:

$$\Gamma(f,gh) = g\Gamma(f,h) + h\Gamma(f,g) \tag{1.115}$$

$$\Gamma(\Phi\circ f, g) = \Phi'(f)\Gamma(f,g) \tag{1.116}$$

$$L(\Phi\circ f) = \Phi'(f)L(f) + \Phi''(f)\Gamma(f). \tag{1.117}$$

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 $x = x_0, x_1, \ldots, x_m = y$ of distinct vertices such that $x_{\nu-1} \sim x_\nu$ for $\nu = 1, \ldots, m$. 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 $x_0, x_1, \ldots, x_m = x_0$ for which all the vertices $x_1, \ldots, x_m$ are distinct. For $m = 3, 4, 5, \ldots$, 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

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

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

$$\Delta : L^2(G) \to L^2(G)$$

$$\Delta v(x) := \frac{1}{d_x}\Big(\sum_{y, y \sim x} v(y) - d_x v(x)\Big) = \frac{1}{d_x} \sum_{y, y \sim x} v(y) - v(x). \tag{1.119}$$

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

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

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

$$Du(x, y) := u(y) - u(x), \tag{1.121}$$

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

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

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

$$(\Delta u, v) = -(Du, Dv) \tag{1.123}$$

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. $(.,.)$:

$$(u, \Delta v) = (\Delta u, v) \tag{1.124}$$

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:

$$(\Delta u, u) \leq 0 \tag{1.125}$$

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

$$\Delta u + \lambda u = 0 \tag{1.126}$$

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,

$$\lambda_k > 0 \tag{1.127}$$

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

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

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,

$$(v_i, v_j) = 0. \tag{1.128}$$

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

$$\sum_x d_x v_i(x) = 0. \tag{1.129}$$

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,

$$\lambda_k = \min_{\substack{u_0,\ldots,u_k \not\equiv 0 \\ (u_i,u_j)=0, \forall i \neq j}} \max_{\substack{u \in \operatorname{span}\{u_0,\ldots,u_k\} \\ u \not\equiv 0}} \frac{(Du, Du)}{(u, u)}. \tag{1.130}$$

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

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

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)

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

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.,

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

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.

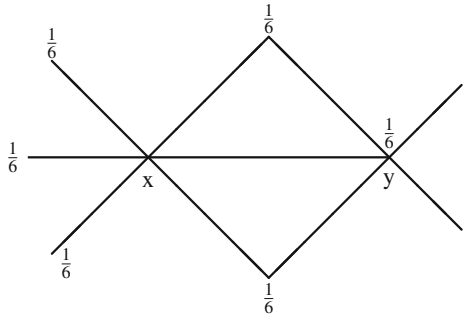


Fig. 1.1 Starting configuration; mass 0 at all vertices without number attached

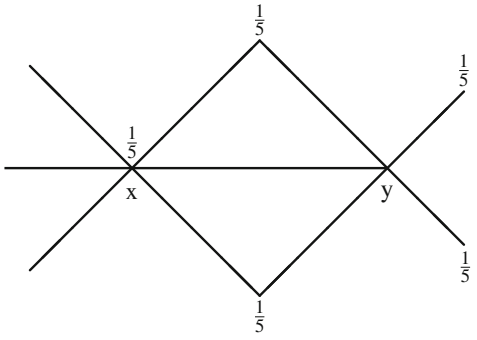


Fig. 1.2 Target configuration

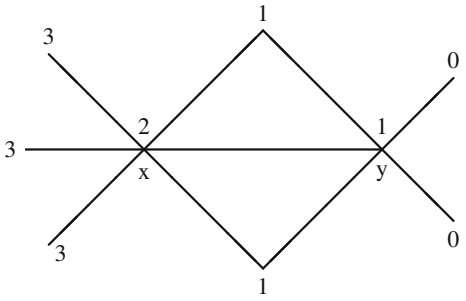


Fig. 1.3 Mass moved from vertices with larger value to those with smaller ones

Similarly, for step (2) we would need

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

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 $\mathscr{K}_n$ ($n \geq 2$) 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,*

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

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

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

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

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

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*

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

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

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

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

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

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,

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

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

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

or equivalently,

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

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*

$$\mu P(\cdot) := \sum_x \mu(x) m_x(\cdot). \tag{1.144}$$

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

$$W_1(\mu P, \nu P) \leq (1-k) W_1(\mu, \nu). \tag{1.145}$$

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

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

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,

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

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* $\bar{x}, \bar{y} \in V$,

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

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

$$w_{xy} = w_{yx} \ge 0. \tag{1.149}$$

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 $x_0 = x, x_1, \ldots, x_n = y$ with $w_{x_{i-1}x_i} > 0$ for $i = 1, \ldots, n$.

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

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

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

$$(u, v) = \sum_x d_x u(x) v(x). \tag{1.151}$$

The Laplacian now is

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

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

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

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

$$\delta_x P^t(y) = \begin{cases} \sum_{x_1,\ldots,x_{t-1}} \frac{w_{xx_1}}{d_x} \frac{w_{x_1x_2}}{d_{x_1}} \cdots \frac{w_{x_{t-1}y}}{d_{x_{t-1}}}, & \text{if } t > 1; \\ \frac{w_{xy}}{d_x}, & \text{if } t = 1. \end{cases} \tag{1.154}$$

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

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

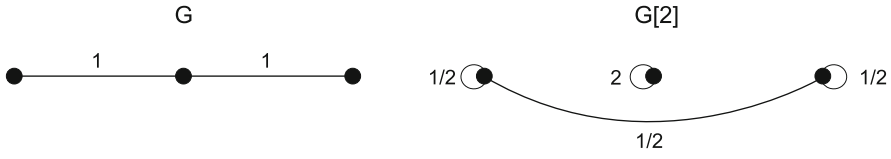
from (1.154).

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

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

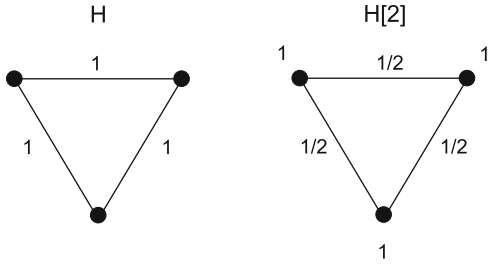
see for instance Grigor'yan [38].

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*

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

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

$$0 = \lambda_0[t] \leq \lambda_1[t] \leq \ldots \leq \lambda_{N-1}[t] \leq 1. \tag{1.157}$$

(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*

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

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

$$d[t](x, y) \leq d(x, y). \tag{1.159}$$

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 $\lambda_1[t] \geq \mathscr{A}[t]$, *then*

$$1 - (1 - \mathscr{A}[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1 - \mathscr{A}[t])^{\frac{1}{t}} \tag{1.160}$$

if t is even and

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

if t is odd.

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

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

for even t, whereas

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

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*

$$\kappa(x,y) \geq k, \tag{1.162}$$

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

$$k \leq \lambda_1 \leq \ldots \leq \lambda_{N-1} \leq 2-k. \tag{1.163}$$

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

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

Then

$$P^t f(x) = \sum_y f(y)\delta_x P^t(y). \tag{1.165}$$

We construct a discrete time heat equation,

$$f(x,t+1) - f(x,t) = \Delta f(x,t), \tag{1.166}$$

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

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

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

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

Here, $\mathrm{Lip}(f)$ 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

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

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

$$\frac{1-k}{|1-\lambda|} \geq 1. \tag{1.168}$$

(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

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

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) $|P^t f(x) - P^t f(y)| \leq (1-k)^t \mathrm{Lip}(f)$ 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*

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

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 $t \to +\infty$ 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*

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

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,*

$$\lambda_k \leq C d_G k^2 \lambda_1, \tag{1.172}$$

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 $S_0, S_2, \ldots, S_k$ of V, such that each S_i has small expansion. Here, the expansion $\phi(S_i)$ of S_i is defined as

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

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

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

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,

$$\frac{\lambda_k}{2} \leq h_k \leq C k^2 \sqrt{\lambda_k}. \tag{1.175}$$

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,*

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

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 $f : V \to \mathbb{R}$ 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

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

and we define

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

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:

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

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

$$\Delta \log u = \frac{\Delta u}{u} - |\nabla \log u|^2 \tag{1.179}$$

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 $t > 0$*

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

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 $g, F : V\times[0,T] \to \mathbb{R}$ be functions. Suppose that $g \geq 0$, and F has a local maximum at $(x^*,t^*) \in V\times]0,T]$. Then*

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

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

Proof (Proof of the Theorem) Let

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

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

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

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

$$2\frac{\partial_t\sqrt{u}}{\sqrt{u}} = \frac{\partial_t u}{u} = \frac{\Delta u}{u}. \tag{1.183}$$

We conclude from (1.182) that

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

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

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

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

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

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

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

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 $\mathscr{G}_p$ 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 $\mathscr{G}_p$ is. In the Riemannian case, the original Li-Yau inequality (1.114) corresponds to $\mathscr{G}_1$. 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 $\mathscr{G}_p$ 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 semiplanar graph. For each vertex $x \in V$, the combinatorial curvature at x is defined as

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

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,

$$2\pi\Phi(x) = 2\pi - \Sigma_x. \tag{1.187}$$

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],

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

whenever $\Sigma_{x \in G: \Phi(x) < 0} \Phi(x)$ converges. Thus, the combinatorial curvature captures a topological property of semiplanar 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 semiplanar graph G has nonnegative combinatorial curvature precisely if the polygonal surface $S(G)$ is an Alexandrov space with nonnegative sectional curvature, i.e. $\mathrm{Sec}\, S(G) \geq 0$ (or $\mathrm{Sec}(G) \geq 0$ 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

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

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 $a_1, \ldots, a_4 \in X$ 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

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

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,

$$B(a_1, r) \cap B(a_2, r) \cap B(a_3, r) \neq \emptyset, \tag{1.189}$$

(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

$$r \leq \bar{r}. \tag{1.190}$$

In more detail, we define the functions

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

and,

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

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

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

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. Bačak, M.: Convex Analysis and Optimization in Hadamard Spaces. De Gruyter, Berlin (2014)
18. Bačak, 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 SchrÃűdinger 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

Chapter 2
Metric Curvatures Revisited: A Brief Overview

Emil Saucan

Abstract We survey metric curvatures, special accent being placed upon the Wald curvature, its relationship with Alexandrov curvature, as well as its application in defining a metric Ricci curvature for *PL* cell complexes and a metric Ricci flow for *PL* surfaces. In addition, a simple, metric way of defining curvature for metric measure spaces is proposed.

2.1 Introduction

We begin with a brief motivation for our interest in the material of this chapter: The reason we study *Metric* Geometry is due to the fact that it provides us with a minimalistic framework that requires no additional smoothness, nor does it impose any supplementary, ad hoc structure upon a given geometric object. It is therefore our firm belief that this property renders the metric method as an approach ideally suited for the study of the various structures and problems encountered in computer Science in general, and in Graphics, Imaging and Vision in particular and that, moreover, this Newtonian stance of "hypotheses non fingo" represents not only the philosophically correct attitude, it also is ideally suited for the truthful intelligence of Digital Spaces. This terse and somewhat vague argument was augmented in [90] with examples drawn from Wavelet Theory (and practice), DNA Microarray Analysis, Imaging and Graphics, and we briefly mention them here, too, as well as a possible application to Networking.

To be sure, the term "Metric Geometry" is extensively used elsewhere and there exists a quite extensive literature on Metric Geometry and its various applications to Computer Science. (See, e.g. [1, 15, 49, 54], to mention only very small number of titles from an impressive literature—without any chance whatsoever of exhausting it.) Unfortunately, the "linear", or "first order Geometry" considered in the articles mentioned above does not fulfill all its expectations, and particularly so in Manifold

E. Saucan (✉)
Electrical Engineering Department, Technion, Haifa 32000, Israel

Mathematics Department, Technion, Haifa 32000, Israel
e-mail: semil@ee.technion.ac.il

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_2

Learning (as well as in Imaging and Graphics). To explain this, we could emphasize the essential importance of "second order Geometry", or succinctly, the Geometry of curvature, in Graphics, Imaging and Manifold Learning. We did this in some detail in our previous work already alluded to, so we refrain repeating ourselves, except to quote again the phrase so aptly coined by the regretted Robert Brooks: "The fundamental notion of differential geometry is the concept of curvature" [17]—And our goal here is to sketch the very basis of a metric ("discrete") Differential Geometry.

Before we conclude the introduction, we wish to outline the structure of the main part of the paper: In Sect. 2.2 we briefly discuss some metric notions of curvature for 1-dimensional geometric objects, i.e. curves. The accent is placed on the so called *Haantjes* (or *Finsler–Haantjes*) *curvature*, partly because we shall make appeal to it in the last section, where we suggest, amongst others, a simple approach to the introduction of a curvature for metric measure spaces. Section 2.3 represents the very nucleus, of this chapter. In its beginning we introduce the "main character" of this paper, namely the *Wald* (or *Wald–Berestovskii*) *curvature*, we discuss its properties and its relationship to the much better known notion of *Alexandrov* curvature and we apply it to develop a metric Ricci flow for *PL* surfaces as well as notion of metric Ricci curvature for cell complexes. In the forth—and last—we recall notion of *snowflaking* and we apply it to the introduction of a simple approach to curvature on metric measure spaces, with applications to Sampling Theory.[1]

2.2 Metric Curvature for Curves

We do not detail here all the basic, simple geometric ideas that reside behind the various notions of metric curvature; the reader who might feel a slower, more graduate and detailed exposition is needed is referred to our previous expositions [86] and [90]. (He might also consult with much benefit [10], and we refer the reader to this book, as well as to [12, 13] for any missing notions in our exposition.) Neither do we (as we already warned the reader) go into details over all the metric curvatures for curves. However, we find impossible to write an overview on the subject of metric curvatures without mentioning them, even if only very briefly.

[1] As the reader will become aware while progressing with this text, we have written previously a book chapters on metric curvatures as well as a largely expository article. However the present paper does not represent a calque of any of these previous ones. For one, the present one is addressed to a much more mathematically literate (not to say "very well educated") audience than the previous expositions. To be sure, certain repetitions are, unfortunately, unavoidable: After all, the same subject represents the common theme of all these three papers. However, we have strived to keep these at an inevitable minimum. Moreover, we did our best to emphasize different aspects (in general, more modern ones) as well as introducing some novel applications.

Fig. 2.1 A standard test image (*left*) and its Menger curvature (*right*)

2.2.1 Menger Curvature

The simplest and most direct version of metric curvature for curves is the so called *Menger curvature* κ_M. It's idea is to mimic the definition of the *osculatory circle*, by first defining the (metric) curvature of triangles (or triples of points), by defining the curvature $K(T)$ of a triangle T to be just $1/R(t)$, where $R(T)$ is the radius of the circle circumscribed to the triangle, then passing to the limit (exactly like in the standard, classical osculatory circle definition). To define $K(T)$, one makes appeal to some very elementary and quite well know formulas of high-school geometry. Unfortunately, this impose on the space under scrutiny an intrinsically Euclidean notion of curvature. Nevertheless, Menger curvature has been employed with considerable success to the study of such problems as finding estimates (obtained via the Cauchy integral) for the regularity of fractals and the flatness of sets in the plane (see [72]). As far as practical implementations are concerned, Menger curvature has been used for curve reconstruction [27]. Also, one might consider its use in for the obvious task of approximating the *principal curvatures*,[2] hence the computation of Gauss and mean curvature, of triangulated (polyhedral) surfaces, and their manifold uses in Graphics, Imaging and other, related fields. Experiments with the Menger curvature (see Fig. 2.1) and with the Haantjes curvature (see below and [92]) have shown, unfortunately, only moderate success.

The reader may have already recognized Menger curvature to be nothing else than Gromov's $\mathbf{K}_3(\{p, q, r\})$, where p, q, r represent the vertices of a triangle, introduced in Mémoli's exposition in the present volume. However, while the general setting of the modern discourse are certainly important, there still is use for the classical, "parochial" Menger curvature, as demonstrated by its many contemporary uses, thus interest and mathematical activity around the "old fashioned" Menger

[2] We know that in mentioning this here we anticipate, somewhat, the reminder of the paper.

curvature still exists, even though the much more general and modern setting of $\mathbf{K}_n(X)$ is much more alluring. Moreover, it should be noted that, beyond this generality there rests a lot of uncertainty, as Gromov himself notes [31].

Remark 2.1 Discussing Menger curvature, and mainly the idea behind it, one can hardly not mention the very recent development [7].[3] Here a kind of "comparison Menger" curvature is introduced. Very loosely formulated, whereas in the classical Menger curvature a specific, Euclidean in nature, formula is developed for the circumradius, here it is only compared to that of a triangle of the same sides, in different model (or gauge) surfaces. This is similar (and presumably inspired by) the Alexandrov curvature (which we shall discuss later on).

Remark 2.2 While again anticipating higher dimensional curvatures (and mainly 2-dimensional ones), we can not leave this short section on Menger curvature without mentioning the *Menger curvature measure*:

$$\mu(\mathscr{T}) = \mu_p(\mathscr{T}) = \sum_{T\in\mathscr{T}} \kappa_M^p(T)(\operatorname{diam} T)^2\,, \tag{2.1}$$

for some $p \geq 1$, where $\kappa_M(T)$ denotes the Menger curvature (of the triangle T).

(See, e.g. [50] and, for a somewhat different approach and another range of problems altogether, [98].)

In the applications range, one possible use of the Menger curvature measure (in its alternative variant) is in the field of Pattern Recognition, for texture segmentation and classification—see [29]. This represents an approach to non local "operators" for Imaging, radically different (being "purely" geometric) from those of Osher et al. [28, 48] and Jost et al. [44, 45].

2.2.2 *Haantjes Curvature*

Having only very briefly presented the Menger curvature, we shall expound in somewhat more detail upon another, less commonly known notion of metric curvature, namely the so called *Haantjes curvature* or *Finsler–Haantjes curvature*.[4] We have chosen to do so not only because this curvature does not mimic $\mathbb{R}^2$ (we already have emphasized this point in [90] and elsewhere), nor because of its adaptivity to applications (again, see [90] and the bibliography within, but see also Remark 2.4 below). The reasons behind our choice are that we both want to present some of Haantjes connections with other notions (not just of curvature), and because

[3] Although, when these notes were started, the mentioned work was still not published.

[4] Named after Haantjes [40], who extended to metric spaces an idea introduced by Finsler in his PhD Thesis.

we wish to suggest it represents a simple and direct alternative—at least in certain applications—of more involved and fashionable concepts.

Definition 2.1 (*Haantjes Curvature*) Let (M, d) be a metric space and let $c : I = [0, 1] \xrightarrow{\sim} c(I) \subset M$ be a homeomorphism, and let $p, q, r \in c(I),\ q, r \neq p$. Denote by $\widehat{qr}$ the arc of $c(I)$ between q and r, and by qr line segment from q to r.

We say that c has *Haantjes curvature* $\kappa_H(p)$ at the point p iff:

$$\kappa_H^2(p) = 24 \lim_{q,r \to p} \frac{l(\widehat{qr}) - d(q,r)}{\left(l(\widehat{qr})\right)^3} \ ; \tag{2.2}$$

where "$l(\widehat{qr})$" denotes the length—in intrinsic metric induced by d—of $\widehat{qr}$.

Remark 2.3 Since for points/arcs where Haantjes curvature exists, $\frac{l(\widehat{qr})}{d(q,r)} \to 1$, as $d(q, r) \to 0$ (see [40]), κ_H can alternatively be defined (see, e.g. [47]) as

$$\kappa_H^2(p) = 24 \lim_{q,r \to p} \frac{l(\widehat{qr}) - d(q,r)}{\left(d(q,r)\right)^3} \ ; \tag{2.3}$$

As it turns out, in applications it is this alternative form of the definition of Haantjes curvature proves itself to be more malleable (see [90] for some details, as well as Fig. 2.2).

The definition of Haantjes curvature (in both its versions) is quite intuitive and even the $\left(d(q, r)\right)^3$ factor is clearly inserted for scaling reasons. Far less intuitive (and somewhat puzzling) is the "24" factor. However, it arises quite naturally in the proof of following basic (and reassuring, so to say[5]) theorem:

Theorem 2.1 *Let $\gamma \in \mathscr{C}^3$ be smooth curve in $\mathbb{R}^3$ and let $p \in \gamma$ be a regular point. Then the metric curvature $\kappa_H(p)$ exists and equals the classical curvature of γ at p.* (For a proof, see [40] or—probably somewhat more easily accessible—[12].)

Moreover, the same result holds for Menger curvature (see [12]). In fact, the two curvatures (when both applicable) coincide, as shown by the next result:

Theorem 2.2 (*Haantjes*) *Let γ be a rectifiable arc in a metric space (M, d), and let $p \in \gamma$. If κ_M and κ_H exist, then they are equal.*

Remark 2.4 With the risk of being somewhat redundant, but to mirror our presentation of Menger curvature, we mention here, in brief, that Haantjes has proved its versatility in such diverse fields as Imaging and Graphics, Wavelets (with applications for texture segmentation) and DNA Microarray Analysis. (For details see again [90] and the bibliography within. See also [62] for an application to the quasi-conformal and quasi-isometric planar representation of (medical) images.)

[5] Since it proves us that, indeed, for smooth curves, Haantjes curvature coincides with the classical notion of curvature.

Fig. 2.2 A natural image (*left*) and its average Haantjes curvature (*right*)

2.2.2.1 Haantjes Curvature and Excess

A metric geometry concept that has been proven itself as both flexible and powerful, in many mathematical settings, and in particular in studying the Global Geometry of Manifolds (see, e.g., [33, 34] and the bibliography therein), is the *excess*:

Definition 2.2 (*Excess*) Given a triangle[6] $T = \triangle(pxq)$ in a metric space (X, d), the *excess* of T is defined as

$$e = e(T) = d(p,x) + d(x,q) - d(p,q). \tag{2.4}$$

A local version of this notion was introduced (seemingly by Otsu [71]), namely the *local excess* (or, more precisely, the *local d-excess*):

$$e_d(X) = \sup_p \sup_{x \in B(p,\rho)} \inf_{q \in S(p,\rho)} \left(e(\triangle(pxq))\right), \tag{2.5}$$

where $\rho \leq \mathrm{rad}(X) = \inf_p \sup_q d(p,q)$, (and where $B(p,\rho), S(p,\rho)$ stand—as they usually do—for the ball and respectively sphere of center p and radius ρ).

In addition, global variations of this quantity have also been defined:

$$e(X) = \inf_{(p,q)} \sup_x \left(e(\triangle(pxq))\right), \tag{2.6}$$

and, the so called *global big excess* (see [71]):

$$E(X) = \sup_q \inf_p \sup_x \left(e(\triangle(pxq))\right). \tag{2.7}$$

[6]Not necessarily geodesic.

Intuitively, it is clear that (local) excess and curvature are closely related concepts, since the geometric "content" of the notion of local excess resides in the fact that, for any $x \in B(p, \rho)$, there exists a (minimal) geodesic γ from p to $S(p, \rho)$ such that γ is close to x. More precisely, we have the following relation between the two notions:

$$\kappa_H^2(T) = \frac{e}{\rho^3}, \tag{2.8}$$

where $\rho = \rho(p, q)$, and where by the curvature of a triangle $T = T(pxq)$ we mean the curvature of the path $\widehat{pxq}$. Here and below we have used a simplified notation and discarded (for sake of simplicity and clarity) the normalizing constant "24".[7] Thus Haantjes curvature can be viewed as a *scaled* version of excess. Keeping this in mind, one can define also a global version of this type of metric curvature, namely by defining, for instance:

$$\kappa_H^2(X) = \frac{E(X)}{\mathrm{diam}^3(X)}, \tag{2.9}$$

or

$$\kappa_H^2(X) = \frac{e(X)}{\mathrm{diam}^3(X)}, \tag{2.10}$$

as preferred.

To be sure, one can proceed in the opposite direction and express the proper (i.e. pointwise) Haantjes curvature by means of the definition (2.5) of local excess, as

$$\kappa_H^2(x) = \lim_{\rho \to 0} e(x). \tag{2.11}$$

Remark 2.5 We should mention that both Menger and Haantjes curvatures have their more modern ("updated" and "sophisticated") respective versions—see [2]. However, let us add here that, at least as far as applications are concerned, we favor the older notions over their more modern "avatars", not solely for their simplicity, but also for a number of reasons, mainly appertaining to their potential:

1. First and foremost, while the Alexander–Bishop variants are more "tight", so to say, they coincide with their classical counterparts on all but the most esoteric spaces;
2. They are applicable to more general settings, fact that represents a further incentive in their application in discrete (i.e. Computer Science driven) settings;

[7] In any case, it is not truly required and, in fact, even cumbersome in practical applications (see [4, 91] for two such examples).

3. In addition, no *a priori* knowledge of the global geometry of the ambient space (i.e. Alexandrov curvature) is presumed, nor is it necessary to first determine the curves of constant curvature (see [2]) in order to compute these curvatures; furthermore
4. They are easy to compute in a direct fashion in the discrete setting (at least amongst those discrete versions we encountered), thus they are simpler and far more intuitive;
5. Last—but certainly not least—they are more ready to lend themselves to discretization, hence admit easy and direct "semi-discrete" (or "semi-continuous") versions, as the one mentioned in Remark 2.4 above. In view of this and their simplicity noted above, they prove to be more conducive towards practical applications.

2.3 Metric Curvature for Surfaces: Wald Curvature

2.3.1 Wald Curvature

We introduce here the main type of metric curvature that we overview in this chapter, namely the so called *Wald curvature*. Wald's seminal idea was to go back to Gauss' original method of defining surface curvature by comparison to a standard, model surface (i.e. the unit sphere in $\mathbb{R}^3$), while extending it to general gauge surfaces, rather than restrict himself to the unit sphere. Moreover, instead of comparing infinitesimal areas (which would be an impossible task in general metric space not endowed with a measure), he compared quadrangles. More precisely, his starting point was the following definition:

Definition 2.3 Let (M, d) be a metric space, and let $Q = \{p_1, \ldots, p_4\} \subset M$, together with the mutual distances: $d_{ij} = d_{ji} = d(p_i, p_j)$; $1 \leq i, j \leq 4$. The set Q together with the set of distances $\{d_{ij}\}_{1 \leq i,j \leq 4}$ is called a *metric quadruple*.

Remark 2.6 The reader has undoubtedly already recognized that the definition above conducts toward $\mathbf{K}_4(X)$. Indeed, we can view, in a sense, this chapter as representing an extended overview of and discussion on $\mathbf{K}_4(X)$, see also Chap. 3.

Remark 2.7 The following slightly more abstract definition can be also considered, one that does not make appeal to the ambient space: a metric quadruple being a 4 point metric space, i.e. $Q = \left(\{p_1, \ldots, p_4\}, \{d_{ij}\}\right)$, where the distances d_{ij} verify the axioms for a metric. However, this comes at a price, as we shall shortly see in Remark 2.8.

Before being able to pass to the next definition we need to introduce some additional notation: S_κ denotes *the complete, simply connected surface of constant Gauss curvature* κ (or *space form*), i.e. $S_\kappa \equiv \mathbb{R}^2$, if $\kappa = 0$; $S_\kappa \equiv \mathbb{S}^2_{\sqrt{\kappa}}$, if $\kappa > 0$; and $S_\kappa \equiv \mathbb{H}^2_{\sqrt{-\kappa}}$, if $\kappa < 0$. Here $S_\kappa \equiv \mathbb{S}^2_{\sqrt{\kappa}}$ denotes the sphere of radius $R = 1/\sqrt{\kappa}$,

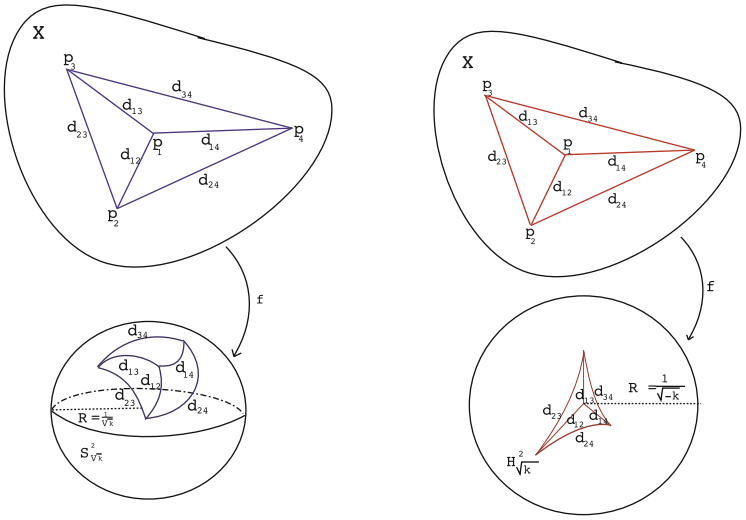


Fig. 2.3 Isometric embedding of a metric quadruple in a gauge surface: $\mathbb{S}^2_{\sqrt{\kappa}}$ (*left*) and $\mathbb{H}^2_{\sqrt{\kappa}}$ (*right*)

and $S_\kappa \equiv \mathbb{H}^2_{\sqrt{-\kappa}}$ stands for the hyperbolic plane of curvature $\sqrt{-\kappa}$, as represented by the Poincaré model of the plane disk of radius $R = 1/\sqrt{-\kappa}$ (Fig. 2.3).

Definition 2.4 The *embedding curvature* $\kappa(Q)$ of the metric quadruple Q is defined to be the curvature κ of the gauge surface S_κ into which Q can be isometrically embedded—if such a surface exists.

Remark 2.8 Even though the basic idea of embedding curvature is, in truth, quite intuitive, care is needed if trying to employ it directly, since there are a number of issues that arise (as we have anticipated in Remark 2.7 above):

1. If one uses the second (abstract) definition of the metric curvature of quadruples, then the very existence of $\kappa(Q)$ is not assured, as it is shown by the following

   Counterexample 2.1 *The metric quadruple of lengths*

   $$d_{12} = d_{13} = d_{14} = 1;\ d_{23} = d_{24} = d_{34} = 2$$

   admits no embedding curvature.
2. Any linear quadruple is embeddable, apart from the Euclidean plane, in all hyperbolic planes (i.e. of any strictly negative curvature), as well as in infinitely many spheres (whose radii are sufficiently large for the quadruple to be realized upon them).

3. Moreover, even if a quadruple has an embedding curvature, it still may be not unique (even if Q is not linear); as it is illustrated by the following examples:

   Example 2.1

   a. For each $\kappa > 0$, each neighborhood of any point $p \in S_\kappa$ contains a non-degenerate quadruple that is also isometrically embeddable in $\mathbb{R}^2$. (For the proof see [12, pp. 372–373].)
   b. The quadruple Q of distances $d_{13} = d_{14} = d_{23} = d_{24} = \pi$, $d_{12} = d_{34} = 3\pi/2$ admits exactly two embedding curvatures: $\kappa_1 = \frac{1}{2}$ and $\kappa_2 \in \left(\frac{1}{4}, \frac{4}{9}\right)$. (See [13].)

We are now able to define the *Wald curvature* [106, 107] (or, more precisely, its modification due to Berestovskii [9]):

Definition 2.5 Let (X, d) be a metric space. An open set $U \subset X$ is called a *region of curvature* $\geq \kappa$ iff any metric quadruple can be isometrically embedded in S_m, for some $m \geq k$. A metric space (X, d) is said to have *Wald-Berestovskii curvature* $\geq \kappa$ iff any $x \in X$ is contained in a region $U = U(x)$ of curvature $\geq \kappa$.

The embedding curvature at a point is now definable most naturally as a limit. However, we need first yet another preparatory definition:

Definition 2.6 (M, d) be a metric space, let $p \in M$ and let N be a neighborhood of p. Then N is called *linear* iff N is contained in a geodesic curve.

Definition 2.7 Let (M, d) be a metric space, and let $p \in M$ be an accumulation point. Then M has (*embedding*) *Wald curvature* $\kappa_W(p)$ at the point p iff

1. Every neighborhood of p is non-linear;
2. For any $\varepsilon > 0$, there exists $\delta > 0$ such that if $Q = \{p_1, \ldots, p_4\} \subset M$ and if $d(p, p_i) < \delta\,, i = 1, \ldots, 4$; then $|\kappa(Q) - \kappa_W(p)| < \varepsilon$.

Fortunately, for "nice" metric spaces—i.e. spaces that are locally sufficiently "plane like"—the embedding curvature exists and it is unique (see, e.g, [12] and, for a briefer but more easily accessible presentation, [86]). Moreover—and this represents a fact that is very important for some of our own goals, as detailed further on (see Sect. 2.6.1)—this embedding curvature coincides with the classical Gaussian curvature. Indeed, one has the following result due to Wald:

Theorem 2.3 (Wald [107]) *Let $S \subset \mathbb{R}^3$, $S \in \mathscr{C}^m$, $m \geq 2$ be a smooth surface. Then, given $p \in S$, $\kappa_W(p)$ exists and $\kappa_W(p) = K(p)$, where $K(p)$ denotes the Gaussian curvature at p.*

Remark 2.9 In the theorem above the metric considered in the computation of Wald curvature is the intrinsic one of the surface. (Indeed, the reciprocal Theorem 2.4 below is formulated, at least *prima facie*, for a much more general class of metric spaces than mere smooth surfaces embedded in Euclidean 3-space.) However, in applications Euclidean (extrinsic) distances are used instead. However, this does not

represent a theoretical obstruction (only, perhaps, a practical one)—see, for instance, [90, Sect. 4.3]and the references therein.[8]

Moreover, Wald also has shown that the following partial reciprocal theorem also holds:

Theorem 2.4 *Let M be a compact and convex metric space. If $\kappa_W(p)$ exists, for all $p \in M$, then M is a smooth surface and $\kappa_W(p) = K(p)$, for all $p \in M$.*

Remark 2.10 Obviously, here the metric considered is the abstract one of the given metric space, that is proven to coincide with the intrinsic one of a smooth surface.

The results above, in conjunction, show that Wald curvature represents, indeed, a proper metrization of the classical (smooth) notion, and not just a mathematical "divertissement", lacking any significant geometric content.

We continue with a definition whose full significance will become more clear in the sequel, where it will be viewed in the correct perspective.

Definition 2.8 A metric quadruple $Q = Q(p_1, p_2, p_3, p_4)$, of distances $d_{ij} = dist(p_i, p_j)$, $i = 1, \ldots, 4$, is called *semi-dependent* (or a *sd-quad*, for brevity), there exist three indices, e.g. 1, 2, 3, such that: $d_{12} + d_{23} = d_{13}$.

Remark 2.11 The condition in the definition above implies, in fact, that the three points in question lie on a common *metric segment* i.e. a subset of a given metric space that is isometric to a segment in $\mathbb{R}$ (see [12, p. 246]).

Perhaps the main advantages of sd-quads stems from in the following fact:

Proposition 2.1 *An sd-quad admits at most one embedding curvature.*

In fact, there also exists a classification criterion—due to Berestovskii [9], see also [79], Theorem 18—for embedding curvature possibilities in the general case:

Theorem 2.5 *Let M, Q be as above. Then one and only one of the following assertion holds:*

1. *Q is linear.*
2. *Q has exactly one embedding curvature.*
3. *Q can be isometrically embedded in some $\mathscr{S}^m_\kappa$, $m \geq 2$; where $\kappa \in [\kappa_1, \kappa_2]$ or $(-\infty, \kappa_0]$, where $\mathscr{S}^m_\kappa \equiv \mathbb{R}^m$, if $\kappa = 0$; $\mathscr{S}^m_\kappa \equiv \mathbb{S}^m_{\sqrt{\kappa}}$, if $\kappa > 0$; and $\mathscr{S}^m_\kappa \equiv \mathbb{H}^m_{\sqrt{-\kappa}}$, if $\kappa < 0$. Moreover, $\kappa \in \{\kappa_0, \kappa_1, \kappa_2\}$. represent the only possible values of* planar *embedding curvatures, i.e. such that $m = 2$. (Here $\mathbb{S}^m_{\sqrt{\kappa}}$ denotes the m-dimensional sphere of radius $R = 1/\sqrt{\kappa}$, and $\mathbb{H}^m_{\sqrt{-\kappa}}$ stands for the m-dimensional hyperbolic space of curvature $\sqrt{-\kappa}$, as represented by the Poincaré model of the ball of radius $R = 1/\sqrt{-\kappa}$.)*
4. *There exist no m and k such that Q can be isometrically embedded in $\mathscr{S}^m_\kappa$.*

[8] The literature on the subject being too vast to even begin and enumerate it here.

2.3.1.1 A Local-to-Global Result

Before passing to the actual computation of Wald curvature, we include here a result who's full importance and relevance will become much clearer later on. More precisely, we bring the fitting version of the Toponogov (or Alexandrov–Toponogov) Comparison Theorem:

Theorem 2.6 (Toponogov's Comparison Theorem for Wald Curvature) *Let (X, d) be an inner metric space of curvature $\geq k$. Then the entire X is a region of Wald curvature $\geq k$.*

Since the proof is somewhat lengthy and technical we do not bring it here—see [75] (see also [79]).

2.3.1.2 Computation of Wald Curvature I: The Exact Formula

A non-negligible part of the attractiveness of Wald curvature does not reside in its simplicity and intuitiveness, but also that it comes endowed, so to say, with a simple formula for its actual computation. (This is in stark contrast with the *Alexandrov* (*comparison*) *curvature* at least in its usual presentation—but we shall elaborate later on on this subject.) More precisely, we have the following formula:

$$\kappa(Q) = \begin{cases} 0 & \text{if } D(Q) = 0\,; \\ \kappa,\ \kappa < 0 & \text{if } det(\cosh \sqrt{-\kappa} \cdot d_{ij}) = 0\,; \\ \kappa,\ \kappa > 0 & \text{if } det(\cos \sqrt{\kappa} \cdot d_{ij}) \text{ and } \sqrt{\kappa} \cdot d_{ij} \leq \pi \\ & \quad \text{and all the principal minors of order 3 are } \geq 0; \end{cases} \tag{2.12}$$

where $d_{ij} = d(x_i, x_j), 1 \leq i, j \leq 4$, $(\cosh \sqrt{-\kappa} \cdot d_{ij})$ is a shorthand for $(\cosh \sqrt{-\kappa} \cdot d_{ij})_{1 \leq i,j \leq n}$, etc., and $D(Q)$ denotes the so called *Cayley–Menger determinant*:

$$D(x_1, x_2, x_3, x_4) = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & d_{12}^2 & d_{13}^2 & d_{14}^2 \\ 1 & d_{12}^2 & 0 & d_{23}^2 & d_{24}^2 \\ 1 & d_{13}^2 & d_{23}^2 & 0 & d_{34}^2 \\ 1 & d_{14}^2 & d_{24}^2 & d_{34}^2 & 0 \end{vmatrix} . \tag{2.13}$$

There is, in fact, nothing mysterious about the formula above. Indeed, it has a very simple geometric meaning ensuing from the following fact:

$$D(p_1, p_2, p_3, p_4) = 8\big(Vol(p_1, p_2, p_3, p_4)\big)^2 , \tag{2.14}$$

where $Vol(p_1, p_2, p_3, p_4)$ denotes the (un-oriented) volume of the parallelepiped determined by the vertices $p_1, \ldots, p_4$ (and with edges $\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_1p_4}$).[9] From here immediately follows that

Proposition 2.2 *The points $p_1, \ldots, p_4$ are the vertices of a nondegenerate simplex in $\mathbb{R}^3$ iff $D(p_1, p_2, p_3, p_4) \neq 0$.*

Clearly, this also implies the opposite assertion, namely that a simplex of vertices $p_1, \ldots, p_4$ is degenerate, i.e. isometrically embeddable in the plane $\mathbb{R}^2 \equiv S_0$.

It is now relatively easy to guess that the expressions appearing in Formula (2.13) for the cases where $\kappa \neq 0$ represent the equivalents of $D(Q)$ in the hyperbolic, respective spherical cases, using the well known fact, that, in the spherical (resp. hyperbolic) metric, the distances d_{ij} are replaced by $\cos d_{ij}$ (resp. $\cosh d_{ij}$). However, the proof of this fact, as well for the analogous formulas and results in higher dimension diverge from the boundaries of this restricted exposition, therefore we refer the reader to [12].

Remark 2.12 A stronger result along these lines also exists. Moreover, it is readily generalized to any dimension. For proofs and further details, see [12].

2.3.1.3 Computation of Wald Curvature II: An Approximation

Unfortunately, using Formula (2.12) for the actual computation of $\kappa(Q)$ is anything but simple, since the equations involved are—apart from the Euclidean case—transcendental, therefore not solvable, in general, using elementary methods. Moreover, they tend to display a numerical instability when solved with computer assisted methods. (See [85, 92] for a more detailed comments and some numerical results.)

Note that Formula (2.12) implies that, in practice, a renormalization might be necessary for some of the vertices of positive Wald–Beretkovskii curvature, which represents yet another impediment in it use.

Fortunate enough, there exists a good approximation result, due to Robinson. Not only does his result give a rational formula for approximating $\kappa(Q)$ and provide good error estimates, it also solves one other problem inherent in the use of the Wald curvature, namely the possible lack of uniqueness of the computed curvature. The way to circumvent this difficulty and the other pitfalls of Formula (2.12) is to make appeal to the simpler geometric configuration of sd-quads:

[9] As a historical note, it is perhaps worthwhile to recall that Formula (2.14) above was proved by Cayley in his very first mathematical paper [21] (published while he was still begrudgingly making his living as a lawyer!...).

Theorem 2.7 ([81]) *Given the metric semi-dependent quadruple* $Q = Q(p_1, p_2, p_3, p_4)$*, of distances* $d_{ij} = d(p_i, p_j)$, $i, j = 1, \ldots, 4$*; the embedding curvature* $\kappa(Q)$ *admits a rational approximation given by:*

$$K(Q) = \frac{6(\cos \measuredangle_0 2 + \cos \measuredangle_0 2')}{d_{24}\left(d_{12} \sin^2(\measuredangle_0 2) + d_{23} \sin^2(\measuredangle_0 2')\right)} \tag{2.15}$$

where: $\measuredangle_0 2 = \measuredangle(p_1 p_2 p_4)$, $\measuredangle_0 2' = \measuredangle(p_3 p_2 p_4)$ *represent the angles of the Euclidean triangles of sides* d_{12}, d_{14}, d_{24} *and* d_{23}, d_{24}, d_{34}*, respectively (see also Fig. 2.4). Moreover the absolute error R satisfies the following inequality:*

$$|R| = |R(Q)| = |\kappa(Q) - K(Q)| < 4\kappa^2(Q)\mathrm{diam}^2(Q)/\lambda(Q)\,, \tag{2.16}$$

where $\lambda(Q) = d_{24}(d_{12} \sin \measuredangle_0 2 + d_{23} \sin \measuredangle_0 2')/S^2$*, and where* $S = Max\{p, p'\}$; $2p = d_{12} + d_{14} + d_{24}$, $2p' = d_{32} + d_{34} + d_{24}$.

We do not bring here the proof of Robinson' result—the interested reader can see [85] and in [92] (and, of course, to Robinson's original paper [81]). However, we would like to underline that basic idea of the proof is to basically calque, in a general metric setting, the original way of defining Gaussian curvature—in this case, rather than accounting for the area distortion, one measures the curvature by the amount of "bending" one has to apply to a general planar quadruple so that it may be "straightened" to a triangle $\triangle(p_1 p_3 p_4)$, with p_2 lying on the edge $p_1 p_3$—i.e. isometrically embedded as a *sd-quad*— in some S_κ.

Remark 2.13 In some special cases (e.g. when $d_{12} = d_{32}$, etc.) simpler formulas are obtained instead of (2.15)—see, again, [81], or [85, 92].

Naturally, there raises the question whether Formula (2.15) (or any of its variations mentioned above) is truly efficient in applications. The following example, due

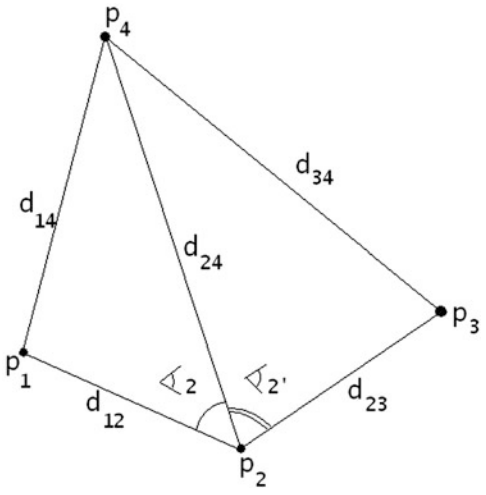


Fig. 2.4 The Euclidean triangles corresponding to an sd-quad

also to Robinson, indicates that, at least in some cases, the actual computed error is far smaller then the theoretical one provided by Formula (2.16).

Example 2.2 ([81]) Let Q_0 be the quadruple of distances $d_{12} = d_{23} = d_{24} = 0.15$, $d_{14} = d_{34}$ and of embedding curvature $\kappa = \kappa(Q_0) = 1$. Then $\kappa S^2 < 1/16$ and $K(Q_0) \approx 1.0030280$, whereas the error computed using Formula (2.16) is $|R| < 0.45$.
For some experimental results and comparison to other metric curvatures for images, see [92]. However, we should emphasize that the results therein do comply to the expectations arising from the following (quite expected, but nevertheless necessary) theorem:

Theorem 2.8 *Let S be a smooth (differentiable) surface. Then, for any point $p \in S$:*

$$K_G(p) = \lim_{n\to 0} K(Q_n)\,;$$

for any sequence $\{Q_n\}$ of sd-quads that satisfy the following conditions:

$$Q_n \to Q = \Box p_1 p p_3 p_4\,;\ \mathrm{diam}(Q_n) \to 0\,.$$

Sketch of Proof Recall that the Gaussian curvature $K_G(p)$ at a point p is given by:

$$K_G(p) = \lim_{n\to 0} \kappa(Q_n)\,;$$

where $Q_n \to Q = \Box p_1 p p_3 p_4\,;\ \mathrm{diam}(Q_n) \to 0$. But, if Q is any sd-quad, then $\kappa^2(Q)\mathrm{diam}^2(Q)/\lambda(Q) < \infty$. Moreover, $|R|$ is small if Q is not close to linearity. In this case $|R(Q)| \sim \mathrm{diam}^2(Q)$, for any given Q (see [81]). The theorem now follows easily. □

Remark 2.14 The convergence result provided in Theorem 2.8 is not just in the sense of measures and errors of different signs do not simply cancel each other. Indeed, $\mathrm{sign}(\kappa(Q)) = \mathrm{sign}(K(Q))$, for any metric quadruple Q.

Wald Curvature and Isometric Embeddings

Proposition 2.2 and Remark 2.12 rise the general problem of the existence of isometric embeddings of generic metric metric spaces into gauge spaces. While in its full generality this is, of course, an unattainable goal, one would still be interested in the much more restricted, but important in the applied setting (Graphics, Imaging, Mathematical Modeling, Networking etc.), problem of isometric embedding of *PL* surfaces in $\mathbb{R}^3$.

A partial result in this direction is a criterion for the local isometric embedding of polyhedral surfaces in $\mathbb{R}^3$, resemblant to the classical Gauss fundamental (compatibility) equation in the classical differential geometry of surfaces, that we

proved in [88]. However, to be able to formulate it we need first some additional notations and results:

First, let us note that, in the context of polyhedral surfaces, the natural choice for the set U required in Definition 2.5 is the *star* of a given vertex v, that is, the set $\{e_{vj}\}_j$ of edges incident to v. Therefore, for such surfaces, the set of metric quadruples containing the vertex v is finite.

Equipped with this quite simple and intuitive choice for U (and in in analogy with Alexandrov spaces—see also Sect. 2.5.1 below) it is quite natural to consider, for *PL* surfaces, the following definition of the *discrete* (*PL*, or "finite scale") Wald curvature $K_W(v)$ at the vertex v:

$$K_W(v) = \min_{v_i, v_j, v_k \in \mathrm{Lk}(v)} K_W^{ijk}(v)\,, \tag{2.17}$$

where $K_W^{ijk}(v) = \kappa(v; v_i, v_j, v_k)$, and where $\mathrm{Lk}(v)$ denotes the *link* of the vertex v.[10] Note that here we consider the (intrinsic) *PL* distance between vertices.

Let $Q = \{x_1, x_2, x_3, x_4\}$ be a *metric quadruple* and let $V_\kappa(x_i)$ be defined as follows:

$$V_\kappa(x_i) = \alpha_\kappa(x_i; x_j, x_l) + \alpha_\kappa(x_i; x_j, x_m) + \alpha_\kappa(x_i; x_l, x_m) \tag{2.18}$$

where $x_i, x_j, x_l, x_m \in Q$ are distinct, and κ is any number, and where the angles $\alpha_i\,, i = 1, 2, 3$ are as in Fig. 2.5.

Proposition 2.3 ([79], Theorem 23) *Let (X, d) be a metric space and let $U \in X$ be an open set. U is a region of curvature $\geq \kappa$ iff $V_\kappa(x) \leq 2\pi$, for any metric quadruple $\{x, y, z, t\} \subset U$.*

We can now state the desired result for local isometric embedding of polyhedral surfaces in $\mathbb{R}^3$: Given a vertex v, with metric curvature $K_W(v)$, the following system of inequalities should hold:

$$\begin{cases} \max A_0(v) \leq 2\pi; \\ \alpha_0(v; v_j, v_l) \leq \alpha_0(v; v_j, v_p) + \alpha_0(v; v_l, v_p), \text{ for all } v_j, v_l, v_p \sim v; \\ V_\kappa(v) \leq 2\pi; \end{cases} \tag{2.19}$$

Here

$$A_0 = \max_i V_0\,; \tag{2.20}$$

[10] Recall that the *link* $\mathrm{lk}(v)$ of a vertex v is the set of all the faces of $\overline{\mathrm{St}}(v)$ that are not incident to v. Here $\overline{\mathrm{St}}(v)$ denotes the *closed star* of v, i.e. the smallest subcomplex (of the given simplicial complex K) that contains $\mathrm{St}(v)$, namely $\overline{\mathrm{St}}(v) = \{\sigma \in \mathrm{St}(v)\} \cup \{\theta \mid \theta \leqslant \sigma\}$, where $\mathrm{St}(v)$ denotes the *star* of v, that is the set of all simplices that have v as a face, i.e $\mathrm{St}(v) = \{\sigma \in K \mid v \leqslant \sigma\}$.

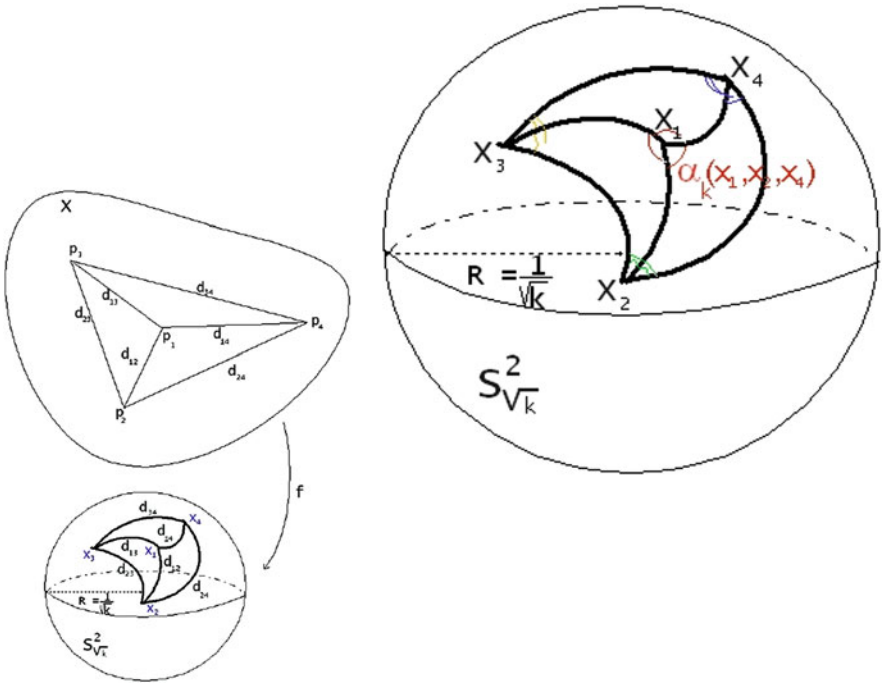


Fig. 2.5 The angles $\alpha_\kappa(x_i, x_j, x_l)$ (*right*), induced by the isometric embedding of a metric quadruple in $\mathbb{S}^2_{\sqrt{\kappa}}$ (*left*)

"$\sim$" denotes incidence, i.e. the existence of a connecting edge $e_i = vv_j$ and, of course, $V_\kappa(v) = \alpha_\kappa(v; v_j, v_l) + \alpha_\kappa(v; v_j, v_p) + \alpha_\kappa(v; v_l, v_p)$, where $v_j, v_l, v_p \sim v$, etc.

Returning to the analogy with the Gauss compatibility equation, the first two inequalities represent the (extrinsic) embedding condition, while the third one represents the intrinsic curvature (of the *PL* manifold) at the vertex v.

Also, for details and a corresponding global embedding criterion see [88].

Remark 2.15 Before passing to more general issues, let us mention here that, precisely as a Menger measure was introduced, one can also consider (and, in fact, much more naturally) a *Wald measure* (for surfaces);

$$\mu_W(v) = K_W(v) \cdot \text{Area}(St(v)), \tag{2.21}$$

where $St(v)$ denotes the star of the vertex v.

We defer the investigation for future study of its usefulness in practice.

2.4 Wald Curvature Under Gromov–Hausdorff Convergence

It is practically impossible, both from a purely mathematical viewpoint, as well as considering the background of this volume as a whole (see mainly Memoli's contribution) to introduce any notion of curvature without discussing its behavior under the Gromov–Hausdorff convergence. We begin with a much more general discussion and continue with some results regarding Wald curvature and Gromov–Hausdorff convergence (see also Chap. 3).

2.4.1 Intrinsic Properties and Gromov–Hausdorff Convergence

Given that many of Graphics and Imaging tasks (and Sampling Theory, as well), reduce, in the end, to better and better approximation by certain nets (or graphs), be they triangular meshes in the first case, or square grids, in the second, it is most natural (and, indeed, necessary) to have a comprehensive, and sound approach to investigating the behavior and convergence, under limits, of the relevant properties.[11] We overview here some significant results regarding convergence of nets in metric spaces (basically due, seemingly, to Gromov).

We begin by reminding the reader the following basic definition:

Definition 2.9 Let (X, d) be a metric space. A set $\{p_1, \ldots, p_m\} \subset X$ is called an ε- *net* on(in) X iff the balls $B(p_k, \varepsilon)$, $k = 1, \ldots, m$ cover X.

It turns out that ε-nets in compact metric spaces have the following important property:

Proposition 2.4 *Let $X, \{X_n\}_{n=1}^{\infty}$ be compact metric spaces. Then $X_n \underset{GH}{\longrightarrow} X$ iff for all $\varepsilon > 0$, there exist* finite *ε-nets $S \subset X$ and $S_n \subset X_n$, such that $S_n \underset{GH}{\longrightarrow} S$ and, moreover, $|S_n| = |S|$, for large enough n.*

The importance of the result above does not reside only in the fact that compact metric spaces can be approximated by finite ε-nets—after all, just the existence of *some* approximation by such sets is hardly surprising—but rather in the fact that, as we shall shortly see, it assures the convergence of geometric properties of S_n to those of S, as $X_n \underset{GH}{\longrightarrow} X$. This would be, in a nutshell, the real significance of the proposition above.

One can also reformulate Proposition 2.4 in an equivalent, in a less concise and elegant manner yet, on the other hand, far more useful in concrete instances (to say nothing of the fact that it is far more familiar in the Applied Mathematics community):

[11] It was, it would appear, Gromov's observation that, in the geometric setting, the relevant convergence is the Gromov–Hausdorff one.

Proposition 2.5 *Let X, Y be compact metric spaces. Then:*

(a) If Y is a (ε, δ)-approximation of X, then $d_{GH}(X, Y) < 2\varepsilon + \delta$.
(b) If $d_{GH}(X, Y) < \varepsilon$, then Y is a 5ε-approximation of X.

Recall that *ε-δ-approximations* are defined as follows:

Definition 2.10 Let X, Y be compact metric spaces, and let $\varepsilon, \delta > 0$. X, Y are called *ε-δ-approximations* (of each other) iff: there exist sequences $\{x_i\}_{i=1}^N \subset X$ and $\{y_i\}_{i=1}^N \subset Y$ such that

(a) $\{x_i\}_{i=1}^N$ is an ε-net in X and $\{y_i\}_{i=1}^N$ is an ε-net in Y;
(b) $|\, d_X(x_i, x_j) - d_(y_i, y_j)\,| < \delta$ for all $i, j \in \{1, \ldots, N\}$.

An $(\varepsilon, \varepsilon)$-approximation is called, for short an *ε-approximation*.

Recall that a metric spaces whose metric d is *intrinsic*, i.e. induced by a *length structure* (i.e. path length) by the ambient metric on a subset of a given metric space is called a *length space*. Such spaces are, for obvious reasons, of special interest in Geometry. (As a basic motivation both theoretical and practical, for considering such spaces, would be that of surfaces in $\mathbb{R}^3$.) The following theorem shows that length spaces are closed in the *Gromov–Hausdorff topology*:

Theorem 2.9 *Let $\{X_n\}_{n=1}^\infty$ be length spaces and let X be a complete metric space such that $X_n \xrightarrow[GH]{} X$. Then X is a length space.*

Using *ε-approximations* one can prove the following theorem and corollary, that are quite important, not only for the specific purpose of this overview, but in a far more general and powerful context (see e.g. [31] and [19]):

Theorem 2.10 (Gromov) *Any compact length space is the GH-limit of a sequence of finite graphs.*

The proof of the theorem above is constructive, therefore potentially adaptable in practical applications (such as those arising in Graphics, Imaging and related fields). For this very reason, and for essential simplicity we bring it below:

Proof Let ε, δ ($\delta \ll \varepsilon$) small enough, and let S be a δ-net in X. Also, let $G = (V, E)$ be the graph with $V = S$ and $E = \{(x, y) \,|\, d(x, y) < \varepsilon\}$. We shall prove that G is an ε-approximation of X, for δ small enough, more precisely, for $\delta < \frac{\varepsilon^2}{4\text{diam}(X)}$.

But, since S is an ε-net both in X and in G, and since $d_G(x, y) \geq d_X(x, y)$, it is sufficient to prove that:

$$d_G(x, y) \leq d_X(x, y) + \varepsilon\,.$$

Let γ be the shortest path between x and y, and let $x_1, \ldots, x_n \in \gamma$, such that $n \leq \text{length}(\gamma)/\varepsilon$ (and $d_X(x_i, x_{i+1}) \leq \varepsilon/2$). Since for any x_i there exists $y_i \in S$, such that $d_X(x_i, y_i) \leq \delta$, it follows that $d_X(y_i, y_{i+1}) \leq d_X(x_i, x_{i+1}) + 2\delta < \varepsilon$.

Therefore, (for $\delta < \varepsilon/4$), there exists an edge $e \in G, e = y_i y_{i+1}$. From this we get the following upper bound for $d_G(x, y)$:

$$d_G(x, y) \leq \sum_{n=0}^{n} d_X(y_i, y_{i+1}) \leq \sum_{n=0}^{n} d_X(x_i, x_{i+1}) + 2\delta n$$

But $n < 2\,\text{length}(\gamma)/\varepsilon \leq 2\text{diam}(X)/\varepsilon$. Moreover: $\delta < \varepsilon^2/4\,\text{diam}(X)$. It follows that:

$$d_G(x, y) \leq d_X(x, y) + \delta\frac{4\text{diam}(X)}{\varepsilon} < d_X(x, y) + \varepsilon\,.$$

Thus, for any $\varepsilon > 0$, there exists a graph an ε-approximation of X by a graph G, $G = G_\varepsilon$. Hence $G_\varepsilon \underset{\varepsilon}{\rightarrow} X$.

In fact, one can strengthen the theorem above as follows:

Corollary 2.1 *Let X be a compact length space. Then X is the Gromov–Hausdorff limit of a sequence $\{G_n\}_{n\geq 1}$ of finite graphs, isometrically embedded in X.*

Remark 2.16 A certain amount of care is needed when applying the theorem above, as the following facts show:

1. If $G_n \underset{\varepsilon}{\rightarrow} X$, $G_n = (V_n, E_n)$. If there exists $N_0 \in \mathbb{N}$ such that

$$(*) \quad |E_n| \leq N_0, \text{ for all } n \in \mathbb{N}\,,$$

   then X is a finite graph.
2. If condition $(*)$ is replaced by:

$$(**) \quad |V_n| \leq N_0, \text{ for all } n \in \mathbb{N}\,,$$

   then X will still be always a graph, but not necessarily finite.

Remark 2.17 Theorem 2.10 can be strengthened as follows: Compact inner metric spaces can be, in fact, Gromov–Hausdorff approximated by smooth surfaces that, moreover, are embedded in $\mathbb{R}^3$, as shown by Cassorla [20] (see also [31, p. 99] and the reference therein). In other words, one can "visualize" in $\mathbb{R}^3$ (up to some predetermined but arbitrarily small error) *any* compact inner metric space. Unfortunately, the genus of the approximating surfaces can not remain bounded. (In consequence, in order that a good approximation even of a simple space be obtained, using the method given in Cassorla's proof, one has to increase the topological complexity of the approximating surface.)

Note also that there is no geometric (curvature) restriction on the approximating surfaces. In fact, it is also stated in [20] that one can approximate the given spaces with a series of smooth surfaces having Gaussian curvature bounded from above by -1 (this being, however a seemingly unpublished result). Unfortunately, to obtain this, one has to abandon the embeddability in $\mathbb{R}^3$ of the approximating surfaces.

We conclude this remark by adding a few words regarding Cassorla's proof: He begins by constructing an approximation by graphs, following Gromov, then he considers the (smooth) boundaries of canonical tube neighborhoods or, in other words, he builds the smooth surfaces having as axes (or nerve) the graph constructed previously.

2.4.2 Wald Curvature and Gromov–Hausdorff Convergence

It turns out (not very surprisingly, in fact, in view of the facts that we shall present in the next section) that it is somewhat naive to hope for a generic result for Wald curvature as such. It turns out that the upper and lower bound for K_W display quite different behaviors. There are very few results we can state, therefore, in the generic case. The basic one is

Lemma 2.1 *Let (X_i, d_i) be compact metric spaces, such that $X_i \xrightarrow[GH]{} X$. If $B_i \subset X_i$, $B_i = B(p_i, r)$ is a region of curvature $\geq k$, for all $i \geq 1$, and if $p_i \xrightarrow[GH]{} p \in X$, then $B = B(p, r)$ is a region of curvature $\geq k$ in X.*
(Recall that by Theorem 2.9 X is also an inner metric.)

In view of Toponogov's Theorem we can now formulate:

Theorem 2.11 *Let (X_i, d_i) be (compact) metric spaces, such that $X_i \xrightarrow[GH]{} X$. If X_i has curvature $\geq k$, for all $i \geq 1$, then X has curvature $\geq k$.*

Remark 2.18 We should note in his context that the diameter function is continuous under the the Gromov–Hausdorff convergence.

2.4.2.1 The Case $K_W \geq K_0$

This is the case where a plethora of powerful results exist, mainly due to Plaut [75, 77, 78].

The first such result represents a generalization of the classical by now compactness theorem of Gromov (see, e.g. [31]). For its formulation we need an additional notation: We denote by $\mathscr{M}(k, n, D)$ the class of all finitely dimensional[12] spaces of curvature $\geq k$, dimension $\leq n$ and diameter $\leq D$. We can now formulate the theorem in question:

Theorem 2.12 (Plaut [79]) *$\mathscr{M}(k, n, D)$ is compact in the Gromov–Hausdorff metric.*

[12]The dimension can be taken as the topological dimension or the Hausdorff dimension—see, e.g. [79].

Remark 2.19 Let us denote by $\mathfrak{M}(k, n, D)$ the set of all Riemannian manifolds satisfying the same conditions as the spaces in $\mathscr{M}(k, n, D)$ (k denoting, in this case, sectional curvature). Then (clearly) $\mathfrak{M}(k, n, D) \subsetneq \mathscr{M}(k, n, D)$.

The results above have quite important consequences for a variety of practical fields (or, at least, for their more theoretical, basic aspects): Since by the now classical Gromov Precompactness Theorem [31], any element in a compact (hence a fortiori precompact) collection of compact metric spaces admits, for any $\varepsilon > 0$, an *ε-net* with at most $N(\varepsilon)$ number of elements, they represent quite general sampling theorems, giving, moreover, a strong upper bound on the number of sampling points—a number that depends, apart from the class on the manifold, only on the quality of the sampling (as given by ε). (The price to be paid, so to say, for the strengths above, is represented by the non-algorithmic nature of these results.)

In fact, even for a larger class of metric spaces, namely $\mathscr{M}(k, n, \varepsilon)$, where $\varepsilon > 0$ denotes a lower bound for the *injectivity radius*[13] some important facts can be asserted, such as the following theorem:

Theorem 2.13 (Plaut [77]) *The elements of $\mathscr{M}(k, n, \varepsilon)$ are smoothable, for any n.*

Again, the theorem above is also highly relevant to Sampling Theory, since it shows that the traditional approach in the field, that is of considering smoothings ("filtrations") of the given signals/images and sampling them according to a more traditional, Gauss curvature based scheme is theoretically valid.

Moreover, the following compactness result also holds:

Theorem 2.14 (Plaut [77]) *$\mathscr{M}^*(k, n, D, \varepsilon)$ is compact in the Gromov–Hausdorff metric, where $\mathscr{M}^*(k, n, D, \varepsilon)$ denotes the class of spaces of dimension equal to n, curvature $\geq k$, diameter $\leq D$ and injectivity radius $\geq \varepsilon$.*

In addition to these compactness results, the following finiteness theorems, representing generalizations of classical results of Cheeger [22], respectively Grove–Petersen [35] and Grove–Petersen–Wu [36] also hold:

Theorem 2.15 (Perelman [73]) *The class $\mathscr{M}^*(k, n, D)$ has finitely many homeomorphism types.*

(We believe that by now the notation must be clear to the reader.)

Theorem 2.16 (Perelman [73]) *$\mathfrak{M}(k, n, D, v)$, where $v > 0$ denotes a lower bound on volume, has finitely many types of homeomorphism for all n, and diffeomorphism, for all $n \neq 4$.*

Remark 2.20 At first glance Theorems 2.15 and 2.16 seemingly are contradicted by the existence of infinitely many homotopy types of lens spaces. However, this is not the case, since they fail to satisfy the conditions of the theorem even as Riemannian manifolds, thus, *a fortiori* as Alexandrov spaces (see [35, p. 196]). For instance, the lens spaces $\mathbb{S}^{2n+1}/\mathbb{Z}_\neg$ have constant curvature equal to 1 and diameter $\pi/2$, but

[13] Without getting into the technical subtleties of the definition of the space of directions S_p at a point p in a space of bounded curvature, the injectivity radius at p is defined as $\inf_{\gamma \in S_p} \sup_t \{\gamma|_{[0,t]} \text{is minimal}\}$.

have no lower bound on the volume (and, indeed, they belong to infinitely many homotopy classes. (Up to a scaling of volume, the same spaces show that the upper bound on diameter is also necessary.)

These last two theorems, as well as the previous, related results of Grove et al. mentioned above have quite practical importance in the Recognition Problem (in Manifold Learning and related fields), in particular in determining the complete so called "Shape DNA".

More geometrically interesting, powerful results exists, but for lack of space and for the sake of cohesiveness of the text, we do not bring here—see [79] and the bibliography cited therein.

2.4.2.2 The Case $K_W \leq K_0$

As already mentioned in the introduction of this section, spaces of Wald curvature bounded from above display a behavior widely divergent from those satisfying the opposite inequality. Most notable they do not satisfy a fitting Toponogov type theorem, the example of the flat torus $T^2 = \mathbb{S}^1 \times \mathbb{S}^1$ being the basic one (see [79, p. 886] for details). However, such spaces still have some very interesting geometric properties, see again [79] as well as [2].

However, the most powerful results, due to Berestovskii [8] and Nikolaev [66–68] (see also [76]) are obtained when combining the lower and upper curvature bounds, the main such theorem being:

Theorem 2.17 *A topological space admits a smooth manifold structure (with or without boundary) iff (i) it is finite dimensional; and (ii) has a metric curvature bounded both from above and from below.*

2.5 Wald and Alexandrov Curvatures Comparison

We have alluded many times to the notion of Alexander curvature. Moreover, it is quite probably that many (if not most) of the readers are quite familiar with this concept permeating modern Mathematics, certainly much more than with the rather esoteric (for some) Wald curvature. It is therefore only fitting that we finally discuss the relationship between these two types of comparison curvature.

2.5.1 Alexandrov Curvature

We begin by reminding the reader the definition of Alexandrov comparison curvature. Before bringing the formal definition, let us just specify the main difference between this approach and Wald curvature: In defining Alexandrov curvature, one makes appeal to *comparison triangles* in the model space (i.e. gauge surface $\mathscr{S}_k$), rather than quadrangles, as in the definition of Wald curvature.

A space is said to be of *Alexandrov curvature* $\geq k$ iff any of the following equivalent conditions holds *locally*, and to be of *Alexandrov curvature* $\leq k$ iff any of the conditions below holds with the opposite inequality:

Definition 2.11 Let X be an inner metric space, let $T = \triangle(p, q, r)$ be a geodesic triangle, of sides $\overline{pq}, \overline{pr}, \overline{qr}$, and let $\widetilde{T}$ denote its(a) representative triangle in $\mathscr{S}_k$.

A0 Given the triangle $T = \triangle(p, q, r)$ and points $x \in \overline{pq}$ and $y \in \overline{pr}$, there exists a representative triangle $\widetilde{T}$ in $\mathscr{S}_k$. Let $\tilde{x}, \tilde{y}$ represent the corresponding points on the sides of $\widetilde{T}$. Then $d(x, y) \geq d(\tilde{x}, \tilde{y})$.

A01 Given the triangles $T = \triangle(p, q, r)$ and $\widetilde{T}$ as above, and a point $x \in \overline{pq}$, $d(x, c) \geq d(\tilde{x}, \tilde{c})$.

A1 Given the triangle $T = \triangle(p, q, r)$, there exists a representative triangle $\widetilde{T}$ in $\mathscr{S}_k$, and $\measuredangle(\overline{pq}, \overline{pr}) \geq \measuredangle(\overline{\tilde{p}\tilde{q}}, \overline{\tilde{p}\tilde{r}})$, where $\measuredangle(\overline{pq}, \overline{pr})$ denotes the angle between $\overline{pq}$ and $\overline{pr}$, etc.

A2 For any hinge $H = (\overline{pq}, \overline{pr})$, there exists a representative hinge $\widetilde{H} = (\overline{\tilde{p}\tilde{q}}, \overline{\tilde{p}\tilde{r}})$ in $\mathscr{S}_k$ and, moreover, $d(p, q) \leq d(\tilde{p}, \tilde{q})$, where a *hinge* is a pair of minimal geodesics with a common end point.

Remark 2.21

1. Axiom $A0$ represents the basic one in defining Alexandrov comparison. It is also the one used by Rinow [80], in his seemingly (semi-)independent development of comparison geometry. (For a shorter presentation of his approach, but somewhat more accessible and in English, see [12, 13] and, for an even briefer one, but easy to reach, [86].) This condition represents nothing more the transformation into an axiom of the following essential geometric fact: Thales Theorem does not hold in Spherical and Hyperbolic Geometry. (In particular—and most spectacular—the line connecting the midpoints of two of the edges of a triangle does not equate half of the third one.[14])
2. Condition $A01$ clearly represents just a particular case of $A0$, by fixing y to be one of the end points of $\overline{pr}$, say $y = r$. However, it is, actually, equivalent to $A0$ and it is, in fact, usually easier to check. It is sometimes used as the basic definition of Alexandrov comparison, saying, for instance that "negatively curved spaces have short ties" (the figure of speech being, we believe, self explanatory)—see, e.g. [105]. The role of *sd*-quads in such fundamental results as Theorem 2.5 becomes now less strange and, in fact, it will become quite clear once the result in the next section is introduced. We anticipate somewhat by adding that now Robinson's method in Theorem 2.7 shows itself as it truly is: A method[15] of approximating the relevant k appearing in Axiom $A01$.

[14]This well known "paradox" of the foundations of Geometry is, unfortunately, generally overlooked in certain applications in Imaging and Graphics, which results in a penalty on the quality of the numerical results.

[15]Developed *avant la lettre*.

3. We have used here (for the most part) Plaut's notation in [79]. For an excellent, detailed, clear (and by now already classical) presentation of the various comparison conditions, see [19].
4. Conditions $A1$ and $A2$ show that one can introduce comparison Geometry via angle comparison. However, we prefer a "purely" metric approach, even if it is somewhat illusory. (See [90] for the application of this approach to a "purely" metric Regge calculus.)

2.5.2 Alexandrov Curvature vs. Wald Curvature

Loosely formulated, the important fact regarding the connection between Wald and Alexandrov curvatures is that *(in the presence of sufficiently many minimal geodesics) Wald curvature is (essentially) equivalent to Alexandrov curvature* or, slightly more precisely, that inner metric spaces with Wald curvature $\geq k$ satisfy the condition of having sufficiently many geodesics. (This fact may be viewed as an extended, weak Hopf–Rinow type theorem.) The formal enouncement of this result requires yet more technical definitions, which we present below for the sake of completeness:

Definition 2.12 Let X be an inner metric space and let $\gamma_{pq} \subset X$ be a minimal geodesic connecting the points p and q. γ is called

1. *extendable* beyond q if there exists a geodesic $\widetilde{\gamma}$, such that $\gamma = \widetilde{\gamma}|_{(p,q)}$ and $q \in \text{int}\widetilde{\gamma}$.
2. *almost extendable* beyond q if for any $\varepsilon > 0$, there exists an $r \in X \setminus \{p, q\}$, such that $\sigma(q; p, r) < \varepsilon$, where $\sigma(q; p, r)$ denotes the *strong excess*

$$\sigma(q; p, r) = \frac{e(T)}{\min d(p, r), d(r, q)}, \tag{2.22}$$

where $T = \Delta(q, p, r)$ (and where $e(T)$, stands, as above, for its excess).

We shall also need the following

Definition 2.13 Let X be as above and let $p \in X$. We denote

$$J_p = \{q \in X \,|\, \exists! \text{ minimal geodesic } \gamma_{pq} \text{ almost extendable beyond } q.\}. \tag{2.23}$$

We have the following

Theorem 2.18 (Plaut [78, 79]) *Let X be an inner metric space of Wald–Berestovskii curvature $\geq \kappa$. Then, for any $p \in X$, J_p contains a dense G_δ set.*

Since, by the Baire Category Theorem, the intersection of countably many dense G_δ sets is a dense G_δ set, we obtain the following corollary:

Corollary 2.2 *Let X be an inner metric space of Wald–Berestovskii curvature $\geq \kappa$, and let $p_1, p_2, \ldots \in X$. Then there exist points $p_1', p_2', \ldots \in X$ such that*

1. *p_i is arbitrarily close to p_i, for all i;*
2. *There exists a unique minimal geodesic connecting p_i and p_i'.*

Moreover, one can take $p_1' = p_1$.

In other words, given any three points p, q, r in X, there exist points p_1, q_1, r_1 arbitrarily close to them (respectively) such that p_1, q_1, r_1 represent the vertices of a triangle whose sides are minimal geodesics or, simply put, one can construct (minimal geodesic) triangles "almost everywhere".

From the corollary above and from the Toponogov Comparison Theorem we obtain the announced theorem that establishes the essential equivalence between Alexandrov and Wald curvature, once the existence of "enough geodesics" is assured:

Theorem 2.19 (Plaut [78]) *Let X be an inner metric space. X is a space of Alexandrov curvature bonded from below iff for every $x \in X$, there exists an open set U, $x \in U$, such that for every $y \in U$ the set J_y contains a dense G_δ set of U.*

(For a different formulation of the results above see [79, Corollary 40].)

In view of the result above, it is easy to recognize finitely dimensional spaces of *Wald* curvature $\geq k$ in the garb that they are widely known in the modern terminology, namely *Alexandrov spaces*. The reader can, therefore, substitute, in the results in the previous section "Alexandrov space" whenever this is possible—this is the form in which many of the theorems in question are better known.

However, Wald curvature allows us to discard conditions as are usually used when employing the Alexandrov triangle comparison, e.g. local compactness, while still being able to obtain many important theorems, such as the Toponogov Theorem and the Hopf–Rinow Theorem that we have discussed above, as well as fitting variants of the Maximal Radius Theorem and of the Sphere Theorems, that we have only alluded too, unfortunately. For details the reader is invited to consult [78].

2.6 A Metric Approach to Ricci Curvature

In this section we concentrate on the application of the metric approach, and more precisely of Wald curvature, to the defining Ricci flow for cell complexes and, in dimension 2, to the development of a fitting metric Ricci flow. Since these problems were studied in detail in our papers [37] and [89], we present only the main ideas, and hope that the interested reader will consult the original papers (especially [89], were a more detailed discussion is contained.[16]) In consequence, we follow here the exposition in the much shorter and restricted proceedings paper. However,

[16] Also, we warn the eventual reader of an unfortunate previously unnoticed typo towards the end of [37].

we emphasize whenever possible those correlations with the theoretical material included in the previous section. Also, a new approach to the problem of the Ricci flow for *PL* surfaces, also concordant to the theory in the previous section is also included.

2.6.1 Metric Ricci Flow for PL Surfaces

Motivated mainly by Perelman's work, the Ricci flow has become lately[17] an object of active interest and research in Graphics and Imaging. Previously, various approaches have been suggested, encompassing such methods as classical approximations of smooth differential operators, as well as discrete, combinatorial methods.

Among these diverse approaches, the most successful so far proved to be the one based on the discrete Ricci flow of Chow and Luo [24], due to Gu (see, e.g. [43] and, for more details, [38]). In truth, the paper [89] was largely motivated by our desire to get a better understanding of the discrete, circle-packing based Ricci flow of Chow and Luo, and its relation with the Ricci flow for smooth surfaces introduced by Hamilton [40] and Chow [23].

2.6.1.1 A Smoothing Based Approach

Our approach, as developed in [89] to this problem is to pass from the discrete context to the smooth one and explore the already classical results known in this setting, by applying Theorem 2.3. To this end we have first to make a few observations: One can pass from the *PL* surfaces to smooth ones by employing *smoothings*, defined in the precise sense of *PL* differential Topology (see [61]). Since, by Munkres [61, Theorem 4.8], such smoothings are δ-*approximations*, and therefore, for δ small enough, also α-approximations of the given piecewise-linear surface S^2_{Pol}, they approximate arbitrarily well both distances and angles on S^2_{Pol}. (Due to space restrictions, we do not bring here these technical definitions, but rather refer the reader to [61].) It should be noted that, while Munkres' results concern *PL* manifolds, they can be applied to polyhedral ones as well, because, by definition, polyhedral manifolds have simplicial subdivisions (and furthermore, such that all vertex links[18] are combinatorial manifolds). Of course, for different subdivisions, one may obtain different polyhedral metrics. However, by the *Hauptvermutung* Theorem in dimension 2 (and, indeed, for smooth triangulations of diffeomorphic

[17] In what would have been probably consider to be a strange—not to say bizarre—development even only a few years ago.

[18] Recall that the *link* $\mathrm{Lk}(v)$ of a vertex v is the set of all the faces of $\overline{\mathrm{St}}(v)$ that are not incident to v. Here $\overline{\mathrm{St}}(v)$ denotes the *closed star* of v, i.e. the smallest subcomplex (of the given simplicial complex K) that contains $\mathrm{St}(v)$, namely $\overline{\mathrm{St}}(v) = \{\sigma \in \mathrm{St}(v)\} \cup \{\theta \mid \theta \leqslant \sigma\}$, where $\mathrm{St}(v)$ denotes the *star* of v, that is the set of all simplices that have v as a face, i.e $\mathrm{St}(v) = \{\sigma \in K \mid v \leqslant \sigma\}$.

manifolds in any dimension) (see e.g. [61] and the references therein), any two subdivisions of the same space will be combinatorially equivalent, hence they will give rise to the same polyhedral metric. It follows from the observations above that *metric quadruples* on S_{Pol} are also arbitrarily well approximated (including their angles) by the corresponding metric quadruples) on the smooth approximating surfaces S_m. But, by Theorem 2.3, $K_{W,m}(p)$—the *Wald metric curvature* of S_m, at a point p—equals the classical (Gauss) curvature $K(p)$. Hence the Gauss curvature of the smooth surfaces S_m approximates arbitrarily well the metric one of S_{PL} (and, as in [16], the smooth surfaces differ from polyhedral one only on (say) the $\frac{1}{m}$-neighborhood of the 1-skeleton of S_{Pol}—see also the discussion below). Moreover, this statement can be made even more precise, by assuring that the convergence is in the Hausdorff metric. This follows from results of Gromov (see e.g. [92] for details).

We can now introduce the *metric* Ricci flow: By analogy with the classical flow

$$\frac{dg_{ij}(t)}{dt} = -2K(t)g_{ij}(t)\,. \tag{2.24}$$

we define the *metric* Ricci flow by

$$\frac{dl_{ij}}{dt} = -2K_i l_{ij}\,, \tag{2.25}$$

where $l_{ij} = l_{ij}(t)$ denote the edges (1-simplices) of the triangulation (*PL* or piecewise flat surface) incident to the vertex $v_i = v_i(t)$, and $K_i = K_i(t)$ denotes the Wald curvature at the same vertex, where, as above, we employ the discrete version of Wald's curvature defined by Formula (2.17).

Remark 2.22 Before continuing further on, it is important to remark the asymmetry in Eq. (2.25), that is caused by the fact that the curvature on two different vertices acts, so to say, on the same edge. However, passing to the smooth case, is that the asymmetry in the metric flow that we observed above disappears automatically via the smoothing process. To this end it is important to note that while the formula of $K_W(v)$ involves the edges incident to v, it is—precisely by this incidence criterion—a curvature attached to the vertex v. (For further details see [90] as well as Sect. 2.6.1.2 below.)

We also consider the close relative of (2.24), the normalized flow

$$\frac{dg_{ij}(t)}{dt} = (K - K(t))g_{ij}(t)\,, \tag{2.26}$$

and its metric counterpart

$$\frac{dl_{ij}}{dt} = (\bar{K} - K_i)l_{ij}\,, \tag{2.27}$$

where $K, \bar{K}$ denote the average classical, respectively Wald, sectional (Gauss) curvature of the initial surface S_0: $K = \int_{S_0} K(t)dA \Big/ \int_{S_0} dA$, and $\bar{K} = \frac{1}{|V|}\sum_{i=1}^{|V|} K_i$, respectively. (Here $|V|$ denotes, as usually, the cardinality of the vertex set of S_{Pol}.)

An Approximation Result

The first result that we can bring is a metric curvature version of classical result of Brehm and Kühnel [16] (where the combinatorial/defect definition of curvature for polyhedral surfaces is used.

Proposition 2.6 *Let S^2_{Pol} be a compact polyhedral surface without boundary. Then there exists a sequence $\{S^2_m\}_{m\in\mathbb{N}}$ of smooth surfaces, (homeomorphic to S^2_{Pol}), such that*

1. *a. $S^2_m = S^2_{Pol}$ outside the $\frac{1}{m}$-neighborhood of the 1-skeleton of S^2_{Pol},*
   b. The sequence $\{S^2_m\}_{m\in\mathbb{N}}$ converges to S^2_{Pol} in the Hausdorff metric;
2. *$K(S^2_m) \to K_W(S^2_{Pol})$, where the convergence is in the weak sense.*

Remark 2.23 As we have already noted above, the converse implication—namely that Gaussian curvature $K(\Sigma)$ of a smooth surface Σ may be approximated arbitrarily well by the Wald curvatures $K_W(\Sigma_{Pol,m})$ of a sequence of approximating polyhedral surfaces $\Sigma_{Pol,m}$—is quite classical.

For a more in-depth discussion and analysis of the convergence rate in the proposition above, see [90].

Remark 2.24 In view of the equivalence of the Alexandrov and Wald curvatures, one can view the result above as a elementary, restricted to dimension 2, but on other hand a more specific and constructive version of Theorem 2.11.

As already stressed, the "good", i.e. metric and curvature, approximations results mentioned above, imply that one can study the properties of the metric Ricci flow via those of its classical counterpart, by passing to a smoothing of the polyhedral surface. The use of the machinery of metric curvature considered has the benefit that, by using it, the "duality" between the combinatorics of the packings (and angles) and the metric disappears: The flow becomes purely metric and, moreover, the curvature at each stage (i.e. for every "t") is given—as in the smooth setting—in an intrinsic manner, that is in terms of the metric alone.

We bring here the most important properties that follow immediately using this approach (for further results and additional details, see [89]).

Existence and Uniqueness

The first result that we should bring here is the following

Proposition 2.7 *Let (S^2_{Pol}, g_{Pol}) be a compact polyhedral 2-manifold without boundary, having bounded discrete Wald curvature. Then there exists $T > 0$ and a smooth family of polyhedral metrics $g(t), t \in [0, T]$, such that*

$$\begin{cases} \frac{\partial g}{\partial t} = -2K_W(t)g(t) \ t \in [0,T]\,; \\ g(0) = g_{Pol}\,. \end{cases} \tag{2.28}$$

(Here $K_W(t)$ denotes the Wald curvature induced by the metric $g(t)$.)

Moreover, both the forwards and the backwards (when existing) Ricci flows have the uniqueness of solutions property, that is, if $g_1(t), g_2(t)$ *are two Ricci flows on* S^2_{Pol}*, such that there exists* $t_0 \in [0, T]$ *such that* $g_1(t_0) = g_2(t_0)$*, then* $g_1(t) = g_2(t)$*, for all* $t \in [0, T]$.
Here, by smooth family of polyhedral surfaces we mean family of polyhedral manifolds, whose edge lengths vary smoothly.

Remark 2.25 An basic critique of the direct approach to the metrization of the Ricci flow adopted in Formula (2.27) is its lacks of symmetry, a deficiency that presumably will not not disappear even when passing to the limit. A natural way of symmetrizing the flow is (using the same notation as before):

$$\frac{dl_{ij}}{dt} = -\frac{K_i + K_j}{2} l_{ij} \,, \tag{2.29}$$

where in this case, K_i, K_j denote the Wald curvatures at the vertices v_i and v_j, respectively. It is worthwhile to remark here that, in fact, this expression appears also in the practical method of computing the combinatorial curvature, where it is derived via the use of a conformal factor (see [38]).

Beyond the theoretical importance, the existence and uniqueness of the backward flow would allow us to find surfaces in the conformal class of a given circle packing (Euclidean or Hyperbolic). More importantly, the use of purely metric, Wald curvature based, approach adopted, rather than the combinatorial (and metric) approach of [24], allows us to give a preliminary and purely theoretical at this point, answer to Question 2, p. 123, of [24], namely whether there exists a Ricci flow defined on the space of all piecewise constant curvature metrics (obtained via the assignment of lengths to a given triangulation of 2-manifold). Since, by the results of Hamilton's [40] and Chow [23], the Ricci flow exists for all compact surfaces, it follows that the fitting metric flow exists for surfaces of piecewise constant curvature. In consequence, given a surface of piecewise constant curvature (e.g. a mesh with edge lengths satisfying the triangle inequality for each triangle), one can evolve it by the Ricci flow, either forward, as in the works discussed above, to obtain, after the suitable area normalization, the polyhedral surface of constant curvature conformally equivalent to it; or backwards (if possible) to find the "primitive" family of surfaces—including the "original" surface obtained via the backwards Ricci flow, at time T—conformally equivalent to the given one.

Convergence Rate

A further type of result, quite important both from the theoretical viewpoint and for computer-driven applications, is that of the *convergence rate* (see [37, 38] for the precise definition).

Since we already know that the solution exists and it is unique (see also the subsection below for the nonformation of singularities), by appealing to the classical

results of [40] and [23], we can control the convergence rate of the curvature, as follows:

Theorem 2.20 *Let (S^2_{Pol}, g_{Pol}) be a compact polyhedral 2-manifold without boundary. Then the normalized metric Ricci flow converges to a polyhedral surface of constant metric curvature. Moreover, the convergence rate is*

1. *exponential, if $\bar{K} = \bar{K}_W < 0$ (i.e. $\chi(S^2_{Pol}) < 0$) ;*
2. *uniform; if $\bar{K} = 0$ (i.e. $\chi(S^2_{Pol}) = 0$);*
3. *exponential, if $\bar{K} > 0$ (i.e. $\chi(S^2_{Pol}) > 0$).*

Here, by a polyhedral surface of constant curvature we mean a polyhedral 2-manifold (without boundary) where curvature is evenly spread among the vertices (i.e. the only significant vertices for the computation of Wald curvature).

Singularities Formation

Another extremely important aspect of the Ricci flow, both smooth or discrete, is that of singularities formation. Again, a certain (theoretical, at least) advantage of the proposed method presents itself. Indeed, by Chow and Luo [24, Theorem 5.1], for compact surfaces of genus ≥ 2, the combinatorial Ricci flow evolves without singularities. However, for surfaces of low genus no such result exists. Indeed, in the case of the Euclidean background metric, that is the one of greatest interest in graphics, singularities do appear. Moreover, such singularities are always combinatorial in nature and amount to the fact that, at some t, the edges of at least one triangle do not satisfy the triangle inequality. These singularities are removed in heuristic manner. However, by Hamilton [40, Theorem 1.1], the smooth Ricci flow exists at all times, i.e. no singularities form. From the considerations above, it follows that the metric Ricci flow also exists at all times without the formation of singularities.

Embeddability in $\mathbb{R}^3$

The importance of the embeddability of the flow is not solely theoretical (e.g. if one considers the problem of the Ricci flow for surfaces of piecewise constant curvature), as it is essential in Imaging (see [3, 99]), and of high importance in Graphics. Indeed, even our very capability of seeing (grayscale) images is nothing but a translation, in the field of vision, of the embeddability of the associated height-surface into $\mathbb{R}^3$. (Or, perhaps one should view the mathematical aspect as a formalization of a physical/biological phenomenon. . .) We should note here that in this respect there exists a certain (mainly theoretical, at this point in time) advantage of our proposed metric flow over the combinatorial Ricci flow [38, 43]. Indeed, in the combinatorial flow, the goal is to produce, via the circle packing metric, a conformal

mapping from the given surface to a surface of constant (Gauss) curvature. Since in the relevant cases (see, e.g. [38]) the surface in question is a planar region (usually a subset of the unit disk), its embeddability (not necessarily isometric) is trivial. Moreover, in the above mentioned works, there is no interest (and indeed, no need) to consider the (isometric) embeddability of the surfaces S_t^2 (see below) for an intermediate time $t \neq 0, T$.

The tool that allows us to obtain this type of results is making appeal (again) to δ-approximations, in combination with classical results in embedding theory. Indeed, by Munkres [61, Theorem 8.8] a δ-approximation of an embedding is also an embedding, for small enough δ. Since, as we have already mentioned, smoothing represent δ-approximations, the possibility of using results regarding smooth surfaces to infer results regarding polyhedral embeddings is proven. (The other direction—namely from smooth to *PL* and polyhedral manifolds—follows from the fact that the *secant approximation* is a δ-approximation if the simplices of the *PL* approximation satisfy a certain nondegeneracy condition—see [61, Lemma 9.3].) We state here the relevant facts:

Let S_0^2 be a smooth surface of positive Gauss curvature, and let S_t^2 denote the surface obtained at time t from S_0^2 via the Ricci flow. (For all omitted background material (proofs, further results, etc.) we refer to [41].)

Proposition 2.8 *Let S_0^2 be the unit sphere $\mathbb{S}^2$, equipped with a smooth metric g, such that $\chi(S_0^2) > 0$. Then the surfaces S_t^2 are (uniquely, up to a congruence) isometrically embeddable in $\mathbb{R}^3$, for any $t \geq 0$.*

In fact, this results can be slightly strengthened as follows:

Corollary 2.3 *Let S_0^2 be a compact smooth surface. If $\chi(S_0^2) > 0$, then there exists some $t_0 \geq 0$, such that the surfaces S_t^2 are isometrically embeddable in $\mathbb{R}^3$, for any $t \geq t_0$.*

In stark contrast with this positive result regarding surfaces uniformized by the sphere, for (complete) surfaces uniformized by the hyperbolic plane we only have the following negative result:

Proposition 2.9 *Let (S_0^2, g_0) be a complete smooth surface, and consider the normalized Ricci flow on it. If $\chi(S_0^2) < 0$, then there exists some $t_0 \geq 0$, such that the surfaces S_t^2 are not isometrically embeddable in $\mathbb{R}^3$, for any $t \geq t_0$.*

2.6.1.2 An Alexandrov Surfaces Based Approach

After a completed version of our paper [89] was essentially finished, we noted that there are other works regarding the Ricci flow on surfaces with conical singularities, and especially Richards paper [80] on the smoothing of (compact) Alexandrov surfaces via the Ricci flow. We would like to stress that our approach as developed is different from Richard's work, being much more direct and, in a sense, more elementary. Moreover, we should also accentuate the fact that our method facilitates

concrete, computational treatment of the flow. On the other hand, Richard's method uses the very Ricci flow for smoothing, and makes no appeal to approximations, making it much more alluring for theoretical ends. However, its proof is far from trivial and we don't even sketch it here, since it would take us to far afield, and the interested reader is invited too study Richard's paper.[19]

More important to our purpose here, Richard's method also provides us with a smoothing of the given *PL* surface, hence all the theoretical results in the previous section also follow via this route.

However, Richard's method of proof seems to be adaptable in order to solve the following

2.1 Devise a purely metric flow.

Surely, such a flow, independent both from smoothing and to the advanced (and somewhat abstract) mathematical apparatus of [80] would provide a powerful and flexible tool for many Imaging and graphics tasks, akin to the one based on Chow and Luo's paper (see the relevant bibliography mentioned above).

2.6.1.3 An Application: Smoothable Metrics on Cube Complexes

We illustrate our belief in the many possible applications of the metric Ricci flow with only one such example (due to space and time restrictions), appertaining to the corpus of "Pure" Mathematics.

The following seemingly well known problem in the theory of cube complexes[20] was posed to the author by Joel Haas [39], together with the basic idea of the first part of the proof, for which the author is deeply grateful.

Let $\mathscr{C}$ be a cube complex, satisfying the following conditions:

1. $\mathscr{C}$ is negatively curved (i.e. such that $\#_v Q \geq 4$, for all vertices v,where $\#_v Q$ denotes the number of cubes incident to the vertex v;
2. The link $\mathrm{lk}(v)$ of any vertex is a *flag complex*, i.e. a simplicial complex such that any 3-arcs closed curve bounds a triangle (2-simplex), i.e. no such curve separates without being a boundary.[21]

Question 1 *Does there exist a Riemannian metric g (on $\mathscr{C}$) such that $K_g \equiv K$, where K denotes the comparison (Alexandrov) curvature of $\mathscr{C}$?*

[19]Note that to apply Richard's result we have only to consider our surfaces as an Alexandrov surface having curvature bounded from below, condition that is, evidently, satisfied. (In this regard and for a discussion on the definition of Wald/Alexandrov curvature for *PL* surfaces, see [89, pp. 26–27].

[20]For a formal definition and more details see, e.g. [82].

[21]Alternatively, this condition may be expressed either as $\mathrm{lk}(v)$ "has no missing simplices (Sageev [82]), or as "a nonsimplex contains a non edge" (W. Dicks, see [11]).

In other words: Does there exist a smoothing of (M, g) (i.e. Riemannian manifold) of a given cube complex $\mathscr{C}$ (that has a manifold structure), such that $K \equiv K_g$? Evidently, an important particular case would be that "cubical version of *PL* approximations"), i.e. that of "cubulations" of a (given) Riemannian manifold.

Remark 2.26 The similar problem can be also posed, of course, for positively curved complexes (i.e. such that $\#_v Q \leq 4$). However, we address here only the negatively curved case. The similar results for polyhedral manifolds of nonnegative curvature was also proved recently—see [51].[22]

Evidently, the answer to Question 1 above is "No", even if $\mathscr{C}$ is actually a manifold, since it is not always possible to recover the Riemannian metric from the discrete ("cubical") one. (Recall that each edge is supposed to be of length 1.) However, in the special case of 3-dimensional cube complexes the question has a positive answer.

We sketch below the proof:

1. *Away from the vertices*, i.e. around the edges,[23] one can use a method developed by Gromov and Thurston [32] to produce a generalized type of branched cover (in any dimension). More precisely, (a) construct negatively curved conical surfaces of revolution, with vertex at a vertex v and with apex angle $\alpha = 2\pi/n$, where $n = \#_v Q$. Each such cone can be canonically mapped upon a Euclidean cone of apex angles $\pi/2$; then (b) glue the outcome of this process to the result of Step (2) below.
2. *Around the vertices* excise an ε-ball neighborhood B_ε of v. On the boundary of B_ε, i.e. on the sphere S_ε one has the natural triangulation by the intersections of S_ε with the cubes of $\mathscr{C}$ incident with v. Moreover, the curvature of the vertices of this triangulation will be $K_\varepsilon \equiv c/\varepsilon^2$, where c is some constant.

   However, while the gluing itself is trivial, one still has to ensure that the result is indeed endowed with a Riemannian metric. For this one has to go through Step 3 of the construction:
3. *Smoothen the ball* B_ε. In general dimension this represents a daunting problem. Indeed, even in dimension 3, Ricci flow—who represents a natural candidate for smoothing with control of curvature—is yet not attainable, since we can offer, at this point in time, no *PL* (metric) Ricci flow. However, due to Perelman's resolution of the Poincaré conjecture, in dimension 3 suffices to smoothen the boundary S_ε. It is at this point where the method described in Sect. 2.3 is applied, producing the required smooth ball $\tilde{S}_\varepsilon$, that has the same curvature as the *PL*[24] one S_ε.

[22] The author would like to thank the anonymous reviewer for bringing to his attention this paper.

[23] Obviously, in the interiors of the faces the metric is already smooth.

[24] But not piecewise Euclidean.

2.6.2 PL Ricci for Cell Complexes

Following [37], we briefly review here a definition of a metric Ricci curvature for *PL* manifolds in dimension higher than 2, as well as its immediate consequence, a method that does not make appeal to smoothings, as we did in the previous section.

2.6.2.1 The Definition

While the results in the preceding sections might be encouraging, one would still like to recover in the metric setting a "full" Ricci curvature, namely one that holds for 3- and higher dimensional manifolds, and not just in the degenerate case of surfaces. Our approach (as developed in [37]) is to start from the following classical formula:

$$\mathrm{Ric}(e_1) = \mathrm{Ric}(e_1, e_1) = \sum_{i=2}^{n} K(e_1, e_i)\,. \tag{2.30}$$

for any orthonormal basis $\{e_1, \cdots, e_n\}$, and where $K(e_1, e_j)$ denotes the sectional curvature of the 2-sections containing the directions e_1.

To adapt this expression for the Ricci curvature to the *PL* case, we first have to be able to define (*variational*) *Jacobi fields*. In this we heavily rely upon Stones's work [101, 102]. Note, however, that we do not need the full strength of Stone's technical apparatus, only the capability determine the relevant two sections and, of course, to decide what a direction at a vertex of a *PL* manifold is.

We start from noting that, in Stone's work, combinatorial Ricci curvature is defined both for the given simplicial complex $\mathscr{T}$, and for its *dual complex* $\mathscr{T}^*$ (see, e.g. [69, pp. 55–56]). For the dual complex, cells—playing here the role of the planes in the classical setting of which sectional curvatures are to be averaged—are considered. Unfortunately, Stone's approach for the given complex, where one computes the Ricci curvature $\mathrm{Ric}(\sigma, \tau_1 - \tau_2)$ of an n-simplex σ in the direction of two adjacent $(n-1)$-faces, τ_1, τ_2, is not natural in a geometric context (even if useful in his purely combinatorial one), except for the 2-dimensional case, where it coincides with the notion of Ricci curvature in a direction. However, passing to the dual complex will not restrict us, since $(\mathscr{T}^*)^* = \mathscr{T}$ and, moreover—and more importantly—considering *thick* triangulations enables us to compute the more natural metric curvature for the dual complex and use the fact that the dual of a thick triangulation is thick (for details, see [37]). Recall that thick (also called *fat*) triangulations are defined as follows:

Definition 2.14 Let $\tau \subset \mathbb{R}^n\,; 0 \leq k \leq n$ be a k-dimensional simplex. The *thickness* (or *fatness*) φ of τ is defined as being:

$$\varphi(\tau) = \frac{\mathrm{dist}(b, \partial\sigma)}{\mathrm{diam}\,\sigma}\,, \tag{2.31}$$

where b denotes the *barycenter* of σ and $\partial\sigma$ represents the standard notation for the *boundary* of σ (i.e the union of the $(n-1)$-dimensional faces of σ). A simplex τ is φ_0-*thick*, for some $\varphi_0 > 0$, if $\varphi(\tau) \geq \varphi_0$. A triangulation (of a submanifold of $\mathbb{R}^n$) $\mathscr{T} = \{\sigma_i\}_{i\in\mathbf{I}}$ is φ_0-*thick* if all its simplices are φ_0-thick. A triangulation $\mathscr{T} = \{\sigma_i\}_{i\in\mathbf{I}}$ is *thick* if there exists $\varphi_0 \geq 0$ such that all its simplices are φ_0-*thick*.

Keeping in mind the notions and facts above, we can now return to the definition of Ricci curvature for simplicial complexes: Given a vertex v_0 in the dual complex, corresponding to a n-dimensional simplicial complex, a *direction* at v_0 is just an oriented edge $e_1 = v_0v_1$. Since there exist precisely n 2-cells, $\mathfrak{c}_1, \ldots, \mathfrak{c}_n$, having e_1 as an edge and, moreover, these cells form part of n relevant variational (Jacobi) fields (see [101]), the Ricci curvature at the vertex v, in the direction e_1 is simply

$$\mathrm{Ric}(v) = \sum_{i=1}^{n} K(\mathfrak{c}_i)\,, \tag{2.32}$$

where we define the sectional curvature of a cell $\mathfrak{c}$ in the following manner:

Definition 2.15 Let $\mathfrak{c}$ be a cell with vertex set $V_{\mathfrak{c}} = \{v_1, \ldots, v_p\}$. The *embedding curvature* $K(\mathfrak{c})$ of $\mathfrak{c}$ is defined as:

$$K(\mathfrak{c}) = \min_{\{i,j,k,l\}\subseteq\{1,\ldots,p\}} \kappa(v_i, v_j, v_k, v_l)\,. \tag{2.33}$$

Remark 2.27 Note that by choosing to work with the dual complex we have restricted ourselves largely to considering solely submanifolds of $\mathbb{R}^N$, for some N sufficiently large. However, in the case of 2-dimensional *PL* manifolds this does nor represent restriction, since, by a result of Burago and Zalgaller [18] (see also [88]) such manifolds admit isometric embeddings in $\mathbb{R}^3$.

Remark 2.28 Evidently, the definition above presumes that cells in the dual complex have at least 4 vertices. However, except for some utterly degenerate (planar) cases, this condition always holds. Still, even in this case Ricci curvature can be computed using a slightly different approach—see the following remark.

Remark 2.29 It is still possible (by dualization) to compute Ricci curvature according, more or less, to Stone's ideas, at least for the 2-dimensional case. Indeed, according to [102],

$$\mathrm{Ric}(\sigma, \tau_1 - \tau_2) = 8n - \sum_{j=1}^{2n-1} \{|\beta_j| \mid \beta_j < \tau_1 \text{ or } \beta_j < \tau_2; \dim \beta_j = n-2\}\,. \tag{2.34}$$

For details and implications of this alternative approach, see [37].

2.6.2.2 Main Results

The first results one wants to ascertain are those ensuring the convergence of the newly defined Ricci curvature. These are quite straightforward, so here we content ourselves with simply stating them (for further details, see [37]).

Theorem 2.21 *Let $\mathscr{T}$ be a thick simplicial complex, and let $\mathscr{T}^*$ denote his dual. Then*

$$\lim_{\text{mesh}(\mathscr{T})\to 0} \text{Ric}(\sigma) = \lim_{\text{mesh}(\mathscr{T}^*)\to 0} C \cdot \text{Ric}^*(\sigma^*)\,, \tag{2.35}$$

where $\sigma \in \mathscr{T}$ and where $\sigma^* \in \mathscr{T}^*$ is (as suggested by the notation) the dual of σ.

Theorem 2.22 *Let M^n be a (smooth) Riemannian manifold and let $\mathscr{T}$ be a thick triangulation of M^n. Then*

$$\text{Ric}_{\mathscr{T}} \to C_1 \cdot \text{Ric}_{M^n},\ \text{as mesh}(\mathscr{T}) \to 0\,, \tag{2.36}$$

where the convergence is the weak convergence (of measures).

Beyond these convergence and approximations results, one would like to address deeper issues. Indeed, having introduced a metric Ricci curvature for *PL* manifolds, one naturally wishes to verify that this represents a proper notion of Ricci curvature, and not just an approximation of the classical notion. According to the synthetic approach to Differential Geometry, a proper notion of Ricci curvature should satisfy adapted versions of the main, essential theorems that hold for the classical notions. The first and foremost among such theorems is the Bonnet–Myers Theorem and, as expected, fitting versions for combinatorial cell complexes and weighted cell complexes were proven by Stone [101–103]. and Forman [26]. Moreover, the Bonnet part of the Bonnet-Myers theorem, that is the one appertaining to the sectional curvature, was also proven for *PL* manifolds, again by Stone—see [100, 103].

In [37] we proved a series of increasingly more general variants of the Bonnet–Myers Theorem, with proofs adapted to the various settings and/or notions of curvature (metric, combinatorial, Alexandrov comparison). Here we bring only two more representative ones.

Theorem 2.23 (*PL* Bonnet–Myers-Metric) *Let M^n_{PL} be a complete, n-dimensional PL, smoothable manifold without boundary, such that*

(i) There exists $d_0 > 0$, such that $\text{mesh}(M^n_{PL}) \leq d_0$*;*
(ii) $K_W(M^n_{PL}) \geq K_0 > 0\,,$
where $K_W(M^n_{PL})$ denotes the sectional curvature of the "combinatorial 2-sections".

Then M^n_{PL} is compact and, moreover

$$\operatorname{diam}(M^n_{PL}) \leq \frac{\pi}{\sqrt{K_0}} \,. \tag{2.37}$$

Unfortunately, determining whether a general *PL* complex has Wald curvature bounded from below can be, in practice, a daunting task. However, in the special case of thick complexes in $\mathbb{R}^N$, for some N one can determine a simple criterion as follows:

Theorem 2.24 (*PL* Bonnet-Myers-Thick Complexes) *Let $M = M^n_{PL}$ be a complete, connected PL manifold thickly embedded in some $\mathbb{R}^N$, such that $K_W(M^2) \geq K_0 > 0$, where M^2 denotes the 2-skeleton of M. Then M^n_{PL} is compact and, moreover*

$$\operatorname{diam}(M^2_{PL}) \leq \frac{\pi}{\sqrt{K_0}} \,. \tag{2.38}$$

Remark 2.30 The embedding condition in the theorem above necessitates, perhaps, further elaboration. One can, for instance, start with a (*PL*-)submanifold of $\mathbb{R}^N$, endowed with a thick triangulation (as it is the case in Graphics and Imaging, for instance). Alternatively, one can begin with an abstract metric *PL* manifold (recall that thickness is a purely metric concept—see Definition 2.14 above and embed it isometrically, or even just quasi-isometrically in $\mathbb{R}^N$. Moreover, one can be given a combinatorial *PL* manifold, i.e. such that the lengths of all the edges equals 1, and consider a quasi-conformal embedding of this object.

2.6.2.3 Scalar Curvature and a Comparison Theorem

Up to this point of we were concerned, in this section, solely with Ricci curvature. However, since Ricci curvature is the mean of sectional curvatures we had to consider them too (and, in fact, even more so in view of our definition of Ricci curvature for *PL* complexes). We did not discuss, however, scalar curvature. It is only fitting, therefore, for us to add a number of observation regarding this invariant, in particular since a immediate, but significant result presents itself.

Of course, we first have to define the scalar curvature $K_W(M)$ of a *PL* manifold M. In light of our preceding discussion and results, the following definition is quite natural:

Definition 2.16 Let $M = M^n_{PL}$ be an n-dimensional *PL* manifold (without boundary). The *scalar metric curvature* scal_W of M is defined as

$$\operatorname{scal}_W(v) = \sum_{\mathfrak{c}} K_W(\mathfrak{c}), \tag{2.39}$$

the sum being taken over all the cells of M^* incident to the vertex v of M^*.

Remark 2.31 Observe that the definition of scalar curvature of M is defined, somewhat counterintuitively, by passing to its dual M^*. However, this is consistent with our approach to Ricci curvature (and also similar to Stone's original approach—see the discussion above).

From this definition and our previous results (see [37]), we immediately[25] obtain, the following generalization of the classical curvature bounds comparison in Riemannian geometry:

Theorem 2.25 (Comparison Theorem) *Let $M = M^n_{PL}$ be an n-dimensional PL manifold (without boundary), such that $K_W(M) \geq K_0 > 0$, i.e. $K(\mathfrak{c}) \geq K_0$, for any 2-cell of the dual manifold (cell complex) M^*. Then*

$$K_W \gtreqless K_0 \Rightarrow \mathrm{Ric}_W \gtreqless nK_0\,. \tag{2.40}$$

Moreover

$$K_W \gtreqless K_0 \Rightarrow \mathrm{scal}_W \gtreqless n(n+1)K_0\,. \tag{2.41}$$

Remark 2.32

1. Inequality (2.41) can be formulated in the seemingly weaker form:

$$\mathrm{Ric}_W \gtreqless nK_0 \Rightarrow \mathrm{scal}_W \gtreqless n(n+1)K_0\,. \tag{2.42}$$

2. Note that in all the inequalities above, the dimension n appears, rather then $n-1$ as in the smooth, Riemannian case (hence, for instance one has in (2.41), $n(n+1)K_0$, instead of $n(n-1)K_0$[26] as in the classical case). This is due to our definition (2.32) of Ricci (and scalar) curvature, via the dual complex of the given triangulation, hence imposing standard and simple combinatorics, at the price of allowing only for such weaker bounds.[27]

2.7 Metric Curvatures for Metric Measure Spaces

While divagating somewhat from our professed goal, namely that of studying metric curvatures, it is impossible, especially in the context of this volume, not to mention the Ricci curvature for metric measure spaces, either as it was developed by Lott, Sturm and Villani [56, 104][28] (and its further elaborations [14]), or in the form

[25] And, in truth rather trivially, since the result holds, regardless of the specific definition for the curvature of a cell.

[26] But, on the other hand, this holds even if $n = 3$!...

[27] Without affecting the analogue of the Bonnet–Myers Theorem—see Sect. 2.2 above.

[28] See also [105].

pioneered by Ollivier [70] (and its further developments, mostly for graphs—due to Yau, Jost and their collaborators [46, 52], but also for polyhedral surfaces [55].[29])

Let us begin with the following observation: However simple and alluring the probabilistic approach may appear, to the geometer it seems somewhat unnatural, and even more so to those whose interest is drove mainly by possible implementations, e.g. people working in information geometry, image processing, manifold learning, etc.

Therefore, without diminishing whatsoever, the extensive theoretical merits of the Lott–Sturm–Villani approach, it still is a natural desire to find a new metric that encapsulates the behaviors of both the original metric and of the given measure. Here the accent should be understood as being placed on "simple", and by simple, we mean a "(geo-)metric", that is a metric that, while incorporating the measure, can still be investigated with very direct geometric methods, such as the ones discussed in detail above, or like the more analytic ones that we shall describe below. This is in some contrast to other means of "coalescing" metric-and-measure into a unique metric, such as the Gromov–Prokhorov and Gromov–Hausdorff–Prokhorov distances (see, e.g. [105] for these and also for some variations, as well as their "practical" versions (such as those in [58, 59]).

2.7.1 The Basic Idea: The Snowflaking Operator

We begin by introducing a number of definitions and facts required in the sequel. (As general bibliographical references for the material in this subsection, including missing proofs, we have used [42, 96, 97].)

2.7.1.1 Quasimetrics

We begin with the following basic definition:

Definition 2.17 Let X be a nonempty set. $q : X \times X \to \mathbb{R}_+$ is called a *K-quasimetric* iff

1. $q(x, y) = 0$ iff $x = y$;
2. $q(x, y) = q(y, x)$, for any $x, y \in X$;
3. $q(x, y) \leq K(q(x, z) + q(z, y))$, for any $x, y, z \in X$.

Remark 2.33 Some authors replace condition (2) above by the following weaker one: There exists $C_0 \geq 1$ such that $q(x, y) \leq C_0 q(y, x)$, for any $x, y \in X$.

[29] When mentioning generalized curvatures for surfaces, one can not fail to mention Morgan's [60] and his students' [25] work on "weighted" surfaces and curves.

Remark 2.34 A number of brief comments:

- A quasimetric is not necessarily a metric (while obviously, any metric is a quasimetric with $K = 1$).

  Counterexample 2.2 *The following counterexample is not only the basic one, it is—as we shall shortly see—very important to us in the sequel:*

$$q_s(x, y) = (d(x, y))^s \,, \tag{2.43}$$

  where d is a metric, is a quasimetric for any $s > 0$, but not, in general, a metric, for $s > 1$ (but it still is for $0 < s < 1$).
- Quasimetric balls can be defined precisely like metric balls, and the constitute the basis for a topology on X.
- For the next remark we need a definition that may appear a bit superfluous at this point, but it will prove to be highly relevant later on:

  Definition 2.18 Let (X, q) and (Y, ρ) be quasimetric spaces, and let $f : X \to Y$ be an injection. f is called *η-quasisymmetric*, where $\eta : [0, \infty) \to [0, \infty)$ is a homeomorphism iff

$$\frac{\rho(f(x), f(a))}{\rho(f(x), f(b))} \leq \eta\left(\frac{q(x, a)}{q(x, b)}\right), \tag{2.44}$$

  for any distinct points $x, a, b \in X$.

  Intuitively, while quasisymmetric mappings may change the size of balls quite dramatically, they do not change very much their shape. This fact is important in the next proposition (see, e.g. [96] for its proof), that shows that whereas, as we noted above, q_s is not a metric, the canonical injection $(X, d) \hookrightarrow (X, q_s)$ is quasisymmetric.

  Proposition 2.10 *Let q be a K-quasimetric on X. Then, there exists $s_0 = s_0(K)$ such that, for any $0 < s \leq s_0$ there exists a metric d_s on X, and a constant $C = C(s, K) \geq 1$, such that*

$$\frac{1}{C} q_s(x, y) \leq d_s(x, y) \leq C q_s(x, y) \,, \tag{2.45}$$

  where q_s is as in (2.43), i.e. $q_s(x, y) = (q(x, y))^s$.

  Remark 2.35 If q is a K-quasimetric ($K \geq 1$), then q_s is bilipschitz equivalent to d_s, for any $s > 0$, such that $(2K)^{2s} \leq 2$, that is for any $s > 0$ such that

$$s \leq \frac{1}{2}(\log_2 K + 1) \,.$$

Moreover, the bilipschitz constant can be chosen to be

$$C = (2K)^{2s}.$$

The importance of the proposition above (augmented by the precise estimates in its succeeding remark) is quite evident, but we would still like to emphasize its relevance for our goal, namely that of combining the metric and measure into a new metric, that is simple yet unifying of the metric and measure properties. What we have succeeded to show so far is that given a quasimetric q_d obtained by snowflaking from a metric d, one can find a metric quantifiable close to it. Therefore, such metric curvatures as, say, Haantjes curvature, can be defined for (curves in) quasimetric spaces via those of the new metric d_s. The properties of the new metric curvature κ_{H,d_s} are clearly close to those of $\kappa_{H,d}$ (where the notation is, we hope, self-explanatory). However, we postpone a more detailed comparative analysis for future work. Also, instead of choosing to incorporate the new metric in the Haantjes curvature (that functions as *geodesic* curvature), one can as well use it to compute a fitting Wald curvature (as a metric analogue of *sectional* curvature).

We still have, however, to be able to produce enough expressive quasimetrics. Here, by "expressive", we mean quasimetrics that not only approximate the original metric, but also incorporate, according the our goal detailed above, as faithful (or significantly) as possible the given measure as well. It turns out that, again, this is quite standard and easy, as we shall see in the next subsection.

2.7.1.2 From Doubling Measures to quasimetrics

We first remind the reader the following basic definition:

Definition 2.19 Let (X, d, μ) be a metric measure space X is called *doubling* iff the measure μ itself is doubling, i.e. iff there exists a constant D such that, for any $x \in X$ and any $r > 0$,

$$\mu\left(B_d[x, 2r]\right) \leq D\mu\left(B_d[x, r]\right). \tag{2.46}$$

(Here $B_d[x, r]$ denotes—as it standardly does—the closed ball of radius r, in the metric d.) A metric measure space (X, d, μ), where μ is doubling is sometimes called *of homogeneous type*.

For the record, a *metric measure space* is a triple $\mathscr{X} = (X, d, \mu)$ where (X, d) is a metric *Polish* space[30] (i.e. complete and having a countable base), and μ is a *Borel measure* on X.

[30] At least, this is the usual convention.

Remark 2.36 If (X, d, μ) is doubling, then it admits *atoms*, i.e. points of positive mass, only at isolated points.

Remark 2.37 The notion of doubling measures is, in fact, intrinsically related to that of curvature, more precisely with that of Ricci curvature: Any Riemannian manifold of nonnegative Ricci curvature is doubling (with respect to the volume measure)—see, e.g. [74]. (Indeed, it may be that this case represents one of the original motivations for studying doubling spaces.) Moreover, this implication is also preserved for the generalized Ricci curvature of Lott-Villani and Sturm (see, [105]).

We have now the necessary ingredient that allows us to construct the desired quasimetric, starting from a metric and a doubling measure: For any $s > 0$, we define the quasimetric $q_{\mu,s}$ as

$$q_{\mu,s}(x, y) = \big(\mu\left(B[x, d(x, y)]\right) + \mu\left(B[y, d(x, y)]\right)\big)^s . \tag{2.47}$$

(This can be written in compact form as $q_{\mu,s}(x, y) = \big(\mu(B_{x,y})\big)^s$, where $B_{x,y} = B[x, d(x, y)] \cup B[y, d(x, y)]$.)

Example 2.3 If $X = \mathbb{R}^n$, with $\mu \equiv Vol_n$, and if $s = 1/n$, then $q_{\mu,s} \equiv \text{const} \cdot d_{Eucl}$. (In particular, for $n = 2$, $q_{\mu,s} = \frac{\sqrt{\pi}}{2} d_{Eucl}$.)

Remark 2.38 For $X = \mathbb{R}^n$, one can define $q_{\mu,s}(x, y)$ simply by

$$q_{\mu,s}(x, y) = \left(\mu\left(B[m, \frac{x + y}{2}]\right)\right)^s ,$$

where m denotes the midpoint of the segment $\overline{xy}$. However, in the general case, and in particular for graphs, one has to use the more general expression (2.47).

Note that, if K is the quasimetric constant of $q_{\mu,s}$, then $K = K(\mu, s)$.

Also, by Proposition 2.10, there exists $s_0 = s_0(\mu) > 0$, such that $q_{\mu,s}$ is bilipschitz equivalent to a metric $d_{\mu,s}$, for any $0 < s \leq s_0$. This fact will play a crucial role in the sequel, as already hinted.

Remark 2.39 Obviously, the geometry induced by the quasimetric $q_{\mu,s}$, and a fortiori by the metric $d_{\mu,s}$, will diverge widely from the geometry given by the original metric d. This is most evident in the properties of the "new" geodesics, in comparison with the "old" ones (e.g. when $X = \mathbb{R}^n$ equipped with the standard Euclidean metric and with μ being the volume element.) However, the deformation of the geometry produced by (2.47) is controlled, and many essential properties are preserved. (For further details, see [96, 97].)

Remark 2.40 To be certain, one would like to explore the relevance of the "snowflaking" above to Imaging, etc. While further applications will be discussed below (see Theorems 2.28 and 2.30), even the very definition might prove to be useful, for instance in texture analysis and segmentation. In the case of images the substrate distance d can be chosen to be the preferred discrete distance (Euclidean, L_1, etc.) For the measure of the balls $B[x, d(x, y)]$ one can settle, of course, just

for the d-area. However, more interesting and relevant for textures measures present themselves, such as the Hausdorff measure (texture are viewed as fractals, sometimes) or the energy (of a *texton*).

As far as the choice of the points x and y is concerned, one possible (typical) choice would be the centers of adjacent neighborhoods or textons.

Clearly, in this case, one take $q_{\mu,s}(x,y)$ to be $\left(\mu\left(B[m, \frac{x+y}{2}]\right)\right)^s$, where m denotes the midpoint of the segment $\overline{xy}$, even if d is not the Euclidean distance. However, this will not be true when working with (communication) networks.

Also, the relevant parameters s have to be chosen such that $s \leq s_0$, where s_0 should be determined from the proof of Proposition 2.10, as restricted to the given specific, concrete context.

Unfortunately, the existence of the doubling measure required in producing the snowflaked quasimetrics may prove to be a quite daunting task (to say nothing about the lesser degree of geometric intuitiveness we are endowed with, in comparison with our grasp of the distance). Luckily enough, a simpler, purely metric condition exists that is, essentially, equivalent to that of doubling measure, at least as far as compete spaces are concerned. More precisely, we have the following

Definition 2.20 A metric space (X,d) is called *doubling* iff there exists $D_1 \geq 1$, such that any ball in X, of radius r, can be covered by at most D_1 balls of radius $r/2$.

(Obviously, there is nothing special about $r/2$, and the metric doubling condition can be formulated in terms of general sets of bounded diameter.)

Remark 2.41 Clearly, the metric and measures arising in Imaging and Vision are (quite trivially) doubling.

As expected (and alluded to above), there exists a connection between the notions of doubling metric and doubling measure. More precisely, we have the following

Lemma 2.2 *Let (X,d) be a metric space such that there exists a doubling measure μ on X. Then (X,d) is doubling (as a metric space).*

(A proof of this fact can be found in [42] or [96].)

The converse statement does not hold in general, a counterexample being $(\mathbb{Q}, d_{eucl})$ (see [42, p. 103]). However, it does hold for the important case of complete spaces:

Theorem 2.26 (Luukkainen–Saksman [57]) *Let (X,d) be a doubling, complete metric space. Then X carries a doubling measure.*

The following consequence since it obviously includes the important particular case of finite graphs:

Corollary 2.4 *Any compact, doubling metric spaces carries a doubling measure μ.*

The corollary above obviously holds for finite graphs.

Before we formulate the important theorem below, we give, for convenience, the following definition:

Definition 2.21 If (X,d) is a metric space, then the metric space (X,d^ε), $0<\varepsilon<1$, is called a *snowflaked version* of (X,d).

Theorem 2.27 (Assouad [5, 6]) *Let (X, d) be a doubling metric space. Then, for each $0 < \varepsilon < 1$, there exists N, such that its ε-snowflaked version is bilipschitz equivalent to a subset of $\mathbb{R}^N$, quantitatively.*

Here, *quantitatively* means that the embedding dimension N and the bilipschitz constant L depend solely on the doubling constant D of X and on the "snowflaking" factor ε, i.e.

$$N = N(D, \varepsilon),\ L = L(D, \varepsilon)\,.$$

Remark 2.42 Assouad's result does not hold, in general, for $\varepsilon = 1$. (For a counterexample, see [42, p. 99]).

From a practical, applicative point of view Assouad's theorem above allows us to "translate" the highly nonintuitive geometry of metric measure spaces to that of the familiar setting of subsets in (some) Euclidean space. In particular, in combination with our geometric, curvature based approach to sampling of images and higher dimensional signals (see [93, 94]), enables us to enunciate the following sampling "meta-theorem":

Theorem 2.28 *Sampling of Ahlfors regular metric measure spaces is quasisymmetrically equivalent, quantitatively, to the sampling of sets in $\mathbb{R}^N$, for some N.*

Before passing further on, let us mention briefly here that the sampling density is, roughly formulated, proportional to $1/K$ (or $1/\mathrm{Ric}$. (For the full details see, for instance, [35, 93, 94] and the references therein.)

Unfortunately, while the theorem above grants the desired framework for geometric sampling of metric measure spaces and, even more, it provides with numerical control over the distortion/error during the embedding in $\mathbb{R}^N$, the constants involved are far from being ideal—see the remark below for a more detailed discussion.

Remark 2.43 The beauty of Assouad's Theorem—and even more so its applicability in the sampling of real data—is marred by the "course of dimensionality": Given that $N = N(D, \varepsilon)$, the fear exist that, as in the case of Nash's Embedding Theorem [64, 65], the embedding dimension is prohibitively high for general manifolds (i.e. data). Obviously, this is even more important if low distortion—i.e. (bi-)lipschitz constant—is an imperative (as it usually is), that is for ε close to 0. And, indeed, Assouad's original construction provides $\lim_{\varepsilon \to 0} N(D, \varepsilon) = \infty$. So it would seem that, the price to pay for low distortion is a high embedding dimension. It is a quite recent result of Naor and Neiman [63] (itself based on ideas of Abraham, Bartal and Neiman [1]), that, in fact, given a (separable) D-doubling metric space, there exist $N = N(D) \in \mathbb{N}$ and $L = L(D, \varepsilon)$, such that for any $\varepsilon \in (0, 1/2)$, the $(1 - \varepsilon)$-snowflaked version of X admits a bilipschitz embedding in $\mathbb{R}^N$, with distortion L. Moreover, specific upper bounds for N and L are given: $N \leq a \log D, L \leq b \left(\frac{\log K}{\varepsilon}\right)^2$, where a and b are constants. So it appears that, at least as far as Assouad's Theorem is concerned, the snowflaking-based embedding is feasible.

At this point, one has to ask oneself whether this result can be improved. The belief in the possibility of such an improvement rests upon the following two facts: On the one hand, Assouad's Theorem assures the existence of a bilipschitz embedding, which represents a much stronger condition then mere quasisymmetry.[31] On the other hand, as we have seen, Ahlfors rigidity is not the most easy property to check directly on a metric measure space, therefore one naturally would wish to find a sampling result similar to Theorem 2.28, that would hold for general doubling spaces. Such a result does exist, and it makes appeal again to the quasimetric $q_{\mu,s}$ as defined by (3.3). However, we have to make an additional assumption, that ensures that $q_{\mu,s}$-lengths of curves in $\mathbb{R}^N$ do not "shrink" too much, due to the presence of the measure μ in the definition of $q_{\mu,s}$ (see [96]). We encode this restriction via

Definition 2.22 A doubling measure μ on $\mathbb{R}^N$ is called a *metric doubling measure* iff there exist a constant C_6, and a metric δ, such that

$$\frac{1}{C_6}\delta(x,y) \leq q_{\mu,\frac{1}{n}} \leq C_6\delta(x,y)\,,$$

for any $x, y \in X$, where $q_{\mu,\frac{1}{n}}$ is associated to μ as in (3.3), with $s = 1/n$.

We can now formulate the desired result, in terms of metric doubling measures:

Theorem 2.29 (Semmes [95], Theorem 1.15; [96], Proposition B 20.2) *Let (X,d) be a doubling metric space. Then there exists a natural number N and a metric doubling measure μ, such that (X,d) is bilipschitz equivalent to a subset of $(\mathbb{R}^N, q_{\mu,\frac{1}{N}})$, where $q_{\mu,\frac{1}{N}}$ is as above.*

Clearly *PL* surfaces (a.k.a. in Graphics as triangular meshes), endowed with a specific additional measure (e.g. luminosity), as well as images satisfy the metric doubling condition. (Even if modelled as fractals, certain textures are not truly such objects, due to the inherent discreteness, hence finiteness, therefore they also can be viewed as metric doubling measure spaces.) Again, this implies, in view of the theorem above, that, at least theoretically, images and meshes can be sampled as "weighted" manifolds, but using classical by now, geometric means of sampling subsets (hypersurfaces) of Euclidean space. The theorem is also relevant for sampling weighted networks, such as communication networks, for instance for choosing positions for routers/sensors [30], load balancing [83], or in-network sensor data storage [84].

Theorem 2.29 above represents a most encouraging result, and the idea of the proof is quite simple: By Assouad's Theorem, $(M, d^{\frac{1}{2}})$ is bilipschitz equivalent to a subset Y of some $\mathbb{R}^N$. The sought for measure on $\mathbb{R}^N$ will be define as $\mu = \text{dist}(x, Y^n)dx$—for details of the proof see [95].

[31] However, quasisymmetry represents a much more flexible analytic tool, than the rigid bilipschitz condition—see [42, 96, 97] for a deeper and far more detailed discussion.

One would naturally would hope that $(\mathbb{R}^n, q_{\mu,\frac{1}{n}})$ can be bilipschitzly embedded in some $\mathbb{R}^N$, for any doubling measure μ. This is a quite ambitious wish and, unfortunately, it is not true in general (see [95]). However, such an embedding exists for "most" metric doubling measures—for a precise formulation and the proof see [95]. Still, in view of the discussion proceeding Theorem 2.28 above, we can formulate the fitting sampling result (recall that given the quasimetric $q_{\mu,s}$, there exists a metric d_s bilipschitz equivalent to it):

Theorem 2.30 *Sampling of doubling metric spaces is bilipschitz equivalent quantitatively to the sampling of sets in* $(\mathbb{R}^N, d_{\frac{1}{N}})$*, for some N, where* $d_{\frac{1}{N}}$ *represents the snowflaked version of d, associated to a certain metric doubling measure* μ.

How relevant this approach to metric curvature for metric measure spaces will turn prove itself to be, besides providing a sound, intuitive and convenient theoretical setting for a wide range of signals is, unfortunately, to early to ascertain. However, in view of the success of the basic snowflaking approach (and related ideas) in solving such problems as the existence of ("enough") Lipschitz functions and Poincaré and Sobolev inequalities (i.e. that of "novel types" of "decent calculus"—see [96, 97]), as well as the existence of fitting versions on metric measure spaces satisfying the $CD(K,N)$ condition (see, for instance, [105]), one can display at least a moderate amount of optimism. For a different approach to sampling spaces satisfying a $CD(K,N)$ condition, as well as an application of the approach exposed in this section to the sampling of weighted graphs/networks, see [87]. Moreover, it turns out that the particular case of $N = +\infty$, when the $CD(K,N)$ condition reduces to the generalized Ricci curvature of Bakry, Emery and Ledoux it is natural and easy to implement of grayscale images, where the density function appearing in the formula of generalized Ricci curvature is nothing the grayscale level—see [53] for further details. We illustrate this new approach to sampling of grayscale images (natural images, but also range images, as well as cartoons) in Fig. 2.6. This is approach is even more natural and having far higher potential benefits in the context of medical images, such as CT and MRI images since, for instance, the density of many types of MRI images is equal to the proton density. A further application of the generalized Ricci curvature is in Graphics, where the density can be taken as, e.g. luminosity or shading of a mesh model. (This represents work in progress.) We should also note that while the generalized Ricci curvature (of metric measure spaces) might appear, prima facie, as an unnecessary complication, it allows not only for the sampling of wider range of images and signals, but it is also an *intrinsic* curvature, i.e. independent of the specific embedding considered. This is not a purely theoretical advantage, for it allows us to dispense with the need to compute the embedding curvature (e.g. tubular radius, reach, etc.), a task that is, in general, both non-trivial and cumbersome.

Fig. 2.6 A standard rest image (above, *left*) as a grayscale surface (above, *right*). One can compare the geometric sampling methods: The Gaussian curvature based one (*bottom*, *left*) and one using the generalized Ricci curvature (*bottom*, *right*)

Acknowledgements Research partly supported by Israel Science Foundation Grants 221/07 and 93/11 and by European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n^{o} [URI-306706].

Part of this work was done while visiting the Max Planck Institute, Leipzig. Their gracious and warm hospitality, as well as their support are gratefully acknowledged.

The author would like to thank to organizers of 2013 CIRM Meeting on Discrete Curvature for the opportunity they gave him to write this book chapter, and especially Pascal Romon for his attentive guidance and support during the process of writing this presentation, as well as of the short conference proceeding notes.

Thanks are also due to the anonymous reviewer for his attentive, insightful and most helpful corrections and suggestions.

References

1. Abraham, I., Bartal, Y., Neiman, O.: Embedding metric spaces in their intrinsic dimension. In: Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 363–372. Society for Industrial and Applied Mathematics, Philadelphia (2008)
2. Alexander, S.B., Bishop, R.L.: Comparison theorems for curves of bounded geodesic curvature in metric spaces of curvature bounded above. Differ. Geom. Appl **6**, 67–86 (1996)
3. Appleboim, E., Saucan, E., Zeevi, Y.Y.: Ricci curvature and flow for image denoising and superesolution. In: Proceedings of EUSIPCO, pp. 2743–2747 (2012)
4. Appleboim, E., Hyams, Y., Krakovski, S., Sageev, C., Saucan, E.: The scale-curvature connection and its application to texture segmentation. Theory Appl. Math. Comput. Sci. **3**(1), 38–54 (2013)
5. Assouad, P.: Étude d'une dimension métrique liée à la possibilité de plongement dans $\mathbb{R}^n$. C. R. Acad. Sci. Paris **288**, 731–734 (1979)
6. Assouad, P.: Plongements lipschitziens dans $\mathbb{R}^n$. Bull. Soc. Math. France **111**, 429–448 (1983)
7. 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)
8. Berestovskii, V.: Introduction of a Riemannian structure in certain metric spaces. Siberian Math. J. **16**, 210–221 (1977)
9. Berestovskii, V.: Spaces with bounded curvature and distance geometry. Siberian Math. J **27(1)**, 8–19 (1986)
10. Berger, M.: A Panoramic View of Riemannian Geometry. Springer, Berlin (2003)
11. Bestvina, M.: Geometric group theory and 3-manifolds hand in hand: the fulfillment of Thurston's vision. Bull. Am. Math. Soc **51**(1), 53–70 (2014)
12. Blumenthal, L.M.: Theory and Applications of Distance Geometry. Claredon Press, Oxford (1953)
13. Blumenthal, L., Menger, K.: Studies in Geometry. A Series of Books in Mathematics, vol. XIV, 512 pp. W.H. Freeman and Company, San Francisco (1970)
14. 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
15. Bourgain, J.: On Lipschitz embedding of finite metric spaces in Hilbert space. Israel J. Math **52**(1), 46–52 (1985)
16. Brehm, U., Kühnel, W.: Smooth approximation of polyhedral surfaces regarding curvatures. Geom. Dedicata **12**, 435–461 (1982)
17. Brooks, R.: Differential Geometry (Lecture Notes). Technion, Haifa (2003)
18. Burago, Y.D., Zalgaller, V.A.: Isometric piecewise linear immersions of two-dimensional manifolds with polyhedral metrics into $\mathbb{R}^3$. St. Petersburg Math. J. **7**(3), 369–385 (1996)
19. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence (2001)
20. Cassorla, M.: Approximating compact inner metric spaces by surfaces. Indiana Univ. Math. **41**, 505–513 (1992)
21. Cayley, A.: On a theorem in the geometry of position. Camb. Math. J. **2**, 267–271 (1841)
22. Cheeger, J.: Finiteness theorems for Riemannian manifolds. Am. J. Math. **90**, 61–74 (1970)
23. Chow, B.: The Ricci flow on the 2-sphere. J. Differ. Geom. **33**(2), 325–334 (1991)
24. Chow, B., Luo, F.: Combinatorial Ricci flows on surfaces. J. Differ. Geom. **63**(1), 97–129 (2003)
25. Corwin, I., Hoffman, N., Hurder, S., Šešum, V., Xu, Y.: Differential geometry of manifolds with density. Rose Hulman Undergraduate J. Math. **7**(1), 1–15 (2006)
26. Forman, R.: Bochner's method for cell complexes and combinatorial Ricci curvature. Discrete Comput. Geom. **29**(3), 323–374 (2003)
27. Giesen, J.: Curve reconstruction, the traveling salesman problem and Menger's theorem on length. In: Proceedings of the 15th ACM Symposium on Computational Geometry (SoCG), pp. 207–216. ACM, New York (1999)

28. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**(3), 1005–1028 (2008)
29. Gilboa, G., Appleboim, E., Saucan, E., Zeevi, Y.Y.: On the role of non-local menger curvature in image processing. In: Proceedings of ICIP 2015, pp. 4337–4341. IEEE (Society) (2015)
30. Goswami, M., Li, S.M., Zhang, J., Gao, J., Saucan, E., Gu, X.D.: Space filling curves for 3d sensor networks with complex topology. In: Proceedings of CCCG 2015, pp. 21–30 (2015)
31. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces. Modern Birkhäuser Classics, 3rd edn. Birkhäuser, Basel (2007)
32. Gromov, M., Thurston, W.: Pinching constants for hyperbolic manifolds. Invent. math. **86**, 1–12 (1987)
33. Grove, K., Markvorsen, S.: Curvature, triameter and beyond. Bull. Am. Math. Soc. **27**, 261–265 (1992)
34. Grove, K., Markvorsen, S.: New extremal problems for the Riemannian recognition program via Alexandrov geometry. J. Am. Math. Soc. **8**, 1–28 (1995)
35. Grove, K., Petersen, P.: Bounding homotopy types by geometry. Ann. Math. **128**, 195–206 (1988)
36. Grove, K., Petersen, P., Wu, J.Y.: Bounding homotopy types by geometry. Invent. Math. **99**(1), 205–213 (1990)
37. Gu, X.D., Saucan, E.: Metric ricci curvature for *pl* manifolds. Geometry **2013**(Article ID 694169), 12 pp. doi:10.1155/2013/694169 (2003)
38. Gu, X.D., Yau, S.T.: Computational Conformal Geometry. International Press, Somerville, MA (2008)
39. Haas, J.: Personal communication (2013)
40. Hamilton, R.S.: The Ricci flow on surfaces. Contemp. Math. **71**, 237–262 (1988)
41. Han, Q., Hong, J.X.: Isometric embedding of Riemannian manifolds in Euclidean spaces. AMS Math. Surv. **130**, 237–262 (2006)
42. Heinonen, J.: Lectures on Analysis on Metric Spaces. Springer, New York (2001)
43. Jin, M., Kim, J., Gu, X.D.: Discrete surface Ricci flow: theory and applications. In: IMA International Conference on Mathematics of Surfaces, pp. 209–232. Springer, Berlin (2007)
44. Jin, Y., Jost, J., Wang, G.: A nonlocal version of the Osher-Sole-Vese model. J. Math. Imaging Vision **44**(2), 99–113 (2012)
45. Jin, Y., Jost, J., Wang, G.: A new nonlocal H^1 model for image denoising. J. Math. Imaging Vision **48**(1), 93–105 (2014)
46. Jost, J., Liu, S.: Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs. Discrete Comput. Geom. **51**, 300–322 (2014)
47. Kay, D.C.: Arc curvature in metric spaces. Geom. Dedicata **9**(1), 91–105 (1980)
48. Kindermann, S., Osher, S., Jones, P.W.: Deblurring and denoising of images by nonlocal functionals. Multiscale Model. Simul. **4**(4), 1091–1115 (2005)
49. Krauthgamer, R., Linial, N., Magen, A.: Metric embeddings – beyond one-dimensional distortion. Discrete Comput. Geom. **31**, 339–356 (2004)
50. Lebedeva, N., Matveev, V., Petrunin, A., Shevchishin, V.: On d-dimensional d-semimetrics and simplex-type inequalities for high-dimensional sine functions. J. Approx. Theory **156**(1), 52–81 (2009)
51. Lebedeva, N., Matveev, V., Petrunin, A., Shevchishin, V.: Smoothing 3-dimensional polyhedral spaces. Electron. Res. Announc. Math. Sci. **22**, 12–19 (2015)
52. Lin, Y., Lu, L., Yau, S.T.: Ricci curvature of graphs. Tohoku Math. J. **63**(4), 605–627 (2011)
53. Lin, S., Luo, Z., Zang, J., Saucan, E.: Generalized Ricci curvature based sampling and reconstruction of images. In: Proceedings of EUSIPCO 2015, pp. 604–608 (2015)
54. Linial, N., London, E., Rabinovich, Y.: The geometry of graphs and some of its algorithmic applications. Combinatorica **15**, 215–245 (1995)
55. 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
56. 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

57. Luukkainen, J., Saksman, E.: Every complete doubling metric space carries a doubling measure. Proc. Am. Math. Soc. **162**(2), 903–991 (2009)
58. Mémoli, F.: On the use of Gromov-Hausdorff distances for shape comparison. In: Proceedings of the Point Based Graphics, Prague (2007)
59. Mémoli, F.: A spectral notion of Gromov–Wasserstein distance and related methods. Appl. Comput. Harmon. Anal. **30**(3), 363–401 (2011)
60. Morgan, F.: Manifolds with density. Not. Am. Math. Soc. **52**, 853–858 (2001)
61. Munkres, J.R.: Elementary Differential Topology (rev. ed.). Princeton University Press, Princeton, NJ (1966)
62. Naitsat, A., Saucan, E., Zeevi, Y.Y.: Volumetric quasi-conformal mappings. In: Proceedings of GRAPP/VISIGRAPP 2015, pp. 46–57 (2015)
63. Naor, A., Neiman, O.: Assouad's theorem with dimension independent of the snowflaking. Rev. Mat. Iberoam. **28**, 1–21 (2012)
64. Nash, J.: $\mathscr{C}^1$ isometric imbeddings. Ann. Math. **60**, 383–396 (1954)
65. Nash, J.: The embedding problem for Riemannian manifolds. Ann. Math. **63**, 20–63 (1956)
66. Nikolaev, I.G.: Parallel translation and smoothness of the metric of spaces of bounded curvature. Math. Dokl. **21**, 263–265 (1980)
67. Nikolaev, I.G.: Parallel displacement of vectors in spaces of Alexandrov two-side-bounded curvature. Siberian Math. J. **24**(1), 106–119 (1983)
68. Nikolaev, I.G.: Smoothness of the metric of spaces with bilaterally bounded curvature in the sense of A. Alexandrov. Siberian Math. J. **24**, 247–263 (1983)
69. Novikov, S.P.: Topology I: General Survey. Springer, Berlin/Heidelberg (1996)
70. Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256**(3), 810–864 (2009). http://dx.doi.org/10.1016/j.jfa.2008.11.001
71. Otsu, Y.: On manifolds with small excess. Am. J. Math. **115**, 1229–1280 (1993)
72. Pajot, H.: Analytic capacity, rectificabilility, menger curvature and the cauchy integral. In: Lecture Notes in Mathematics (LNM) 1799. Springer, Berlin (2002)
73. Perelman, G.: Alexandrov's spaces with curvature bounded from below II. Tech. rep., Leningrad Department of the Steklov Institute of Mathematics (LOMI) (1991)
74. Petersen, P.: Riemannian Geometry. Springer, New York (1998)
75. Plaut, C.: Almost Riemannian spaces. J. Differ. Geom. **34**, 515–537 (1991)
76. Plaut, C.: A metric characterization of manifolds with boundary. Compos. Math. **81**(3), 337–354 (1992)
77. Plaut, C.: Metric curvature, convergence, and topological finiteness. Duke Math. J. **66**(1), 43–57 (1992)
78. Plaut, C.: Spaces of Wald-Berestowskii curvature bounded below. J. Geom. Anal. **6**(1), 113–134 (1996)
79. Plaut, C.: Metric spaces of curvature $\geq k$. In: Daverman, R.J., Sher, R.B. (eds.) Handbook of Geometric Topology, pp. 819–898. Elsevier, Amsterdam (2002)
80. Richard, T.: Canonical smoothing of compact Alexandrov surfaces via Ricci flow. Tech. rep., Université Paris-Est Créteil (2012). ArXiv:1204.5461
81. Robinson, C.V.: A simple way of computing the Gauss curvature of a surface. Reports of a Mathematical Colloquium (Second Series) 5–6(1), pp. 16–24 (1944)
82. Sageev, M.: Cat(0) cube complexes and groups. In: Geometric Group Theory, IAS/Park City Mathematics Series, vol. 21, pp. 7–53. AMS and IAS/PCMI (2014)
83. Sarkar, R., Yin, X., Gao, J., Luo, F., Gu, X.D.: Greedy routing with guaranteed delivery using Ricci flows. In: Proceedings of IPSN'09, pp. 121–132 (2009)
84. Sarkar, R., Zeng, W., Gao, J., Gu, X.D.: Covering space for in-network sensor data storage. In: Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks, pp. 232–243. ACM, New York (2010)
85. Saucan, E.: Surface triangulation – the metric approach. Tech. rep., Technion (2004). Arxiv:cs.GR/0401023
86. Saucan, E.: Curvature – smooth, piecewise-linear and metric. In: What is Geometry? pp. 237–268. Polimetrica, Milano (2006)

87. Saucan, E.: A simple sampling method for metric measure spaces. preprint (2011). ArXiv:1103.3843v1 [cs.IT]
88. Saucan, E.: On a construction of Burago and Zalgaller. Asian J. Math. **16**(5), 587–606 (2012)
89. Saucan, E.: A metric Ricci flow for surfaces and its applications. Geom. Imaging Comput. **1**(2), 259–301 (2014)
90. Saucan, E.: Metric curvatures and their applications. Geom. Imaging Comput. **2**(4), 257–334 (2015)
91. Saucan, E., Appleboim, E.: Curvature based clustering for dna microarray data analysis. In: Lecture Notes in Computer Science, vol. 3523, pp. 405–412. Springer, New York (2005)
92. Saucan, E., Appleboim, E.: Metric methods in surface triangulation. In: Lecture Notes in Computer Science, vol. 5654, pp. 335–355. Springer, New York (2009)
93. Saucan, E., Appleboim, E., Zeevi, Y.: Sampling and reconstruction of surfaces and higher dimensional manifolds. J. Math. Imaging Vision **30**(1), 105–123 (2008)
94. Saucan, E., Appleboim, E., Zeevi, Y.: Geometric approach to sampling and communication. Sampl. Theory Signal Image Process. **11**(1), 1–24 (2012)
95. Semmes, S.: Bilipschitz mappings and strong a^{∞} weights. Acad. Sci. Fenn. Math. **18**, 211–248 (1993)
96. Semmes, S.: Metric spaces and mappings seen at many scales. In: Metric Structures for Riemannian and non-Riemannian Spaces. Progress in Mathematics, vol. 152, pp. 401–518. Birkhauser, Boston (1999)
97. Semmes, S.: Some Novel Types of Fractal Geometry. Clarendon Press, Oxford (2001)
98. Shephard, G.C.: Angle deficiences of convex polytopes. J. Lond. Math. Soc. **43**, 325–336 (1968)
99. Sonn, E., Saucan, E., Appelboim, E., Zeevi, Y.Y.: Ricci flow for image processing. In: Proceedings of IEEEI 2014. IEEEI (2014)
100. Stone, D.A.: Sectional curvatures in piecewise linear manifolds. Bull. Am. Math. Soc. **79**(5), 1060–1063 (1973)
101. Stone, D.A.: A combinatorial analogue of a theorem of Myers. Ill. J. Math. **20**(1), 12–21 (1976)
102. Stone, D.A.: Correction to my paper: "A combinatorial analogue of a theorem of Myers" (Illinois J. Math. **20**(1), 12–21 (1976). Illinois J. Math. **20**(3), 551–554 (1976)
103. Stone, D.A.: Geodesics in piecewise linear manifolds. Trans. Am. Math. Soc. **215**, 1–44 (1976)
104. 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
105. Villani, C.: Optimal transport, Old and new. In: 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
106. Wald, A.: Sur la courbure des surfaces. C. R. Acad. Sci. Paris **201**, 918–920 (1935)
107. Wald, A.: Begründung einer koordinatenlosen Differentialgeometrie der Flächen. Ergebnisse eines Math. Kolloquiums, 1. Reihe **7**, 24–46 (1936)

Chapter 3
Distances Between Datasets

Facundo Mémoli

Abstract We overview the construction and quantitative aspects of the Gromov–Hausdorff and Gromov–Wasserstein distances.

3.1 Introduction

Modeling datasets as metric spaces seems to be natural for some applications and concepts revolving around the Gromov–Hausdorff distance—a notion of distance between compact metric spaces—provide a useful language for expressing properties of data and shape analysis methods. In many situations, however, this is not enough, and one must incorporate other sources of information into the model, with "weights" attached to each point being one of them. This gives rise to the idea of representing data as metric measure spaces, which are metric spaces endowed with a probability measure. In terms of a distance, the Gromov–Hausdorff metric is replaced with the Gromov–Wasserstein metric.

3.1.1 Notation and Background Concepts

The book by Burago et al. [4] is a valuable source for many concepts in metric geometry. We refer the reader to that book for any concepts not explicitly defined in these notes.

We let $\mathscr{M}$ denote the collection of all compact metric spaces and by $\mathscr{M}^{\text{iso}}$ the collection of all isometry classes of $\mathscr{M}$. Recall that for a given metric space $(X, d_X) \in \mathscr{M}$, its diameter is defined as $\operatorname{diam} X := \max_{x,x' \in X} d_X(x, x')$. Similarly, the radius of X is defined as $\mathbf{rad}\,(X) := \min_{x \in X} \max_{x' \in X} d_X(x, x')$.

F. Mémoli (✉)
Department of Mathematics, The Ohio State University, Columbus, OH, USA
e-mail: memoli@math.osu.edu

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_3

For a fixed metric space (Z, d_Z), we let d_{H}^Z denote the Hausdorff distance between (closed) subsets of Z: for $A, B \subset Z$ closed,

$$d_{\mathrm{H}}^Z\big(A, B\big) := \inf\big\{\varepsilon > 0 | \, A \subset B^\varepsilon, \text{ and } B \subset A^\varepsilon\big\}.$$

Here, $A^\varepsilon := \{z \in Z|, d_Z(z, A) \leq \varepsilon\}$. For $\varepsilon \geq 0$, and ε-net of a the metric space Z is any subset $A \subset Z$ such that $d_{\mathrm{H}}^Z\big(A, Z\big) \leq \varepsilon$.

We will often refer to a metric space (X, d_X) by only X, but the notation for the underlying metric will be implicitly understood to be d_X. Recall that a map $\varphi : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is an isometric embedding if $d_Y(\varphi(x), \varphi(x')) = d_X(x, x')$ for all $x, x' \in X$. The map φ is an isometry if it is a surjective isometric embedding.

Recall that given measurable spaces (X, Σ_X) and (Y, Σ_Y), a measure μ on (X, Σ_X) and a measurable map $f : X \to Y$, the push-forward measure $f_\# \mu$ on (Y, Σ_Y) acts according to $f_\# \mu(B) = \mu(f^{-1}(B))$ for any $B \in \Sigma_Y$.

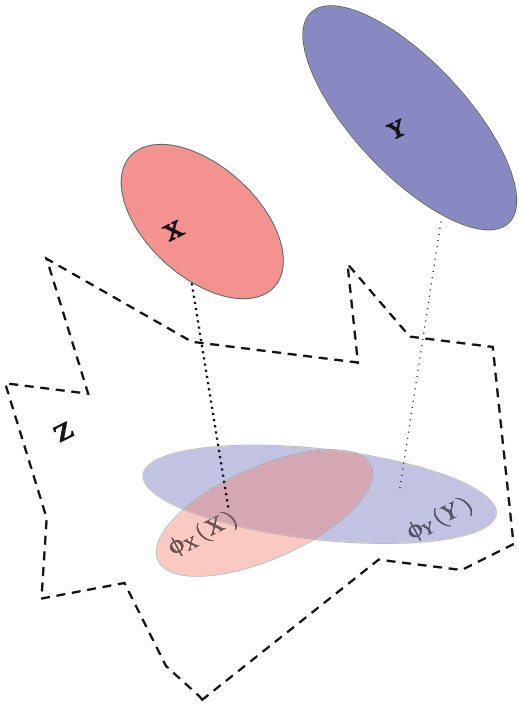
A metric measure space (mm-space for short) is a triple (X, d_X, μ_X) where (X, d_X) is a compact metric space and μ_X is a Borel probability measure with full support: $\mathrm{supp}\,(\mu_X) = X$. We denote by $\mathscr{M}^w$ the collection of all mm-spaces. An isomorphism between $X, Y \in \mathscr{M}^w$ is any isometry $\Psi : X \to Y$, such that $\Psi_\# \mu_X = \mu_Y$.

3.2 The Gromov–Hausdorff Distance

The goal is to measure distance between two given abstract compact metric spaces. In general, these two spaces may not be readily given as subsets of a common metric space. In this case, the following construction by Gromov [7] applies.

Given (X, d_X) and (Y, d_Y) in $\mathscr{M}$ one considers any "sufficiently rich" third metric space (Z, d_Z) inside which one can find isometric copies of X and Y and measures the Hausdorff distance in Z between these copies. Finally, one minimizes over the choice of the isometric copies and Z. Formally, let Z, $\phi_X : X \to Z$ and $\phi_Y : Y \to Z$ be respectively a metric space and isometric embeddings of X and Y into Z. Then, the Gromov-Hausdorff distance between X and Y is defined as

$$d_{\mathrm{GH}}(X, Y) := \inf_{Z, \phi_X, \phi_Y} d_{\mathrm{H}}^Z\big(\phi_X(X), \phi_Y(Y)\big)\,. \tag{3.1}$$



Theorem 3.1 ([7]) *d_{GH} is a legitimate distance on the collection of isometry classes of $\mathscr{M}$.*

From the practical point of view this definition might not look appealing. As we recall below, there are other more computational suggestive equivalent definitions whose implementation has been explored. But now we try to interpret the definition we have given so far.

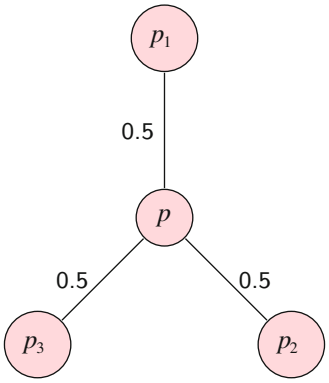
3.2.1 An Example

Consider the metric spaces X consisting exactly of three points at distance 1 from each other, and Y consisting of exactly one point. Notice that X and Y can be *simultaneously* embedded into $\mathbb{R}^2$ in an isometric way so that $Z = \mathbb{R}^2$ is a valid choice in (3.1) above. The maps ϕ_X and ϕ_Y represent the relative positions of X and Y in the plane.

By homogeneity, we can assume that the the embedding of X is fixed. When choosing ϕ_Y one notices that the optimal relative position of $q := \phi_Y(Y)$ with respect to $\Delta := \phi_X(X)$ happens when q is the center of the (equilateral) triangle Δ. In that case, the Hausdorff distance in (3.1) is $\delta_0 := \frac{1}{\sqrt{3}}$ and we conclude that $d_{\mathrm{GH}}(X, Y) \leq \delta_0$. One would be tempted to think that δ_0 is in fact *equal* to Gromov-Hausdorff distance between X and Y but this is not the case!

The same construction that we did above for $\mathbb{R}^2$ can in fact be done on the model hyperbolic two-dimensional space $\mathbb{H}_\kappa$ of curvature $-\kappa$ for *any* $\kappa \leq 0$. As $\kappa \to -\infty$, the (geodesic interpolation of the) triangle Δ becomes 'thinner' and intuitively, the Hausdorff distance δ_κ between the optimal embeddings in $\mathbb{H}_\kappa$ will decrease as κ decreases.

One can in fact consider the following target metric space: Z_∞ consists of four points p_1, p_2, p_3, and p such that $d_Z(p_i, p_j) = 1$ for $i \neq j$ and $d_Z(p_i, p) = \frac{1}{2}$ for all i. This metric space with four points can be regarded as a subset of the *real tree* (geodesic) metric space below:



This metric space can be regarded as an extremal case of the construction involving the $\mathbb{H}_\kappa$ that was described above. The interesting fact is that if we let $\phi_X(X) = \{p_1, p_2, p_3\}$ and $\phi_Y(Y) = \{q\}$, then $\delta_\infty := d_{\mathrm{H}}^{Z_\infty}\big(\phi_X(X), \phi_Y(Y)\big) = \frac{1}{2}$ which is strictly smaller than δ_0! and thus proves that

$$d_{\mathrm{GH}}(X, Y) \leq \frac{1}{2} < \frac{1}{\sqrt{3}}.$$

One can in fact check that $\delta_\infty < \delta_\kappa \leq \delta_0$ for all $\kappa \in [0, \infty)$. In any case, as we recall in Corollary 3.2 below, $d_{\mathrm{GH}}(X, Y)$ is always bounded below by $\frac{1}{2}|\operatorname{diam} X - \operatorname{diam} Y|$. Since in the present case $\operatorname{diam} X = 1$ and $\operatorname{diam} Y = 0$, we obtain that $d_{\mathrm{GH}}(X, Y) \geq \frac{1}{2}$ which together with the reverse inequality obtained above implies that in fact, for the example under consideration, $d_{\mathrm{GH}}(X, Y) = \frac{1}{2}$!

3.2.2 A Simplification

Kalton and Ostrovskii [8] observed that one can *equivalently* define the Gromov-Hausdorff distance between X and Y by considering Z in (3.1) to be the disjoint union $X \sqcup Y$ together with *any* metric d such that $\mathrm{d}_{|X\times X} = d_X$ and $\mathrm{d}_{|Y\times Y} = d_Y$. Let

$\mathscr{D}(d_X, d_Y)$ denote the set of *all* such metrics on $X \sqcup Y$. Then, they observe that

$$d_{\mathrm{GH}}(X, Y) = \inf_{\mathrm{d} \in \mathscr{D}(d_X, d_Y)} d_{\mathrm{H}}^{(X \sqcup Y, \mathrm{d})}\big(X, Y\big). \tag{3.2}$$

This expression for the Gromov-Hausdorff distance seems more appealing for the computationally minded: imagine that X and Y are finite, then the *variable* d in the underlying optimization problem can be regarded as a matrix in $\mathbb{R}^{|X| \times |Y|}$. If we assume that $|X| = |Y| = n$ then the number of *linear constraints* that each d in $\mathscr{D}(d_X, d_Y)$ must satisfy is of order n^3 (all triangle inequalities). Even more explicitly, the optimization problem over $\mathscr{D}(d_X, d_Y)$ that one must solve in practice is (cf. [11]) $\min_{\mathrm{d}} J(\mathrm{d})$ where

$$J(\mathrm{d}) := \max\left(\max_{x \in X} \min_{y \in Y} \mathrm{d}(x, y), \max_{y \in Y} \min_{x \in X} \mathrm{d}(x, y) \right).$$

The complexity from the original definition (3.1) is now hidden in the fact that $J(\cdot)$ is highly non-linear.

Going back to the example discussed in 3.2.1, one can state that in the context of (3.2), the optimal metric on $X \sqcup Y$ is

$$\mathrm{d}^* := \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

3.2.3 Another Expression

One says that a subset $R \subset X \times Y$ is a correspondence between sets X and Y whenever $\pi_1(R) = X$ and $\pi_2(R) = Y$, where $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are the canonical projections. Let $\mathscr{R}(X, Y)$ denote the set of all correspondences between X and Y.

The Gromov–Hausdorff (GH) distance between compact metric spaces (X, d_X) and (Y, d_Y) can be proved to be equal to [8]:

$$d_{\mathrm{GH}}(X, Y) := \frac{1}{2} \inf_{R} \sup_{(x,y),(x',y') \in R} \left| d_X(x, x') - d_Y(y, y') \right| \tag{3.3}$$

where R ranges over $\mathscr{R}(X, Y)$.

Example 3.1 The GH distance between any compact metric space X and the space with exactly one point is equal to $\frac{1}{2} \operatorname{diam} X$.

It turns out that $(\mathscr{M}, d_{\mathscr{GH}})$ is a nice space in that is has many compact subclasses.

Theorem 3.2 ([4]) *Let $N : [0, +\infty) \to \mathbb{N}$ be a bounded function and $D > 0$. Let $\mathscr{F}(N, D) \subset \mathscr{M}$ be any family of compact metric spaces, such that* $\operatorname{diam} X \leq D$ *for all $X \in \mathscr{F}(N, D)$, such that for any $\varepsilon > 0$, any $X \in \mathscr{F}(N, \varepsilon)$ admits an ε-net with at most $N(\varepsilon)$ elements. Then, $\mathscr{F}(N, D)$ is pre-compact in the Gromov–Hausdorff topology.*

Example 3.2 An important example of families, such as the above, is given by those closed n-dimensional Riemannian manifolds $(X, g_X) \in \mathscr{M}(n, \kappa, D)$ with the diameter bounded by $D > 0$ and the Ricci curvature bounded below by κ.

Theorem 3.3 ([19]) *The space $(\mathscr{M}^{\mathrm{iso}}, d_{\mathrm{GH}})$ is complete.*

It then follows from the two theorems above that classes $\mathscr{F}(N, D)$, such as above, are totally bounded for the Gromov–Hausdorff distance. This means that such classes are easy to organize in the sense of clustering or databases.

3.2.4 The Case of Subsets of Euclidean Space

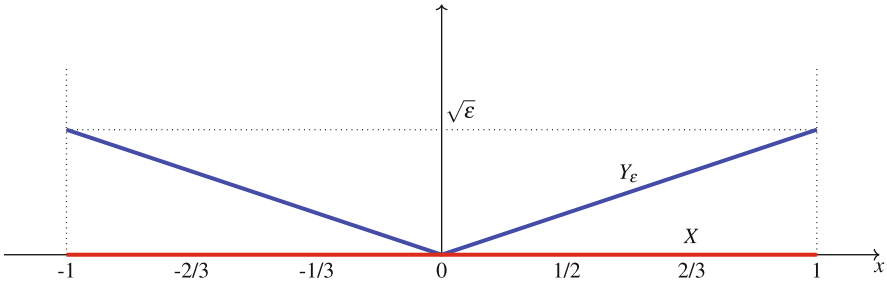
Even if we saw in Sect. 3.2.1 above that when X and Y are subsets of $\mathbb{R}^d$ the optimal Z in (3.1) may not be $\mathbb{R}^d$, one can still relate $d_{\mathrm{GH}}(X, Y)$ with some natural notion of distance for subsets of Euclidean space. Doing this provides more insight into as to how the Gromov-Hausdorff distance operates in situations for which we already have a well developed intuition.

An *intrinsic approach* to comparing two subsets X and Y of $\mathbb{R}^d$ would be to regard them as metric spaces by endowing them with the restriction of the ambient space metric: $d_X(\cdot, \cdot) = \| \cdot - \cdot \|$ etc. So, one can consider $d_{\mathrm{GH}}(X, Y)$ as a possible notion of dissimilarity between X and Y.

Another notion of dissimilarity that is frequently considered in shape and data analysis arises from the Hausdorff distance modulo rigid isometries and constitutes an *extrinsic approach*: let $E(d)$ denote the group of isometries of $\mathbb{R}^d$ and define

$$d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y) := \inf_{T \in E(d)} d_{\mathrm{H}}^{\mathbb{R}^d}\big(X, T(Y)\big).$$

Since in this case, one can always choose $Z = \mathbb{R}^d$ in (3.1) above, one immediately sees that $d_{\mathrm{GH}}(X, Y) \leq d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y)$. Even if we already saw in Sect. 3.2.1 that the equality cannot take place in general, one could hope that for some suitable $C > 0$, $d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y) \leq C \cdot d_{\mathrm{GH}}(X, Y)$ for all $X, Y \subset \mathbb{R}^d$ compact. Interestingly, however, this cannot happen! Consider $X = [-1, 1]$. Fix $0 < \varepsilon \ll 1$ and let $f_\varepsilon(x) := |x| \cdot \sqrt{\varepsilon}$. Let Y_ε be the set $\{(x, f_\varepsilon(x)); x \in [-1, 1]\}$. Notice that $\mathbf{rad}\,(X) = 1$ and $\mathbf{rad}\,(Y_\varepsilon) = \sqrt{1 + \varepsilon}$.



In any case, it is clear that for $\varepsilon > 0$ small enough, $d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y_\varepsilon) = \frac{\sqrt{\varepsilon}}{2}$. However, since by Proposition 3.1 and Corollary 3.1 below,

- $d_{\mathrm{GH}}(X, Y_\varepsilon) \geq \frac{1}{2}|\mathbf{rad}\,(X) - \mathbf{rad}\,(Y_\varepsilon)\,| = \frac{1}{2}(\sqrt{1+\varepsilon} - 1) \geq \frac{\varepsilon}{2+2\sqrt{2}}$ and
- $d_{\mathrm{GH}}(X, Y_\varepsilon) \leq \frac{1}{2}\sup_{|x|\neq|x'|} |x - x'| \cdot \left(\sqrt{1 + \varepsilon \cdot \left(\frac{|x|-|x'|}{x-x'}\right)^2} - 1\right) \leq \varepsilon$, since $\big||x| - |x'|\big| \leq |x - x'|$ for all $x, x' \in X$.

It follows that $d_{\mathrm{GH}}(X, Y_\varepsilon)$ is of order ε and therefore no constant $C > 0$ will guarantee that $C \cdot d_{\mathrm{GH}}(X, Y_\varepsilon) \geq d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y_\varepsilon)$ for all $1 \gg \varepsilon > 0$!

What does hold for this construction is that $C \cdot \big(d_{\mathrm{GH}}(X, Y_\varepsilon)\big)^{1/2} \geq d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y_\varepsilon)$ for some constant $C > 0$. It turns out that this is not an isolated phenomenon:

Theorem 3.4 ([10]) *For each natural number $d \geq 2$ there exists $c_d > 0$ such that for all $X, Y \in \mathbb{R}^d$ one has*

$$d_{\mathrm{GH}}(X, Y) \leq d_{\mathrm{H}}^{\mathbb{R}^d,\mathrm{rigid}}(X, Y) \leq c_d \cdot M^{1/2} \cdot \big(d_{\mathrm{GH}}(X, Y)\big)^{1/2},$$

where $M = \max(\operatorname{diam} X, \operatorname{diam} Y)$.

3.2.5 Another Expression and Consequences

For two sets X and Y let $\mathscr{R}(X, Y)$ denote the set of all *correspondences* between X and Y, that is, sets $R \subseteq X \times Y$ such that $\pi_1(R) = X$ and $\pi_2(R) = Y$. In general, we will refer to any non-empty set R of $X \times Y$ as a *relation* between X and Y. Obviously, all correspondences are relations.

The *distortion* of a relation R between the metric spaces (X, d_X) and (Y, d_Y) is defined as the number

$$\operatorname{dis}(R) := \sup_{(x,y),(x',y')\in R} \big|d_X(x, x') - d_Y(y, y')\big|.$$

Notice that given a function $\varphi : X \to Y$ one can define the relation $R_\varphi := \{(x, \varphi(x)); x \in X\}$, and in that case we write $\mathrm{dis}(\varphi) := \mathrm{dis}(R_\varphi) = \sup_{x,x' \in X} \left| d_X(x, x') - d_Y(\varphi(x), \varphi(x')) \right|$. Similarly, when $\psi : Y \to X$ is given, it induces the relation $R_\psi := \{(\psi(y), y); y \in Y\}$. Note that the structure of R_φ is different from the structure of R_ψ.

Now, when a map $\varphi : X \to Y$ and a map $\psi : Y \to X$ are both specified, we consider the relation $R_{\varphi,\psi} := R_\varphi \bigcup R_\psi$ and note that in fact $R_{\varphi,\psi}$ is actually a correspondence between X and Y.

Furthermore, one can explicitly compute that

$$\mathrm{dis}(R_{\varphi,\psi}) = \max\big(\mathrm{dis}(\varphi), \mathrm{dis}(\psi), C(\varphi, \psi)\big),$$

where $C(\varphi, \psi) := \sup_{x \in X, y \in Y} \left| d_X(x, \psi(y)) - d_Y(\varphi(x), y) \right|$. Notice that if $C(\varphi, \psi) < \eta$ for some $\eta > 0$, then $\left| d_X(x, \psi(y)) - d_Y(\varphi(x), y) \right| < \eta$ for all $(x, y) \in X \times Y$. In particular, for $x = \psi(y)$, it follows that $d_Y(\varphi \circ \psi(y), y) < \eta$ for all $y \in Y$. Similarly one can obtain $d_X(x, \psi \circ \varphi(x)) < \eta$ for all $x \in X$. These two conditions are often interpreted as meaning that φ and ψ are close to being inverses of each other. This proximity is quantified by η.

An interesting and useful characterization of the Gromov-Hausdorff distance based on optimization over correspondences is the following:

Theorem 3.5 ([8]) *For all $X, Y \in \mathscr{M}$ one has that*

$$d_{\mathrm{GH}}(X, Y) \overset{\text{(I)}}{=} \frac{1}{2} \inf_{R \in \mathscr{R}(X,Y)} \mathrm{dis}(R) \overset{\text{(II)}}{=} \frac{1}{2} \inf_{\varphi,\psi} \mathrm{dis}(R_{\varphi,\psi}).$$

Corollary 3.1 *Let X be a set and d and d' be any two metrics on X. Then,*

$$d_{\mathrm{GH}}((X, d), (X, d')) \leq \frac{1}{2} \sup_{x,x' \in X} \left| d(x, x') - d'(x, x') \right|.$$

The theorem above is significant for several reasons. First of all, (I) indicates that solving for the Gromov-Hausdorff distance between two finite metric spaces is an instance of a well known combinatorial optimization problem called the *bottleneck quadratic assignment problem* or bQAP. The bQAP is NP-Hard and furthermore, computing any $(1 + \varepsilon)$ of the optimal solution is also NP-Hard for any $\varepsilon > 0$ [18]. See [3, 14, 15] for some heuristic approaches.

A second observation stemming from the equality (II) in the theorem is the fact that since the term $C(\varphi, \psi)$ acts as a *coupling term* in the optimization

$$d_{\mathrm{GH}}(X, Y) = \frac{1}{2} \inf_{\varphi,\psi} \max\big(\mathrm{dis}(\varphi), \mathrm{dis}(\psi), C(\varphi, \psi)\big),$$

one could conceive of dropping it from the expression above yielding

$$d_{\mathrm{GH}}(X,Y) \geq \frac{1}{2}\max\left(\inf_{\varphi}\mathrm{dis}(\varphi), \inf_{\psi}\mathrm{dis}(\psi)\right) =: \widehat{d}_{\mathrm{GH}}(X,Y)\,.$$

It is important to notice that computing $\widehat{d}_{\mathrm{GH}}(X,Y)$, which we call the *modified Gromov-Hausdorff distance* [13], leads to solving two *decoupled* optimization problems, a feature which is desirable in applications. However, the computational complexity of the problems of the type $\inf_{\varphi}\mathrm{dis}(\varphi)$ could still be high. We will explore some interesting structure that arises from this modified definition in the next section but for now we will make one more observation based on the expression given by Theorem 3.5.

From equality (I) it follows that the Gromov-Hausdorff distance between any compact metric space and the metric space consisting of exactly one point is $d_{\mathrm{GH}}(X,*) = \frac{1}{2}\operatorname{diam} X$. As a corollary from Theorem 3.1 and this observation one has

Corollary 3.2 *For all $X, Y \in \mathscr{M}$, $d_{\mathrm{GH}}(X,Y) \geq \frac{1}{2}\left|\operatorname{diam} X - \operatorname{diam} Y\right|$.*

Proof The inequality $d_{\mathrm{GH}}(X,Y) \geq \left|d_{\mathrm{GH}}(X,*) - d_{\mathrm{GH}}(Y,*)\right|$ is guaranteed by the triangle inequality for the Gromov-Hausdorff distance. The remark preceding the statement completes the proof.

A similar lower bound for the Gromov-Hausdorff distance arises from considering the radius of metric spaces:

Proposition 3.1 ([13]) *For all $X, Y \in \mathscr{M}$, $d_{\mathrm{GH}}(X,Y) \geq \frac{1}{2}\left|\mathbf{rad}\,(X) - \mathbf{rad}\,(Y)\right|$.*

3.3 The Modified Gromov-Hausdorff and Curvature Sets

It could appear plausible that by dropping the coupling term $C(\varphi,\psi)$ in the optimization above one might have lost some of the nice theoretical properties enjoyed by the Gromov-Hausdorff distance. This is not the case, and in fact the modified Gromov-Hausdorff retains many of these good properties:

Theorem 3.6 ([13]) *The modified Gromov-Hausdorff distance satisfies:*

1. *$\widehat{d}_{\mathrm{GH}}$ is a legitimate metric on the isometry classes of $\mathscr{M}$.*
2. *$d_{\mathrm{GH}}(X,Y) \geq \widehat{d}_{\mathrm{GH}}(X,Y)$ for all $X, Y \in \mathscr{M}$.*
3. *d_{GH} and $\widehat{d}_{\mathrm{GH}}$ are topologically equivalent within d_{GH}-precompact families of $\mathscr{M}$.*

It is however interesting that the equality in item (2) *does not take place in general*. In fact, [13] provides a counterexample.

3.3.1 Curvature Sets

Gromov [7] defines for each $n \in \mathbb{N}$ the curvature sets of $X \in \mathscr{M}$ in the following way: let $\Psi_X^{(n)} : X^{\times n} \to \mathbb{R}^{n \times n}$ be the matrix valued map defined by $(x_1, \ldots, x_n) \mapsto ((d_X(x_i, x_j)))_{i,j=1}^n$. This map simply assigns to each n-tuple of points its *distance matrix*: the matrix arising from restricting the metric on X to the given n-tuple. Then, the n-th curvature set of X is

$$\mathbf{K}_n(X) := \left\{ \Psi_X^{(n)}(x_1, \ldots, x_n); (x_1, \ldots, x_n) \in X^{\times n} \right\}.$$

In colloquial terms, curvature sets are just 'bags' containing all the possible distance matrices of a given size arising from points sampled from X.

For example, when $n = 2$, $\mathbf{K}_2(X)$ contains the same information as $\{d_X(x, x'); x, x' \in X\} \subset \mathbb{R}_+$. In contrast, $\mathbf{K}_3(X)$ contains all 'triangles' from X and this particular case suggest one possible justification for the name 'curvature sets'. Indeed, let X be a smooth planar curve. Consider any three points x_1, x_2 and x_3 on X close to each other. Then, if $a = \|x_2 - x_1\|$, $b = \|x_1 - x_3\|$, and $c = \|x_1 - x_2\|$, the inverse of the radius R of the circle circumscribed to the triangle $\Delta x_1 x_2 x_3$ admits an explicit expression in terms of a, b and c: $R^{-1} = \frac{4\, S(a,b,c)}{a\, b\, c}$ where $S(a, b, c)$ is the area of the triangle as given by Heron's formula.[1] The crucial observation is that R can be computed exclusively from the information contained in $\mathbf{K}_3(X)$. Now, by an argument involving a series expansion [5], as $a, b, c \to 0$ R^{-1} converges to the curvature κ of X at the point of coalescence of x_1, x_2, x_3.

Curvature sets absorb all the information that one needs in order to determine whether two compact metric spaces are isometric or not.

Theorem 3.7 ([7]) *Let $X, Y \in \mathscr{M}$. Then, X and Y are isometric if and only if $\mathbf{K}_n(X) = \mathbf{K}_n(Y)$ for all $n \in \mathbb{N}$.*

Constructions similar to curvature sets have also been considered by Peter Olver in the context of subsets of Euclidean space [16]. The Reader should compare with the metric quadruples of Chap. 2.

3.3.1.1 An Example: Curvature Sets of Spheres

We illustrate the definition with an example from [13]. Consider first the case of the standard circle S^1 endowed with the angular distance. We will exactly characterize $\mathbf{K}_3(S^1)$. For that purpose first consider any embedding of S^1 into $\mathbb{R}^2$ and observe that for any three points on S^1 exactly one the following two conditions holds: (a) there exists a line through the center of the circle such that the three points are contained on one side of the line; (b) no such line exists.

[1] $S(a, b, c) = \frac{1}{4}\big((a + b + c)(a - b + c)(a + b - c)(-a + b + c)\big)^{1/2}$.

Case (a) means that one of the three distances defined by the three points must forcibly be equal to the sum of the other two distances. Case (b) implies that the sum of the three distances is exactly 2π. Also note that, by symmetry, case (a) unrolls into three different cases depending on the identity of the distance that is equal to the sum of the other two. Each of these four situations gives a *linear* relation between the three distances! Thus, we obtain that $\mathbf{K}_3(S^1)$ is isomorphic to the tetrahedron with vertices $(0,0,0)$, $(0,\pi,\pi)$, $(\pi,0,\pi)$, and $(\pi,\pi,0)$.

The case of S^2, when endowed with the standard geodesic distance, is similar and one can prove that $\mathbf{K}_3(S^2)$ is the *convex hull* of $\mathbf{K}_3(S^1)$.

3.3.2 Comparing Curvature Sets?

An interesting property of curvature sets is that they are *isometry invariants* of metric spaces which 'live' in fixed target spaces. More precisely, for any $X, Y \in \mathscr{M}$, $\mathbf{K}_n(X)$ and $\mathbf{K}_n(Y)$ are both subsets of $\mathbb{R}^{n\times n}$.

With the purpose of discriminating X and Y one may conceive of comparing $\mathbf{K}_n(X)$ and $\mathbf{K}_n(Y)$. Since they are both (compact) subsets of $\mathbb{R}^{n\times n}$ one could compute the Hausdorff distance between them. For this we first endow $\mathbb{R}^{n\times n}$ with the distance $d_{\ell\infty}(A,B) := \max_{i,j} |a_{i,j} - b_{i,j}|$ for $A = ((a_{i,j}))$ and $B = ((b_{i,j}))$ in $\mathbb{R}^{n\times n}$. Then, we compute

$$d_n(X,Y) := \frac{1}{2}\, d_{\mathrm{H}}^{\mathbb{R}^{n\times n}}\big(\mathbf{K}_n(X), \mathbf{K}_n(Y)\big),$$

and use this number as an indication of how similar X and Y are. The best possible measure of dissimilarity that this sort of idea suggests is to consider

$$d_\infty(X,Y) := \sup_{n\in\mathbb{N}} d_n(X,Y).$$

Theorem 3.7 guarantees that d_∞ defines a legitimate metric on $\mathscr{M}$ modulo isometries.

Interestingly, one has the following 'structural theorem' for the modified Gromov-Hausdorff distance in terms of curvature sets:

Theorem 3.8 ([13]) *For all $X, Y \in \mathscr{M}$, $\widehat{d}_{\mathrm{GH}}(X,Y) = d_\infty(X,Y)$.*

This theorem provides a useful path for computing estimates to the Gromov-Hausdorff distance. Furthermore, the theorem suggests a way of 'slicing' the computation/approximation of the Gromov-Hausdorff distance between finite metric spaces, since one might want to consider computing d_n for a fixed n and hope that this provides enough information for discriminating spaces within a given family. For finite spaces, the computation of d_n would incur a polynomial cost, albeit of

a high order. There are some known classes of metric spaces $\mathscr{C} \subset \mathscr{M}$ that are characterized up to isometry by $\mathbf{K}_n(\cdot)$ for some finite $n = n(\mathscr{C})$, see [13].

3.3.2.1 A Lower Bound for $d_{\mathrm{GH}}(S^1, S^2)$

Theorems 3.6 item (2) and 3.8 then guarantee that

$$d_{\mathrm{GH}}(S^1, S^2) \geq d_3(S^1, S^2) = \frac{1}{2} d_{\mathrm{H}}^{\mathbb{R}^{3\times 3}}\left(\mathbf{K}_3(S^1), \mathbf{K}_3(S^2)\right) =: \xi.$$

Since $\mathbf{K}_3(S^2)$ is the convex hull of $\mathbf{K}_3(S^1)$, $\mathbf{K}_3(S^1) \subset \mathbf{K}_3(S^2)$, and therefore,

$$\xi = \frac{1}{2} \max_{p\in\mathbf{K}_3(S^2)} \min_{q\in\mathbf{K}_3(S^1)} \|p-q\|_\infty = \min_{q\in\mathbf{K}_3(S^1)} \|g-q\|,$$

where $g = \frac{\pi}{2}(1,1,1)$ is the center of $\mathbf{K}_3(S^2)$. But now, the center $c = \frac{2\pi}{3}(1,1,1)$ of the face of $\mathbf{K}_3(S^1)$ determined by $\pi\,(0,1,1)$, $\pi\,(1,0,1)$, and $\pi\,(1,1,0)$ is at minimal ℓ^∞ distance from g so that $\xi = \frac{1}{2}|\frac{\pi}{2} - \frac{2\pi}{3}| = \frac{\pi}{12}$, and we find the lower bound $d_{\mathrm{GH}}(S^1, S^2) \geq \frac{\pi}{12}$.

3.3.3 *Asking for More*

In many practical applications, one would like to take into account "weights" attached to points in a dataset. For example, the two metric spaces with the weights below are isometric, but not isomorphic in the sense that no isometry respects the weights:

The idea is that weights represent how much we trust a given "measurement" in practical applications. This leads to considering a more general collection of datasets and, in turn, an adapted notion of equality and a compatible metric over them. This naturally leads to regarding datasets as mm-spaces and then finding a notion of distance on $\mathscr{M}^w$ compatible with isomorphism of mm-spaces.

3.4 A Metric on $\mathscr{M}^{\mathbf{w}}$

Let (X, d_X, μ_X) and (Y, d_Y, μ_Y) be two given mm-spaces. In our path to defining a distance between mm-spaces, we emulate the construction of the Gromov–Hausdorff distance and start by identifying a notion of correspondence between mm-spaces.

A probability measure μ over $X \times Y$ is called a coupling between μ_X and μ_Y if $(\pi_1)_\# \mu = \mu_X$ and $(\pi_2)_\# \mu = \mu_Y$. We denote by $\mathscr{U}(\mu_X, \mu_Y)$ the collection of all couplings between μ_X and μ_Y.

Example 3.3 When $Y = \{p\}$, $\mu_Y = \delta_p$, and thus, there is a unique coupling between X and Y: $\mathscr{U}(\mu_X, \mu_Y) = \{\mu_X \otimes \delta_p\}$.

Example 3.4 Consider for example the spaces with two points each that we depicted above. In that case, μ_X can be identified with the vector $(\frac{1}{2}, \frac{1}{2})$ and μ_Y with the vector $(\frac{3}{4}, \frac{1}{4})$. In this case, one sees that the matrix:

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & 0 \end{bmatrix}$$

induces a valid coupling.

Now, given $p \geq 1$, consider the function

$$(x, y, x', y') \mapsto \left| d_X(x, x') - d_Y(y, y') \right|^p$$

and pick any $\mu \in \mathscr{U}(\mu_X, \mu_Y)$. One then integrates this function against the measure $\mu \otimes \mu$ and infimizes over the choice of $\mu \in \mathscr{U}(\mu_X, \mu_Y)$ to define the Gromov–Wasserstein distance of order p [11] (cf. the definition of the Wasserstein distance in Chaps. 1 and 5):

$$d_{\mathrm{GW},p}(X, Y) := \frac{1}{2} \inf_{\mu} \left(\iint |d_X(x, x') - d_Y(y, y')|^p \, \mu(dx \times dy) \, \mu(dx' \times dy') \right)^{1/p}$$

Remark 3.1 This is an L^p analogue of Eq. (3.1).

Theorem 3.9 ([11]) *The Gromov–Wasserstein distance of order $p \geq 1$ defines a proper distance on the collection of isomorphism classes of mm-spaces.*

By standard compactness arguments, one can prove that the infimum above is always attained [11]. Let $\mathscr{U}_p^{\mathrm{opt}}(X, Y)$ denote the set of all the couplings in $\mathscr{U}(\mu_X, \mu_Y)$ that achieve the minimum. The structure of the former set depends not only on μ_X and μ_Y, but also on d_X, d_Y and p.

Example 3.5 Consider the mm-space with exactly one point: $(\{*\}, (0), \delta_*)$. Then,

$$d_{\mathrm{GW},p}(X, \{*\}) = \frac{1}{2} \left(\iint \left(d_X(x, x') \right)^p \mu_X(dx) \, \mu_X(dx') \right)^{1/p}$$

and we define $\mathbf{diam}_p(X)$—the p-statistical diameter of X—as twice the right-hand side. Notice that $\lim_{p\to\infty} \mathbf{diam}_p(X)$ is equal to the usual diameter of X (as a metric space).

Question 3.1 To what extent are we able to replicate the nice properties of $(\mathscr{M}, d_{\mathrm{GH}})$ in the context of $(\mathscr{M}^w, d_{\mathrm{GW},p})$? In particular, it is of interest to investigate whether this new space of datasets is complete and whether one can easily identify rich pre-compact classes.

Remark 3.2 A different version of the Gromov–Wasserstein distance was considered by Sturm in [21]. In his construction, in analogy with (3.1), Sturm embeds mm-spaces X and Y into a common third metric space Z via isometric embeddings ϕ_X and ϕ_Y. Then, one computes the Wasserstein distance between the push-forward measures $(\phi_X)_\# \mu_X$ and $(\phi_Y)_\# \mu_Y$ in Z. Finally, one infimizes the resulting quantity over all possible choices of ϕ_X, ϕ_Y, and Z. It is important to remark that Sturm's version and the one discussed here *do not agree* in general. See [11] for details.

3.4.1 Pre-compactness

Theorem 3.10 ([11]) *For a non-decreasing function* $\rho : [0,\infty) \to [0,1]$, *such that* $\rho(\varepsilon) > 0$ *for* $\varepsilon > 0$ *and* $D > 0$, *let* $\mathscr{F}^w(\rho, D) \subset \mathscr{M}^w$ *denote the set of all mm-spaces* X, *such that* $\operatorname{diam} X \leq D$ *and* $\inf_x \mu_X\big(B_\varepsilon(x)\big) \geq \rho(\varepsilon)$ *for all* $\varepsilon > 0$. *Then,* $\mathscr{F}^w(\rho, D)$ *is pre-compact for the Gromov–Wasserstein topology, for any* $p \geq 1$.

Remark 3.3 Recall Example 3.2, where closed n-dimensional Riemannian manifolds were regarded as metric spaces. One can, all the same, embed closed Riemannian manifolds into $\mathscr{M}^w$ via $(X, g_X) \mapsto (X, d_X, \mu_X)$, where d_X is the geodesic distance induced by the metric tensor g_X and μ_X stands for the normalized volume measure on X. It is well known [20] that for $\varepsilon > 0$ small, $\mu_X\big(B_\varepsilon(x)\big) = \frac{c_n}{\mathrm{vol}(X)}\varepsilon^n\big(1 - \frac{s_X(x)}{6(n+1)}\varepsilon^2 + O(\varepsilon^4)\big)$, where $s_X(x)$ is the scalar curvature of X at x, and $\mathrm{vol}(X)$ is the total volume of X. Thus, a lower bound on $\mu_X(B_\varepsilon(x))$ somehow plays the role of a proxy for an upper bound on curvature.

3.4.2 Completeness

The Space $\mathscr{M}^w$ with any p-Gromov–Wasserstein distance is not complete. Indeed, consider the following family of mm-spaces: $\Delta_n \in \mathscr{M}^w$, where Δ_n consists of $n \in \mathbb{N}$ points at distance one from each other, and all with weights $1/n$.

Claim For all $n, m \geq 1$, $d_{\mathrm{GW},p}(\Delta_n, \Delta_m) \leq \frac{1}{2}\left(n^{-1/p} + m^{-1/p}\right)$.

The claim will follow from the following claim and triangle inequality for $d_{\mathrm{GW},p}$:

Claim For all $n, m \geq 1$, $d_{\mathrm{GW},p}(\Delta_n, \Delta_{n\cdot m}) \leq \frac{1}{2}\, n^{-1/p}$.
In order to verify the claim, we denote by $\{x_1, x_2, \ldots, x_n\}$ the points of Δ_n and label the points in $\Delta_{n\cdot m}$ by $\{y_{11}, \ldots, y_{1m}, y_{21}, \ldots, y_{2m}, \ldots\ldots, y_{n1}, \ldots, y_{nm}\}$. Consider the following coupling between μ_n and $\mu_{n\cdot m}$, the reference measures on Δ_n and $\Delta_{n\cdot m}$:

$$\mu(x_i, y_{kj}) := \frac{1}{n \cdot m}\delta_{ik}, \text{ for all } i, k \in \{1, \ldots, n\} \text{ and } j \in \{1, \ldots, m\}$$

It is clear that this defines a valid coupling between μ_n and $\mu_{n\cdot m}$.
Now, note that

$$\begin{aligned} J(\mu) :=& \sum_{i,i'} \sum_{(k,j),(k',j')} |d_{\Delta_n}(x_i, x_{i'}) - d_{\Delta_{n\cdot m}}(y_{kj}, y_{k'j'})|^p \mu(x_i, y_{kj})\ \mu(x_{i'}, y_{k'j'}) \\ =& \frac{1}{(n \cdot m)^2} \sum_{i,i'} \sum_{j,j'} |d_{\Delta_n}(x_i, x_{i'}) - d_{\Delta_{n\cdot m}}(y_{ij}, y_{i'j'})|^p \\ =& \frac{1}{(n \cdot m)^2} \sum_{i} \sum_{j,j'} |1 - \delta_{jj'}|^p \\ =& \frac{m-1}{n \cdot m} \\ \leq& n^{-1} \end{aligned}$$

Now, by definition, $d_{\mathrm{GW},p}(\Delta_m, \Delta_{n\cdot m}) \leq \frac{1}{2}\big(J(\mu)\big)^{1/p}$, so the claim follows.
Claim 3.4.2 indicates that $\{\Delta_n\}_{n\in\mathbb{N}}$ constitutes a Cauchy sequence in $\mathscr{M}^w$. However, a potential limit object for this sequence will have countably infinitely many points at distance one from each other. This space is not compact, thus $d_{\mathrm{GW},p}$ is not a complete metric.

3.4.3 Other Properties: Geodesics and Alexandrov Curvature

Recently, Sturm [22] pointed out that $\mathscr{M}^w$ is a geodesic space when endowed with any $d_{\mathrm{GW},p}$, $p \geq 1$. This means that given any two spaces X_0, X_1 in $\mathscr{M}^w$, one can find a curve $[0, 1] \ni t \mapsto X_t \in \mathscr{M}^w$, such that

$$d_{\mathrm{GW},p}(X_t, X_s) = |t - s|\, d_{\mathrm{GW},p}(X_0, X_1),\ \ s, t \in [0, 1].$$

Proposition 3.2 ([22]) *For each $p \geq 1$, the space $(\mathscr{M}^w, d_{\mathrm{GW},p})$ is geodesic. Furthermore, for $p > 1$, the following curves on $\mathscr{M}^w$ define geodesics between (X_0, d_0, μ_0) and (X_1, d_1, μ_1) in $\mathscr{M}^w$:*

$$[0, 1] \ni t \mapsto (X_0 \times X_1, d_t, \mu)$$

where $d_t\big((x_0,x_1),(x_0',x_1')\big) := (1-t)\, d_0(x_0,x_0') + t\, d_1(x_1,x_1')$ *for* $(x_0,x_1),(x_0',x_1') \in X_0 \times X_1$ *and* $\mu \in \mathscr{U}_p^{\mathrm{opt}}(X,Y)$. *Furthermore, for* $p > 1$, *all geodesics are of this form.*

Sturm further proved that the completion $\overline{\mathscr{M}^w}$ of the space $\mathscr{M}^w$ with metric $d_{\mathrm{GW},2}$ satisfies:

Theorem 3.11 ([22]) *The metric space* $\big(\overline{\mathscr{M}^w}, d_{\mathrm{GW},2}\big)$ *is an Alexandrov space of curvature* ≥ 0.

Amongst the consequences of this property is the fact that one can conceive of gradient flows on the space of all mm-spaces [22]. See also Chap. 5.

3.4.4 The Metric $d_{\mathrm{GW},p}$ in Applications

Applications of the notion of Gromov-Wasserstein distance arise in shape and data analysis. In shape analysis, the main application is shape matching under invariances. Many easily computable lower bounds for the GW distance have been discussed in [9, 11]. All of them lead to solving linear programming optimization problems (for which there are polynomial time algorithms) or can be computed via explicit formulas. As an example, consider the following invariant of an mm-space (X, d_X, μ_X):

$$H_X : [0,\infty) \to [0,1],\ \ t \mapsto \mu_X \otimes \mu_X\big(\{(x,x') |\, d_X(x,x') \leq t\}\big)$$

This invariant simply encodes the distribution of pairwise distances on the dataset X, and it is defined by analogy with the so-called shape distributions that are well known in computer graphics [17]. Then, one has:

Proposition 3.3 ([9, 11]) *Let* $X, Y \in \mathscr{M}^w$ *be any two mm-spaces and* $p \geq 1$. *Then,*

$$d_{\mathrm{GW},p}(X,Y) \geq \frac{1}{2}\int_0^\infty |H_X(t) - H_Y(t)|\, dt$$

Remark 3.4 This invariant is also related to the work of Boutin and Kemper [1] and Brinkman and Olver [2].

Other lower bounds which can be computed in time polynomial with respect to the number of points in the underlying mm-spaces have been reported in [11]. As a primary example, one has that the *local shape distributions* of shapes provide a lower bound which is strictly stronger than the ones in the Proposition above. In more detail, consider for a given mm-space (X, d_X, μ_X) the invariant:

$$h_X : X \times [0,\infty) \to [0,1],\ \ (x,t) \mapsto \mu_X\Big(\overline{B_t(x)}\Big).$$

Then, for mm-spaces X and Y consider the *cost* function $c_{X,Y} : X \times Y \to \mathbb{R}^+$ given by:

$$c_{X,Y}(x, y) := \int_0^\infty \left| h_X(x, t) - h_Y(y, t) \right| dt.$$

One then has:

Proposition 3.4 ([9, 11]) *Let $X, Y \in \mathscr{M}^w$ be any two mm-spaces and $p \geq 1$. Then,*

$$d_{\mathrm{GW},p}(X, Y) \geq \frac{1}{2} \inf_{\mu} \iint c_{X,Y}(x, y)\, \mu(dx \times dy),$$

where μ ranges in $\mathscr{U}(\mu_X, \mu_Y)$.

Remark 3.5 Solving for the infimum above leads to a mass transportation problem for which there exists efficient linear programming techniques. See discussion in [6].

Remark 3.6 It is possible to define a notion of *spectral* Gromov-Wasserstein distance which operates at the level of compact Riemannian manifolds without boundaries, and is based on the comparison of heat-kernels. This notion permits inter-relating many pre-existing shape matching methods and suggests some others [12].

3.5 Discussion and Outlook

The Gromov–Hausdorff distance offers a useful language for expressing different tasks in shape and data analysis. Its origins are in the work of Gromov on synthetic geometry. For finite metric spaces, the Gromov–Hausdorff distance leads to solving NP-hard combinatorial optimization problems. Related to construction is Gromov–Wasserstein distances that operate on metric measure spaces [11, 21]. In contrast to the Gromov–Hausdorff distance, the computation of Gromov–Wasserstein distances leads to solving quadratic optimization problems on continuous variables. The space of all metric measures spaces endowed with a certain variant of the Gromov–Wasserstein distance [11] enjoys nice theoretical properties [22]. It seems of interest to develop provably correct approximations to these distances when restricted to some suitable subclasses of finite metric spaces. Other aspects of the Gromov–Wasserstein distance are discussed in [10–12, 21, 22].

References

1. Boutin, M., Kemper, G.: On reconstructing n-point configurations from the distribution of distances or areas. Adv. Appl. Math. **32**(4), 709–735 (2004)
2. Brinkman, D., Olver, P.J.: Invariant histograms. Am. Math. Mon. **119**(1), 4–24 (2012)

3. Bronstein, A.M., Bronstein, M.M., Kimmel, R.: Efficient computation of isometry-invariant distances between surfaces. SIAM J. Sci. Comput. **28**(5), 1812–1836 (2006)
4. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence (2001)
5. Calabi, E., Olver, P.J., Shakiban, C., Tannenbaum, A., Haker, S.: Differential and numerically invariant signature curves applied to object recognition. Int. J. Comput. Vis. **26**(2), 107–135 (1998)
6. Cuturi, M.: Sinkhorn distances: lightspeed computation of optimal transport. In: Advances in Neural Information Processing Systems, pp. 2292–2300 (2013)
7. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces, 3rd edn. Modern Birkhäuser Classics. Birkhäuser, Boston (2007).
8. Kalton, N.J., Ostrovskii, M.I.: Distances between Banach spaces. Forum Math. **11**(1), 17–48 (1999)
9. Mémoli, F.: On the use of Gromov-Hausdorff distances for shape comparison. In: Proceedings of the Point Based Graphics. Prague, Czech Republic (2007)
10. Mémoli, F.: Gromov-Hausdorff distances in Euclidean spaces. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. pp. 1–8. IEEE, New York (2008)
11. Mémoli, F.: Gromov-Hausdorff distances and the metric approach to object matching. Found. Comput. Math. **11**(4), 417–487 (2011)
12. Mémoli, F.: A spectral notion of Gromov–Wasserstein distance and related methods. Appl. Comput. Harmon. Anal. **30**(3), 363–401 (2011)
13. Mémoli, F.: Some properties of Gromov–Hausdorff distances. Discrete Comput. Geom. **48**(2), 416–440 (2012)
14. Mémoli, F., Sapiro, G.: Comparing point clouds. In: Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pp. 32–40. ACM, New York (2004)
15. Mémoli, F., Sapiro, G.: A theoretical and computational framework for isometry invariant recognition of point cloud data. Found. Comput. Math. **5**(3), 313–347 (2005)
16. Olver, P.J.: Joint invariant signatures. Found. Comput. Math. **1**(1), 3–68 (2001)
17. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. ACM Trans. Graph. **21**(4), 807–832 (2002)
18. Pardalos, P.M., Wolkowicz, H., et al.: Quadratic Assignment and Related Problems: DIMACS Workshop, May 20–21, 1993, vol. 16. American Mathematical Society, Providence (1994)
19. Petersen, P.: Gromov-Hausdorff convergence of metric spaces. In: Differential geometry: Riemannian geometry, Proceedings of the Symposium in Pure Mathematics, 8–18 July 1990, vol. 54, pp. 489–504. American Mathematical Society, Los Angeles, US (1993).
20. Sakai, T.: Riemannian geometry. Translations of Mathematical Monographs, vol. 149. American Mathematical Society, Providence (1996)
21. 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
22. Sturm, K.T.: The space of spaces: curvature bounds and gradient flows on the space of metric measure spaces. arXiv preprint arXiv:1208.0434 (2012)

Chapter 4
Inference of Curvature Using Tubular Neighborhoods

Frédéric Chazal, David Cohen-Steiner, André Lieutier, Quentin Mérigot, and Boris Thibert

Abstract Geometric inference deals with the problem of recovering the geometry and topology of a compact subset K of $\mathbb{R}^d$ from an approximation by a finite set P. This problem has seen several important developments in the previous decade. Many of the proposed constructions share a common feature: they estimate the geometry of the underlying compact set K using offsets of P, that is r-sublevel set of the distance function to P. These offset correspond to what is called tubular neighborhoods in differential geometry. First and second-order geometric quantities are encoded in the tube K^r around a manifold. For instance, the classical tube formula asserts that it is possible to estimate the curvature of a compact smooth submanifold K from the volume of its offsets. One can hope that if the finite set P is close to K in the Hausdorff sense, some of this geometric information remains in the offsets of P. In this chapter, we will see how this idea can be used to infer generalized notions of curvature such as Federer's curvature measures.

F. Chazal
Inria Saclay, Palaiseau, France
e-mail: Frederic.Chazal@inria.fr

D. Cohen-Steiner
Inria Sophia-Antipolis, Sophia-Antipolis, France
e-mail: David.Cohen-Steiner@inria.fr

A. Lieutier
Dassault Systèmes, Aix-en-Provence, France
e-mail: Andre.Lieutier@3ds.com

Q. Mérigot (✉)
Laboratoire de Mathématiques d'Orsay, University of Paris-Sud, CNRS, Université Paris-Saclay, Orsay, France
e-mail: Quentin.Merigot@math.u-psud.fr

B. Thibert
Laboratoire Jean Kuntzmann, Université Grenoble-Alpes, CNRS, Saint-Martin-d'Hères, France
e-mail: Boris.Thibert@imag.fr

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_4

4.1 Introduction

Geometric inference deals with the problem of recovering the geometry and topology of a compact subset K of $\mathbb{R}^d$ from an approximation by a finite set P. This problem has seen several important developments in the previous decade. Many of the proposed constructions share a common feature: they estimate the geometry of the underlying compact set K using *offsets* of P,

$$P^r := \bigcup_{p \in P} \mathrm{B}(p, r), \tag{4.1}$$

which can also be seen as the r-sublevel set of the distance function to P. These offset correspond to what is called *tubular neighborhoods* in differential geometry. First and second-order geometric quantities are encoded in the tube K^r around a manifold. For instance, the classical tube formula asserts that it is possible to estimate the curvature of a compact smooth submanifold K from the volume of its offsets. One can hope that if the finite set P is close to K in the Hausdorff sense, some of this geometric information remains in the offsets of P. In this chapter, we will see how this idea can be used to infer generalized notions of curvature such as Federer's curvature measures. This approach based on tubular neighborhoods also lies at the root of Chap. 8, and, more remotely, Chap. 9, albeit in a more specific setting.

Notation The space $\mathbb{R}^d$ is equipped with the canonical dot product $\langle .|. \rangle$ and the induced norm $\|.\|$

4.2 Distance Function and Sets with Positive Reach

Throughout this chapter, K will denote a compact set in the Euclidean d-space $\mathbb{R}^d$, usually with no additional regularity assumption. The *distance function* to K, denoted d_K, is defined by

$$\mathrm{d}_K : x \in \mathbb{R}^d \mapsto \min_{p \in K} \|x - p\| \tag{4.2}$$

In this short section, we review some regularity properties of the distance function to a compact set, which we will use to establish stability results for generalized notions of curvature. We also introduce the class of sets with positive *reach* and sets with positive μ*-reach*, for which it is possible to define and infer generalized notions of curvature.

Definition 4.1 (Offset) The *r-offset of* K, also called tubular neighborhood in geometry, is the set K^r of points at distance at most r of K, or equivalently the sublevel set $K^r := \{x \in \mathbb{R}^d; \mathrm{d}_K(x) \leq r\}$.

Definition 4.2 (Hausdorff Distance) The Hausdorff distance between two compact subsets K and P of $\mathbb{R}^d$ can be defined in term of offsets:

$$\mathrm{d_H}(K,P) := \min\{r \geq 0 \text{ s.t. } K \subseteq P^r \text{ and } P \subseteq K^r\} \tag{4.3}$$

Loosely speaking, a finite set P is within Hausdorff distance r from a compact set K if it is sampled close to K ($P \subseteq K^r$) and densely in K ($K \subseteq P^r$). An alternative characterization of the Hausdorff distance is given by the following equality, where $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} \|f(x)|$.

$$\mathrm{d_H}(K,K') := \|\mathrm{d}_K - \mathrm{d}_{K'}\|_\infty \tag{4.4}$$

4.2.1 Gradient of the Distance and Sets with Positive Reach

4.2.1.1 Projection Function, Medial Axis and Gradient

The semi-concavity of the squared distance function to a compact set has been remarked and used in different contexts [3, 14, 17, 21]. We will use the fact that for any compact subset $K \subseteq \mathbb{R}^d$, the square of the distance function to K is 1-semiconcave.

Definition 4.3 A function ϕ defined on a subset Ω of $\mathbb{R}^d$ is λ-concave if and only if the function $\phi - \lambda \|.\|^2$ is concave.

It is easy to see that the distance function to a compact set is 1-Lipschitz. By Rademacher's theorem, this implies that this function is differentiable almost everywhere. The next proposition shows that the squared distance function to a compact set has the same regularity as a concave (or convex) function. In particular, Alexandrov's theorem implies that distance functions to compact set are twice differentiable almost everywhere.

Definition 4.4 (Projection Function and Medial Axis) A point p of K that realizes the minimum in the definition (4.2) of the distance function $\mathrm{d}_K(x)$ is called a *projection* of x on K. The set of such projections is denoted $\mathrm{proj}_K(x)$, and is always non-empty by compactness of K. The *medial axis* of K, denoted $\mathrm{Med}(K)$ is the set of points x in $\mathbb{R}^d$ that have more than one projection on K. On the complement of the medial axis, points have a single projection on K, allowing one to define a map $\mathrm{p}_K : \mathbb{R}^d \setminus \mathrm{Med}(K) \to K$ called the *projection function* on K.

Proposition 4.1 *The squared distance function to a compact subset K of $\mathbb{R}^d$ is 1-semiconcave and differentiable on $\mathbb{R}^d \setminus \mathrm{Med}(K)$. For every $x \notin \mathrm{Med}(K) \cup K$, one has*

$$\nabla_x \mathrm{d}_K^2 = 2(x - \mathrm{p}_K(x)) \qquad \nabla_x \mathrm{d}_K = \frac{x - \mathrm{p}_K(x)}{\|x - \mathrm{p}_K(x)\|} \tag{4.5}$$

Proof The function $\mathrm{d}_K(.)^2 - \|.\|^2$ is a minimum of linear functions:

$$\begin{aligned}\mathrm{d}_K(x)^2 - \|x\|^2 &= \min_{p\in P} \|x-p\|^2 - \|x\|^2 \\ &= \min_{p\in P} \|p\|^2 - 2\langle x|p\rangle\end{aligned}$$

and is therefore concave. A concave function is differentiable almost everywhere, and one has $\nabla_x[\mathrm{d}_K(.)^2 - \|.\|^2] = -2\mathrm{p}_K(x)$ at points of differentiability. This implies the desired formulas. □

4.2.1.2 Sets with Positive Reach

In his seminal article on curvature measures [12], Federer introduced the class of sets with positive reach, a class which generalizes both convex subsets and compact smooth submanifolds of $\mathbb{R}^d$.

Definition 4.5 (Reach) The *reach* of a compact set K, denoted by $\mathrm{reach}(K)$ is the minimum distance between K and its medial axis.

Example 4.1 It is well known that the projection to a closed convex set $K \subseteq \mathbb{R}^d$ is uniquely defined on the whole space, so that $\mathrm{reach}(K) = +\infty$. The reciprocal of this statement is a theorem of Motzkin [20]: if $\mathrm{reach}(K) = +\infty$, then K is convex.

Example 4.2 The tubular neighborhood theorem implies that a smooth compact submanifold of $\mathbb{R}^d$ always has positive reach. In the case of a smooth compact hypersurface, this follows from the proof of Proposition 4.3. In addition, the reach of submanifold M is always less than the minimum radius of curvature of M.

Example 4.3 Note however, that the reach is a global quantity, and cannot be lower bounded by any function of the minimum curvature radius. For instance, consider a compact set consisting of two spheres of radius R at distance ε. Then, the reach of the union of those two sphere is $\frac{\varepsilon}{2}$ while the minimum curvature radius remains constant and equal to R. It is also possible to construct similar examples involving connected manifolds.

The definition of curvature measures of sets with positive reach relies on the fact that the boundary of small tubular neighborhoods around those sets are hypersurface of class $\mathscr{C}^{1,1}$, that is $\mathscr{C}^1$ hypersurface with a Lipschitz normal vector field (see [12, Theorem 4.8]).

Proposition 4.2 ([12]) *Let $K \subseteq \mathbb{R}^d$ be a set with positive reach. Then, for any r in $(0, \mathrm{reach}(K))$ the restriction of the projection function to K^r is Lipschitz. In particular, the level set $\partial K^r = \mathrm{d}_K^{-1}(\{r\})$ is a $\mathscr{C}^{1,1}$ hypersurface.*

4.2.2 Generalized Gradient and Sets with Positive μ-Reach

The reach of a compact set a very unstable quantity. For instance, the reach of a triangulation is always zero, whereas smooth surfaces, which have positive reach, can be approximated arbitrarily well by triangulations. In this section, we will see how to define a relaxed notion of reach using a generalized gradient of the distance function.

4.2.2.1 Generalized Gradient

The distance function to a compact set is differentiable everywhere but on the medial axis. However, the semiconcavity property allows one to define a generalized gradient of the distance function everywhere. This generalized gradient coincides with the usual gradient of the distance function when it is differentiable.[1]

Definition 4.6 Let K be a compact subset of $\mathbb{R}^d$ and let x be a point in $\mathbb{R}^d \setminus K$. We define the generalized gradient of the distance function to K at x by:

$$\nabla \mathrm{d}_K(x) = \frac{x - \widetilde{p}_K(x)}{\mathrm{d}_K(x)}, \tag{4.6}$$

where $\widetilde{p}_K(x)$ is the center of the (uniquely defined) smallest ball containing the set of projections $\mathrm{proj}_K(x)$.

In particular, the norm $\|\nabla \mathrm{d}_K(x)\|$ equals one if and only if $\mathrm{proj}_K(x)$ is contained in a ball with zero radius, i.e. if it is a singleton. In other words,

$$\mathrm{Med}(K) = \{x \in \mathbb{R}^d \setminus K; \|\nabla \mathrm{d}_K(x)\| < 1\}. \tag{4.7}$$

4.2.2.2 Sets with Positive μ-Reach

The μ-reach is a relaxed version of the reach, which had originally been introduced because of its better stability properties with respect to perturbations of the compact set [6]. By Eq. (4.7), the reach of a compact set K is equal to the maximal radius r such that $\|\nabla \mathrm{d}_K(x)\| = 1$ on the offset K^r:

$$\mathrm{reach}(K) = \sup\{r \geq 0; \forall x \in K^r, \|\nabla \mathrm{d}_K(x)\| = 1\}. \tag{4.8}$$

This remark suggests a parameterized notion of reach, called the μ-reach.

[1] This generalized gradient coincides with the orthogonal projection of the origin on the supdifferential of the distance function [5, Lemma 5.2].

Definition 4.7 (μ-Reach) Let $\mu \in (0,1)$. The μ-reach of a compact set K is the maximal radius r such that $\|\nabla \mathrm{d}_K(x)\| \geq \mu$ on the offset K^r:

$$\mathrm{reach}_\mu(K) = \sup\{r \geq 0;\ \forall x \in K^r,\ \|\nabla \mathrm{d}_K(x)\| \geq \mu\}. \tag{4.9}$$

With $\mu = 1$, we recover the notion of reach introduced earlier.
In addition to smooth manifolds and convex sets, the class of compact sets with positive μ-reach also contains triangulations and non-convex polyhedra. Offsets of sets with positive μ-reach are not smooth in the sense of Proposition 4.2, but they still possess some regularity properties.

4.2.2.3 Offsets of Compact Sets with Positive μ-Reach

Let $K \subset \mathbb{R}^d$ be a compact set with positive μ-reach. A theorem of Fu [14, Corollary 3.4] implies that for any radius r in $(0, \mathrm{reach}_\mu(K))$, the closure of the complement of the tubular neighborhood K^r has positive reach. This lower bound was made quantitative in [4].

Theorem 4.1 ([4]) *Let $K \subseteq \mathbb{R}^d$ be a set with positive μ-reach. Then, for any radius r in $(0, \mathrm{reach}_\mu(K))$, one has*

$$\mathrm{reach}\left(\overline{\mathbb{R}^d \setminus K^r}\right) \geq \mu r. \tag{4.10}$$

From Proposition 4.2, this implies that offsets of $\overline{\mathbb{R}^d \setminus K^r}$ are of class $\mathscr{C}^{1,1}$. These so-called *double offsets* are used in computer aided design to smoothen a surface. The (r,t)-double offset of K is the set of points that are at distance t of the complementary of K^r:

$$K^{r,t} := \overline{\mathbb{R}^d \setminus K^r}^t = \left\{x \in \mathbb{R}^d;\ \mathrm{d}\left(x, \overline{\mathbb{R}^d \setminus K^r}\right) \leq t\right\} \tag{4.11}$$

The following theorem is thus a direct consequence [4].

Theorem 4.2 (Double Offset Theorem) *Let $K \subseteq \mathbb{R}^d$ be a set with positive μ-reach. Then, for any radius r in $(0, \mathrm{reach}_\mu(K))$ and every $t < \mu r$ the hypersurface $\partial K^{r,t}$ is $C^{1,1}$-smooth. In addition,*

$$\mathrm{reach}(\partial K^{r,t}) \geq \min(t, \mu r - t). \tag{4.12}$$

This implies in particular that the smallest of the principal radii of curvature at any point of $\partial K^{r,t}$ is at least $\min(t, \mu r - t)$.

4.3 Boundary Measures and Federer's Curvature Measures

In this section, we introduce Federer's curvature measures, which apply to a large class of compact subsets embedded in the Euclidean space. The main objective here is to prove a stability theorem for these curvature measures which implies that it is possible to estimate them from point cloud data. As a first step, we consider a simpler notion called *boundary measures* and introduced in [8] for the purpose of geometric inference.

4.3.1 Boundary Measures

Loosely speaking, the boundary measure associated to a surface encodes the distribution of normals to the surface *at a certain scale*, and can be used to detect sharp edges, or highly curved areas, where the concentration of normals is much higher.

Definition 4.8 (Boundary Measures) If K is a compact subset and E a domain of $\mathbb{R}^d$, the *boundary measure* $\mu_{K,E}$ is defined as follows: for any subset $B \subseteq \mathbb{R}^d$, $\mu_{K,E}(B)$ is the d-volume of the set of points of E whose projection on K is in B, i.e.

$$\mu_{K,E}(B) = \mathscr{H}^d(p_K^{-1}(B \cap K) \cap E). \tag{4.13}$$

Here and in the following, $\mathscr{H}^d$ denotes the d-dimensional Hausdorff measure. By construction, the total mass of this measure is equal to $\mathscr{H}^d(E)$. The construction of the boundary measure is illustrated on Fig. 4.1.

Example 4.4 Let S be a unit-length segment in the plane with endpoints a and b. The set S^r is the union of a rectangle of dimension $1 \times 2r$ whose points project on the segment and two half-disks of radius r whose points are projected on a and b. It follows that

$$\mu_{S,S^r} = 2r\, \mathscr{H}^1\big|_S + \frac{\pi}{2} r^2 \delta_a + \frac{\pi}{2} r^2 \delta_b \tag{4.14}$$

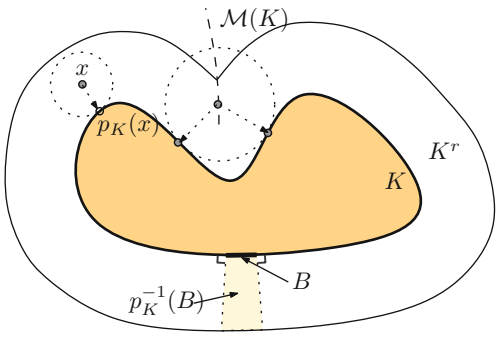


Fig. 4.1 Boundary measure of $K \subset \mathbb{R}^d$. The medial axis $\mathrm{Med}(K)$ of K is the *dashed line*. Remark that the boundary of the offset ∂K^r is smooth everywhere but at its point of intersection with the medial axis

Example 4.5 If P is a convex solid polyhedron of $\mathbb{R}^3$, F its faces, E its edges and V its vertices, then one can see that:

$$\mu_{P,P^r} = \mathscr{H}^3\big|_P + r\sum_{f\in F}\mathscr{H}^2\big|_f + r^2\sum_{e\in E}H(e)\,\mathscr{H}^1\big|_e + r^3\sum_{v\in V}K(v)\delta_v \tag{4.15}$$

where $H(e)$ is the angle between the normals of the faces adjacent to the edge e and $K(v)$ the solid angle formed by the normals of the faces adjacent to v. As shown by Steiner and Minkowski, for general convex polyhedra the measure μ_{K,K^r} is a sum of weighted Hausdorff measures supported on the i-skeleton of K, and whose local density is the local external dihedral angle.

4.3.2 Stability of Boundary Measures

In this section, we suppose that E is a fixed open set with rectifiable boundary, and we obtain a quantitative stability theorem for the map $K \mapsto \mu_{K,E}$. What we mean by *stable* is that if the Hausdorff distance between two compact sets K and P is small, then the *bounded-Lipschitz* distance between the boundary measures $\mu_{K,E}$ and $\mu_{P,E}$ is also small.

Definition 4.9 (Bounded-Lipschitz Distance) The *bounded-Lipschitz distance* between two signed measures μ, ν with finite total mass is

$$\mathrm{d_{bL}}(\mu,\nu) = \sup_f \left| \int f\mathrm{d}\mu - \int f\mathrm{d}\nu \right|, \tag{4.16}$$

where the supremum is over all 1-Lipschitz function in $\mathbb{R}^d$ s.t. $\|f\|_\infty \leq 1$.

Theorem 4.3 *If E is a fixed open subset of $\mathbb{R}^d$ with rectifiable boundary, for each compact $K \subseteq \mathbb{R}^d$, then*

$$\mathrm{d_{bL}}(\mu_{K,E},\mu_{K',E}) \leq \mathrm{const}(E,K,d)\mathrm{d_H}(K,K')^{1/2} \tag{4.17}$$

Moreover the constant only depends on the diameter of K.

Before proving this theorem, we will first show that the exponent $1/2$ in the right-hand side of Eq. (4.17) is optimal.

Lemma 4.1 *There exists a sequence of compact subsets K_n of $\mathbb{R}^d$ converging to a compact set K and a domain E, such that*

$$\mathrm{d_{bL}}(\mu_{K,E},\mu_{K_n,E}) \geq \mathrm{const}\cdot\mathrm{d_H}(K,K_n)^{1/2} \tag{4.18}$$

Proof Let K be the closed unit disk in the plane, K_n be the regular polygon with n edges inscribed in K, and let E be the annulus $\mathrm{B}(0,2) \setminus \mathrm{B}(0,1)$. Denote ℓ the

edge length of K_n. Pythagoras theorem can be used to bound the Hausdorff distance between K and K_n in term of ℓ_n: $\mathrm{d_H}(K, K_n) \leq \mathrm{const} \cdot \ell_n^2$. The measure $\mu = \mu_{K,E}$ is proportional to the uniform (lineic) measure on the unit circle. On the other hand, the map p_{K_n} projects a constant fraction of the mass of E onto the vertices V_n of K_n. The cost of spreading out the mass concentrated on these vertices to get a uniform measure on the circle is proportional to the distance between consecutive vertices, so that $\mathrm{d_{bL}}(\mu_{D,E}, \mu_{P_\ell,E}) \geq \mathrm{const} \cdot \ell_n$. □

The next lemma shows that Theorem 4.3 can be deduced from a L^1 stability result for projection functions.

Lemma 4.2 $\mathrm{d_{bL}}(\mu_{K,E}, \mu_{L,E}) \leq \|\mathrm{p}_K - \mathrm{p}_L\|_{\mathrm{L}^1(E)}$

Proof Consider a 1-Lipschitz function f on $\mathbb{R}^d$. Then, by the change of variable formula, and using the Lipschitz property,

$$\begin{aligned}\int_{\mathbb{R}^d} f(p)\mathrm{d}(\mu_{K,E}(p) - \mu_{L,E}(p)) &= \int_E f(\mathrm{p}_K(x)) - f(\mathrm{p}_L(x))\mathrm{d}\mathscr{H}^d(x) \\ &\leq \int_E \|\mathrm{p}_K(x) - \mathrm{p}_L(x)\|\, \mathrm{d}\mathscr{H}^d(x) = \|p_K - \mathrm{p}_K\|_{\mathrm{L}^1(E)}\end{aligned}$$

Taking the maximum over 1-Lipschitz functions bounded by 1 gives the desired bound. □

Proposition 4.1 allows us to rewrite the projection function p_K as the gradient of a convex function: setting $\psi_K(x) := \frac{1}{2}(\|x\|^2 - \mathrm{d}_K(x)^2)$, one has $\mathrm{p}_K = \nabla\psi_K$. This rewriting recasts a difficult geometric question into a seemingly easier analytical question, namely a L^1-stability of gradients of convex functions. This is the object of the next theorem. The proof presented here is different from the original proof in [8] and gives a slightly better constant.

Theorem 4.4 ([8]) *Let $\phi, \psi : \mathbb{R}^d \to \mathbb{R}$ be two convex functions and E be a bounded domain with rectifiable boundary. Then,*

$$\|\nabla\phi - \nabla\psi\|^2_{\mathrm{L}^2(E)} \leq 2\,\|\phi - \psi\|_{\mathrm{L}^\infty(E)}\,(\|\nabla\phi\|_{\mathrm{L}^\infty(E)} + \|\nabla\psi\|_{\mathrm{L}^\infty(E)})\mathscr{H}^{d-1}(\partial E) \tag{4.19}$$

Proof Note that if the inequality (4.19) holds for smooth convex functions, then it also holds for any convex function. Indeed, it suffices to remark that any convex function ϕ can be approximated by a sequence of smooth convex functions (ϕ_n) such that $\nabla\phi_n$ converges uniformly to $\nabla\phi$ on any compact domain. By another approximation argument, it is possible to replace E by a domain with smooth boundary. From now on, we will assume that ϕ, ψ and ∂E are smooth. We use Stokes theorem to get:

$$\int_E \|\nabla\phi - \nabla\psi\|^2 = \int_{\partial E} (\phi - \psi)\,\langle\nabla\phi - \nabla\psi|\mathbf{n}_E\rangle - \int_E (\phi - \psi)\Delta(\phi - \psi) \tag{4.20}$$

The first term of this sum is easily bounded, using the fact that an integral is bounded by the maximum of the integrand multiplied by the measure of the domain, and the Cauchy-Schwarz inequality:

$$\left|\int_{\partial E}(\phi-\psi)\,\langle\nabla\phi-\nabla\psi|\mathbf{n}_E\rangle\right| \leq \|\phi-\psi\|_{\mathrm{L}^\infty(E)}\,(\|\nabla\phi\|_{\mathrm{L}^\infty(E)}+\|\nabla\psi\|_{\mathrm{L}^\infty(E)})\mathscr{H}^{d-1}(\partial E) \tag{4.21}$$

We can bound the second term similarly:

$$\left|\int_E(\phi-\psi)\Delta(\phi-\psi)\right| \leq \|\phi-\psi\|_{\mathrm{L}^\infty(E)}\int_E(|\Delta\phi|+|\Delta\psi|) \tag{4.22}$$

We now use the convexity hypothesis, which implies that $\Delta\phi$ is non-negative, which allows us to apply Stokes theorem again:

$$\int_E|\Delta\phi| = \int_E\Delta\phi = \int_{\partial E}\langle\nabla\phi|\mathbf{n}_E\rangle \leq \|\nabla\phi\|_{\mathrm{L}^\infty(E)}\,\mathscr{H}^{d-1}(\partial E) \tag{4.23}$$

Combining Equations (4.22)–(4.23), we get

$$\left|\int_E(\phi-\psi)\Delta(\phi-\psi)\right| \leq \|\phi-\psi\|_{\mathrm{L}^\infty(E)}\,(\|\nabla\phi\|_{\mathrm{L}^\infty(E)}+\|\nabla\psi\|_{\mathrm{L}^\infty(E)})\mathscr{H}^{d-1}(\partial E) \tag{4.24}$$

Finally, Equations (4.20), (4.21) and (4.24) give the desired inequality (4.19). □

Proof (Proof of Theorem 4.3) We introduce the functions $\psi_K(x) := \frac{1}{2}(\|x\|^2 - \mathrm{d}_K(x)^2)$, which is convex, and we define $\psi_{K'}$ similarly. Now, thanks to Lemma 4.2 and using the Cauchy-Schwarz inequality, we have

$$\mathrm{d_{bL}}(\mu_{K,E},\mu_{K',E}) \leq \|\mathrm{p}_K-\mathrm{p}_{K'}\|_{\mathrm{L}^1(E)} \leq \mathscr{H}^d(E)\,\|\mathrm{p}_K-\mathrm{p}_{K'}\|_{\mathrm{L}^2(E)}\,. \tag{4.25}$$

We are now ready to apply Theorem 4.4. Without loss of generality, we assume that the Hausdorff distance between K and K' is bounded by $\mathrm{diam}(K)$ and that K contains the origin. Using $\|\mathrm{p}_K\|_\infty = \max_{p\in K}\|p\| \leq \mathrm{diam}(K)$,

$$\|\nabla\psi_K\|_{\mathrm{L}^\infty(E)}+\|\nabla\psi_{K'}\|_{\mathrm{L}^\infty(E)} \leq \|\mathrm{p}_K\|_\infty+\|\mathrm{p}_{K'}\|_\infty \leq 3\,\mathrm{diam}(K). \tag{4.26}$$

Finally, we need to control the quantity $\|\phi-\psi\|_{\mathrm{L}^\infty(E)}$. For this purpose, we use the relation $\|\mathrm{d}_K-\mathrm{d}_{K'}\|_\infty = \mathrm{d_H}(K,K')$:

$$\begin{aligned}\|\phi-\psi\|_{\mathrm{L}^\infty(E)} &= \left\|\mathrm{d}_K^2-\mathrm{d}_{K'}^2\right\|_{\mathrm{L}^\infty(E)}\\ &\leq (\|\mathrm{d}_K\|_{\mathrm{L}^\infty(E)}+\|\mathrm{d}_{K'}\|_{\mathrm{L}^\infty(E)})\mathrm{d_H}(K,K')\\ &\leq \mathrm{const}(K,E)\mathrm{d_H}(K,K'),\end{aligned} \tag{4.27}$$

where the last inequality follows from the assumption that $\mathrm{d_H}(K, K') \leq \mathrm{diam}(K)$, which implies that $\|\mathrm{d}_{K'}\|_{\mathrm{L}^\infty(E)} \leq \|\mathrm{d}_K\|_{\mathrm{L}^\infty(E)} + \mathrm{diam}(K)$. The stability inequality (4.17) then follows from (4.25), Theorem 4.4 and Eqs. (4.26)–(4.27). □

4.3.3 Tube Formulas and Federer's Curvature Measures

We start with the tube formulas of Steiner, Minkowski and Weyl, before turning to the more precise tube formula of Federer, which is then used to define the curvature measures of a large class of compact sets.

4.3.3.1 Tube Formulas

A tube formula for a compact set K in Euclidean space asserts that the Lebesgue volume of the tubular neighborhoods K^r is a degree d polynomial in r on a certain interval. The first tube formula is due to Steiner and shows that if K is a convex polygon in the Euclidean plane, the function $r \mapsto \mathscr{H}^d(K^r)$ is a polynomial of degree two, and more precisely,

$$\mathscr{H}^d(K^r) = \mathscr{H}^2(K) + r\mathscr{H}^1(\partial K) + \pi r^2 \tag{4.28}$$

The proof of this fact is (almost) contained in Fig. 4.2: every vertex with exterior angle α_i contributes a volume of $\alpha_i r^2$ to K^r, while every segment contributes $r \times \ell_j$. Summing these up on every segment and vertex yields the $2D$ Steiner formula. Minkowski proved a similar polynomial behavior for the volume of the offsets of any compact convex set in $\mathbb{R}^d$.

Weyl [24] proved that the polynomial behavior for $r \mapsto \mathscr{H}^d(K^r)$ also holds for small values of r when K is a compact smooth submanifold of $\mathbb{R}^d$. He also proved that the coefficients of this polynomial can be computed from the second

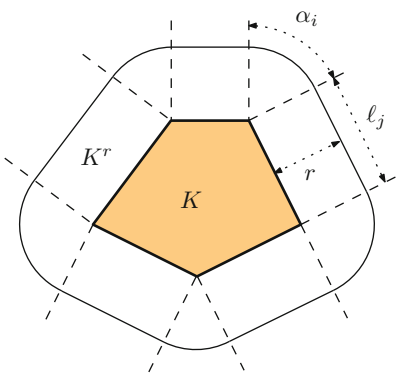


Fig. 4.2 Offset of a polygon in the Euclidean plane

fundamental form of K. The following proposition deals with the (simple) case of an hypersurface bounding a compact domain.

Proposition 4.3 *Let K be a bounded domain of $\mathbb{R}^d$ with smooth boundary M. Then, for sufficiently small values of $r > 0$,*

$$\mathscr{H}^d(K^r) = \mathscr{H}^d(K) + \sum_{k=1}^{d-1} \text{const}(d,k) r^{k+1} \int_M \left[\sum_{i_1 < \cdots < i_k} \kappa_{i_1} \ldots \kappa_{i_k}(p) \right] \mathrm{d}p \tag{4.29}$$

where $\kappa_1(p), \ldots, \kappa_d(p)$ are the principal curvatures at point p of $\partial K = M$.

Proof Let $\mathbf{n}$ be an oriented normal field on M. The map $\Phi : M \times \mathbb{R} \to \mathbb{R}^d$, $(p,t) \mapsto p + t\mathbf{n}$ is locally injective; by compactness of M, it is also injective on $M \times [0,r]$ for r small enough. One has $\mathrm{d}_{(p,t)}\Phi = \mathrm{id}_{\mathrm{T}_pM} + t\mathrm{d}_p\mathbf{n} + \mathbf{n}$, i.e. $\left|\det(\mathrm{d}_{(p,t)}\Phi)\right| = \left|\det(\mathrm{id} + t\mathrm{d}_p\mathbf{n})\right|$. For $t = 0$, $\det(\mathrm{d}_{(p,t)}\Phi) = 1 > 0$ at any point $p \in M$; as a consequence, and by compactness of M again, this determinant remains positive for small enough values of t. This allows us to apply a following change-of-variable formula for small values of r:

$$\begin{aligned} \mathscr{H}^d(K^r) &= \mathscr{H}^d(K) + \int_{K^r \setminus K} 1 \mathrm{d}x \\ &= \mathscr{H}^d(K) + \int_M \int_0^r \det(\mathrm{id} + t\mathrm{d}_p\mathbf{n}) \mathrm{d}t \mathrm{d}\mathbf{n} \end{aligned} \tag{4.30}$$

The eigenvalues of the map $\mathrm{d}_p\mathbf{n}$ are the d principal curvatures of M at p, which means:

$$\det(\mathrm{id} + t\mathrm{d}_p\mathbf{n}) = \prod_{i=1}^{d-1}(1 + t\kappa_i(p)) = \sum_{k=1}^{d} t^k \left[\sum_{i_1 < \cdots < i_k} \kappa_{i_1}(p) \ldots \kappa_{i_k}(p) \right] \tag{4.31}$$

We conclude the proof by putting Eq. (4.31) in Eq. (4.30). □

4.3.3.2 Federer's Curvature Measures

The contribution of Federer to the theory of tube formulas is twofold. First, Federer defines the class of compact sets with positive reach, which includes both compact convex sets and compact smooth submanifolds of $\mathbb{R}^d$. The *reach* of a compact subset of $\mathbb{R}^d$ is an interesting quantity, because it gives a lower bound on the largest R such that the map $r \mapsto \mathscr{H}^d(K^r)$ is a polynomial of degree d on $[0,R]$. Second, Federer associates to any set with positive reach a family of $d+1$ curvature measures, which allow one to recover local curvature information. The construction of these curvature measures rely on a local version of the tube formula [12].

As mentioned in Sect. 4.2.1.1, sets with positive reach generalize both convex sets and compact smooth submanifolds. In order to introduce the *curvature measures* of a set K with positive reach, Federer first proves that for any positive r in $(0, \text{reach}(K)]$, the boundary of the offset ∂K^r is a hypersurface of class $\mathscr{C}^{1,1}$, i.e. ∂K^r is a $\mathscr{C}^1$ hypersurface with a Lipschitz normal field. Federer's then extends the tube formula presented in Proposition 4.3 to this less smooth setting. Finally, the existence of curvature measures for K, as well as the polynomial behavior for the volume of the offsets is obtained by approximation, by letting r go to zero. We will not prove these facts, but only quote Federer's existence result. As usual, ω_k denotes the volume of the k-dimensional sphere in $\mathbb{R}^{k+1}$.

Theorem 4.5 (Federer) *For any compact set $K \subseteq \mathbb{R}^d$ with positive reach, there exists $d+1$ signed measures $\Phi_{K,0}, \ldots, \Phi_{K,d}^d$ such that for $r \leq \text{reach}(K)$,*

$$\mu_{K,K^r} = \sum_{i=0}^{d} \omega_{d-i} \Phi_{K,i} r^i. \tag{4.32}$$

Definition 4.10 The measures $\Phi_{K,0}, \ldots, \Phi_{K,d}$ introduced in Theorem 4.5 are called *Federer's curvature measures* of K.

4.3.4 Stability of Federer's Curvature Measures

The purpose of this section is to show that Federer's curvature measures of a compact set with positive reach can be estimated from a Hausdorff approximation of this set, without any regularity hypothesis on the approximating set. In order to do so, we explain how to associate an ersatz of Federer's curvature measures to any compact set, which coincide with the original curvature measures when the set has a sufficiently large reach.

Definition 4.11 (Approximate Curvature Measures) Let $R > 0$ and r be a family of numbers $(r_i)_{0 \leq i \leq d}$ such that $0 < r_0 < \ldots < r_d$. Given any compact set K in $\mathbb{R}^d$ and any Borel set B, we define the *approximate curvatures* $(\Phi_{K,i}^{(r)}(B))_{0 \leq i \leq d}$ as the set of coefficients which satisfy the interpolation equations:

$$\forall 0 \leq i \leq d-1,\ \mu_{K,K^{r_i}}(B) = \sum_{j=0}^{d} \omega_{d-j} \Phi_{K,j}^{(r)}(B) r_i^j \tag{4.33}$$

Since the numbers (r_i) are distinct, this define the approximate curvature measures uniquely. More precisely, by Lagrange interpolation, there exist real coefficients (L_{ij}) such that

$$\forall 0 \leq i \leq d-1,\ \Phi_{K,j}^{(r)} = \sum_{0 \leq i \leq d} L_{ij} \mu_{K,K^{r_i}}. \tag{4.34}$$

This shows that the approximate curvature measure $\Phi_{K,j}^{(r)}$ is a signed measure. Moreover, one recovers Federer's curvature measures when the reach of K is sufficiently large. More precisely, if $r_d \leq \text{reach}(K)$, then $\Phi_{K,j}^{(r)} = \Phi_{K,j}$.

We are now able to state the following stability theorem from [8].

Theorem 4.6 *Given any compact set $K \subseteq \mathbb{R}^d$ and (r) as in Definition 4.11, there exist a constant $C = \text{const}(K, r, d)$ such that for any compact set K' sufficiently close to K,*

$$\mathrm{d_{bL}}\left(\Phi_{K,j}^{(r)}, \Phi_{K',j}^{(r)}\right) \leq C\mathrm{d_H}(K, K')^{1/2}$$

Moreover, if $r_d \leq \text{reach}(K)$ then one can estimate Federer's curvature measures of K from K':

$$\mathrm{d_{bL}}\left(\Phi_{K,j}^{(r)}, \Phi_{K',j}\right) \leq C\mathrm{d_H}(K, K')^{1/2}.$$

We will see first how Theorem 4.6 can be deduced from the following stability result for boundary measures.

Theorem 4.7 *Given any compact set $K \subseteq \mathbb{R}^d$ and r a positive number, there exist a constant $C = \text{const}(K, r, d)$ such that for any compact set K' sufficiently close to K,*

$$\mathrm{d_{bL}}\left(\mu_{K,K^r}, \mu_{K',K'^r}\right) \leq C\mathrm{d_H}(K, K')^{1/2}.$$

Proof (Proof of Theorem 4.7) Let E be the symmetric difference between K^r and K'^r. Then, by the triangle inequality for the bounded-Lipschitz distance,

$$\mathrm{d_{bL}}\left(\mu_{K,K^r}, \mu_{K',K'^r}\right) \leq \mathrm{d_{bL}}\left(\mu_{K,E}, \mu_{K',E}\right) + \mathrm{d_{bL}}\left(\mu_{K,K^r}, \mu_{K,E}\right) + \mathrm{d_{bL}}\left(\mu_{K',K'^r}, \mu_{K',E}\right)$$

Theorem 4.3 implies that the first term of the right-hand side is of order $\mathrm{O}(\mathrm{d_H}^{1/2}(K, K'))$. We only need to deal with the two last terms. Given any 1-Lipschitz function f with $\|f\|_\infty \leq 1$, the change-of-variable formula implies

$$\begin{aligned}\int_{\mathbb{R}^d} f(x)\mathrm{d}(\mu_{K,K^r} - \mu_{K,E}) &\leq \int_{K^r} f(\mathrm{p}_K(x))\mathrm{d}\mathscr{H}^d(x) - \int_E f(\mathrm{p}_K(x))\mathrm{d}\mathscr{H}^d(x) \\ &\leq \mathscr{H}^d(K^r \setminus E)\end{aligned}$$

Taking the maximum over such functions gives a bound on the bounded-Lipschitz distance between these measures. Overall, we have

$$\mathrm{d_{bL}}\left(\mu_{K,K^r}, \mu_{K',K'^r}\right) \leq C\mathrm{d_H}^{1/2}(K, K') + \mathscr{H}^d(K^r \setminus E) + \mathscr{H}^d(K'^r \setminus E).$$

Defining $\varepsilon = \mathrm{d_H}(K, K')$, one has the inclusion $K^r \setminus E \subseteq K^r \setminus K^{r-\varepsilon}$. This term can be shown to be of order $\mathrm{O}(\mathrm{d_H}(K, K'))$ using Proposition 4.2 from [8] and the coarea formula. □

Proof (Proof of Theorem 4.6) Thanks to Eq. (4.34), one has for any 1-Lipschitz function f such that $\|f\|_\infty \leq 1$,

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f \mathrm{d}(\Phi_{K,j}^{(r)} - \Phi_{K',j}^{(r)})\right| &\leq \sum_{0\leq i\leq d} \left|L_{ij}\right| \left|\int_{\mathbb{R}^d} f \mathrm{d}(\mu_{K,K^{r_i}} - \mu_{K,K^{r_i}})\right| \\ &\leq \sum_{0\leq i\leq d} \left|L_{ij}\right| \mathrm{d_{bL}}(\mu_{K,K^{r_i}}, \mu_{K,K^{r_i}}) \\ &\leq \mathrm{const}(K, (r), d)\mathrm{d_H}(K, K')^{1/2},\end{aligned}$$

where the second inequality follows from the definition of the bounded-Lipschitz distance, and the third inequality comes from Theorem 4.7. To conclude, it suffices to apply the definition of the bounded-Lipschitz distance once again. □

4.3.5 Computation of Boundary Measures and Visualization

We explain briefly how to compute the boundary measure of a point cloud, that is a finite subset P of $\mathbb{R}^d$. The Federer's curvature measures can be recovered from the boundary measures through polynomial fitting. The computation relies on the *Voronoi diagram* of P, which is a practical way of encoding the distance function to P in low dimension.

Definition 4.12 (Voronoi Diagram) Let P be a point cloud of $\mathbb{R}^d$. The *Voronoi diagram* of P is a decomposition of the space into convex polyhedra called *Voronoi cells*. The Voronoi cell of the point p in P is defined by:

$$\mathrm{Vor}_P(p) = \{x \in \mathbb{R}^d; \forall q \in P, \ \|x - p\| \leq \|x - q\|\}. \tag{4.35}$$

Lemma 4.3 *Let P be a point cloud of $\mathbb{R}^d$. The boundary measure of P with respect to a domain E is given by*

$$\mu_{P,E} = \sum_{p\in P} \mathscr{H}^d(\mathrm{Vor}_P(p) \cap E)\delta_p. \tag{4.36}$$

In addition, when $E = P^r$ is an offset of P, one has

$$\mu_{P,P^r} = \sum_{p\in P} \mathscr{H}^d(\mathrm{Vor}_P(p) \cap \mathrm{B}(p, r)). \tag{4.37}$$

Proof Equation (4.36) follows from the fact that a point x belongs to the Voronoi cell of p in P if and only if $\mathrm{d}_P(x) = p$ and iff $\mathrm{p}_P(x) = p$. To prove the second equation, we only need to remark that

$$\begin{aligned}\mathrm{Vor}_P(p) \cap P^r &= \{x \in \mathbb{R}^d; \mathrm{d}_P(x) = \|x-p\| \text{ and } \mathrm{d}_P(x) \leq r\} \\ &= \{x \in \mathbb{R}^d; \mathrm{d}_P(x) = \|x-p\| \text{ and } \|x-p\| \leq r\} \\ &= \mathrm{Vor}_P(p) \cap \mathrm{B}(p,r)\end{aligned}$$

□

Thanks to this lemma, computing boundary measures amounts to evaluating the volume of intersections of Voronoi cells with balls. This leads to a practical algorithm in dimension 2 and 3 [19]. This approach becomes too costly in higher dimensions due to the exponential complexity of the Voronoi diagram, but it is possible to resort to a Monte-Carlo method [8].

4.3.5.1 Visualization

It is not trivial how to visualize a finitely supported measure μ_{P,P^r} on $\mathbb{R}^d$, even when the ambient dimension is $d = 2, 3$. The obvious idea is to display a ball whose volume is proportional to the mass of the Dirac at each point in the support of the measure. This is however not satisfactory as two measures which are close for the bounded-Lipschitz distance could lead to very different visualizations. Indeed, consider $\mu = \delta_x$ and $\mu_\eta = (\delta_x + \delta_y)/2$ where $\|x-y\| \leq \eta$: μ would be displayed as a single ball $\mathrm{B}(x,r)$ while the nearby measure μ_η would be displayed as two overlapping balls of smaller radius $\mathrm{B}(x,\alpha r) \cup \mathrm{B}(x,\alpha r)$ with $\alpha = 2^{-1/d}$.

In order to construct a visualization which is stable with respect to the bounded-Lipschitz distance, we therefore convolve each boundary measure with a fixed bounded-Lipschitz function χ.

Definition 4.13 (Convolution) The convolution of a finite measure μ on $\mathbb{R}^d$ with a bounded measurable function χ is the function $\mu * \chi : \mathbb{R}^d \to \mathbb{R}$ defined by

$$[\mu * \chi](x) := \int_{\mathbb{R}^d} \chi(y-x)\mathrm{d}\mu(y) \tag{4.38}$$

The convolved measure $\mu * \chi$ is stable with respect to the bounded-Lipschitz distance, by its very definition. More precisely,

Lemma 4.4 *Let χ be a function on $\mathbb{R}^d$ such that* $\mathrm{Lip}\chi \leq 1$ *and* $\|\chi\|_\infty \leq 1$. *Then for finite measures μ, ν,*

$$\|\mu * \chi - \nu * \chi\|_\infty \leq \mathrm{d_{bL}}(\mu, \nu) \tag{4.39}$$

In practice, we choose the convolution kernel to be the "hat function" $\chi_\varepsilon(y) = \max(\varepsilon - \|x-y\|, 0)$, and we display at each point in P a ball whose volume is proportional to the value of the function $\mu_{P,P^r} * \chi_\varepsilon$. Figure 4.3 shows the convolved boundary measures of point clouds sampling piecewise-smooth surfaces in $\mathbb{R}^3$.

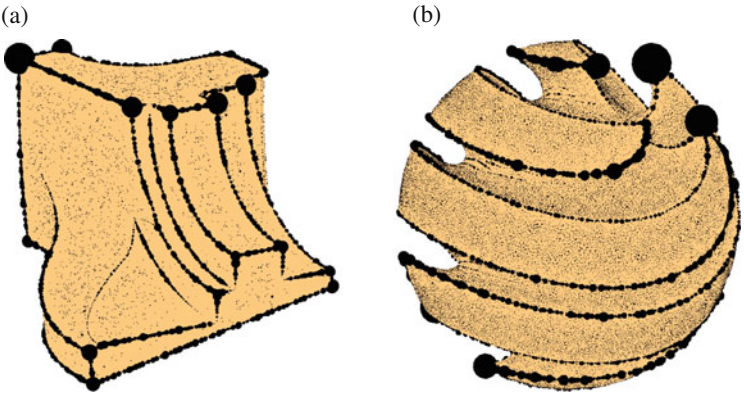


Fig. 4.3 Convolved boundary measures of 100k point clouds sampled from point clouds sampled from the fandisk and sharp sphere models and rescaled so as to have unit diameter. The offset radius is set to $r = 0.05$ and $r = 0.1$ respectively and the convolution radius is given by $\varepsilon = 0.02$ and $\varepsilon = 0.03$. (**a**) Fandisk; (**b**) Sharp sphere

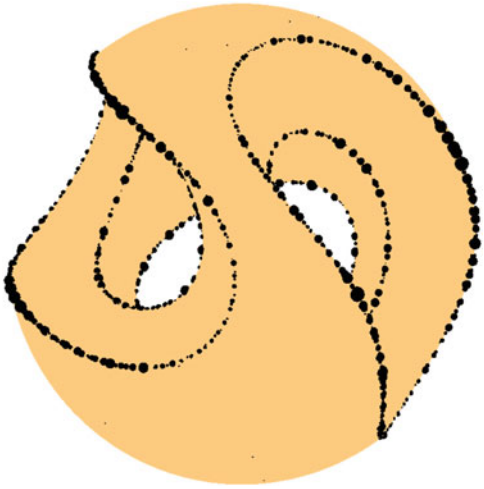


Fig. 4.4 Feature points extracted from a point cloud sampling of a CSG model by thresholding low values of the convolved boundary measure

4.3.5.2 Sharp Features Extraction

Extracting the set of sharp features of a compact sets known through a finite point cloud sampling is of interest in many geometry processing applications. Figure 4.3 suggests that the sharp corners carry more mass than the points on the sharp edges, which again carry more mass than the smooth points. This observation can be turned into more quantitative estimations, see e.g. [18, Chap. 3]. In Fig. 4.4 we display

points whose boundary measures carries more mass than some threshold, or more precisely (for the same reasons as in the previous paragraph), where the values of the function $\mu_{K,K^r} * \chi_\varepsilon$ are higher than some threshold.

4.4 Voronoi Covariance Measures and Local Minkowski Tensors

In some applications, such as feature-aware surface reconstruction, scalar quantities such as those encoded in Federer's curvature measures are not sufficient, and one also wishes to recover directional information such as principal curvature directions or the direction of a sharp edge in a piecewise smooth surface.

4.4.1 Covariance Matrices of Voronoi Cells

Voronoi-based normal estimation [1, 2] rely on the intuition that for a noiseless sampling of a smooth surface the Voronoi cells are elongated in the direction of the normal to the surface. For instance, in Fig. 4.5, the Voronoi cell of the red point on upper face is an elongated cylinder, and it is possible to estimate the normal to that face by analyzing the shape of this Voronoi cell. A practical tool for estimating the direction in which a domain is elongated is the notion of covariance matrix.

Definition 4.14 (Covariance Matrix) The *covariance matrix* of a bounded domain $V \subseteq \mathbb{R}^d$ is the symmetric matrix, or *tensor*, defined by:

$$\mathrm{cov}(V, p) = \int_V (x-p) \otimes (x-p) \mathrm{d}x$$

where $v \otimes w$ denotes the $n \times n$ matrix defined by $[v \otimes w]_{i,j} = w_i v_j$.

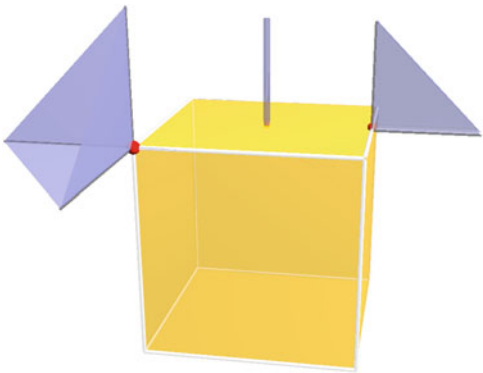


Fig. 4.5 The Voronoi cell of a point x on a cube is pencil, triangle or cone-shaped depending on the dimension of the normal cone

The covariance matrix of a domain is also known as the *inertia matrix* in solid mechanics and its eigenvectors capture the principal axes of the domain V with respect to the base point p. In the case of a noiseless sampling of a smooth surface, one can approximate the normal direction at each sample by the eigenvector corresponding to the largest eigenvalue of the covariance matrix of the corresponding Voronoi cell intersected with a bounding box of the point set [1]. Under strong noise, individual Voronoi cells can become ill shaped, but it is possible to average the covariance matrices of nearby Voronoi cells to recover the correct normal directions.

Note, however, that the shapes of Voronoi cells provide more information than just the normal direction. When the underlying surface is not smooth, some of its points will have normal cones rather than single normal directions. Nevertheless, even in this case, the shapes of Voronoi cells accurately reflect the shapes of the underlying normal cones. Some geometric properties of these normal cones can then be estimated using the covariance matrices of the Voronoi cells.

4.4.2 Voronoi Covariance Measure

It is possible to mimic the definition of the boundary measure to construct a tensor-valued measure which summarizes and extends the covariance matrices of Voronoi cells.

Definition 4.15 The *Voronoi covariance measure* (also called VCM) of a compact subset K of $\mathbb{R}^d$ with respect to a bounded domain E a tensor-valued measure denoted by $\mathscr{V}_{K,R}$. This measure maps every (Borel) subset B of $\mathbb{R}^d$ to the symmetric matrix defined by:

$$\mathscr{V}_{K,E}(B) := \int_{E \cap \mathrm{p}_K^{-1}(B \cap K)} (x - \mathrm{p}_K(x)) \otimes (x - \mathrm{p}_K(x)) \mathrm{d}x \tag{4.40}$$

Example 4.6 The Voronoi covariance measure of a point cloud $P \subseteq \mathbb{R}^d$ summarizes the covariance matrices of Voronoi cells. More precisely, the Voronoi covariance measure of P with respect to a bounded domain E is a sum of Dirac masses. The weight in front of each Dirac is the covariance matrix of the corresponding Voronoi cell. More precisely,

$$\mathscr{V}_{K,E} = \sum_{p \in P} \mathrm{cov}(\mathrm{Vor}_P(p), p) \delta_p. \tag{4.41}$$

These quantities can be computed efficiently, provided that one is able to compute (or approximate) the intersection between the Voronoi cell and E.

Definition 4.16 The bounded-Lipschitz distance between two measures μ, ν taking values in a normed vector space $(X, \|.\|)$ is defined by

$$\mathrm{d_{bL}}(\mu, \nu) = \sup_f \left\| \int f \mathrm{d}\mu - \int f \mathrm{d}\nu \right\|. \tag{4.42}$$

For tensor-valued measure, the vector space is the space of symmetric matrices and $\|.\|$ is the operator norm.

The Hausdorff stability result for boundary measures (Theorems 4.3 and 4.7) can be generalized to Voronoi covariance measures. The VCM can be used for the estimation of the location and direction of sharp features, leading to a practical and efficient algorithm [19].

Theorem 4.8 *Given any compact set $K \subseteq \mathbb{R}^d$ and r a positive number, there exist a constant $C = \mathrm{const}(K, r, d)$ such that for any compact set K' sufficiently close to K,*

$$\mathrm{d_{bL}}\left(\mathscr{V}_{K,K^r}, \mathscr{V}_{K',K'^r}\right) \leq C \mathrm{d_H}(K, K')^{1/2}.$$

4.4.2.1 Extensions

The Voronoi covariance measure is closely related to the notion of local Minkowski tensor, which was recently introduced by Hug and Schneider [15]. Theorems 4.8 and 4.6 have been extended to this setting by Hug et al. [16]. A robust variant of the Voronoi covariance measure is introduced and studied in the PhD thesis of Cuel [10].

4.5 Stability of Anisotropic Curvature Measures

In this last section, we consider the question of approximating anisotropic curvature measures of a compact set from a Hausdorff-approximation. Here, we assume that the unknown compact set K has positive μ-reach: this include smooth manifolds, convex domains and triangulations (see Sect. 4.2.2). We show that it is possible to approximate the anisotropic curvature measures of K^r from those of K'^r, where K' is a Hausdorff approximation of K and where r lies in some range. In practical applications, the second set K' is a point cloud and its offset K'^r is a finite union of balls, whose anisotropic curvature measures can be computed.

4.5.1 Anisotropic Curvature Measures of Sets with Positive Reach

Let V be a compact set with positive reach and let t in $(0, \mathrm{reach}(V))$. Since the hyper-surface ∂V^t is of class $C^{1,1}$, the second fundamental form and the principal

curvatures of ∂V^t are defined almost everywhere. One can therefore define the anisotropic curvature measure introduced in [9].

Definition 4.17 Let V be a compact set with positive reach. The anisotropic curvature measure of V associates to any Borel set B the matrix

$$\mathrm{H}_V(B) = \lim_{t\to 0} \int_{\partial V^t \cap \mathrm{p}_V^{-1}(B)} H_{\partial V^t}(p)dp,$$

where $H_{\partial V^t}$ is a matrix-valued function defined on $\mathbb{R}^d$ that coincides with the second fundamental form of ∂V^t on the tangent space, and vanishes on the orthogonal component.

Remark that this definition is coherent with the Federer curvature measures. Indeed, the kth Federer curvature measures satisfies for every Borel subset B of $\mathbb{R}^d$:

$$\Phi_{V,k}(B) = \lim_{t\to 0} \int_{\partial V^t \cap \mathrm{p}_V^{-1}(B)} s^k(p)dp,$$

where p_V is the projection onto V and s^k is the k-th elementary symmetric polynomial of the principal curvatures $\lambda_1,\ldots,\lambda_{d-1}$ of ∂V^t.

Now, let K be a compact set whose μ-reach is greater than $r > 0$. Then $V = \mathbb{R}^d \setminus K^r$ has a reach greater than μr. It is then possible to define the curvature measures of K^r [22, 23] by:

$$\Phi^k_{K^r}(B) = (-1)^k \Phi^k_V(B) \quad \text{and} \quad \mathrm{H}_{K^r}(B) = -\mathrm{H}_V(B).$$

4.5.2 Stability of the Curvature Measures of the Offsets

The following theorem states that if a compact set K is close in the Hausdorff sense to a compact set K' with positive μ-reach, then the Federer curvature measures and anisotropic curvature measure of the offsets K^r and K'^r are close for the bounded-Lipschitz distance. This result is similar but not equivalent to Theorem 4.6. The result in Theorem 4.6 is limited to the Federer curvature measures but it derives from Theorem 4.7, which holds without any assumptions on the underlying compact set, whereas Theorem 4.9 requires to assume a lower bound on the μ-reach. We recall that the bounded-lipschitz distance is given in Definition 4.16.

Theorem 4.9 *Let $r > 0$, K and K' be two compact subsets of $\mathbb{R}^d$ such that $\mathrm{reach}_\mu(K') > r$. We suppose that the Hausdorff distance $\varepsilon = d_H(K, K')$ between K and K' is such that $\varepsilon < \frac{\mu^2}{60+9\mu^2} r$. Then one has*

$$\mathrm{d_{bL}}(\Phi_{K^r,k}, \Phi_{K'^r,k}) \leq \mathrm{const}(r, \mu, d, K)\ \sqrt{\epsilon}$$

and

$$\mathrm{d_{bL}}(\mathrm{H}_{K^r}, \mathrm{H}_{K'^r}) \leq \mathrm{const}(r, \mu, d, K)\ \sqrt{\epsilon},$$

where $\mathrm{const}(r, \mu, d, K)$ *depends on r, μ, d and the diameter of K.*

This theorem implies that one can estimate locally the curvature measures of a compact set K' with positive μ-reach from an Hausdorff approximation. If $f : \mathbb{R}^d \to \mathbb{R}$ is measurable function, the previous theorem implies that

$$\left| \int f \mathrm{d}\Phi_{K^r,k} - \int f \mathrm{d}\Phi_{K'^r,k} \right| \leq \mathrm{const}(r, \mu, d, K)\ \|f\|_{\mathrm{BL}}\ \sqrt{\epsilon}, \tag{4.43}$$

where $\|f\|_{\mathrm{BL}} = \mathrm{Lip}(f) + \|f\|_\infty\|$ is the bounded-lipschitz norm of f. The similar inequality holds for the anisotropic curvature measure. In practice, we take the hat function $f(x) = \max(1 - \|x - c\|/r, 0)$ equal to 1 at a point $c \in \partial K'^r$. Then we can retrieve local information about the curvature of K'^r from the curvature of K^r in the neighborhood of c.

As illustrated in Fig. 4.6, the upper bound of Theorem 4.9 is tight. However the constant $\mathrm{const}(r, \mu, d, K)$ in Eq. (4.43) can be localized. It does not have to depend on the whole compact set K, but can only depend on the diameter of the support $\mathrm{spt}f = \{x \in \mathbb{R}^d, f(x) \neq 0\}$ of f. See [7] for more details.

4.5.3 Computation of the Curvature Measures of 3D Point Clouds

When the compact set K is a finite point set in $\mathbb{R}^3$ it is possible to provide explicit formula for the curvature measures. The boundary of K^r is a spherical polyhedron: its faces are spherical polygons; its edges are circle arcs contained in the intersection of pairs of spheres of radius r with centers in K; its vertices belong to the intersection of three spheres of radius r with centers in K. One has explicit integral formula

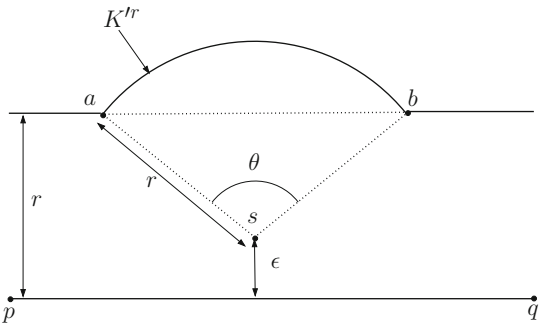


Fig. 4.6 Tightness of the bound of Theorem 4.9. We consider compact sets $K = [p, q]$ and $K' = [p, q] \cup \{s\}$, where s is at a distance ϵ from K. We have $d_H(K, K') = \epsilon$ and the total curvature θ of K'_r between a and b satisfies $\theta = 2 \arccos\left(\frac{r-\epsilon}{r}\right) = O\left(\sqrt{\epsilon}\right)$

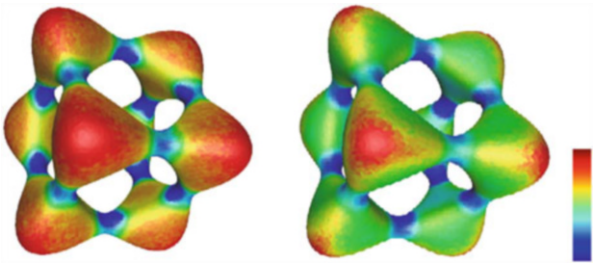


Fig. 4.7 The Gauss (*left*) and mean (*right*) curvatures computed on the offset of a point set sampled around a smooth surface. The colors are related to the values of the curvature according to the colorbar on the right, the *blue color* corresponding to the lowest values

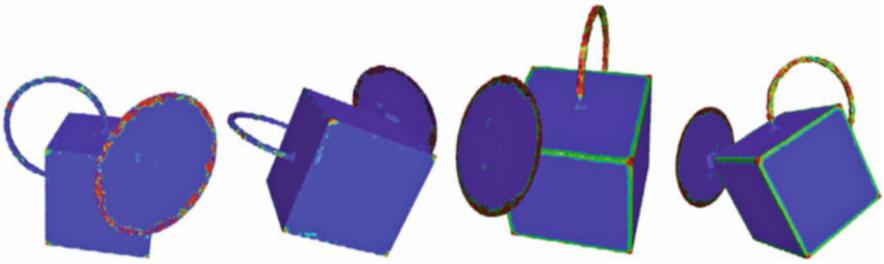


Fig. 4.8 The Gauss (*first and second image*) and mean (*third and fourth image*) curvatures computed on the offset of a point cloud sampled around a non-manifold set union of a cube with a disc and a circle. As expected, the vertices and the boundary of the disc have a large Gaussian curvature

for the curvature measures for each vertex/edge/face of the spherical polygon [7]. Moreover, the combinatorial structure of ∂K^r is in one-to-one correspondence with the boundary of the α-shape of K [11].

In Figs. 4.7 and 4.8 below, the curvatures are represented on the boundary of the α-shape (for $\alpha = r$) of the point clouds where each triangle is colored according to the curvature value of its corresponding vertex in ∂K^r and to the colorbar on the right of Fig. 4.7. Note that the color values are different for the different examples (since the extrema values are different). This algorithm can be easily adapted to calculate the anisotropic curvature measures for a finite set of points. In particular, this allows to estimate the principal curvatures and principal directions from a point set.

4.5.4 *Sketch of Proof of Theorem 4.9*

We first need to introduce the notion of normal cycle for sets with positive reach. Let V be a set with positive reach. We define the set:

$$S(V) = \{(p, n) \in \mathbb{R}^d \times \mathbb{S}^{d-1},\ p \in \partial V \text{ and } n \in \mathrm{Nor}_V(p)\},$$

where $\mathrm{Nor}_V(p) := \{v \in \mathbb{R}^d; \exists t > 0, \mathrm{p}_V(p + tv) = p\}$ is the normal cone of V at p. One can show that $S(V)$ is a Lipschitz (d-1)-manifold. The normal cycle $N(V)$ of V is then by definition the (d-1)-current on $\mathbb{R}^d \times \mathbb{R}^d$ defined for every (d-1)-differential form ω by:

$$N(V)(\omega) = \int_{S(V)} \omega.$$

The normal cycle contains in fact all the curvature information and allows to define the curvature measures [9]. For every k, the curvature measure $\Phi_{V,k}(B)$ is given by $N(V)(\bar{f}\omega_k)$ where $\bar{f}(p,n) = \mathbf{1}_B(p)$ and ω_k is a $d-1$-differential form on $\mathbb{R}^d \times \mathbb{R}^d$ that does not depend on V. Similarly, the anisotropic curvature measure $\overline{H}_V(X,Y)$ applied in the directions X and Y is given by $N(V)(\bar{f}\omega_H^{X,Y})$, where $\omega_H^{X,Y}$ is a $d-1$-differential form that depends on X and Y.

The proof can now be divided into three steps: in the first step, we show that the problem can be carried onto the double offsets (that are smooth); in a second step, we compare the normal cycles of the double offsets; in the last step, we combine Step 1 and Step 2 to show that the curvature measures of the two offsets are close. Let K and K' be two compact sets with positive μ-reach that satisfy all the assumptions of Theorem 4.9.

Step 1: Carrying the problem into the double offsets

First note that $\overline{\mathbb{R}^d \setminus K^r}$ and $\overline{\mathbb{R}^d \setminus K'^r}$ have positive reach. We introduce the map:

$$\begin{array}{rcl} F_{-t} : \mathbb{R}^d \times \mathbb{R}^d & \to & \mathbb{R}^d \times \mathbb{R}^d \\ (p,n) & \mapsto & (p - tn, n) \end{array}.$$

If V is any compact set with positive reach, the map F_{-t} induces naturally a one-to-one correspondence between the support of the normal cycle of the offset V^t and the support of the normal cycle of V. In particular, this map allows to send simultaneously the normal cycles of $K^{r,t}$ and $K'^{r,t}$ to respectively the normal cycles of K^r and K'^r. More precisely, one has:

$$N(\overline{\mathbb{R}^d \setminus K^r}) - N(\overline{\mathbb{R}^d \setminus K'^r}) = F_{-t\sharp}(N(K^{r,t}) - N(K^{r,t})),$$

where $F_{-t\sharp}$ denotes the push-forward for currents. Therefore, in order to compare the normal cycles of $\overline{\mathbb{R}^d \setminus K^r}$ and $\overline{\mathbb{R}^d \setminus K'^r}$, it is sufficient to compare the normal cycles of the double offsets $K^{r,t}$ and $K'^{r,t}$.

Step 2: Comparison of the normal cycles of the double offsets

In order to compare the normal cycles of $K^{r,t}$ and $K'^{r,t}$, we first need to compare their (geometrical) supports in $\mathbb{R}^d \times \mathbb{R}^d$. One first shows that the Hausdorff distance between $\partial K^{r,t}$ and $\partial K'^{r,t}$ is less than ϵ/μ. Using a result of [5] one also shows that the difference between the normals of $\partial K^{r,t}$ and $\partial K'^{r,t}$ is bounded by $30\sqrt{\epsilon/(\mu t)}$. Hence the (geometrical) supports of $N(K^{r,t})$ and $N(K'^{r,t})$ are close to each other. Let us take $t = \mu r/2$. Since the reach of $\partial K^{r,t}$ is larger than t, the projection map $\mathrm{p}_{\partial K^{r,t}}$

onto $\partial K^{r,t}$ is then defined on the offset $U_t := (\partial K^{r,t})^t$. Since $\partial K'^{r,t} \subset U_t$, the map $\mathrm{p}_{\partial K^{r,t}}$ induces a one-to-one map between $\partial K'^{r,t}$ and $\partial K^{r,t}$. We now define

$$\begin{array}{rcl} \psi : U_t \times \mathbb{R}^d \to & \mathrm{spt}\,(N\,(K_{r,t})) \\ (x,n) & \mapsto (\,\mathrm{p}_{\partial K^{r,t}}(x), n_{\mathrm{p}_{\partial K^{r,t}}(x)}) \end{array}.$$

Using the affine homotopy between ψ and the identity, the homotopy Lemma [13, 4.1.9, pp. 363–364] allows to show that $N(K^{r,t})$ and $N(K'^{r,t})$ are close. More precisely

$$N(K^{r,t}) - N(K'^{r,t}) = \partial R, \tag{4.44}$$

where ∂R is the boundary of a particular d-rectifiable R current whose support $\mathrm{spt}(R)$ has a d-volume bounded by $\mathscr{H}^d(\mathrm{spt}(R)) \leq k(r,\mu,d)\mathscr{H}^{d-1}(\partial K'^{r,t})\sqrt{\varepsilon}$, where $\mathscr{H}^k$ denotes the k-dimensional Hausdorff measure and $k(r,\mu,d)$ is a constant that only depends on r, μ and d.

Step 3

Here, instead of applying the measures to a Borel set, we apply them to a function $f : \mathbb{R}^d \to \mathbb{R}$ (that can be for example indicatrix of Borel sets). Let us take an differential form ω_k associated to a given curvature measure. We could also consider the form $\omega_{\mathrm{H}}^{X,Y}$ associated to the anisotropic curvature measure. By combining previous equations, one has:

$$\Phi_{\overline{\mathbb{R}^d \setminus K^r},k}(f) - \Phi_{\overline{\mathbb{R}^d \setminus K'^r},k}(f) = N(\overline{\mathbb{R}^d \setminus K^r})(\bar{f}\omega_k) - N(\overline{\mathbb{R}^d \setminus K'^r})(\bar{f}\omega_k) = F_{-t\sharp}\partial R(\bar{f}\omega_k).$$

We show that F_{-t} is $\sqrt{1+t^2}$-Lipschitz. Since ω_k and $\mathrm{d}\omega_k$ are uniformly bounded by a constant depending on the dimension, $\mathrm{Lip}(\bar{f}) = \mathrm{Lip}(f)$, one gets by Stokes theorem:

$$|\Phi_{\overline{\mathbb{R}^d \setminus K^r},k}(f) - \Phi_{\overline{\mathbb{R}^d \setminus K'^r},k}(f)| \leq k(r,\mu,d)\ \|f\|_{\mathrm{BL}}\ \mathscr{H}^{d-1}(\partial K'^{r,t})\ \sqrt{\epsilon},$$

The previous inequality still holds for K^r and K'^r. To conclude the proof, we use the bound on $\mathscr{H}^{d-1}(\partial K'^{r,t})$ [8] and also use the critical function stability result to get rid of the assumption on the μ-reach of K.

References

1. Alliez, P., Cohen-Steiner, D., Tong, Y., Desbrun, M.: Voronoi-based variational reconstruction of unoriented point sets. In: Proceedings of the Eurographics Symposium on Geometry Processing, vol. 7, pp. 39–48 (2007)
2. Amenta, N., Bern, M.: Surface reconstruction by Voronoi filtering. Discret. Comput. Geom. **22**(4), 481–504 (1999)

3. Cannarsa, P., Sinestrari, C.: Semiconcave functions, Hamilton-Jacobi equations, and optimal control. In: Progress in Nonlinear Differential Equations and their Applications, vol. 58. Birkhäuser Boston Inc., Boston (2004)
4. Chazal, F., Cohen-Steiner, D., Lieutier, A., Thibert, B.: Shape smoothing using double offsets. In: Proceedings of the 2007 ACM Symposium on Solid and Physical Modeling, pp. 183–192. ACM (2007)
5. Chazal, F., Cohen-Steiner, D., Lieutier, A.: Normal cone approximation and offset shape isotopy. Comput. Geom. **42**(6), 566–581 (2009)
6. Chazal, F., Cohen-Steiner, D., Lieutier, A.: A sampling theory for compact sets in Euclidean space. Discret. Comput. Geom. **41**(3), 461–479 (2009)
7. Chazal, F., Cohen-Steiner, D., Lieutier, A., Thibert, B.: Stability of curvature measures. Comput. Graphics Forum **28**, 1485–1496 (2009)
8. Chazal, F., Cohen-Steiner, D., Mérigot, Q.: Boundary measures for geometric inference. Found. Comput. Math. **10**(2), 221–240 (2010)
9. Cohen-Steiner, D., Morvan, J.M.: Second fundamental measure of geometric sets and local approximation of curvatures. J. Differ. Geom. **74**(3), 363–394 (2006)
10. Cuel, L., Lachaud, J.O., Mérigot, Q., Thibert, B.: Robust geometry estimation using the generalized Voronoi covariance measure. SIAM J. Imag. Sci. **8**(2), 1293–1314 (2015)
11. Edelsbrunner, H.: The union of balls and its dual shape. In: Proceedings of the Ninth Annual Symposium on Computational Geometry, pp. 218–231. ACM (1993)
12. Federer, H.: Curvature measures. Trans. Am. Math. Soc. **93**, 418–491 (1959)
13. Federer, H.: Geometric Measure Theory. Springer, Berlin (1969)
14. Fu, J.: Tubular neighborhoods in Euclidean spaces. Duke Math. J. **52**(4), 1025–1046 (1985)
15. Hug, D., Schneider, R.: Local tensor valuations. Geom. Funct. Anal. **24**(5), 1516–1564 (2014)
16. Hug, D., Kiderlen, M., Svane, A.M.: Voronoi-based estimation of Minkowski tensors from finite point samples. Discret. Comput. Geom. **57**(3), 545–570 (2017). https://link.springer.com/article/10.1007%2Fs00454-016-9851-x
17. Lieutier, A.: Any open bounded subset of $\mathbb{R}^n$ has the same homotopy type as its medial axis. Comput. Aided Geom. Des. **36**(11), 1029–1046 (2004)
18. Mérigot, Q.: Geometric structure detection in point clouds. Theses, Université Nice Sophia Antipolis (2009). https://tel.archives-ouvertes.fr/tel-00443038
19. Mérigot, Q., Ovsjanikov, M., Guibas, L.: Voronoi-based curvature and feature estimation from point clouds. IEEE Trans. Vis. Comput. Graph. **17**(6), 743–756 (2011)
20. Motzkin, T.: Sur quelques propriétés caractéristiques des ensembles convexes. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. **21**, 562–567 (1935)
21. Petrunin, A.: Semiconcave functions in Alexandrov's geometry. In: Surveys in Differential Geometry, vol. XI, pp. 137–201. Int. Press, Somerville (2007)
22. Rataj, J., Zähle, M.: Normal cycles of Lipschitz manifolds by approximation with parallel sets. Differ. Geom. Appl. **19**(1), 113–126 (2003)
23. Rataj, J., Zähle, M.: General normal cycles and Lipschitz manifolds of bounded curvature. Ann. Global Anal. Geom. **27**(2), 135–156 (2005)
24. Weyl, H.: On the volume of tubes. Am. J. Math. **61**(2), 461–472 (1939)

Chapter 5
Entropic Ricci Curvature for Discrete Spaces

Jan Maas

Abstract We give a short overview on a recently developed notion of Ricci curvature for discrete spaces. This notion relies on geodesic convexity properties of the relative entropy along geodesics in the space of probability densities, for a metric which is similar to (but different from) the 2-Wasserstein metric. The theory can be considered as a discrete counterpart to the theory of Ricci curvature for geodesic measure spaces developed by Lott–Sturm–Villani.

5.1 Ricci Curvature Lower Bounds for Geodesic Measure Spaces

In the last decade, ideas from the theory of optimal transport have led to significant progress in the analysis and geometry on non-smooth spaces. The starting point for these developments were the independent works by Lott and Villani [20] and Sturm [29], who introduced a notion of Ricci curvature for metric measure spaces based on a beautiful connection between optimal transport and entropy that originates in McCann's pioneering work [22].

Let $(\mathscr{X}, d)$ be a complete and separable metric space and let $\mathscr{P}(\mathscr{X})$ be the space of Borel probability measures on $\mathscr{X}$. For $1 \leq p < \infty$ and $\mu, \nu \in \mathscr{P}(\mathscr{X})$ we consider the *Monge-Kantorovich metric* W_p (often called Wasserstein metric or Wasserstein distance), defined by

$$W_p(\mu,\nu) := \inf_{\Gamma \in \Pi(\mu,\nu)} \left(\int_{\mathscr{X}\times\mathscr{X}} d(x,y)^p \mathrm{d}\Gamma(x,y) \right)^{1/p} \qquad \text{for } \mu, \nu \in \mathscr{P}(\mathscr{X}) \,,$$

where $\Pi(\mu,\nu)$ denotes the set of all couplings of μ and ν, i.e., all probability measures $\Gamma \in \mathscr{P}(\mathscr{X} \times \mathscr{X})$ satisfying

$$\Gamma(A \times \mathscr{X}) = \mu(A) \quad \text{and} \quad \Gamma(\mathscr{X} \times A) = \nu(A)$$

J. Maas (✉)
Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria
e-mail: jan.maas@ist.ac.at

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_5

for all Borel sets $A \subseteq \mathscr{X}$. Loosely speaking, $W_p(\mu_0, \mu_1)^p$ is the minimal transportation cost required to transport an amount of mass from its initial configuration μ to a prescribed final configuration ν, at a cost of $d(x, y)^p$ per unit. It can be shown that W_p defines a metric on $\mathscr{P}_p(\mathscr{X})$, the space of probability measures with finite p-th moment. Moreover, if $(\mathscr{X}, d)$ is a geodesic space (i.e., every pair of points $x_0, x_1 \in \mathscr{X}$ can be joined by a curve $\gamma : [0, 1] \to \mathscr{X}$ such that $d(\gamma(s), \gamma(t)) = |s - t| d(x_0, x_1)$ for all $s, t \in [0, 1]$), then the 2-Wasserstein space $(\mathscr{P}_2(\mathscr{X}), W_2)$ is a geodesic space as well.

Given a reference measure $\nu \in \mathscr{P}(\mathscr{X})$, the relative entropy with respect to ν is defined by

$$\mathrm{Ent}_\nu(\mu) := \int_{\mathscr{X}} \rho(x) \log \rho(x) \mathrm{d}\nu(x)$$

whenever $\mu \in \mathscr{P}(\mathscr{X})$ is absolutely continuous with density $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$, provided that the integral is well-defined. If ν is a probability measure, this quantity takes values in $[0, +\infty]$.

The following result, proved in [9, 27, 28], characterizes Ricci curvature lower bounds on Riemannian manifolds in terms of convexity properties of the relative entropy (with respect to the volume measure) and optimal transport.

Theorem 5.1 (Characterization of Ricci Lower Bounds on Riemannian Manifolds) *Let $\kappa \in \mathbb{R}$. For a complete Riemannian manifold $\mathscr{M}$, the following assertions are equivalent:*

1. $\mathrm{Ric} \geq \kappa$ *on* $\mathscr{M}$.
2. *Each pair of probability measures $\mu_0, \mu_1 \in \mathscr{P}_2(\mathscr{M})$ can be connected by a constant speed W_2-geodesic $(\mu_t)_{t \in [0,1]}$ along which the entropy satisfies the κ-convexity inequality*

$$\mathrm{Ent}_{\mathrm{Vol}}(\mu_t) \leq (1 - t)\, \mathrm{Ent}_{\mathrm{Vol}}(\mu_0) + t\, \mathrm{Ent}_{\mathrm{Vol}}(\mu_1) - \frac{\kappa}{2} t(1 - t) W_2(\mu_0, \mu_1)^2 .$$

While the definition of the Ricci tensor requires a Riemannian structure on the underlying space, the second condition makes sense in much greater generality: the sole requirements are a metric (to define the Wasserstein metric) and a measure (to define the relative entropy). Therefore the following definition makes sense:

Definition 5.1 (Lott–Sturm–Villani) Let $\kappa \in \mathbb{R}$. A metric measure space $(\mathscr{X}, d, \nu)$ is said to have "Ricci curvature bounded from below by κ" if every pair of probability measures $\mu_0, \mu_1 \in \mathscr{P}_2(\mathscr{X})$ can be connected by a constant speed W_2-geodesic $(\mu_t)_{t \in [0,1]}$ along which the relative entropy Ent_ν satisfies the κ-convexity inequality

$$\mathrm{Ent}_\nu(\mu_t) \leq (1 - t)\, \mathrm{Ent}_\nu(\mu_0) + t\, \mathrm{Ent}_\nu(\mu_1) - \frac{\kappa}{2} t(1 - t) W_2(\mu_0, \mu_1)^2 .$$

This definition has become the starting point for many interesting developments in the analysis and geometry on metric measure spaces. A large number of geometric and functional inequalities with sharp constants can be derived from this definition. A crucial feature of the theory is its robustness, i.e., the *stability* of lower Ricci curvature bounds under convergence of metric measure space in the sense of measured Gromov–Hausdorff convergence. A stronger version of Definition 5.1, a curvature-dimension criterion, has been introduced in [20, 29] as well. We refer to [30] for an overview of this theory. More recently, important refinements of the curvature-dimension criterion have been introduced in [2, 13].

5.1.1 Discrete Spaces

If $(\mathscr{X}, d)$ is a geodesic space, the Lott–Sturm–Villani criterion is non-trivial, since we have already mentioned that the 2-Wasserstein space $(\mathscr{P}_2(\mathscr{X}), W_2)$ is a geodesic space as well. If $(\mathscr{X}, d)$ is *discrete*, the situation turns out to be completely different. Indeed, suppose that $(\mu_t)_{t\in[0,1]}$ is W_2-Lipschitz, i.e., $W_2(\mu_s, \mu_t) \le L|s-t|$ for some $L < \infty$. Fix a point $x \in \mathscr{X}$ and set $m(t) := \mu_t(\{x\})$. Since $\mathscr{X}$ is discrete, there exists $\delta > 0$ such that all other points are at least at distance δ from x. As a consequence, $W_2(\mu_s, \mu_t) \ge \delta\sqrt{|m(t) - m(s)|}$. Combining the estimates above, we infer that $t \mapsto m(t)$ is Hölder continuous with exponent 2, hence m is constant. Since x was arbitrary, we conclude that every W_2-Lipschitz curve is constant. In particular, there are no W_2-geodesics, so that the Lott–Sturm–Villani criterion is (trivially) not satisfied for any $\kappa \in \mathbb{R}$.

While the curvature concept based on geodesic convexity of the entropy has been extremely powerful in the continuous (geodesic) setting, these observations show that the 2-Wasserstein metric is not the appropriate object for an analogous discrete theory.

To motivate the definition of a suitable discrete counterpart of W_2, we shall describe a seminal result in the continuous setting in which the Wasserstein metric plays a central role.

5.2 The Heat Flow as Gradient Flow of the Entropy

It is a classical fact that the heat equation $\partial_t u = \Delta u$ can be regarded as the gradient flow equation in $L^2(\mathbb{R}^n)$ for the Dirichlet energy $\mathscr{E} : L^2(\mathbb{R}^n) \to \mathbb{R} \cup \{+\infty\}$ given by

$$\mathscr{E}(u) = \begin{cases} \frac{1}{2}\int_{\mathbb{R}^n} |\nabla u(x)|^2 \mathrm{d}x\,, & u \in H^1(\mathbb{R}^n)\,, \\ +\infty\,, & \text{otherwise.} \end{cases}$$

A quite different gradient flow structure for the heat equation, physically very appealing, was discovered by Jordan et al. [18] at the end of the 1990s.

Theorem 5.2 (Heat Flow is Gradient Flow of the Entropy) *The heat equation* $\partial_t \rho = \Delta \rho$ *is the gradient flow equation for the Boltzmann-Shannon entropy in the* 2*-Wasserstein space* $(\mathscr{P}_2(\mathbb{R}^n), W_2)$.

Note that the meaning of this statement is not clear a priori . While gradient flows are traditionally considered in smooth (e.g. Riemannian or Hilbertian) settings, the statement of this result involves the notion of a gradient flow in a metric space, in which the notion of a gradient is not defined.

In the original paper [18], the result was made rigorous using the discrete minimizing movement scheme given by

$$\mu^{(0)} := \mu_0 \,, \qquad \mu^{(k+1)} := \operatorname{argmin}_{\sigma \in \mathscr{P}(\mathbb{R}^n)} \left(\operatorname{Ent}(\sigma) + \frac{1}{2h} W_2(\mu^{(k)}, \sigma)^2 \right)$$

The authors showed that the piecewise constant functions (μ_h) defined by

$$\mu_h(t) := \mu^{(k)} \,, \qquad t \in [kh, (k+1)h) \,,$$

converge, as $h \downarrow 0$, to the solution of the heat equation $\partial_t \mu = \Delta \mu$ with initial condition $\mu|_{t=0} = \mu_0$. It is also possible to interpret Theorem 5.2 using the theory of gradient flows in metric spaces, which has been systematically developed in [1].

Another interpretation of this result can be given in terms of a formal infinite-dimensional Riemannian structure on the space of probability measures [26]. This structure is closely related to the so-called Benamou–Brenier Formula, which provides a dynamical characterization of the 2-Wasserstein metric inspired by fluid mechanics.

5.2.1 The Benamou-Brenier Formula

Fix two probability measures $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{R}^n)$. Instead of considering couplings as in the Monge–Kantorovich problem, we consider continuous-time interpolations $(\mu_t)_{t \in [0,1]}$ between μ_0 and μ_1. Under mild regularity conditions, such interpolations satisfy the continuity equation (in the sense of distributions)

$$\partial_t \mu_t + \nabla \cdot (\mu_t v_t) = 0 \tag{5.1}$$

for a suitable time-dependent velocity vector field $v : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$. The remarkable Benamou–Brenier formula [4] asserts that the squared Wasserstein distance is obtained by minimizing the total kinetic energy among all solutions to the continuity equation. More precisely,

$$W_2(\mu_0, \mu_1)^2 = \inf_{(\mu_t, v_t)_t} \left\{ \int_0^1 \int_{\mathbb{R}^n} |v_t(x)|^2 \, \mathrm{d}\mu_t(x) \mathrm{d}t \right\} \,, \tag{5.2}$$

where the infimum runs over all sufficiently regular solutions $(\mu_t, v_t)_{t\in[0,1]}$ to (5.1) with boundary conditions $\mu|_{t=0} = \mu_0$ and $\mu|_{t=1} = \mu_1$.

Given a sufficiently regular curve of probability measures $(\mu_t)_{t\in(-\epsilon,\epsilon)}$, there are many different velocity vector fields v that satisfy the continuity equation (5.1) at $t = 0$. However, it turns out that only one of those vector fields is a *generalized gradient*, in the sense that it belongs to the closure of $\{\nabla\psi : \psi \in C_c^\infty(\mathbb{R}^n)\}$ in $L^2(\mathbb{R}^n, \mu_0; \mathbb{R}^n)$. This vector field minimizes the kinetic energy $\int_{\mathbb{R}^n} |v(x)|^2 \mathrm{d}\mu_0(x)$ among all vector fields v satisfying the continuity equation at $t = 0$, see, e.g., [1, Sect. 8] for details.

It was realized by Otto [26] that the Benamou–Brenier formula can be regarded as the Riemannian distance formula associated to a formal infinite-dimensional Riemannian structure on $\mathscr{P}_2(\mathbb{R}^n)$, which can be described as follows:

- Given a smooth curve $(\mu_t)_{t\in(-\epsilon,\epsilon)}$ in $\mathscr{P}_2(\mathbb{R}^n)$ with $\mu_0 = \mu$, let $v = \nabla\psi$ be the unique (generalized) gradient vector field satisfying the continuity equation $\partial_t\mu_t + \nabla\cdot(\mu\nabla\psi) = 0$ at time 0. The velocity vector field $\nabla\psi$ is regarded as the tangent vector at $\mu \in \mathscr{P}_2(\mathbb{R}^n)$ associated to the curve $(\mu_t)_{t\in(-\epsilon,\epsilon)}$. Thus, by means of the continuity equation, the tangent space at μ is identified with the space of generalized gradients.
- Taking this identification into account, the tangent space at $\mu \in \mathscr{P}(\mathbb{R}^n)$ will be endowed with the $L^2(\mu)$-scalar product:

$$\langle\nabla\psi_1, \nabla\psi_2\rangle_{T_\mu} := \int_{\mathbb{R}^n} \langle\nabla\psi_1(x), \nabla\psi_2(x)\rangle \mathrm{d}\mu(x) \,.$$

It is then clear that the Benamou–Brenier formula is precisely the formula for the Riemann distance induced by this formal infinite-dimensional Riemannian structure.

5.3 A Gradient Flow Structure for Reversible Markov Chains

We shall now describe a (standard) discrete framework in which an analogue of Theorem 5.2 has been obtained.

Let $\mathscr{L} : \mathbb{R}^{\mathscr{X}} \to \mathbb{R}^{\mathscr{X}}$ be the generator of a continuous time Markov chain on a finite set $\mathscr{X}$. Thus $\mathscr{L}$ is given by

$$(\mathscr{L}\psi)(x) := \sum_{y\in\mathscr{X}} Q(x,y)(\psi(y) - \psi(x)) \,,$$

where $Q(x, y) \geq 0$ denotes the transition rate from x to y. We shall use the convention that $Q(x, x) = 0$ for all $x \in \mathscr{X}$. Assuming that the Markov chain is irreducible, there

exists a unique invariant probability measure π on $\mathscr{X}$, which means that

$$(\mathscr{L}^*\pi)(x) := \sum_{y\in\mathscr{X}} \pi(y)Q(y,x) - \pi(x)Q(x,y) = 0\,, \qquad x \in \mathscr{X}\,.$$

We shall assume in addition that π is reversible, i.e., the detailed balance equations

$$\pi(x)Q(x,y) = \pi(y)Q(y,x)$$

hold for all $x, y \in \mathscr{X}$. This assumption implies that $\mathscr{L}$ is selfadjoint as an operator on $L^2(\mathscr{X},\pi)$. The associated Markov semigroup $(e^{t\mathscr{L}})_{t\geq 0}$ may be regarded as a discrete analogue of the heat semigroup. Note that the set of probability densities

$$\mathscr{D}(\mathscr{X}) := \Big\{\rho : \mathscr{X} \to \mathbb{R}_+ \mid \sum_{x\in\mathscr{X}} \rho(x)\pi(x) = 1\Big\}$$

is invariant under the action of the semigroup $(e^{t\mathscr{L}})_{t\geq 0}$.

Let $\mathscr{H} : \mathscr{D}(\mathscr{X}) \to \mathbb{R}$ be the relative entropy functional given by

$$\mathscr{H}(\rho) := \sum_{x\in\mathscr{X}} \rho(x)\log\rho(x)\;\pi(x)\,.$$

It is straightforward to check that $\mathscr{H}$ decreases along solutions to the discrete "heat equation" $\partial_t\rho = \mathscr{L}\rho$. In view of Theorem 5.2 one might wonder whether this equation has a gradient flow structure for the relative entropy with respect to the 2-Wasserstein metric (for a suitable metric on $\mathscr{X}$). However, the discussion at the end of Sect. 5.1 demonstrates that this is not the case: since the 2-Wasserstein space does not contain any Lipschitz curves, there aren't any gradient flows at all!

It turns out to be possible to construct a different metric on $\mathscr{P}(\mathscr{X})$, which allows one to prove a discrete analogue of the JKO-Theorem. The construction is inspired by the Benamou–Brenier formula (5.2).

5.3.1 Discrete Transport Metrics

In order to define suitable discrete transport metrics, it will be necessary to state a discrete version of the continuity equation (5.1). For this purpose, let $\mathscr{E} := \{(x,y) : Q(x,y) > 0\}$ be the set of edges in the incidence graph induced by Q, and set $w(x,y) := Q(x,y)\pi(x)$. The *discrete gradient* is given by

$$\nabla : L^2(\mathscr{X},\pi) \to L^2(\mathscr{E},w)\,, \qquad \nabla\psi(x,y) := \psi(y) - \psi(x)\,.$$

This operator maps functions (defined on $\mathscr{X}$) to vector fields (defined on $\mathscr{E}$). The negative of its adjoint is the *discrete divergence* given by

$$\nabla\cdot : L^2(\mathscr{E}, w) \to L^2(\mathscr{X}, \pi) , \qquad (\nabla \cdot V)(x) := \frac{1}{2}\sum_{y\in\mathscr{X}} Q(x,y)(V(x,y) - V(y,x)) .$$

Note however that a problem arises if one attempts to formulate a discrete analogue of the continuity equation (5.1). Namely, the divergence term contains the product of the density ρ (defined on $\mathscr{X}$) and the velocity vector field V (defined on $\mathscr{E}$). Since these objects are defined on different spaces, there is no canonical way to multiply them, hence there is no canonical discretization of the continuity equation.

In fact, there is additional freedom that we need to exploit: given a probability density $\rho \in \mathscr{D}(\mathscr{X})$, we introduce the quantity $\hat{\rho} : \mathscr{E} \to \mathbb{R}$ by

$$\hat{\rho}(x,y) := \theta(\rho(x), \rho(y)) , \tag{5.3}$$

where $\theta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is a symmetric continuous function, smooth and strictly positive on $(0,\infty)\times(0,\infty)$, which will be carefully chosen below. In many applications, $\theta(r,s)$ will be a suitable mean of r and s. Having introduced this quantity, the natural discrete analogue of the continuity equation (5.1) is given by $\partial_t\rho + \nabla\cdot(\hat{\rho}V) = 0$, or more explicitly,

$$\partial_t\rho_t(x) + \frac{1}{2}\sum_{y\in\mathscr{X}} \big(V_t(x,y) - V_t(y,x)\big)\widehat{\rho_t}(x,y)Q(x,y) = 0 \tag{5.4}$$

The following definition is now a natural discrete analogue of the Benamou–Brenier formula.

Definition 5.2 (Discrete Transport Metric) For $\rho_0, \rho_1 \in \mathscr{D}(\mathscr{X})$ we define

$$\mathscr{W}(\rho_0,\rho_1)^2 = \inf_{\rho,V}\left\{\frac{1}{2}\int_0^1 \sum_{x,y\in\mathscr{X}} |V_t(x,y)|^2\widehat{\rho_t}(x,y)Q(x,y)\pi(x)\mathrm{d}t\right\} , \tag{5.5}$$

where the infimum runs over all pairs $(\rho_t, V_t)_{t\in[0,1]}$ solving the continuity equation (5.4) with boundary conditions $\rho|_{t=0} = \rho_0$ and $\rho|_{t=1} = \rho_1$.

Note that the definition depends on the Markov triple $(\mathscr{X}, Q, \pi)$ as well as on the choice of the function θ. As in the continuous setting, it is possible to show that one may restrict the infimum to vector fields $(V_t)_{t\in[0,1]}$ of gradient type, i.e., one may assume that $V_t(x,y) = \nabla\psi_t(x,y)$ for some function $\psi_t : \mathscr{X} \to \mathbb{R}$. One can show that $\mathscr{W}(\rho_0,\rho_1) < \infty$ if ρ_0, ρ_1 are everywhere strictly positive. If ρ_0 or ρ_1 vanishes somewhere, then $\mathscr{W}(\rho_0,\rho_1)$ may be finite or infinite, depending on the choice of θ.

5.3.2 A Riemannian Structure on the Space of Probability Measures

The distance $\mathscr{W}$ is induced by a Riemannian structure on $\mathscr{P}_+(\mathscr{X})$, the space of strictly positive probability densities on $\mathscr{X}$. This structure is a natural discrete analogue of the one described in Sect. 5.2.1. Indeed, given a smooth curve $(\rho_t)_{t\in(-\epsilon,\epsilon)}$ in $\mathscr{P}(\mathscr{X})S$ with $\rho_0 = \rho$, there exists a unique discrete gradient $\nabla\psi$ satisfying the discrete continuity equation

$$\partial_t \rho_t(x) + \sum_{y\in\mathscr{X}} \big(\psi_t(y) - \psi_t(x)\big)\widehat{\rho}_t(x,y)Q(x,y) = 0 \tag{5.6}$$

at time 0. As before, we shall regard $\nabla\psi$ as the tangent vector at ρ associated to the curve $(\rho_t)_{t\in(-\epsilon,\epsilon)}$. We thus identify the tangent space at ρ with the set of discrete gradients. Under this identification, we define a scalar product on the tangent space at $\rho \in \mathscr{P}(\mathbb{R}^n)$ by

$$\langle \nabla\psi_1, \nabla\psi_2\rangle_{T_\rho} := \frac{1}{2}\sum_{x,y\in\mathscr{X}} \big(\psi_1(x) - \psi_1(y)\big)\big(\psi_2(x) - \psi_2(y)\big)\hat{\rho}(x,y)Q(x,y)\pi(x)\ .$$

The induced Riemannian distance coincides with (5.5), since it is not hard to show that the minimizing vector field v in (5.5) is a discrete gradient for every $t \in (0,1)$.

5.3.3 The Discrete JKO-Theorem

We are now in a position to obtain an analogue of Theorem 5.2 in the discrete setting. Let us first compute the gradient of the relative entropy $\mathscr{H}$ in the Riemannian structure described above. For this purpose, take a smooth curve $(\rho_t)_t$ in $\mathscr{P}_+(\mathscr{X})$ satisfying the discrete continuity equation (5.6). Using the detailed balance assumption, it follows that

$$\begin{aligned}
\partial_t \mathscr{H}(\rho_t) &= \sum_{x\in\mathscr{X}} \big(1 + \log\rho_t(x)\big)\partial_t\rho_t(x)\pi(x) \\
&= -\sum_{x,y\in\mathscr{X}} \big(1 + \log\rho_t(x)\big)\big(\psi_t(y) - \psi_t(x)\big)\widehat{\rho}_t(x,y)Q(x,y)\pi(x) \\
&= \frac{1}{2}\sum_{x,y\in\mathscr{X}} \big(\log\rho_t(x) - \log\rho_t(y)\big)\big(\psi_t(x) - \psi_t(y)\big)\widehat{\rho}_t(x,y)Q(x,y)\pi(x) \\
&= \langle \nabla\log\rho_t, \nabla\psi_t\rangle_{T_{\rho_t}}
\end{aligned}$$

On the other hand, by the definition of the gradient in the Riemannian structure inducing $\mathscr{W}$, we have $\partial_t \mathscr{H}(\rho_t) = \langle \mathrm{grad}_{\mathscr{W}} \mathscr{H}(\rho_t), \nabla\psi_t\rangle_{T_{\rho_t}}$. Since the computation above holds for any curve $(\rho_t)_t$ (hence for any vector field $\nabla\psi_t$), we infer that

$$\mathrm{grad}_{\mathscr{W}} \mathscr{H}(\rho) = \nabla \log \rho \tag{5.7}$$

for $\rho \in \mathscr{P}(\mathscr{X})S$. It follows from this identity that the gradient flow equation for $\mathscr{H}$ (in the Riemannian structure associated to $\mathscr{W}$) is given by the discrete continuity equation (5.6) together with the additional equation $\nabla\psi_t = -\nabla \log \rho_t$. More explicitly, we arrive at

$$\partial_t \rho_t(x) - \sum_{y\in\mathscr{X}} \big(\log \rho_t(y) - \log \rho_t(x)\big)\widehat{\rho}_t(x,y)Q(x,y) = 0\ .$$

Of course, this equation depends on the choice of θ in (5.3), since the Riemannian metric depends on θ. Note however that this equation reduces to the discrete heat equation $\partial_t \rho = \mathscr{L}\rho$ if we choose θ in such a way that $(\log \rho(y) - \log \rho(x))\hat{\rho}(x,y) = \rho(y) - \rho(x)$. In other words, if θ is the *logarithmic mean* defined by

$$\theta_{\log}(r,s) := \frac{r-s}{\log r - \log s} = \int_0^1 r^{1-p}s^p \mathrm{d}p\ ,$$

we obtain the following discrete JKO-Theorem:

Theorem 5.3 (Discrete JKO) *The gradient flow equation for the relative entropy $\mathscr{H}$ with respect to $\mathscr{W}$ is given by the discrete heat equation $\partial_t \rho = \mathscr{L}\rho$, provided $\theta = \theta_{\log}$.*

This result has been obtained in the independent papers [21] (in the setting of Markov chains) and [23] (in the setting of reaction-diffusion systems). Related gradient flow structures for Fokker-Planck equations on graphs have been discovered in [8]. A modification of the proof above shows that the discrete porous medium equations $\partial_t \rho = \Delta\varphi(\rho)$ can be formulated as gradient flow for the entropy functional $\mathscr{F}(\rho) = \sum_{x\in\mathscr{X}} f(\rho(x))\pi(x)$ with respect to the metric $\mathscr{W}$, provided that $\theta(r,s) = \frac{\varphi(r)-\varphi(s)}{f'(r)-f'(s)}$ for some increasing function φ and some convex function f, cf. [12].

5.4 Discrete Entropic Ricci Curvature

In view of the gradient flow result of Theorem 5.3, the metric $\mathscr{W}$ (with $\theta = \theta_{\log}$) can be viewed as a natural discrete counterpart of the Wasserstein metric W_2. The following definition, proposed in [21], is a discrete analogue of Definition 5.1 by Lott–Sturm–Villani.

Definition 5.3 (Discrete Entropic Ricci Curvature) Let $\kappa \in \mathbb{R}$. A reversible Markov chain $(\mathscr{X}, Q, \pi)$ is said to have "Ricci curvature bounded from below by κ" if every pair of probability densities $\rho_0, \rho_1 \in \mathscr{D}_+(\mathscr{X})$ can be connected by a constant speed $\mathscr{W}$-geodesic $(\rho_t)_{t\in[0,1]}$ along which the relative entropy $\mathscr{H}$ satisfies the κ-convexity inequality

$$\mathscr{H}(\rho_t) \leq (1-t)\mathscr{H}(\rho_0) + t\mathscr{H}(\rho_1) - \frac{\kappa}{2}t(1-t)\mathscr{W}(\rho_0,\rho_1)^2 .$$

We shall use the notation $\mathrm{Ric}(\mathscr{X}, Q, \pi) \geq \kappa$ for brevity.

The restriction to probability densities $\rho_0, \rho_1 \in \mathscr{D}(\mathscr{X})S$ is not essential: if $\theta = \theta_{\log}$ it can be shown that $\mathscr{W}$ is finite on the full space $\mathscr{D}(\mathscr{X})$, and the condition in Definition 5.3 may be imposed for all $\rho_0, \rho_1 \in \mathscr{D}(\mathscr{X})$ without changing the definition.

It has been shown in [11] that several discrete analogues of classical continuous results can be obtained from this definition. The following result is a discrete analogue of a classical result by Bakry and Émery [3].

Theorem 5.4 (Discrete Bakry–Émery) *If* $\mathrm{Ric}(\mathscr{X}, Q, \pi) \geq \kappa$ *for some* $\kappa > 0$, *then the modified logarithmic Sobolev inequality*

$$\mathscr{H}(\rho) \leq \frac{1}{2\kappa}\mathscr{I}(\rho) \tag{MLSI(κ)}$$

holds for all $\rho \in \mathscr{D}_+(\mathscr{X})$, *where*

$$\mathscr{I}(\rho) := \frac{1}{2}\sum_{x,y\in\mathscr{X}} \big(\log\rho(x) - \log\rho(y)\big)\big(\rho(x)-\rho(y)\big)Q(x,y)\pi(x)$$

is a discrete version of the Fisher information.

There are different (non-equivalent) versions of the logarithmic Sobolev inequality that appear in discrete settings. The relevance of the inequality (MLSI(κ)) is due to the fact that it implies the exponential convergence result $\mathscr{H}(\rho_t) \leq e^{-2\kappa t}\mathscr{H}(\rho_0)$ for solutions to the discrete heat equation $\partial_t\rho = \mathscr{L}\rho$.

The following result from [11] is a discrete version of a celebrated result by Otto and Villani [27].

Theorem 5.5 (Discrete Otto–Villani) *If* $(\mathscr{X}, Q, \pi)$ *satisfies* (MLSI(κ)) *for some* $\kappa > 0$, *then the modified Talagrand inequality*

$$\mathscr{W}(\rho, \mathbf{1}) \leq \sqrt{\frac{2}{\kappa}\mathscr{H}(\rho)} \tag{T$_{\mathscr{W}}$(κ)}$$

holds for all $\rho \in \mathscr{D}(\mathscr{X})$.

In this result, $\mathbf{1}$ denotes the constant probability density corresponding to the invariant measure π. The analogous inequality with $\mathscr{W}$ replaced by W_2 has been

extensively studied in continuous settings, but it is not hard to see that it can never hold in the discrete case. The inequality ($\mathrm{T}_{\mathscr{W}}(\kappa)$) provides a natural discrete substitute, which captures two different phenomena. On the one hand, ($\mathrm{T}_{\mathscr{W}}(\kappa)$) contains spectral information, since it implies the Poincaré inequality (or spectral gap inequality)

$$\|\varphi\|^2_{L^2(\mathscr{X},\pi)} \leq \frac{1}{\kappa}\mathscr{D}(\varphi) \tag{$\mathrm{P}(\kappa)$}$$

where $\mathscr{D}(\varphi) = \frac{1}{2}\sum_{x,y\in\mathscr{X}} \big(\varphi(x)-\varphi(y)\big)^2 Q(x,y)\pi(x)$ denotes the Dirichlet energy. On the other hand, ($\mathrm{T}_{\mathscr{W}}(\kappa)$) yields the T_1-transport inequality

$$W_1(\rho,\mathbf{1}) \leq \sqrt{\frac{2}{\kappa'}\mathscr{H}(\rho)}$$

for a possibly different constant κ', where W_1 is the 1-Wasserstein distance induced by the graph distance on the incidence graph of $(\mathscr{X},Q,\pi)$. The latter inequality implies an exponential concentration inequality for π with an explicit rate, cf. [11].

5.4.1 Discrete Spaces with Lower Ricci Curvature Bounds

As discrete entropic Ricci curvature bounds have significant consequences for the Markov chains under consideration, it is of interest to obtain sharp bounds for the Ricci curvature in concrete examples. Since such bounds are defined in terms of convexity properties of the relative entropy $\mathscr{H}$, one needs to calculate the Hessian of $\mathscr{H}$ in the Riemannian structure inducing $\mathscr{W}$. Therefore, we would like to compute the second derivative of the entropy along $\mathscr{W}$-geodesics. The geodesic equations for $\mathscr{W}$ are given by the continuity equation (5.6) in conjunction with the equations

$$\partial_t\psi_t(x) + \frac{1}{2}\sum_{y\in\mathscr{X}} \big(\psi_t(x)-\psi_t(y)\big)^2 \partial_1\theta(\rho(x),\rho(y))Q(x,y) = 0\,, \qquad x\in\mathscr{X}\,. \tag{5.8}$$

This equation is reminiscent of the Hamilton-Jacobi equation $\partial_t\psi + \frac{1}{2}|\nabla\psi|^2 = 0$, which appears in the description of 2-Wasserstein geodesics in continuous settings. Note however that the discrete equation (5.8) depends both on ρ and ψ, which is a source of additional difficulties in the discrete setting. An explicit computation based on the Eqs. (5.4) and (5.8) shows that the second derivative of the entropy

along a unit speed $\mathscr{W}$-geodesics $(\rho_t)_t$ is given by

$$\begin{aligned}\partial_t^2\big|_{t=0}\mathscr{H}(\rho_t) &= \frac{1}{4}\sum_{x,y,z\in\mathscr{X}}\big(\psi(x)-\psi(y)\big)^2\Big(\partial_1\theta\big(\rho(x),\rho(y)\big)\big(\rho(z)-\rho(x)\big)Q(x,z)\\ &\qquad +\partial_2\theta\big(\rho(x),\rho(y)\big)\big(\rho(z)-\rho(y)\big)Q(y,z)\Big)Q(x,y)\pi(x)\\ &= -\frac{1}{2}\sum_{x,y,z\in\mathscr{X}}\Big(Q(x,z)\big(\psi(z)-\psi(x)\big)-Q(y,z)\big(\psi(z)-\psi(y)\big)\Big)\\ &\qquad \times\big(\psi(x)-\psi(y)\big)\hat{\rho}(x,y)Q(x,y)\pi(x) =: \mathscr{B}(\rho,\psi)\ ,\end{aligned}$$

Proving a bound of the form $\mathrm{Ric}(\mathscr{X},Q,\pi)\geq\kappa$ thus corresponds to showing that

$$\mathscr{B}(\rho,\psi)\geq\frac{\kappa}{2}\sum_{x,y}(\psi(x)-\psi(y))^2\hat{\rho}(x,y)Q(x,y)\pi(x)\ . \tag{5.9}$$

The right-hand side above may be seen as a discrete version of the continuous formula $\int\big(\frac{1}{2}|\nabla\psi|^2\Delta\rho-\langle\nabla\Delta\psi,\nabla\psi\rangle\rho\big)\mathrm{d}x$. On a Riemannian manifold, such an expression can be simplified after integration by parts with the help of Bochner's identity, which asserts that

$$\frac{1}{2}\Delta\big(|\nabla\psi_t|^2\big)-\langle\nabla\psi_t,\nabla\Delta\psi_t\rangle=|\mathrm{D}^2\psi|^2+\mathrm{Ric}(\nabla\psi,\nabla\psi)\ .$$

Hence a lower bound on the Ricci curvature yields a corresponding bound for the Hessian of the entropy. In the discrete case there doesn't seem to be an exact analogue of Bochner's formula, and the challenge is to get around this difficulty in concrete examples of interest.

5.4.2 Examples

In recent years, discrete entropic Ricci curvature bounds have been obtained in several concrete examples.

A result by Mielke [24] asserts that every reversible Markov chain on a finite state space $\mathscr{X}$ satisfies $\mathrm{Ric}(\mathscr{X},Q,\pi)\geq\kappa$ for some (possibly negative) $\kappa\in\mathbb{R}$. Even though the Riemannian manifold $\mathscr{P}_+(\mathscr{X})$ is finite-dimensional, this result is non-trivial (and its proof is rather delicate), since the Riemannian metric inducing $\mathscr{W}$ is degenerate at the boundary of $\mathscr{P}_+(\mathscr{X})$.

If the state space $\mathscr{X}$ has a one-dimensional structure, the situation is quite well understood. Consider a birth-death chain (on a finite set $\{0,1,\dots,N\}$ for some $N\geq 1$) with transition rates $Q(n,n+1)=a(n)$ and $Q(n,n-1)=b(n)$. In this situation,

Mielke [24] obtained the following lower bounds on the Ricci curvature for birth-death processes (see also [16] for a different proof).

Theorem 5.6 (Ricci Bounds for Birth-Death Chains) *Let $\kappa \in [0, \infty)$. Assume that the rate of birth a is non-increasing, and that the rate of death b is non-decreasing. Assume moreover that*

$$\frac{1}{2}\big(a(n) - a(n+1) + b(n+1) - b(n)\big) + \frac{1}{2}\Theta\big(a(n) - a(n+1), b(n+1) - b(n)\big) \geq \kappa \tag{5.10}$$

for all $n = 0, \ldots, N-1$, where

$$\Theta(\alpha, \beta) = \inf_{s,t>0} \theta(s,t)\left(\frac{\alpha}{s} + \frac{\beta}{t}\right),$$

and $\theta(s,t)$ is the logarithmic mean of s and t. Then the birth-death process has Ricci curvature bounded from below by κ.

In general this result does not give the optimal constants, but the result is asymptotically sharp in the sense that one recovers the optimal Ricci curvature bounds for one-dimensional Fokker-Planck equations by passing to the continuum limit after a suitable rescaling, cf [24].

The following result, taken from [11], asserts that Ricci bounds are preserved under taking product chains. If $\mathscr{L}_i$ generates a Markov chain on $\mathscr{X}_i$ with reversible measure π_i for $i = 1, 2$, the corresponding product chain on $\mathscr{X}_1 \times \mathscr{X}_2$ is the Markov chain with transition semigroup $e^{t\mathscr{L}} := e^{t\mathscr{L}_1} \otimes e^{t\mathscr{L}_2}$. Note that the generator is given by

$$\mathscr{L}^{\otimes} = \mathscr{L}_1 \otimes I + I \otimes \mathscr{L}_2 .$$

We shall write $Q^{\otimes}$ accordingly. The product measure $\pi_1 \otimes \pi_2$ is the reversible invariant measure for $\mathscr{L}^{\otimes}$.

Theorem 5.7 (Tensorisation of Discrete Entropic Ricci Curvature) *For $i = 1, 2$, let $(\mathscr{X}_i, Q_i, \pi_i)$ be reversible Markov chains with lower Ricci bounds $\kappa_i \in \mathbb{R}$. Then the associated product chain satisfies* $\mathrm{Ric}(\mathscr{X}_1 \times \mathscr{X}_2, Q^{\otimes}, \pi_1 \otimes \pi_2) \geq \min\{\kappa_1, \kappa_2\}$.

This result allows one to consider product chains in arbitrarily high dimension, without any loss in the Ricci curvature constant. As a consequence, it can be proved that the optimal Ricci curvature bound for simple random walk on the discrete cube $\{0,1\}^N$ is given by $\frac{2}{N}$. Similarly, it follows that the Ricci curvature for any regular rectangular lattice in any dimension is non-negative.

Much less is known for Markov chains in multiple dimensions which do not have a product structure, but there has been some recent progress in this direction. Using combinatorial methods, entropic Ricci curvature bounds have been obtained for the Bernoulli-Laplace model, which describes a simple random walk on a slice

of the discrete cube, as well as for the random transposition model, which describes a random walk on the group of permutations [14]. A more systematic method for proving discrete Ricci bounds, inspired by the work of Caputo et al. [6] on functional inequalities, has been developed in [16] and applied to zero-range processes on the complete graph.

We conclude by mentioning that there are several other interesting notions of discrete Ricci curvature that have been actively studied in recent years. Important examples include Ollivier's Ricci curvature [25] and the Bakry-Émery criterion [3] and its variants. It is not hard to show that the Bakry-Émery criterion can be obtained from the inequality (5.9) that characterizes discrete entropic Ricci curvature, by replacing the logarithmic mean $\theta_{\log}$ by the arithmetic mean. Since the proofs in [11, 14, 16] remain valid for the arithmetic mean, these papers also yield curvature bounds in the sense of Bakry-Émery. We refer to other chapters in this volume, in particular the contribution by F. Bauer, B. Hua, J. Jost, S. Liu and G. Wang (Chap. 1) for a discussion of Ollivier's Ricci curvature and the Bakry-Émery condition.

A discretization of the Lott-Sturm-Villani theory based on the notion of approximate W_2-midpoints (different from the theory described in this note) was given by Bonciocat and Sturm [5]. Convexity of the entropy in discrete settings has been investigated along other interesting curves of probability measures in [17] and in [19].

Note added: After this short paper was written, several papers dealing with discrete entropic Ricci curvature have appeared. In particular, Erbar and Fathi [10] obtained various interesting functional inequalities under the assumption of non-negative entropic Ricci curvature. Moreover, entropic Ricci curvature bounds have been proved for several weakly interacting Markov chains (including Glauber dynamics for a class of spin systems) by Erbar et al. [15]. Related functional inequalities have been investigated by Che et al. [7].

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2nd edn. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008)
2. Ambrosio, L., Gigli, N., Savaré, G.: Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163**(7), 1405–1490 (2014). http://dx.doi.org/10.1215/00127094-2681605.
3. Bakry, D., Émery, M.: Diffusions hypercontractives. In: Séminaire de probabilités, XIX, 1983/1984. Lecture Notes in Mathematics, vol. 1123, pp. 177–206. Springer, Berlin (1985). http://dx.doi.org/10.1007/BFb0075847
4. Benamou, J.D., Brenier, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000). http://dx.doi.org/10.1007/s002110050002
5. Bonciocat, A.I., Sturm, K.T.: Mass transportation and rough curvature bounds for discrete spaces. J. Funct. Anal. **256**(9), 2944–2966 (2009). http://dx.doi.org/10.1016/j.jfa.2009.01.029

6. Caputo, P., Dai Pra, P., Posta, G.: Convex entropy decay via the Bochner-Bakry-Emery approach. Ann. Inst. Henri Poincaré Probab. Stat. **45**(3), 734–753 (2009). http://dx.doi.org/10.1214/08-AIHP183
7. Che, R., Huang, W., Li, Y., Tetali, P.: Convergence to global equilibrium for Fokker-Planck equations on a graph and Talagrand-type inequalities. J. Differ. Equ. **261**(4), 2552–2583 (2016). http://dx.doi.org/10.1016/j.jde.2016.05.003
8. Chow, S.N., Huang, W., Li, Y., Zhou, H.: Fokker-Planck equations for a free energy functional or Markov process on a graph. Arch. Ration. Mech. Anal. **203**(3), 969–1008 (2012). http://dx.doi.org/10.1007/s00205-011-0471-6
9. Cordero-Erausquin, D., McCann, R.J., Schmuckenschläger, M.: A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. Invent. Math. **146**(2), 219–257 (2001). http://dx.doi.org/10.1007/s002220100160
10. Erbar, M., Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. arXiv preprint arXiv:1612.00514 (2016)
11. Erbar, M., Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206**(3), 997–1038 (2012). http://dx.doi.org/10.1007/s00205-012-0554-z
12. Erbar, M., Maas, J.: Gradient flow structures for discrete porous medium equations. Discrete Contin. Dyn. Syst. **34**(4), 1355–1374 (2014). http://dx.doi.org/10.3934/dcds.2014.34.1355
13. Erbar, M., Kuwada, K., Sturm, K.T.: On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201**(3), 993–1071 (2015). http://dx.doi.org/10.1007/s00222-014-0563-7
14. Erbar, M., Maas, J., Tetali, P.: Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models. Ann. Fac. Sci. Toulouse Math. (6) **24**(4), 781–800 (2015). http://dx.doi.org/10.5802/afst.1464
15. Erbar, M., Henderson, C., Menz, G., Tetali, P.: Ricci curvature bounds for weakly interacting Markov chains. arXiv preprint arXiv:1602.05414 (2016)
16. Fathi, M., Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. Ann. Appl. Probab. **26**(3), 1774–1806 (2016). http://dx.doi.org/10.1214/15-AAP1133
17. Gozlan, N., Roberto, C., Samson, P.M., Tetali, P.: Displacement convexity of entropy and related inequalities on graphs. Probab. Theory Relat. Fields **160**(1–2), 47–94 (2014). http://dx.doi.org/10.1007/s00440-013-0523-y
18. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)
19. Léonard, C.: Lazy random walks and optimal transport on graphs. Ann. Probab. **44**(3), 1864–1915 (2016). http://dx.doi.org/10.1214/15-AOP1012
20. Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. (2) **169**(3), 903–991 (2009). http://dx.doi.org/10.4007/annals.2009.169.903
21. Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011). http://dx.doi.org/10.1016/j.jfa.2011.06.009
22. McCann, R.J.: A convexity principle for interacting gases. Adv. Math. **128**(1), 153–179 (1997). http://dx.doi.org/10.1006/aima.1997.1634
23. Mielke, A.: A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. Nonlinearity **24**(4), 1329–1346 (2011). http://dx.doi.org/10.1088/0951-7715/24/4/016
24. Mielke, A.: Geodesic convexity of the relative entropy in reversible Markov chains. Calc. Var. Partial Differ. Equ. **48**(1–2), 1–31 (2013). http://dx.doi.org/10.1007/s00526-012-0538-8
25. Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256**(3), 810–864 (2009). http://dx.doi.org/10.1016/j.jfa.2008.11.001
26. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. Commun. Partial Differ. Equ. **26**(1–2), 101–174 (2001). http://dx.doi.org/10.1081/PDE-100002243
27. Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**(2), 361–400 (2000). http://dx.doi.org/10.1006/jfan.1999.3557

28. Renesse, M.K.v., Sturm, K.T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature. Commun. Pure Appl. Math. **58**(7), 923–940 (2005). http://dx.doi.org/10.1002/cpa.20060
29. Sturm, K.T.: On the geometry of metric measure spaces. I and II. Acta Math. **196**(1), 65–177 (2006). http://dx.doi.org/10.1007/s11511-006-0003-7
30. Villani, C.: Optimal Transport, Old and New. Grundlehren der Mathematischen Wissenschaften, vol. 338. Springer, Berlin (2009). http://dx.doi.org/10.1007/978-3-540-71050-9

Chapter 6
Geometric and Spectral Consequences of Curvature Bounds on Tessellations

Matthias Keller

Abstract This chapter focuses on geometric and spectral consequences of curvature bounds. Several of the results presented here have analogues in Riemannian geometry but in some cases one can go even beyond the Riemannian results and there also striking differences. The geometric setting of this chapter are tessellations and the curvature notion arises as a combinatorial quantity which can be interpreted as an angular defect and goes back to Descartes. First, we study the geometric consequences of curvature bounds. Here, a discrete Gauss–Bonnet theorem provides a starting point from which various directions shall be explored. These directions include analogues of a theorem of Myers, a Hadamard–Cartan theorem, volume growth bounds, strong isoperimetric inequalities and Gromov hyperbolicity. Secondly, we investigate spectral properties of the Laplacian which are often consequences of the geometric properties established before. For example we present analogues to a theorem of McKean about the spectral gap, a theorem by Donnelly-Li about discrete spectrum, we discuss the phenomena of compactly supported eigenfunctions and briefly elaborate on stability of the ℓ^2 spectrum for the Laplacian on ℓ^p.

6.1 Introduction

This chapter focuses on the theme of geometric and spectral consequences of curvature bounds. The geometric setting are tessellations of surfaces with finite and zero genus. Our main focus is on infinite tessellations. Several of the results presented here have analogues in Riemannian geometry. In some cases one can go even beyond the Riemannian results and there also striking differences which shall be highlighted.

M. Keller (✉)
Institut für Mathematik, Universität Potsdam, 14476 Potsdam, Germany
e-mail: matthias.keller@uni-potsdam.de

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_6

The curvature under investigation arises as a combinatorial quantity with the geometric interpretation of an angular defect. For a vertex v the curvature is given by

$$\Phi(v) = 1 - \frac{d_v}{2} + \sum_{f \text{ face with } v \in f} \frac{1}{\deg(f)} = \frac{1}{2\pi}\Big(2\pi - \sum_{f \text{ face with } v \in f} \beta(f)\Big),$$

where d_v is the vertex degree, $\deg(f)$ the number of vertices contained in a face f and $\beta(f) = (\deg(f) - 2)/2\deg(f)$ the inner angle of f considered as a regular $\deg(f)$-gon. We introduce this concept in detail and discuss the basic features below which is the first part of the chapter.

In the second part of the chapter, in Sect. 6.2, geometric consequences of curvature bounds are discussed. Here, the discrete Gauss–Bonnet Theorem provides a starting point from which various directions shall be explored. First, it is discussed that positivity of the curvature yields finiteness of the graph which is an analogue to a theorem of Myers from Riemannian geometry. In contrast, non-positive curvature yields infinite tessellations for which we study further geometric properties. Specifically, a Hadamard–Cartan theorem is presented which states that geodesics can be continued indefinitely as a consequence of non-positive curvature. Furthermore, volume growth bounds are derived from upper and lower curvature bounds. Finally, hyperbolicity properties such as strong isoperimetric inequalities and Gromov hyperbolicity as consequences of negative curvature are studied.

In the third part of the chapter, in Sect. 6.3, we focus on spectral properties of the combinatorial Laplacian

$$\Delta\varphi(v) = \sum_{w \sim v} (\varphi(w) - \varphi(v))$$

on $\ell^2(V)$. Let us stress that this combinatorial Laplacian differs from the normalized Laplacian introduced in the previous chapter (which is also referred to as the Tutte Laplacian or harmonic Laplacian). The main reason is that the operator above captures phenomena related to unbounded geometry for infinite graphs much better.

Many of the spectral properties of Δ are consequences of the geometric properties which were established before. First, we discuss whether the combinatorial Laplacian admits a spectral gap due to the curvature being non-negative or negative which includes an analogue to a classical theorem of McKean. Secondly, we focus on the case of uniformly decreasing curvature which was studied in the Riemannian setting by Donnelly/Li. Here, we not only characterize the situation, where the combinatorial Laplacian has only discrete eigenvalues as spectrum but also determine the asymptotics of eigenvalues as well as the decay of eigenfunctions. Here, it appears that the eigenvalue asymptotics behave differently than in the case of manifolds. Thirdly, we explore the phenomena of compactly supported eigenfunctions which is often studied under the term unique continuation of eigenfunctions. It came as a surprise that in contrast to the continuous setting such

compactly supported eigenfunctions appear in the discrete setting. However, they can be excluded in situations of non-positive curvature. Finally, we briefly consider the ℓ^p-spectrum of the combinatorial Laplacian and ask when it differs from the ℓ^2 spectrum, which are discrete analogues of results of Sturm.

In the outlook section, which forms the fourth part, we discuss how the results of the first parts can be generalized to planar graphs or more general polygonal complexes including two dimensional buildings.

All the results are discussed in the light of the present literature which reaches back from the seventies to the very present and has contributions from various schools of researchers. We sketch the key ideas of the proofs without going into too much detail.

6.1.1 Graphs

Let a connected simple graph (V, E) be given. We call two vertices $x, y \in V$ which are connected by an edge *adjacent* and denote $x \sim y$. We call a sequence $(x_r)_{r\geq 0}$ of subsequently adjacent and pairwise distinct vertices a *path*. For a finite path $(x_0, \ldots, x_n)$, we call n the *length* of the path. For two vertices $x, y \in V$, we define the *distance* $d(x, y)$ of x and y to be the length of the shortest path connecting x and y. We call a path (x_r) a geodesic if $d(x_0, x_n) = n$ for all indices n of the path.

The *sphere* $S_r = S_r(o)$ of radius r about a fixed vertex o (which is mostly suppressed in notation) includes the vertices with distance r to o. The *ball* $B_r = B_r(o)$ of radius r about o is the union of all spheres of distance less or equal to r. Denote the cardinality of a set A by $\#A$ or $|A|$.

6.1.2 Tessellations of Surfaces

In this chapter we consider tessellations. From a purely graph-theoretic approach this might seem rather restrictive but it is natural from a discretization and geometric point of view and it allows to prove finer results.

We assume that (V, E) can be *embedded* in an oriented topological surface $\mathscr{S}$ without self intersections. That means that the vertices V are identified with points in $\mathscr{S}$ and the edges E are identified with simple curves connecting its end vertices such that two different edges intersect only in the vertices they have in common. Throughout this chapter we will not distinguish between the graph and its embedding.

We always assume that the embedding is *locally compact*, i.e., every compact $K \subseteq \mathscr{S}$ intersects only finitely many edges.

A graph is called *planar* if there is a locally compact embedding into a surface $\mathscr{S}$ homeomorphic to $\mathbb{R}^2$ and *spherical* if $\mathscr{S}$ can be chosen to be homeomorphic to the sphere $\mathbb{S}^2$.

For an embedded graph (V, E), we let the set of faces F consist of the closures of the connected components of

$$\mathscr{S} \setminus \bigcup E.$$

We denote $G = (V, E, F)$ and following [5, 6] we call G a *tessellation* or a *tiling* if

(T1) every edge is included in two faces,
(T2) every two faces are either disjoint or intersect in one vertex or one edge,
(T3) every face is homeomorphic to a closed disc.

Let us shortly discuss the local compactness assumption in the light of tessellations. This assumption can be violated for two reasons. The first is that the point where the assumption fails is a vertex. This, however, implies that infinitely many edges emanate from this vertex. So, a locally compact embedding first of all implies that the graph is locally finite, i.e., each vertex has only finitely many neighbors. Secondly, the assumption can fail for a point which is not a vertex. Removing all such points from the surface, annihilates the problem and the embedding becomes locally compact. So, in this case the graph can be embedded locally compactly in a different surface.

6.1.3 Curvature

To define a curvature function, we introduce the vertex degree and the face degree. First we let the *degree* of a vertex $v \in V$ be defined as

$$d_v = \#\text{edges emanating from } v.$$

The *face degree* of a face $f \in F$ is defined as

$$\deg(f) = \#\text{boundary edges of } f = \#\text{boundary vertices of } f.$$

The *vertex curvature* $\Phi : V \to \mathbb{R}$ is defined as

$$\Phi(v) = 1 - \frac{d_v}{2} + \sum_{f \in F, v \in f} \frac{1}{\deg(f)}.$$

This idea is already found in the works of Descartes, see e.g. [21]. Later it was introduced in the above form by Stone, [51], where he refers ideas going back to Alexandrov. The notion reappeared in various settings [25, 31] and was studied intensively since then [5, 6, 15, 28, 30, 32, 33, 35, 48, 57, 60].

This notion of curvature is motivated by an angular defect as follows (cf. Chap. 2 and 1): Think of a face f as a regular polygon. Then, the interior angles of f are

equal to

$$\beta(f) = 2\pi \frac{\deg(f) - 2}{2\deg(f)}.$$

This formula is easily derived as going around f results in an angle of 2π, while going around the $\deg(f)$ corners of f one takes a turn by an angle of $\pi - \beta(f)$ each time. With this interpretation the curvature of a vertex $v \in V$ can be rewritten as

$$2\pi\Phi(v) = 2\pi - \sum_{f\in F, v\in f} \beta(f).$$

Despite of this interpretation, it should be stressed that the mathematical nature of Φ is purely combinatorial. However, when we assume an embedding such that every polygon is regular we have the geometric interpretation above. Furthermore, the curvature satisfies a Gauss-Bonnet formula which is found in the next section in Theorem 6.1.

Before we come to this, we introduce a finer notion of curvature. To this end, we let the set of *corners* $C(G)$ of the graph G be given by the set of pairs $(v,f) \in V \times F$ such that v is contained in f. One can picture the corners around a vertex as the connected components of a small neighborhood of v in $\mathscr{S}$ after removing the edges emanating from v. The *corner curvature* $\Phi_C : C(G) \to \mathbb{R}$ is given by

$$\Phi_C(v,f) = \frac{1}{d_v} - \frac{1}{2} + \frac{1}{\deg(f)}.$$

This quantity was first introduced in [5, 6] for tessellations and in [33] for general planar graphs (cf. Sect. 6.4.1). The corner curvature measures the contribution of every corner of a vertex to the vertex curvature, i.e., summing over all corners of a vertex, we get

$$\Phi(v) = \sum_{(v,f)\in C(G)} \Phi_C(v,f).$$

With the interpretation of the angular defect above the corner curvature can be expressed as

$$2\pi\Phi_C(v,f) = \frac{1}{d_v}(2\pi - d_v\beta(f)), \quad (v,f) \in C(G).$$

We say a tessellation is *flat* if $\Phi \equiv 0$. Let us discuss a few examples.

Example 6.1 We say a tessellation of $\mathscr{S}$ is (p,q)-regular if q regular polygons of degree p meet in every vertex. Then, the sign of vertex curvature is already determined by the sign of the corner curvature which is constantly $1/q - 1/2 + 1/p$. Let us come to more specific examples.

(a) The examples of (p, q)-regular tessellations of $\mathbb{S}^2$ of positive curvature are the platonic solids, i.e., the tetrahedron $(3, 3)$, the octahedron $(3, 4)$, the icosahedron $(3, 5)$, the cube $(4, 4)$ and the dodecahedron $(5, 3)$.
(b) Flat (p, q)-regular tessellations of $\mathbb{R}^2$ are exactly either the square tessellation associated to $\mathbb{Z}^2$ of type $(4, 4)$, the flat triangular tessellation $(3, 6)$ and the hexagonal honeycomb tessellation $(6, 3)$. Furthermore, any large enough rectangular subset of $\mathbb{Z}^2$ can be realized as a tessellation of the torus in the usual way by identifying opposite sides. This results again flat curvature. Similar constructions can be realized for the other two flat tessellations.
(c) Planar (p, q)-regular tessellations of the type $(4, 5)$, $(5, 5)$, $(5, 6)$, $(5, 4)$, $(6, 5)$ or such that $p \geq 7$ or $q \geq 3$ give rise to tessellations of the hyperbolic plane with negative curvature.

See Fig. 6.1 for examples of (a), (b) and (c).

Next, we discuss examples of non definite curvature, see Fig. 6.2 as well.

Example 6.2 The so called Cairo tiling consists of pentagons. For each pentagon there are two non adjacent vertices that are contained in four other pentagons. This results in negative corner and vertex curvature. The other three vertices are contained in three pentagons resulting in positive corner and vertex curvature.

Example 6.3 The Penrose tiling consists of squares. There are vertices adjacent to five and six squares resulting in negative curvature, vertices adjacent to four squares

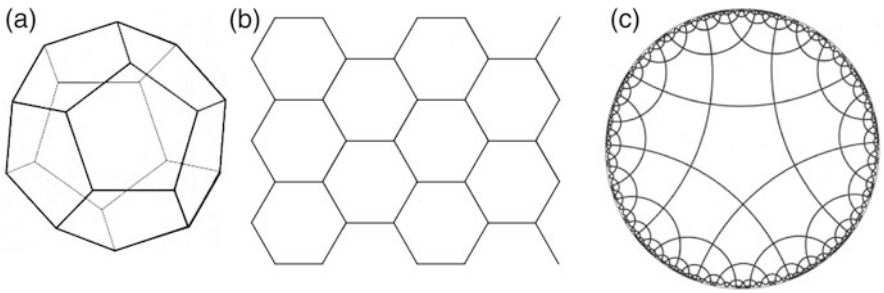


Fig. 6.1 (**a**) The dodecahedron, (**b**) the hexagonal tiling and (**c**) the $(5, 4)$ tessellation embedded in the Poincaré disc

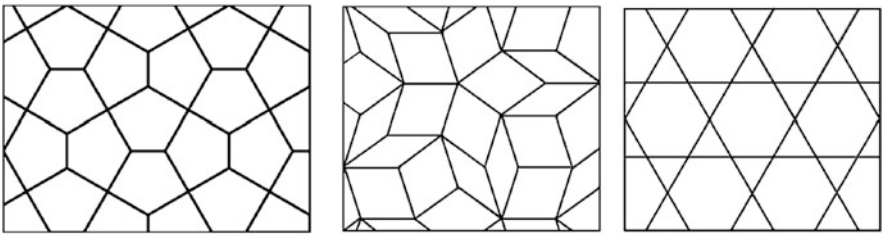


Fig. 6.2 The Cairo tiling, the Penrose tiling and the Kagome lattice

resulting in zero curvature and vertices adjacent to three squares resulting in positive curvature. This holds for the corner curvatures as well as for the vertex curvature.

Finally, we discuss examples of non-positive and negative vertex curvature which have yet non definite corner curvature.

Example 6.4 The class of example we discuss next consists of regular polygons of degree $p \geq 6$ and of triangles. In each vertex two triangles and two p-gons meet such that faces of the same type are opposite to each other. The vertex curvature in a vertex is then $1/p - 1/6 \leq 0$. However, the curvature in corners of a triangle is $1/12$ and in the corners of a p-gon is $1/p - 1/2 < 0$. Hence, the corner curvature is non-definite. A special case with $q = 6$ is called the trihexagonal tiling or the Kagome lattice or the hexadeltille.

6.1.4 Duality

We make some brief remarks about duality. To every tessellation $G = (V, E, F)$ we can relate a dual tessellation $G^* = (V^*, E^*, F^*)$ in the following way: To define the vertex set we chose for each face in F a point in its interior which we refer to as a vertex in V^*. If two faces in F intersect in an edge we connect the corresponding vertices in V^* by an edge. This gives a graph embedded in the same surface and it can be seen that the faces of (V^*, E^*) can be related one-to-one to the vertices in V. Let $C(G^*)$ be the corners of G^*. We see that, there are bijective maps from V to F^*, E to E^*, F to V^* and $C(G)$ to $C(G^*)$. Denote the bijective maps $V \to F^*$, $v \mapsto v^*$ and $F \to V^*, f \mapsto f^*$ and observe that

$$d_v = \deg(v*) \quad \text{and} \quad \deg(f) = d_{f^*}.$$

We stress that d_{v^*} is a face and f^* is a vertex in G^*. We denote the corner curvature of G^* by Φ_C^* and observe that

$$\Phi_C(v,f) = \frac{1}{d_v} - \frac{1}{2} + \frac{1}{\deg(f)} = \frac{1}{\deg(v^*)} - \frac{1}{2} + \frac{1}{d_{f^*}} = \Phi_C^*(f^*, v^*).$$

Moreover, defining the face curvature $\Phi_F^* : F^* \to \mathbb{R}$ on G^* viz

$$\Phi_F^*(f) = \sum_{(v,f)\in C(G^*)} \Phi_C^*(v,f) = 1 - \frac{\deg(f)}{2} + \sum_{v\in V^*, v\in f} \frac{1}{d_v},$$

for $f \in F^*$, we see that

$$\Phi(v) = \sum_{(v,f)\in C(G)} \Phi_C(v,f) = \sum_{(f^*,v^*)\in C^*(G)} \Phi_C^*(f^*, v^*) = \Phi_F^*(v^*),$$

for $v \in V$. Similarly, if we define a face curvature on G, then the vertex curvature on G^* can be related analogously. We summarize that for each tessellation G the dual graph G^* gives rise to a tessellation. Moreover, these two graphs have the same corner curvature with respect to the canonical bijection. The vertex curvature in G translate into a face curvature in G^* and vice versa.

6.2 Geometry

6.2.1 Gauss–Bonnet Theorem

In Riemannian geometry the Gauss–Bonnet Theorem is a link between the geometry and topology of a surface. In particular, it relates the curvature of a surface to its Euler characteristic. An analogous theorem holds true in the discrete setting. Various versions of the theorem below are found in [5, 13–15, 33].

For a surface $\mathscr{S}$ the genus g is defined to be the largest number of nonintersecting simple closed curves that can be drawn on the surface without separating it. It can be thought as the number of holes in a surface. The Euler characteristic $\chi(\mathscr{S})$ of $\mathscr{S}$ is given as

$$\chi(\mathscr{S}) = 2 - 2g.$$

Theorem 6.1 (Gauss–Bonnet Theorem) *Let $G = (V, E, F)$ be a tessellation embedded locally compactly in a compact oriented surface $\mathscr{S}$ of finite genus. Then,*

$$\sum_{v \in V} \Phi(v) = \chi(\mathscr{S}).$$

Proof (Sketch of the proof) Since the embedding is locally compact, the tessellation G embedded in a compact surface is finite. We calculate

$$\sum_{v \in V} \Phi(v) = \sum_{v \in V} \sum_{(v,f) \in C(G)} \Big(\frac{1}{d_v} - \frac{1}{2} + \frac{1}{\deg(f)} \Big).$$

Using $\#\{(v,f) \in C(G)\} = d_v$ for fixed $v \in V$, we arrive at

$$\ldots = |V| - \frac{1}{2} \sum_{v \in V} d_v + \sum_{f \in F} \sum_{(v,f) \in C(G)} \frac{1}{\deg(f)}.$$

Now, employing $\sum_{v \in V} d_v = 2|E|$ and $\#\{(v,f) \in C(G)\} = \deg(f)$ for fixed $f \in F$, yields

$$\ldots = |V| - |E| + |F|.$$

The statement now follows from the identity $|V|-|E|+|F| = \chi(\mathscr{S})$ which is known as Poincaré formula. This formula is proven by induction over g. The base case $g = 0$ is known as the Euler-Descartes formula. The induction step for a tessellation $G_g = (V_g, E_g, F_g)$ embedded in a surface $\mathscr{S}_g$ of genus g works by cutting along a path of edges that separates the surface. Say the length of the path is p. The cut surface has now a boundary consisting of two paths of edges of length p. In each of these boundary paths one glues a polygon of degree p to "close up" the surface in order to get a surface of genus $g - 1$. For this surface we know the formula by the induction hypothesis. Now by cutting and gluing, we increased the number of vertices by p, we also increased the number of edges also by p and we increased the number of faces by 2. This results in the formula $|V_g| - |E_g| + |F_g| = \chi(S_g)$. □

6.2.2 Approximating Flat and Infinite Curvature

From the definition of the curvature it follows that Φ takes only values in $\mathbb{Q}$. So a question is which values can actually be obtained or approximated. Here, we discuss that Φ is bounded from above but not from below. Moreover, we find that flat curvature can be approximated from above but not from below.

Clearly, the minimal vertex and face degree is 3 by assumption. Hence, the maximal value that can be assumed by Φ is $3/2$.

On the other hand, Φ is not necessarily bounded from below as it can be seen by the next proposition.

Proposition 6.1 *Let $G = (V, E, F)$ be a tessellation. Then, for all $v \in V$*

$$-\frac{d_v}{2} \leq \Phi(v) \leq 1 - \frac{d_v}{6}.$$

In particular, $\Phi(v) \leq 0$ if $d_v \geq 6$ and for (v_n) we have $d_{v_n} \to \infty$ if and only if $\Phi(v_n) \to -\infty$ for $n \to \infty$.

Proof The lower bound is immediate and the upper bound follows as $\deg(f) \geq 3$ for all $f \in F$ and the number of corners about a vertex equals the vertex degree. □

Let us now turn to flat curvature. We discuss that the value zero can be approximated from below but not from above.

For positive curvature we consider so called prisms or antiprisms, see Fig. 6.3. In the figure, the unbounded face outside arises from a face in the sphere.

A *prism* is a tessellation of the sphere. It is given by two polygons of degree p that are surrounded by p squares. In particular, each vertex is adjacent to two squares and one of the two p-gons. The curvature of any vertex v is given by

$$\Phi(v) = 1 - \frac{3}{2} + \left(\frac{1}{p} + 2\frac{1}{4}\right) = \frac{1}{p}.$$

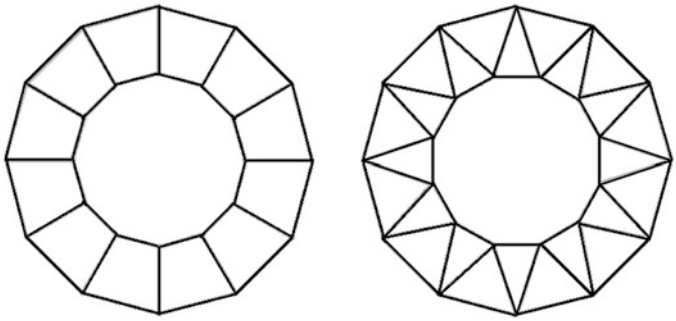


Fig. 6.3 A prism and an antiprism projected from the sphere to the Euclidean plane

Similarly, an *antiprism* consists of two p-gons that have triangles glued along the boundary such that in the resulting tessellation every vertex is adjacent to three triangles and one p-gon. The curvature of any vertex v is given by

$$\Phi(v) = 1 - \frac{4}{2} + \left(\frac{1}{p} + 3\frac{1}{3}\right) = \frac{1}{p}.$$

In the case when p tends to ∞, the curvature of prisms and antiprisms tends to 0 from above.

In the case of negative curvature there is a uniform lower bound which was shown by Higuchi [28], via listing all essential cases.

Theorem 6.2 ([28, Proposition 2.1.]) *Let G be a tessellation. If $\Phi < 0$, then $\Phi \leq -1/1806$. This maximum is can be assumed at a vertex v with $d_v \geq 3$ and which is adjacent to faces f_1, f_2, f_3 such that $(|f_1|, |f_2|, |f_3|) = (3, 7, 43)$.*

The theorem implies that flat curvature can not be approximated from below.

6.2.3 Finiteness

Next, we turn to question how the sign of the curvature determines finiteness or infiniteness of the tessellation. A rough synopsis is that positive curvature implies finiteness and in the case of planar tessellations non-positive curvature implies infiniteness.

In Riemannian geometry a theorem that positive curvature implies compactness of the manifold goes back to Myers, [46]. For graphs this question was first studied in [51, 52] and came up again later as Higuchi's conjecture [28, Conjecture 3.2.]. A definite answer was given by DeVos/Mohar [15, Theorem 1.7] after first steps were taken in [14, 54].

Theorem 6.3 (Myers Theorem for Tessellations, [15]) *Let $G = (V, E, F)$ be a tessellation embedded locally compactly in a surface such that $\Phi > 0$. Then, the graph is finite.*

The arguments of the proof use an asymptotic Gauss–Bonnet formula for infinite graphs with non-negative curvature. Furthermore, it is then shown that *big faces* of degree larger than 42 can not be too close. In particular, if an edge connects two big faces by its end vertices then the graph is a prism or antiprism which is essentially proven by listing all possible cases. Then, a lower bound on the sum of the curvature is derived in terms of the number of vertices for graphs which are neither prisms or antiprisms.

The upper bound for the number of vertices of a graph with positive curvature given in [15] for graphs that are not prisms or antiprisms is 3444. This was later improved by Zhang [59] to 580 vertices and by Oh [10] to 380 vertices while the largest known graphs with positive curvature has 208 vertices and was constructed by Nicholson and Sneddon [47].

Next, we turn to results which complement the result above. An interpretation of the theorem below is that non-positive curvature implies that a planar tessellation is infinite.

Theorem 6.4 *Let $G = (V, E, F)$ be a tessellation embedded locally compactly in $\mathbb{S}^2$. Then, the graph admits some positive curvature.*

Proof The statement follows directly from the Gauss–Bonnet theorem. □

We end this section by a result of Chen [13] that states that an infinite tessellation with non-negative curvature can have at most finitely many vertices with positive curvature.

Theorem 6.5 ([13, Theorem 3.5.]) *Let G be a tessellation such that $\Phi \geq 0$. Then, the number of vertices with non-vanishing curvature is finite.*

6.2.4 Absence of Cut Locus

In Riemannian geometry the Hadamard–Cartan theorem states that on the universal cover of a complete manifold with non-positive sectional curvature all geodesics can be continued indefinitely.

For graphs such a theorem can be proven under the assumption of non-positive corner curvature. Below we will discuss examples how such a statement may fail for non-positive (or even negative) vertex curvature.

For a vertex v the *cut locus* is the set of vertices, for which the distance function $d(v, \cdot)$ assumes a local maximum. We say the graph has *empty cut locus* if the cut locus of every vertex is empty. This is equivalent to the fact that one can continue every finite geodesic indefinitely.

Theorem 6.6 ([6, Theorem 1]) *Let G be a planar tessellation such that $\Phi_C \leq 0$. Then, the graph has empty cut locus.*

The proof of the theorem involves an analysis of the boundary structure of distance balls of tessellations with non-positive corner curvature. In [6] it is shown that the balls satisfy a subtle notion of convexity which is referred to as admissibility.

The statement of the theorem can easily be seen to fail if one only assumes $\Phi \leq 0$. An example is the Kagome lattice of Example 6.4. This tessellation satisfies $\Phi \leq 0$ but has corners of positive curvature, so, it does not satisfy the assumption of Theorem 6.6. Indeed, for an arbitrary vertex o there are two vertices in B_3 that have only neighbors in B_3—these are the opposite vertices in the hexagons that contain o.

For negative vertex curvature, consider a tessellation with $2p$-gons, $p \geq 4$, instead of hexagons as also mentioned in Example 6.4. Then this tessellation satisfies $\Phi < 0$ but also has corners of positive curvature. Indeed, every vertex o has vertices in B_p that have no neighbor in S_{p+1}.

6.2.5 *Volume Growth*

We turn to implications of the sign of the curvature on the volume growth of planar tessellations. The rough synopsis of the exposition below is that non-negative curvature implies polynomial growth and for negative curvature implies at least exponential growth.

In [30] Hua, Jost and Liu showed that a tessellation of non-negative curvature does not grow faster than quadratically.

Theorem 6.7 ([30, Theorem 1.1]) *Let G be a tessellation embedded locally compactly in a surface of finite genus such that $\Phi \geq 0$. Then, there is a constant C depending only on the maximal face degree such that for all $r \geq 0$*

$$\#B_r \leq Cr^2.$$

The strategy of the proof is to relate volume growth of the tessellation to the one of the corresponding Alexandrov space in which it is canonically embedded.

Next, we turn to the case of negative curvature. We first discuss lower bounds on the volume growth that are proven in [5].

Theorem 6.8 ([5, Corollary 5.2.]) *Let $G = (V, E, F)$ be a planar tessellation that has empty cut locus and satisfies $\Phi \leq -k < 0$. Then,*

$$\#B_r \geq (1 + 2Ck)^r,$$

where $C = p/(p-1)$ if $p = \sup_{f \in F} \deg(f) < \infty$ and $C = 1$ otherwise.

The assumption of emptiness of cut locus is for example implied by non-positive corner curvature, Theorem 6.6.

Next, we turn to upper bounds. There is a rough immediate result by comparing the tessellation with a tree. Note that for a regular trees with vertex degree q, the number of vertices in S_r is exactly $(q-1)^r$. We estimate the volume growth of a

tessellation by removing edges within spheres and splitting geodesics from o that intersect in a vertex into two different ones. This results in the theorem below.

Theorem 6.9 ([35, Theorem 4]) *Let $G = (V, E, F)$ be a planar tessellation such that $q = \sup_{v \in V} d_v < \infty$. Then,*

$$\#B_r \leq (q-1)^r.$$

For tessellations whose faces have fixed degree p, one can give an explicit recursion formula for the size of the distances spheres. The recursion is expressed in terms of *normalized average curvatures* over spheres

$$\overline{\Phi}_r := \overline{\Phi}(S_r) := \left(\frac{2p}{p-2}\right) \frac{1}{\#S_r} \sum_{v \in S_r} \Phi(v).$$

Recall that the constant $2\pi(p-2)/2p$ is the internal angle $\beta(f)$ of a regular p-gon f.

Similar to the lower bound above the growth formula can be proven for tessellations without *cut locus* which is implied by non-positive corner curvature, Theorem 6.6. In the case of face regular graphs non-positive corner curvature is equivalent to non-positive curvature. However, the theorem below is not restricted to the non-positive curvature case.

For $3 \leq p < \infty$, let $N = \frac{p-2}{2}$ if p is even and $N = p-2$ if p is odd, and

$$b_l = \begin{cases} \frac{4}{p-2} - 2 & : \text{ if } p \text{ is odd and } l = \frac{N-1}{2}, \\ \frac{4}{p-2} & : \text{ else,} \end{cases}$$

for $0 \leq l \leq N-1$.

Theorem 6.10 ([35, Theorem 2]) *Let $G = (V, E, F)$ be a p-face regular tessellation with empty cut locus. Then the following $(N+1)$-step recursion formulas for $r \geq 1$ holds*

$$\#S_{r+1} = \begin{cases} \sum_{l=0}^{r-1} (b_l - \overline{\Phi}_{r-l}) \#S_{r-l} + \#S_1 & : \text{ if } r < N, \\ \sum_{l=0}^{N-1} (b_l - \overline{\Phi}_{N-l}) \#S_{N-l} & : \text{ if } r = N, \\ \sum_{l=0}^{N-1} (b_l - \overline{\Phi}_{r-l}) \#S_{r-l} - \#S_{r-N} & : \text{ if } r > N. \end{cases}$$

The $(N+1)$-step recursion formula gives rise to a *recursion matrix* M_r, $r \geq 0$, mapping $\mathbb{R}^N$ to $\mathbb{R}^N$ such that $M_r(\#S_{r-N}, \ldots, \#S_r) = (\#S_{r-N+1}, \ldots, \#S_{r+1})$.

In the special case when also the vertex degree is constant, say q, we have a (p, q)-regular tessellation. Then, the constant $b_l - \overline{\Phi}_k$ is equal to $q-2$, except for $l = (N-1)/2$ and q odd in which case we have $q-4$. In particular, there is a matrix M such that $M = M_r$ for all $r \geq 0$. The characteristic polynomial of M is then given

by the complex polynomial

$$g_{p,q}(z) = 1 - (q-2)z - \cdots - (q-2)z^N + z^{N+1},$$

if p is even, and

$$g_{p,q}(z) = 1 - (q-2)z - \cdots - (q-4)z^{\frac{N+1}{2}} - \cdots - (q-2)z^N + z^{N+1},$$

if p is odd. By Bartholdi and Ceccherini-Silberstein [2] and Cannon and Wagreich [11], $g_{p,q}$ is a reciprocal Salem polynomial, i.e., its roots lie on the complex unit circle except for two positive reciprocal real zeros

$$\frac{1}{x_{p,q}} < 1 < x_{p,q} < p - 1.$$

This yields

$$\limsup_{t\to\infty} \frac{1}{r} \log \#S_r = \log x_{p,q}$$

in the special case of a (p,q)-regular tessellation. In particular, the considerations above recover the results of Cannon and Wagreich [11] and Floyd and Plotnick [22, Sect. 3] that the growth function is a rational function.

To relate growth rate of an arbitrary face regular tessellation to a (p,q)-regular tessellation we present a volume growth comparison theorem which is an analogue to the Bishop-Guenther-Gromov comparison theorem from Riemannian geometry.

Theorem 6.11 (Theorem 3 in [35]) *Let $G = (V,E,F)$ and $\widetilde{G} = (\widetilde{V},\widetilde{E},\widetilde{F})$ be two p-face regular tessellations with non-positive vertex curvature and let $S_r \subset V$ and $\widetilde{S}_r \subset \widetilde{V}$ be spheres with respect to the centers $o \in V$ and $\widetilde{o} \in \widetilde{V}$, respectively. Assume that the normalized average spherical curvatures satisfy*

$$\overline{\Phi}(\widetilde{S}_r) \leq \overline{\Phi}(S_r) \leq 0, \quad r \geq 0.$$

Then the difference sequence $(\#\widetilde{S}_r - \#S_r)$ satisfies $\#\widetilde{S}_r - \#S_r \geq 0$, $r \geq 0$, and is monotone non-decreasing.

Proof (Idea of proof) The proof given in [35] depends heavily on the assumption of constant face degree. Let us briefly illustrate the underlying idea. We count the number of faces c_j^r that intersect a ball B_r in $1 \leq j \leq p$ vertices, where p is the constant face degree. If $j \leq p-2$, then this number equals the number of faces c_{j+2}^{r+1} that intersect the ball B_{r+1} in $j+2$ vertices. Finally, one can relate the numbers c_j^r to the number of vertices in a sphere S_r. □

An important quantity to relate volume growth to spectral theory is the exponential growth rate. For a set $W \subseteq V$, define

$$\mathrm{vol}(W) = \sum_{v \in W} d_v.$$

Let E_W be the set of edges that have both vertices in W. We find

$$\mathrm{vol}(W) = 2\#E_W + \#\partial W$$

which can be interpreted as twice the number of edges within W and once the number of edges leaving W. The *exponential growth rate* is defined as

$$\mu = \limsup_{r \to \infty} \frac{1}{r} \log \mathrm{vol}(B_r).$$

For connected graphs μ does not depend on the choice of the center o of the balls.

In order to relate the results above to μ we need the following lemma.

Lemma 6.1 *Let G be a tessellation of a surface of finite genus. Then,*

$$\mu = \limsup_{r \to \infty} \frac{1}{r} \log \#B_r = \limsup_{r \to \infty} \frac{1}{r} \log \#S_r.$$

Proof Let g be the genus of the surface. Using the fact that every face is included in two edges and every face includes at least three vertices we derive from Poincaré's formula for finite $W \subseteq V$

$$\#E_W \leq C\#W$$

with $C = \max\{2g + 1, 3\}$, (for details see [8, Lemma 6.2.]). We derive for $r \geq 1$

$$\#B_r \leq \mathrm{vol}(B_r) = 2\#E_{B_r} + \#\partial B_r \leq 2\#E_{B_{r+1}} \leq 2C\#B_{r+1}.$$

This yields the first equality. To see the second equality we observe first that $\#S_r \leq \#B_r$. Furthermore, let R_k be a subsequence realizing the lim sup. Let $0 \leq r_k \leq R_k$ such that $\#S_{r_k} \geq \#S_j$, $0 \leq j \leq R_k$. Then,

$$\frac{1}{R_k} \log \#B_{R_k} = \frac{1}{R_k} \log \left(\#S_{r_k} \sum_{j=0}^{R_k} \frac{\#S_j}{\#S_{r_k}} \right) \leq \frac{1}{r_k} \log \#S_{r_k} + \frac{1}{R_k} \log R_k.$$

Hence, the assertion follows by taking limits. □

6.2.6 Isoperimetric Inequalities

The isoperimetric problem goes back to a tale about Dido of Carthage. Here, we consider such a problem in the graph case and want to minimize the ratio of the boundary area and the volume of finite sets. The ratio is called the isoperimetric constants and it has many applications in mathematics and computer science, as e.g. applications to mixing times of Markov chains which is relevant in simulation, [41, 45].

A rough synopsis of this section is that non-negative curvature implies zero isoperimetric constant, while negative curvature implies positive isoperimetric constant.

In the negative curvature case, we further discuss an explicit formula for (p, q)-regular tessellations, give lower bounds as results of upper bounds on the curvature and give a criterion for positivity by asymptotically negative curvature. As a corollary we show that this implies Gromov hyperbolicity of the tessellation. Finally, we discuss an isoperimetric constant at infinity.

Let us be more precise. We let

$$\partial W = \{(v, w) \in W \times (V \setminus W) \mid v \sim w\}$$

for a finite set $W \subseteq V$. This set can be interpreted as the edges leaving W and the area of the boundary will be number of elements in ∂W. Furthermore, as in the previous section we let the volume of a finite set W be given by

$$\mathrm{vol}(W) = \sum_{v \in W} d_v.$$

We define the *isoperimetric constant* to be

$$\alpha = \inf_{W \subset V \text{finite}} \frac{\#\partial W}{\mathrm{vol}(W)}.$$

By the formula $\mathrm{vol}(W) = 2\#E_W + \#\partial W$ we conclude

$$0 \leq \alpha < 1.$$

One can also define the volume by the number of vertices of the set and define the isoperimetric constant accordingly. However, measuring volume by counting edges as above proves to be more effective for spectral estimates as the estimate does not become trivial for unbounded vertex degree; compare the results of [16] and [17]. Another choice for the area of the boundary is to involve so called intrinsic metrics, see [4]. Although the corresponding isoperimetric constant works well for spectral estimate, the constant is hard to estimates combinatorially in the tessellation case.

We first start with a theorem which is a corollary of Theorem 6.7.

Theorem 6.12 *Let G be a tessellation embedded in a surface of finite genus such that $\Phi \geq 0$. Then,*

$$\alpha = 0$$

Proof First of all notice that $\Phi \leq 0$ implies $d_v \leq 6$ by Proposition 6.1. Thus, $\#\partial B_r \leq 6\#S_r$. As $\mathrm{vol}(B_r) \geq \#B_r \geq \#S_r + \#S_{r-1}$, we derive

$$\alpha \leq \frac{\#\partial B_r}{\mathrm{vol}(B_r)} \leq \frac{6\#S_r}{\#S_r + \#S_{r-1}}.$$

We set $a = \frac{\alpha/6}{1-\alpha/6}$ and observe that $a > 1$ whenever $\alpha > 0$. We conclude by iteration $a^r \leq \#S_r$ which stands in contradiction to $\#B_r \leq Cr^2$ of Theorem 6.7. □

We consider now graphs with negative curvature. First, we mention that the isoperimetric constant can be calculated explicitly for (p, q)-regular tessellations. These formulas were independently obtained by Häggström et al. [26] and Higuchi and Shirai [29] by quite different techniques.

Theorem 6.13 ([26, 29]) *Let G be a (p, q)-regular tessellation such that $\Phi_C = \frac{1}{p} - \frac{1}{2} + \frac{1}{q} \leq 0$. Then,*

$$\alpha = \frac{q-2}{q}\sqrt{1 - \frac{4}{(p-2)(q-2)}}.$$

While [29] give a rather explicit construction of the minimizing sets, the authors of [26] use duality. We give a rough sketch of the proof.

Proof (Idea of proof [26]) Recall that E_W are the edges with both vertices in W and consider

$$\beta(G) = \inf_{W \subseteq V \text{ finite}} \frac{\#W}{\#E_W},$$
$$\delta(G) = \inf_{W \subseteq V \text{ finite}} \frac{\#W}{\#E_W + \#\partial W}.$$

It can be observed that

$$\beta(G) = \frac{2}{q(1-\alpha)} \quad \text{and} \quad \delta(G) = \frac{2}{q(1+\alpha)}.$$

For the proof of the theorem one needs to show that the minimizing sets for α, β and δ have to grow whenever $\alpha > 0$. Knowing this the major step in the proof is to show the relation

$$\beta(G) + \delta(G^*) = 1,$$

where G^* is the dual tessellation. Resolving this formula with the equations for β and δ above yields the statement. □

We now look at the non-regular case, where we have a uniform upper bound on the curvature. In this case we get explicit estimates on α. Although, these estimates are not sharp for planar tessellations, they are sharp if one considers sequences of tessellations where the face degrees grow to infinity. This is discussed in [35].

Whenever the face degree is bounded by some p and the vertex degree is bounded by some q the following constant $C_{p,q} \geq 1$ will enter the estimate of the isoperimetric constant below

$$C_{p,q} := \begin{cases} 1 & : \text{ if } p = \infty, \\ 1 + \frac{2}{p-2} & : \text{ if } p < \infty \text{ and } q = \infty, \\ \left(1 + \frac{2}{p-2}\right)\left(1 + \frac{2}{(p-2)(q-2)-2}\right) & : \text{ if } p, q < \infty. \end{cases}$$

Theorem 6.14 ([35, Theorem 1]) *Let G be a planar tessellation such that $\deg(f) \leq p$ for all $f \in F$ and $d_v \leq q$ for all $v \in V$ with $p, q \in [3, \infty]$. Assume $\Phi < 0$ and let $K := \inf_{v \in V} -\frac{1}{d_v}\Phi(v)$. Then*

$$\alpha \geq 2C_{p,q}K.$$

A key insight for the proof is a formula which is attributed in [5] to Harm Derksen. It is an immediate consequence of the Gauss–Bonnet theorem and direct calculation. It can be interpreted that the formula for the curvature of a simply connected finite set has the same form as the curvature of a vertex. Here, we call a set $W \subseteq V$ *simply connected* if W and $V \setminus W$ are connected.

Lemma 6.2 ([5, Proposition 2.1.]) *Let $W \subseteq V$ be a finite simply connected subset of a planar tessellation. Then,*

$$\sum_{v \in W} \Phi(v) = 1 - \frac{\#\partial W}{2} + \sum_{f \in F, f \cap W \neq \emptyset, f \cap (V \setminus W) \neq \emptyset} \frac{\#(f \cap W)}{\deg(f)}.$$

From this lemma, one immediately derives the estimate

$$\frac{\#\partial W}{\mathrm{vol}(W)} \geq \frac{-2\sum_{v \in W} \Phi(v)}{\mathrm{vol}(W)} \geq 2K.$$

To obtain the statement of Theorem 6.14 on still has to squeeze in $C_{p,q} \geq 1$. This needs some further rather subtle considerations.

Next, we turn to a result of Woess [57]. It states that if the asymptotic curvature defined as

$$\overline{\Phi} = \lim_{n \to \infty} \sup_{W \subseteq V, \#W \geq n} \frac{1}{\#W} \sum_{v \in W} \Phi(v)$$

is negative, then the isoperimetric constant is positive. The limit exists or is $-\infty$ as the sequence is monotone decreasing.

Theorem 6.15 ([57, Theorem A]) *Let G be a planar tessellation such that $\overline{\Phi} < 0$. Then, $\alpha > 0$.*

Again the result can be proven by the formula in Lemma 6.2 above which is however not stated explicitly in [57]. Preceding the result above, Dodziuk proved in [16] that planar graphs with $d_v \geq 7$, $v \in V$, satisfy $\alpha > 0$. In particular, this assumption implies $\Phi < 0$ by Proposition 6.1.

By considerations of Oh [48], we see that planar tessellations satisfying $\overline{\Phi} < 0$ are Gromov hyperbolic provided there is an upper bound on the face degree. A metric space is called *Gromov hyperbolic* if there is $\delta > 0$ such that every geodesic triangle T is δ-thin, i.e., any side of T is contained in the δ-neighborhood of the union of the other two sides. For graphs we consider the combinatorial graph distance d as a metric.

Corollary 6.1 ([48]) *Let G be a planar tessellation such that $\overline{\Phi} < 0$ and assume there is an upper bound on the face degree. Then, (G, d) is Gromov hyperbolic.*

Proof By Oh [48, Theorem 6] we see that a planar graph is Gromov hyperbolic if a constant called $\Phi(G)$, which is defined in [48], is positive. By Oh [48, Theorem 1] this constant $\Phi(G)$ is positive if and only if our isoperimetric constant α is positive. This equivalence holds true for all planar graphs whose faces are homeomorphic to a closed disk, an assumption which is satisfied for tessellations. □

To complement this result we remark that by a result of Cao [12] one can check that Gromov hyperbolicity of a planar tessellation with bounded vertex degree implies $\alpha > 0$. For details see [36, Proof of Theorem 3.8].

We end this section by discussing an asymptotic isoperimetric constant called the isoperimetric constant at infinity. This constant is discussed in [24, 32, 44]. We define the *isoperimetric constant at infinity* to be

$$\alpha_\infty = \sup_{K \subset V \text{finite}} \inf_{W \subset V \setminus K \text{ finite}} \frac{\#\partial W}{\mathrm{vol}(W)}.$$

Clearly, $\alpha \leq \alpha_\infty$. Furthermore, it is discussed in [3, Sect. 6.3] that $\alpha > 0$ if and only if $\alpha_\infty > 0$. The arguments employed there use basic spectral theory.

Accordingly, we define an upper bound on the *curvature at infinity* by

$$\Phi_\infty = \inf_{K \subset V \text{finite}} \sup_{v \in V \setminus K} \Phi(v).$$

For the curvature at infinity and the isoperimetric constant at infinity can be related as a consequence of [32] and Proposition 6.1.

Theorem 6.16 ([32, Proposition 6]) *Let G be a planar tessellation such that $\Phi_\infty = -\infty$. Then, $\alpha_\infty = 1$.*

6.3 Spectral Theory

In this section we study the spectral theory of the combinatorial Laplacian on tessellations. Many of the results are based on the geometric insights gathered in the previous section.

Since the definition of the combinatorial Laplacian does not depend on the tessellating structure, we introduce it on general graphs. It shall be stressed that although the Laplace operator considered here has the same quadratic form as the normalized Laplacian introduced in the previous Chap. 1, it is defined on a different Hilbert space. This arises from a different notion of volume. Indeed, there are two canonical ways to measure volume on graphs, either by counting vertices or edges. The volume considered by Jost is the function vol used in the previous section and is associated to counting edges. Here, we use the counting measure which means volume is determined by counting vertices.

Both viewpoints have there merits. The normalized Laplacian considered by Jost based on the edge volume captures perfectly the metric features which come from the combinatorial graph distance, see e.g. [34, Sect. 3.2.2.]. Moreover, this operator has the advantage that it is always bounded which avoids various technicalities. On the other hand, phenomena which are related to unbounded geometry are much better captured by the combinatorial Laplacian we study here. In summary both Laplacians arise from natural geometric considerations and each is suitable for the study of certain phenomena. Specifically, we study here discreteness of spectrum in the case of uniformly decreasing curvature together with eigenvalue asymptotics and decay properties of eigenfunctions. We also study spectral bounds implied by geometric data and consider unique continuation properties of eigenfunctions.

6.3.1 The Combinatorial Laplacian

In this section we introduce the combinatorial Laplacian on graphs. This operator has many applications in various fields of mathematics. We introduce the operator for general graphs since the restriction to tessellations yields no additional basic information or properties.

Let (V, E) be a graph. Then, the combinatorial Laplacian Δ acts on functions $\varphi : V \to \mathbb{R}$ as

$$\Delta\varphi(v) = \sum_{w \sim v} (\varphi(w) - \varphi(v)), \qquad v \in V.$$

In order to study spectral theory, we restrict Δ to a space of functions with more structure. We refer to [55] for a background on operator theory. Let $\ell^2(V)$ be the

space of square summable real valued functions, i.e.,

$$\ell^2(V) = \{\varphi : V \to \mathbb{R} \mid \sum_{v \in V} |\varphi(v)|^2 < \infty\}.$$

The space $\ell^2(V)$ equipped with the scalar product

$$\langle \varphi, \psi \rangle = \sum_{v \in V} \varphi(v)\psi(v)$$

is a Hilbert space. We denote the corresponding norm by $\| \cdot \|$.

One can check easily that Δ is a bounded operator on $\ell^2(V)$ if and only if

$$\sup_{v \in V} d_v < \infty.$$

In the case, where Δ is unbounded we have to restrict Δ to a dense subspace of $\ell^2(V)$ to define a selfadjoint operator. It was first shown in [58, Theorem 1.3.1.] that Δ is a selfadjoint on

$$D(\Delta) = \{\varphi \in \ell^2(V) \mid \Delta\varphi \in \ell^2(V)\}.$$

From now on we refer to Δ restricted to $D(\Delta)$ as the *combinatorial Laplacian*.

It can be shown that the functions of finite support $C_c(V)$ are dense in $D(\Delta)$ with respect to the graph norm. (Note that the graph norm is a functional analytic quantity referring to the norm $\|\varphi\| + \|\Delta\varphi\|$ on $D(\Delta)$.) Since

$$\langle \Delta\varphi, \varphi \rangle = -\frac{1}{2} \sum_{v,w \in V, v \sim w} (\varphi(v) - \varphi(w))^2,$$

the operator $-\Delta$ is positive. In what follows we study the spectral theory of $-\Delta$. Clearly, $D(\Delta) = D(-\Delta)$.

We denote the spectrum of $-\Delta$ by

$$\sigma(-\Delta) = \{z \in \mathbb{C} \mid -\Delta - z\mathrm{Id} \text{ has no bounded inverse}\},$$

where Id is the identity operator on $\ell^2(V)$. Since $-\Delta$ is selfadjoint and positive, we have by general theory

$$\sigma(-\Delta) \subseteq [0, \infty).$$

In the case $\sup_v d_v \leq D$, we even have

$$\sigma(-\Delta) \subseteq [0, 2D].$$

For finite graphs the spectrum only consists of eigenvalues of Δ. If the graph is infinite the space $\ell^2(V)$ is infinite dimensional and, therefore, there might be values $\lambda \in \sigma(\Delta)$ which do not allow for an $\ell^2(V)$ eigenfunction. However, by a criterion of Weyl one still has approximate eigenfunctions.

We denote the *bottom of the spectrum* of $-\Delta$ by $\lambda_0 = \lambda_0(-\Delta) = \min \sigma(-\Delta)$. We have by the Rayleigh-Ritz characterization and the density of $C_c(V)$ in $\ell^2(V)$

$$\lambda_0(-\Delta) = \inf_{\varphi \in C_c(V),\, \|\varphi\|=1} \langle (-\Delta)\varphi, \varphi \rangle.$$

In the case where $\lambda_0 > 0$ one says that $-\Delta$ has a *spectral gap*.

For finite graphs the constant functions are in $\ell^2(V)$ and are, therefore, eigenfunctions to the eigenvalue 0. For infinite graphs the constant functions are never in $\ell^2(V)$. However, in some cases it is still possible that the constant functions can be approximated with respect to graph norm by functions in $\ell^2(V)$ in which case there is no spectral gap.

It is said that $-\Delta$ has *pure discrete spectrum* if $\sigma(-\Delta)$ consists only of discrete eigenvalues of finite multiplicity. This implies that they only accumulate at infinity. Clearly, this can only happen if Δ is unbounded. In the case of pure discrete spectrum we denote the eigenvalues of $-\Delta$ by λ_n, $n \geq 0$, in increasing order counted with multiplicity.

The part of the spectrum which are no discrete eigenvalues of finite multiplicity is called the *essential spectrum* and is denoted by $\sigma_{\mathrm{ess}}(-\Delta)$. Furthermore, we let

$$\lambda_0^{\mathrm{ess}}(-\Delta) = \min \sigma_{\mathrm{ess}}(-\Delta).$$

6.3.2 Bottom of the Spectrum

Let us recall some well known results from discrete spectral geometry to estimate the bottom of the spectrum. These results are classically proven for general graphs and the normalized Laplacian which we denote here by $\widetilde{\Delta}$ (confer Chap. 1). We relate the bottom of spectra $\lambda_0(-\widetilde{\Delta})$ to $\lambda_0(-\Delta)$ and $\lambda_0^{\mathrm{ess}}(-\widetilde{\Delta})$ to $\lambda_0^{\mathrm{ess}}(-\Delta)$ as it was shown by very elementary arguments in [32]

$$m\lambda_0(-\widetilde{\Delta}) \leq \lambda_0(-\Delta) \leq \lambda_0^{\mathrm{ess}}(-\Delta) \leq M_\infty \lambda_0^{\mathrm{ess}}(-\widetilde{\Delta})$$

with $m = \min_{v \in V} d_v$ and $M_\infty = \inf_{K \subseteq V,\, \text{finite}} \sup_{v \in V \setminus K} d_v$ whenever $M_\infty < \infty$.

Recall the definition of the volume growth rate

$$\mu = \limsup_{r \to \infty} \frac{1}{r} \log \mathrm{vol}(B_r)$$

from Sect. 6.2.5 and the isoperimetric constant

$$\alpha = \inf_{W \subset V \text{finite}} \frac{\#\partial W}{\mathrm{vol}(W)}.$$

from Sect. 6.2.6. Then, by results of Fujiwara [23, 24]

$$m\big(1 - \sqrt{1-\alpha^2}\big) \leq \lambda_0(-\Delta) \leq \lambda_0^{\mathrm{ess}}(-\Delta) \leq M_\infty\Big(1 - \frac{2e^{\mu/2}}{e^\mu + 1}\Big).$$

The lower bound was preceded by a result of Dodziuk and Kendall [17] and is found in a similar form in the work of Mohar [43]. The upper bound is preceded by results of Dodziuk and Karp [18] and Ohno and Urakawa [49].

We first use the upper bound in the case of non-negative curvature. We conclude by the estimate above and Theorem 6.7.

Corollary 6.2 *Let G be a tessellation embedded in a surface of finite genus such that $\Phi \geq 0$. Then, $\lambda_0(-\Delta) = \lambda_0^{\mathrm{ess}}(-\Delta) = 0$.*

Furthermore, we use Theorem 6.15 to derive the following.

Theorem 6.17 *Let G be a planar tessellation such that $\overline{\Phi} < 0$. Then, $\lambda_0 > 0$.*

Finally, we combine both estimates above with Theorem 6.11 and Theorem 6.12.

Theorem 6.18 *Let G be a planar tessellation such that $\deg(f) \leq p$ for all $f \in F$ and $d_v \leq q$ for all $v \in V$ with $p, q \in [3, \infty]$. Assume $\Phi < 0$ and let $K = \inf_{v \in V} -\frac{1}{d_v}\Phi(v)$. Then,*

$$2mK^2 \leq m(1 - \sqrt{1 - 4C_{p,q}^2 K^2}) \leq \lambda_0(-\Delta),$$

where $C_{p,q}$ is defined in Sect. 6.2.6. If additionally $p, q < \infty$, then

$$\lambda_0^{\mathrm{ess}}(-\Delta) \leq M_\infty\Big(1 - \frac{2x_{p,q}^{1/2}}{x_{p,q} + 1}\Big),$$

where $x_{p,q}$ is the largest real zero of $g_{p,q}$ defined in Sect. 6.2.5.

The lower bound above can be considered as a discrete analogue to a theorem of McKean [42]. McKean proved for a n-dimensional complete Riemannian manifold M with upper sectional curvature bound $-k$ that the bottom of the spectrum of the Laplace-Beltrami $-\Delta_M$ satisfies

$$(n-1)^2 k/4 \leq \lambda_0(-\Delta_M).$$

It shall be noted that by Theorem 6.2 the assumption $\Phi < 0$ implies $K > 0$ and, therefore, $\lambda_0(-\Delta) > 0$ in this case.

6.3.3 Discrete Spectrum, Eigenvalue Asymptotics and Decay of Eigenfunctions

We now turn to a characterization of pure discrete spectrum. The theorem below is an analogue of a theorem of Donnelly/Li [20] from Riemannian geometry. Recall the definition of the upper asymptotic curvature bound

$$\Phi_\infty = \sup_{K \subset V \text{ finite}} \inf_{v \in V \setminus K} \Phi(v)$$

Theorem 6.19 ([32, Theorem 3]) *Let G be a planar tessellation. Then the spectrum of $-\Delta$ is purely discrete if and only if $\Phi_\infty = -\infty$.*

Proof (Idea of proof) If the spectrum of $-\Delta$ is purely discrete, then, by a criterion attributed to Perrson in the continuous case, $\langle(-\Delta)\varphi_n, \varphi_n\rangle \to \infty$ for every normalized sequence (φ_n) in $\ell^2(V)$ which converges weakly to zero. For a sequence (v_n) of vertices and the delta functions δ_{v_n}, we have

$$\langle(-\Delta)\delta_{v_n}, \delta_{v_n}\rangle = d_{v_n} \leq -2\Phi(v_n),$$

where the last inequality follows from Proposition 6.1. This implies $\Phi_\infty = -\infty$.

On the other hand, assume $\Phi_\infty = -\infty$. By Theorem 6.14 we infer that outside of large enough finite sets the isoperimetric constant is uniformly positive. Moreover, by an argument as in Theorem 6.18 the bottom of the spectrum of the operator restricted to functions supported outside of larger and larger finite sets converges to ∞. This, however, is equivalent to pure discrete spectrum, as the restricted operators are finite rank perturbations of the original operator. □

In [24] Fujiwara proved a similar statement as the theorem above for the normalized Laplacian on trees, namely that spectrum is discrete except for the point 1, where the discrete eigenvalues accumulate. Furthermore, Wojciechowski [58] showed discreteness of the spectrum of $-\Delta$ on general graphs in terms of a different curvature quantity sometimes referred to as a mean curvature.

We observe the following standard fact. Whenever, the spectrum of $-\Delta$ is purely discrete, then $\lambda_0(-\Delta) > 0$: If 0 was in the spectrum, then it must be a discrete eigenvalue. However, the only eigenfunctions φ to 0 which have finite energy, (i.e., $\sum_{v\sim w}(\varphi(v)-\varphi(w))^2 < \infty$) are the constant functions which are not in $\ell^2(V)$ in the case of infinite graphs).

In the case of discrete spectrum we next present asymptotics for eigenvalue λ_n of $-\Delta$, i.e.,

$$\sigma(-\Delta) = \{0 < \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_n \leq \ldots\}$$

For the case $\Phi_\infty = -\infty$, it follows from Proposition 6.1 that the vertices can be ordered $V = \{v_n\}$ such that

$$d_{v_n} \leq d_{v_{n+1}}, \qquad n \geq 0.$$

Theorem 6.20 ([7, Theorem 1.6.]) *Let G be a planar tessellation such that $\Phi_\infty = -\infty$. Then,*

$$d_{v_n} - 2\sqrt{d_{v_n}} \lesssim \lambda_n \lesssim d_{v_n} + 2\sqrt{d_{v_n}},$$

that is

$$\lim_{n\to\infty} \frac{\lambda_n}{d_{v_n}} = 1$$

and

$$-2 \leq \liminf_{n\to\infty} \frac{\lambda_n - d_{v_n}}{\sqrt{d_{v_n}}} \leq \limsup_{n\to\infty} \frac{\lambda_n - d_{v_n}}{\sqrt{d_{v_n}}} \leq 2.$$

Proof (Idea of proof) The proof in [7] is divided into two steps. First one shows that every tessellation whose curvature decreases to $-\infty$ allows for a spanning tree such that the combinatorial Laplacian on the tessellation and the tree are bounded perturbations of each other. Equivalently this means that there is a number N such that from each vertex at most N adjacent edges are canceled to obtain the spanning tree. By the Min-Max-Principle both operators share the same eigenvalue asymptotics.

Secondly, one shows the corresponding eigenvalue for the operator on the tree. The proof uses isoperimetric techniques and again a version of the Min-Max-Principle. □

In the case of planar tessellations with constant face degree one can show bounds with an even more geometric flavor. Assume all faces have degree p. Then, the internal angle of every face f is given by

$$\beta(f) = \beta(p) = 2\pi \frac{(p-2)}{p}.$$

Corollary 6.3 ([7, Corollary 1.8.]) *Let G be a planar tessellation such that $\Phi_\infty = -\infty$. Suppose the face degree is constantly $p \geq 3$ outside of some finite set. Then, for large n*

$$-\frac{2\pi\Phi(v_n)}{\beta(p)} - 2\sqrt{-\frac{2\pi\Phi(v_n)}{\beta(p)}} \lesssim \lambda_n \lesssim -\frac{2\pi\Phi(v_n)}{\beta(p)} + 2\sqrt{-\frac{2\pi\Phi(v_n)}{\beta(p)}},$$

that is

$$-\frac{2\pi\Phi(v_n)}{\lambda_n} \to \beta(p),$$

and

$$-\frac{\beta(p)}{2} \leq \liminf_{n\to\infty} \frac{\sqrt{-2\pi\Phi(v_n)}}{\lambda_n + 2\pi\Phi(v_n)/\beta(p)} \leq \limsup_{n\to\infty} \frac{\sqrt{-2\pi\Phi(v_n)}}{\lambda_n + 2\pi\Phi(v_n)/\beta(p)} \leq \frac{\beta(p)}{2}.$$

The eigenvalue asymptotics in the theorem and the corollary above present a case where phenomena in the discrete and continuous world drift apart. In particular, for Riemannian manifolds one expects upper bounds of eigenvalues λ_k by some constant multiplied by k^2, see [38].

Next, we come to the decay of eigenfunctions. It turns out that eigenfunctions decay exponentially in an ℓ^2 sense.

Theorem 6.21 ([37]) *If $\Phi_\infty = -\infty$ and $\varphi_n \in D(\Delta)$, $n \geq 0$, are eigenfunctions, i.e.,*

$$\Delta\varphi_n = \lambda_n\varphi_n,$$

then, for any $\beta < e^{-1}$ and $o \in V$,

$$|\Phi|^{\frac{1}{2}} e^{\beta d(o,\cdot)}\varphi_n \in \ell^2(V),$$

where $d(\cdot,\cdot)$ is the natural graph distance.

The proof is based on ideas of Agmon for Schrödinger operators in $\mathbb{R}^n$. The somewhat curious constant e^{-1} comes in via an optimization that is necessary by the non-locality of the graph Laplacian in contrast to the strongly local Laplace operator on $\mathbb{R}^n$.

6.3.4 Unique Continuation of Eigenfunctions

Under the phenomena unique continuation one understands that a function which is zero outside of a compact set must also be zero on the compact set. Such unique continuation properties for eigenfunctions hold in great generality for local elliptic operators. Often very strong quantitative statements, called Carleman estimates, can be proven which all have the basic corollary that there are no eigenfunctions supported on a compact subset of the space. However, for graphs such eigenfunctions can be produced rather easily. This was observed by many authors, see e.g. [1, 9, 19, 39]. Below we will also discuss examples. The purpose

of this section is to discuss that non-positive corner curvature excludes such a phenomena, i.e., there are no compactly eigenfunctions.

Klassert et al. [40] proved a unique continuation result for tessellations with non-positive corner curvature. This result was later generalized to planar graphs in [33] with a different proof.

Theorem 6.22 ([40, Theorem 4]) *Let G be a planar tessellation such that $\Phi_C \leq 0$. Then, there are no eigenfunctions of compact support.*

Proof (Idea of proof) The proof given in [33] works by a polar decomposition of the Laplacian into $\Delta = E^\top + D + E$, where D is the combinatorial Laplacian on the spheres with Dirichlet boundary conditions and E is the operator with matrix elements $E(v, w) = 1$ if $v \in S_{r+1}$, $w \in S_r$, $v \sim w$ for all $r \geq 0$ and zero otherwise. Hence, the operator E encodes the edges reaching a sphere from the previous one. Denote the restriction of E and D to the functions supported on S_r by E_r and D_r. For an eigenfunction φ let φ_r be the restriction of φ to S_r. Then, the eigenvalue equation $\Delta\varphi = \lambda\varphi$ reads in polar decomposition as

$$E_r\varphi_r + (D - \lambda)\varphi_{r+1} + E_{r+1}^\top\varphi_{r+2} = 0, \qquad r \geq 0.$$

Now, by subtle geometric arguments it is shown by induction over r that E_r is injective. Therefore, if φ vanishes outside of B_r, the above equation reads as $E_r\varphi_r = 0$ which implies $\varphi_r = 0$ by injectivity of E_r. □

Let us discuss, that the statement of the theorem fails, if one only assumes $\Phi \leq 0$.

Example 6.5 Consider the Kagome lattice of Example 6.4. Pick a hexagon and denote the vertices of the adjacent triangles which are not in the hexagon by $v_1, \ldots, v_6$. Now let $\varphi : V \to \{-1, 1\}$ be supported on $v_1, \ldots, v_6$ such that $\varphi(v_j) = (-1)^j$, $j = 1, \ldots, 6$. Then it follows that $\Delta\varphi = 6\varphi$ and φ is therefore an eigenfunction of compact support. The same idea works if one considers the other examples of Example 6.4 when we replace the hexagon with a $2p$-gon with $p > 3$. In this case we even have $\Phi < 0$.

A question that arises is whether it is sufficient that the corner curvature is non positive outside of a finite set in order to have at most finitely many compactly supported eigenfunctions. Still this is not the case as the example in Fig. 6.4 shows.

However, if the curvature is sufficiently large outside of a finite set a unique continuation result can be proven.

Theorem 6.23 ([7, Theorem 1.9.]) *Let G be a planar graph. Assume $\Phi_\infty = -\infty$. Then, outside of a finite set there are no eigenfunctions of compact support of Δ. In particular, there are at most finitely many linearly independent eigenfunctions of finite support.*

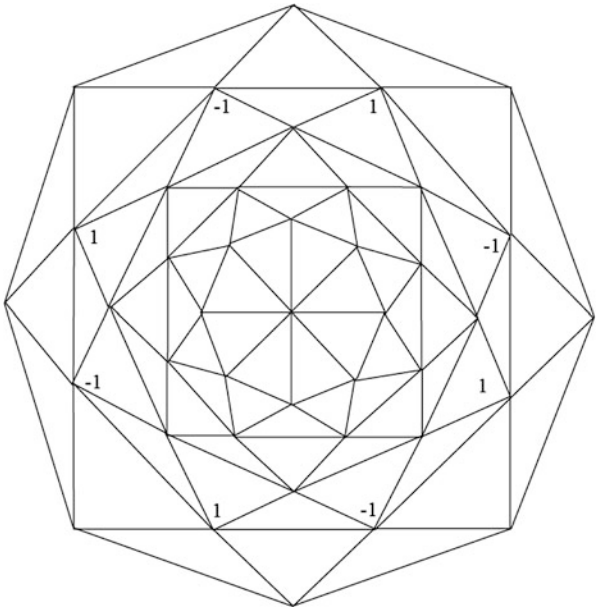


Fig. 6.4 The first six distance spheres of a tessellation that admits infinitely many linearly independent compactly supported eigenfunctions and has $\Phi_C \leq 0$ outside of B_2

6.3.5 ℓ^p Spectrum

In this section we study the spectral theory of $-\Delta$ as an operator on the Banach spaces $\ell^p(V)$, $p \in [1, \infty]$. Denote the restriction of Δ to

$$D(\Delta_p) = \{f \in \ell^p(V) \mid \Delta f \in \ell^p(V)\}$$

by Δ_p and let $\Delta_\infty = \Delta_1^*$. Clearly, $D(\Delta_p) = D(-\Delta_p)$.

A question asked by Simon [50] and affirmatively answered by Hempel and Voigt [27] for Schrödinger operators is whether the spectrum depends on the underlying Banach space. Sturm, [53], addressed this question in the setting of uniformly elliptic operators on manifolds. As a special case, he considers consequences of curvature bounds. For the tessellation case the following result is found in [3].

Theorem 6.24 ([3, Theorem 7.1., 7.2. and 7.3.]) *Let G be a planar tessellation.*

(a) If $\Phi \geq 0$, then $\sigma(-\Delta_2) = \sigma(-\Delta_p)$ for $p \in [1, \infty]$.
(b) If $-K \leq \Phi < 0$, then $\lambda_0(-\Delta_2) \neq \lambda_0(-\Delta_1)$, i.e., $\sigma(-\Delta_2) \neq \sigma(-\Delta_1)$.
(c) If $\Phi_\infty = -\infty$, then $\sigma(-\Delta_2) = \sigma(-\Delta_p)$ for all $p \in (1, \infty)$.

6.4 Extensions to More General Graphs

In this final section we discuss how the results of the previous section can be generalized to planar graphs and more general polygonal complexes.

6.4.1 Curvature on Planar Graphs

First we address general planar graphs as it was studied in [33]. The first step is to extend the notion of curvature. Secondly, we show that non-positive already implies that the graph looks locally like a tessellation. Such graphs allow for a suitable embedding into a tessellation which is used to extend the results for tessellations to general planar graphs.

Let (V, E) be a planar graph which is embedded locally compactly into a surface homeomorphic to $\mathbb{R}^2$ which gives rise to faces F as above. To define curvature on planar graphs, we have to extend the definitions of degrees of faces and vertices. In order to do so, we introduce the degree of a corner. As above the corners $C(G)$ are the pairs $(v,f) \in V \times F$ such that $v \in f$. For a corner $(v,f) \in C(G)$ we define the *degree* $|(v,f)|$ by the minimal number of times the vertex v is met by a boundary walk of f. Roughly speaking the degree of the corner (v,f) is the number of times f touches v.

This gives rise to a definition of the *degree* of $v \in V$ and $f \in F$ by

$$d_v = \sum_{(v,g)\in C(G)} |(v,g)| \quad \text{and} \quad \deg(f) = \sum_{(w,f)\in C(G)} |(w,f)|.$$

In a tessellation the degree of a corner is always one. So, these definitions indeed extend the ones of tessellations.

A face f is called *unbounded* if $\deg(f) = \infty$. Moreover, a planar graph is simple if and only if $\deg(f) \geq 3$ for all $f \in F$.

We define the *corner curvature* $\Phi_C : C(G) \to \mathbb{R}$ by

$$\Phi_C(v,f) = \frac{1}{d_v} - \frac{1}{2} + \frac{1}{\deg(f)}$$

and the *vertex curvature* by $\Phi : V \to \mathbb{R}$ by

$$\Phi(v) = \sum_{(v,f)\in C(G)} |(v,f)|\Phi_C(v,f).$$

These definitions are consistent with the definition of Φ_C and Φ on tessellations. Furthermore, they allow to show a Gauss–Bonnet formula, [33, Proposition 1].

It turns out that non-negative curvature already has strong consequences on the structure of the graph. To this end, we look at a generalization of tessellations. We call a face a *polygon* if it is homeomorphic to an open disc and we call it an *infinigon* if it is homeomorphic to the upper half space in $\mathbb{R}^2$. For example the faces of a tessellation are all polygons by (T3) and the faces of a tree are infinigons.

We call a planar graph *locally tessellating* if it satisfies the following properties:

(T1) Every edge is contained in two faces.
(T2*) Two faces are either disjoint or intersect in a vertex or in a path of edges. If this path consists of more than one edge then both faces are unbounded.
(T3*) Every face is a polygon or an infinigon.

The assumption (T1) is the same as in the definition of tessellations.

Tessellations, trees as well as hybrids of both of these are examples of locally tessellating graphs. As mentioned above non-positive curvature on planar graphs implies a graph is almost a tessellation.

Theorem 6.25 ([33, Theorem 1]) *Let G be a connected planar graph. If $\Phi_C \leq 0$ or if G is simple with $\Phi \leq 0$ then G is locally tessellating and infinite.*

Proof (Idea of proof) To prove this theorem one isolates finite subgraphs with simply closed boundary of the tessellation on which some of the assumptions (T1), (T2*), (T3*) fail. Then one copies this subgraph finitely many times and pastes the copies along their boundary paths to be finally embedded into the unit sphere. Now, the Gauss–Bonnet theorem implies that there must be some positive curvature. □

By Keller [33, Theorem 2] locally tessellating graphs can be embedded into tessellations in a way that approximates the original curvature arbitrarily close.

Theorem 6.26 *Let G be a locally tessellating graph that satisfies $\Phi \leq 0$. Let $W \subset V$ be a finite and simply connected set and $\varepsilon \in (0, 1/1806)$. Then, there is a tessellation G' which is a supergraph of G such that the following properties hold.*

(a) The embedding of G into the supergraph G' is a graph isomorphism of the subgraphs G_W and G'_W and the embedding is an isometry on W.
(b) If $\Phi_C \leq 0$, then $\Phi'_C \leq \Phi_C + \varepsilon$ and if $\Phi \leq 0$, then $\Phi' \leq \Phi + \varepsilon$.

By doing so, one can carry over results from tessellations to locally tessellating graphs and by Theorem 6.26 above to planar graphs in the case of non-positive curvature.

Among the geometric applications are the following:

- Absence of cut locus for non-positive corner curvature [33, Theorem 3].
- Bounds for the growth of distance balls for negative vertex curvature [33, Theorem 5].
- Positivity and bounds for the isoperimetric constant for negative vertex curvature [33, Theorem 6].
- Gromov hyperbolicity for negative corner curvature [33, Theorem 7].

There are also applications in spectral theory.

- The geometric bounds yield bounds on the bottom of the spectrum.
- Uniformly decreasing curvature is equivalent to purely discrete spectrum [33, Theorem 8].
- The same eigenvalue asymptotics as in Theorem 6.20 [7], and the same decay of eigenfunctions as in Theorem 6.21 [37].
- Unique continuation results for eigenfunctions [33, Theorem 9].

6.4.2 *Sectional Curvature for Polygonal Complexes*

To end this chapter we discuss generalizations to non planar graphs for which we can define a notion of sectional curvature. In [56] such a program was undertaken to study questions in group theory. In [36] another somewhat more restrictive approach was taken to define curvature for polygonal complexes which however allows to prove various results in geometry and spectral theory.

Let us be more precise. In [36] a *polygonal complex* $G = (V, E, F)$ is said to have *planar substructures* if there is a system $\mathscr{A}$ of subcomplexes called the *apartments* that satisfy the following axioms:

(PCPS1) For every two faces there is an apartment which contains both of them.
(PCPS2) The apartments are convex, that is every geodesic of faces, which starts and ends in an apartment, stays completely in the apartment.
(PCPS3) The apartments are planar or spherical tessellations.

An important example of such polygonal complexes are two dimensional buildings.

We define the degree $\deg(f)$ of a face $f \in F$ as in the case of tessellation by the number of edges included in f. Moreover, we denote the degree of a vertex v with respect to an apartment Σ by $d_v^{(\Sigma)}$. The corners of an apartment are denote by $C_\Sigma(G)$. For an apartment $\Sigma = (V_\Sigma, E_\Sigma, F_\Sigma)$, we define the *sectional corner curvature* $\Phi_C : C_\Sigma(G) \to \mathbb{R}$ with respect to Σ by

$$\Phi_C^{(\Sigma)}(v,f) = \frac{1}{d_v^\Sigma} - \frac{1}{2} + \frac{1}{\deg(f)}$$

and the *sectional face curvature* by $\Phi_F^{(\Sigma)} : F_\Sigma \to \mathbb{R}$ by

$$\Phi_F^{(\Sigma)}(f) = \sum_{(v,f)\in C_\Sigma(G)} \Phi_C^{(\Sigma)}(v,f) = 1 - \frac{\deg(f)}{2} + \sum_{v\in V_\Sigma, v\in f} \frac{1}{d_v^\Sigma}.$$

With this definition one can prove various of the results of Sects. 6.2 and 6.3. Here, we refrain from stating the results precisely but only mention them and refer

to [36] for details:

- Finiteness and infiniteness depending on the sign of the curvature, [36, Theorem 3.13.]
- Absence of cut locus for non-positive sectional corner curvature [36, Theorem 3.1.]
- Positivity and bounds for an isoperimetric constant for negative sectional face curvature, [36, Theorems 3.8. and 3.11.]
- Gromov hyperbolicity for negative sectional corner curvature, [36, Theorem 3.6.]

To study spectral theory one considers the combinatorial Laplacian on functions defined on faces. This is due to the fact that the geometric assumptions for the polygonal complexes are made with respect to the faces. For a function $\varphi : F \to \mathbb{R}$, we define

$$\Delta\varphi(f) = \sum_{g \in F,\ g \sim f} (\varphi(g) - \varphi(f)),$$

where $g \sim f$ means that f and g share an edge. Restricting Δ to

$$D(\Delta) = \{\varphi \in \ell^2(F) \mid \Delta\varphi \in \ell^2(F)\}$$

gives rise to a selfadjoint operator. For this operator one can prove:

- Discreteness of spectrum and eigenvalue asymptotics under the assumption of uniform decreasing corner curvature, [36, Theorem 4.1.]
- Unique continuation of eigenfunctions, [36, Theorem 4.3.].

Acknowledgements MK enjoyed the hospitality of C.I.R.M. and acknowledges the financial support of the German Science Foundation (DFG).

References

1. Aizenman, M., Warzel, S.: The canopy graph and level statistics for random operators on trees. Math. Phys. Anal. Geom. **9**(4), 291–333 (2007) (2006). doi:10.1007/s11040-007-9018-3. http://dx.doi.org/10.1007/s11040-007-9018-3
2. Bartholdi, L., Ceccherini-Silberstein, T.G.: Salem numbers and growth series of some hyperbolic graphs. Geom. Dedicata **90**, 107–114 (2002). http://dx.doi.org/10.1023/A:1014902918849
3. Bauer, F., Hua, B., Keller, M.: On the l^p spectrum of Laplacians on graphs. Adv. Math. **248**, 717–735 (2013). http://dx.doi.org/10.1016/j.aim.2013.05.029
4. Bauer, F., Keller, M., Wojciechowski, R.K.: Cheeger inequalities for unbounded graph Laplacians. J. Eur. Math. Soc. (JEMS) **17**(2), 259–271 (2015). http://dx.doi.org/10.4171/JEMS/503
5. Baues, O., Peyerimhoff, N.: Curvature and geometry of tessellating plane graphs. Discrete Comput. Geom. **25**(1), 141–159 (2001). http://dx.doi.org/10.1007/s004540010076

6. Baues, O., Peyerimhoff, N.: Geodesics in non-positively curved plane tessellations. Adv. Geom. **6**(2), 243–263 (2006). http://dx.doi.org/10.1515/ADVGEOM.2006.014
7. Bonnefont, M., Golénia, S., Keller, M.: Eigenvalue asymptotics and unique continuation of eigenfunctions on planar graphs (2014). Preprint
8. Bonnefont, M., Golénia, S., Keller, M.: Eigenvalue asymptotics for Schrödinger operators on sparse graphs. Ann. Inst. Fourier (Grenoble) **65**(5), 1969–1998 (2015). http://aif.cedram.org/item?id=AIF_2015__65_5_1969_0
9. Breuer, J., Keller, M.: Spectral analysis of certain spherically homogeneous graphs. Oper. Matrices **7**(4), 825–847 (2013)
10. Byung-Geun, O.: On the number of vertices of positively curved planar graphs. Discret. Math. **340**(6), 1300–1310 (2017). doi:10.1016/j.disc.2017.01.025, ISSN:0012-365X, http://dx.doi.org/10.1016/j.disc.2017.01.025
11. Cannon, J.W., Wagreich, P.: Growth functions of surface groups. Math. Ann. **293**(2), 239–257 (1992). http://dx.doi.org/10.1007/BF01444714
12. Cao, J.: Cheeger isoperimetric constants of Gromov-hyperbolic spaces with quasi-poles. Commun. Contemp. Math. **2**(4), 511–533 (2000). http://dx.doi.org/10.1142/S0219199700000232
13. Chen, B.: The Gauss-Bonnet formula of polytopal manifolds and the characterization of embedded graphs with nonnegative curvature. Proc. Am. Math. Soc. **137**(5), 1601–1611 (2009). http://dx.doi.org/10.1090/S0002-9939-08-09739-6
14. Chen, B., Chen, G.: Gauss-Bonnet formula, finiteness condition, and characterizations of graphs embedded in surfaces. Graphs Combinatorics **24**(3), 159–183 (2008). http://dx.doi.org/10.1007/s00373-008-0782-z
15. 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
16. Dodziuk, J.: Difference equations, isoperimetric inequality and transience of certain random walks. Trans. Am. Math. Soc. **284**(2), 787–794 (1984). http://dx.doi.org/10.2307/1999107
17. Dodziuk, J., Kendall, W.S.: Combinatorial Laplacians and isoperimetric inequality. In: From Local Times to Global Geometry, Control and Physics (Coventry, 1984/85). Pitman Research Notes in Mathematics Series, vol. 150, pp. 68–74. Longman Sci. Tech., Harlow (1986)
18. Dodziuk, J., Karp, L.: Spectral and function theory for combinatorial Laplacians. In: Geometry of Random Motion (Ithaca, N.Y., 1987). Contemporary Mathematics, vol. 73, pp. 25–40. American Mathematical Society, Providence (1988). http://dx.doi.org/10.1090/conm/073/954626
19. Dodziuk, J., Linnell, P., Mathai, V., Schick, T., Yates, S.: Approximating L^2-invariants and the Atiyah conjecture. Commun. Pure Appl. Math. **56**(7), 839–873 (2003). http://dx.doi.org/10.1002/cpa.10076. Dedicated to the memory of Jürgen K. Moser
20. Donnelly, H., Li, P.: Pure point spectrum and negative curvature for noncompact manifolds. Duke Math. J. **46**(3), 497–503 (1979). http://projecteuclid.org/euclid.dmj/1077313570
21. Federico, P.J.: Descartes on polyhedra. Sources in the History of Mathematics and Physical Sciences, vol. 4. Springer, New York/Berlin (1982). A study of the ıt De solidorum elementis
22. Floyd, W.J., Plotnick, S.P.: Growth functions on Fuchsian groups and the Euler characteristic. Invent. Math. **88**(1), 1–29 (1987). http://dx.doi.org/10.1007/BF01405088
23. Fujiwara, K.: Growth and the spectrum of the Laplacian of an infinite graph. Tohoku Math. J. (2) **48**(2), 293–302 (1996). http://dx.doi.org/10.2748/tmj/1178225382
24. Fujiwara, K.: The Laplacian on rapidly branching trees. Duke Math. J. **83**(1), 191–202 (1996). http://dx.doi.org/10.1215/S0012-7094-96-08308-8
25. Gromov, M.: Hyperbolic groups. In: Essays in Group Theory. Mathematical Sciences Research Institute Publications, vol. 8, pp. 75–263. Springer, New York (1987). http://dx.doi.org/10.1007/978-1-4613-9586-7_3
26. Häggström, O., Jonasson, J., Lyons, R.: Explicit isoperimetric constants and phase transitions in the random-cluster model. Ann. Probab. **30**(1), 443–473 (2002). http://dx.doi.org/10.1214/aop/1020107775

27. Hempel, R., Voigt, J.: The spectrum of a Schrödinger operator in $L_p(\mathbf{R}^\nu)$ is p-independent. Commun. Math. Phys. **104**(2), 243–250 (1986). URL http://projecteuclid.org/euclid.cmp/1104115001
28. Higuchi, Y.: Combinatorial curvature for planar graphs. J. Graph Theory **38**(4), 220–229 (2001). http://dx.doi.org/10.1002/jgt.10004
29. Higuchi, Y., Shirai, T.: Isoperimetric constants of (d,f)-regular planar graphs. Interdiscip. Inform. Sci. **9**(2), 221–228 (2003). http://dx.doi.org/10.4036/iis.2003.221
30. 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
31. Ishida, M.: Pseudo-curvature of a graph. Lecture at 'Workshop on topological graph theory', Yokohama National University (1990)
32. Keller, M.: The essential spectrum of the Laplacian on rapidly branching tessellations. Math. Ann. **346**(1), 51–66 (2010). http://dx.doi.org/10.1007/s00208-009-0384-y
33. Keller, M.: Curvature, geometry and spectral properties of planar graphs. Discrete Comput. Geom. **46**(3), 500–525 (2011). http://dx.doi.org/10.1007/s00454-011-9333-0
34. Keller, M.: Intrinsic metrics on graphs – A survey. In: Mugnolo, D. (ed.) Mathematical Technology of Networks (Proc. Bielefeld 2013). Proceedings in Mathematics and Statistics. Springer, New York (2014)
35. Keller, M., Peyerimhoff, N.: Cheeger constants, growth and spectrum of locally tessellating planar graphs. Math. Z. **268**(3–4), 871–886 (2011). http://dx.doi.org/10.1007/s00209-010-0699-0
36. Keller, M., Peyerimhoff, N., Pogorzelski, F.: Sectional curvature of polygonal complexes with planar substructures. Adv. Math. **307**, 1070–1107 (2017)
37. Keller, M., Lenz, D.: Agmon type estimates and purely discrete spectrum for graphs. Preprint
38. Keller, M., Liu, S., Peyerimhoff, N.: A note on eigenvalue bounds for non-compact manifolds, Preprint 2017, arXiv:1706.02437
39. Klassert, S., Lenz, D., Stollmann, P.: Discontinuities of the integrated density of states for random operators on Delone sets. Commun. Math. Phys. **241**(2–3), 235–243 (2003)
40. Klassert, S., Lenz, D., Peyerimhoff, N., Stollmann, P.: Elliptic operators on planar graphs: unique continuation for eigenfunctions and nonpositive curvature. Proc. Am. Math. Soc. **134**(5), 1549–1559 (2006). http://dx.doi.org/10.1090/S0002-9939-05-08103-7
41. Levin, D.A., Peres, Y., Wilmer, E.L.: Markov Chains and Mixing Times. American Mathematical Society, Providence (2009). With a chapter by James G. Propp and David B. Wilson
42. McKean, H.P.: An upper bound to the spectrum of Δ on a manifold of negative curvature. J. Differ. Geom. **4**, 359–366 (1970)
43. Mohar, B.: Isoperimetric inequalities, growth, and the spectrum of graphs. Linear Algebra Appl. **103**, 119–131 (1988). http://dx.doi.org/10.1016/0024-3795(88)90224-8
44. Mohar, B.: Some relations between analytic and geometric properties of infinite graphs. Discrete Math. **95**(1–3), 193–219 (1991). http://dx.doi.org/10.1016/0012-365X(91)90337-2. Directions in infinite graph theory and combinatorics (Cambridge, 1989)
45. Montenegro, R., Tetali, P.: Mathematical aspects of mixing times in Markov chains. Found. Trends Theor. Comput. Sci. **1**(3), x+121 (2006)
46. Myers, S.B.: Riemannian manifolds with positive mean curvature. Duke Math. J. **8**, 401–404 (1941)
47. Nicholson, R., Sneddon, J.: New graphs with thinly spread positive combinatorial curvature. N. Z. J. Math. **41**, 39–43 (2011)
48. Oh, B.G.: Duality properties of strong isoperimetric inequalities on a planar graph and combinatorial curvatures. Discrete Comput. Geom. **51**(4), 859–884 (2014). http://dx.doi.org/10.1007/s00454-014-9592-7
49. Ohno, Y., Urakawa, H.: On the first eigenvalue of the combinatorial Laplacian for a graph. Interdiscip. Inform. Sci. **1**(1), 33–46 (1994). http://dx.doi.org/10.4036/iis.1994.33

50. Simon, B.: Brownian motion, L^p properties of Schrödinger operators and the localization of binding. J. Funct. Anal. **35**(2), 215–229 (1980). http://dx.doi.org/10.1016/0022-1236(80)90006-3
51. Stone, D.A.: A combinatorial analogue of a theorem of Myers. Ill. J. Math. **20**(1), 12–21 (1976)
52. Stone, D.A.: Correction to my paper: A combinatorial analogue of a theorem of Myers (Ill. J. Math. **20**(1), 12–21 (1976)). Ill. J. Math. **20**(3), 551–554 (1976)
53. Sturm, K.T.: On the L^p-spectrum of uniformly elliptic operators on Riemannian manifolds. J. Funct. Anal. **118**(2), 442–453 (1993). http://dx.doi.org/10.1006/jfan.1993.1150
54. Sun, L., Yu, X.: Positively curved cubic plane graphs are finite. J. Graph Theory **47**(4), 241–274 (2004). http://dx.doi.org/10.1002/jgt.20026
55. Weidmann, J.: Linear Operators in Hilbert Spaces. Graduate Texts in Mathematics, vol. 68. Springer, New York/Berlin (1980). Translated from the German by Joseph Szücs
56. Wise, D.T.: Sectional curvature, compact cores, and local quasiconvexity. Geom. Funct. Anal. **14**(2), 433–468 (2004). http://dx.doi.org/10.1007/s00039-004-0463-x
57. Woess, W.: A note on tilings and strong isoperimetric inequality. Math. Proc. Camb. Philos. Soc. **124**(3), 385–393 (1998). http://dx.doi.org/10.1017/S0305004197002429
58. Wojciechowski, R.K.: Stochastic completeness of graphs. Thesis (Ph.D.)–City University of New York. ProQuest LLC, Ann Arbor, MI (2008). http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3310649.
59. Zhang, L.: A result on combinatorial curvature for embedded graphs on a surface. Discrete Math. **308**(24), 6588–6595 (2008). http://dx.doi.org/10.1016/j.disc.2007.11.007
60. Żuk, A.: On the norms of the random walks on planar graphs. Ann. Inst. Fourier (Grenoble) **47**(5), 1463–1490 (1997). http://www.numdam.org/item?id=AIF_1997__47_5_1463_0

Chapter 7
The Geometric Spectrum of a Graph and Associated Curvatures

Paul Baird

Abstract We approach the problem of defining curvature on a graph by attempting to attach a 'best-fit polytope' to each vertex, or more precisely what we refer to as a configured star. How this should be done depends upon the global structure of the graph which is reflected in its geometric spectrum. Various curvatures naturally arise from local liftings of the graph into a suitable Euclidean space.

7.1 Introduction

One of the challenges of graph theory is to define notions of curvature purely in terms of combinatorial structure without recourse to a predefined metric structure. One would like to see geometry and curvature *emerge* from combinatorial structure rather than being imposed upon it. Different approaches go back to classical work of Descartes in the context of 3-dimensional polyhedra. In this volume, Chap. 6, M. Keller discusses a notion of combinatorial curvature that arises when there exists an embedding of a graph into a surface. Defined at each vertex, this doesn't depend on any metric structure, but only on the number of vertices (the face degree) of incident faces. The Ricci curvature of Ollivier, studied in Chap. 1 of this volume, is defined in terms of optimal transport of local measures. Its relation to the clustering coefficient, a measure of local connectedness, has been explored by Jost and Liu [19], see also Chap. 1. Common themes occur in this chapter: on the one hand local embeddings of a graph arise from the geometric spectrum leading to notions of curvature; on the other hand connectedness appears to have a crucial influence on the nature of the geometric spectrum, and so on the local geometry that arises.

Our approach is to appeal directly to the way sense data correlates with our brains to infer 3-dimensionality, just as we visualize a 3D-cube in the planar graph illustrated in Fig. 7.1.

P. Baird (✉)
Laboratoire de Mathématiques de Bretagne Atlantique, Université de Bretagne Occidentale, 6 av. Victor Le Gorgeu – CS 93837, 29238 Brest Cedex, France
e-mail: Paul.Baird@univ-brest.fr

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_7

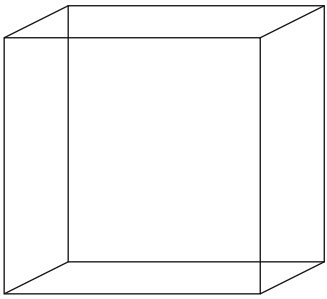


Fig. 7.1 A planar graph can produce the mental image of a 3D-object

Table 7.1 The geometric spectral values associated to the convex regular polyhedra

Polyhedron	γ
Tetrahedron	$3/4$
Cube	0
Octahedron	$1/2$
Icosahedron	$\frac{2-\sqrt{5}}{3-\sqrt{5}} < 0$
Dodecahedron	$\frac{3(1-\sqrt{5})}{2(3-\sqrt{5})} < 0$

More specifically, we would like metric structure and curvature to emerge from combinatorial structure, but in a way that derives from potential geometric realizations. In order to achieve this end, we introduce what we refer to as the *geometric spectrum* of a graph. An element of this spectrum is a real-valued function defined on the vertices which need not be constant, and occurs as the parameter γ in the quadratic difference equation:

$$\gamma(\Delta\phi)^2 = (\nabla\phi)^2, \tag{7.1}$$

where ϕ is a complex-valued function on the vertices (for definitions, see below), thought of as an orthogonal projection to the complex plane of a (local) realization of the graph in a Euclidean space. A requirement is that γ should remain invariant under any similarity transformation of this realization. In this way, the geometry that arises is implicit in the graph, rather than being imposed by the way it is embedded into Euclidean space. As an example, the convex regular polytopes in Euclidean 3-space satisfy (7.1) with γ constant as given in Table 7.1.

To a solution ϕ, we cannot in general hope for a global realization of the graph, just a local one of each vertex and its neighbors. However, this is sufficient to enable one to develop notions of (relative) edge length and curvature. More precisely, a non-trivial solution allows one to fit a *configured star* to each vertex and its neighbors, giving an invariant lifting into Euclidean space. A configured star generalizes the vertex figure of a regular polytope, so an intuitive picture to have in mind, is that of attempting to attach a 'best-fit' regular polytope to each vertex, just as in smooth geometry we can approach curvature via 'best-fit' circles, or spheres. How we do this depends on the global structure of the graph as reflected

in its geometric spectrum. Such lifts to a configured star in Euclidean 3-space are essentially unique up to a sign ambiguity, however, if we wish to extend the lifting beyond the immediate neighbors of a vertex, then in general one needs to relax this condition and allow more general invariant stars.

An embedded graph is one that can be embedded in a surface. This allows one to associate faces to the graph, which in turn determines a combinatorial quantity called *face degree*, defined to be the number of boundary edges (or vertices) of each face. In this volume, Chap. 6, M. Keller discusses a natural combinatorial curvature associated to an embedded graph, and in particular, its relation to the so-called physical Laplacian (rather than the harmonic, or Tutte Laplacian that appears in (7.1) above). In view of Theorem 7.1 below, which relates our combinatorial construction to smooth submanifolds, there should be strong connections between these two approaches.

We begin in Sect. 7.2 by defining the geometric spectrum and in Sect. 7.3 we derive the lifting properties discussed above. For an arbitrary graph, the geometric spectrum may be quite complicated and difficult to compute. We indicate how this may be done using Gröbner bases in Appendix 1 and give some examples. Also, in this appendix we introduce a polynomial graph invariant, the γ-polynomial, whose roots determine the constant elements of the spectrum. At present we know of no two non-isomorphic graphs with the same γ-polynomial. Connections with vertex colorings of a graph are discussed in respect of a particular example.

The most immediate consequence of the lift to a configured star, explored in Sect. 7.4, is the designation of a Gauss map at each vertex given by the axis of the star. If the lift is into $\mathbb{R}^3$, then this defines a point on the 2-sphere, which we can connect to other points according to whether or not they are connected by an edge in the original graph. The geodesic distance between two such points now gives an *edge* curvature on the graph. Another consequence of the lift is to endow each edge with a length, which in turn allows us to define the distance between two points, for which we establish a triangle inequality. The graph now has the structure of a path metric space and we can explore notions of curvature in the sense of Alexandrov, by comparing triangles with those in a 2-dimensional space form (see also Chap. 2 in this book). To establish curvature bounds in the general case requires further investigation, but we indicate how this may be done with an example.

Our main goal is to define curvature in a way that depends only on the element γ of the geometric spectrum and not on the solution ϕ to (7.1). For Gaussian curvature, defined in terms of angular deficit at a vertex, this is possible provided we accept an approximation to well-known classical theorems for polytopes (Sect. 7.5). In Sect. 7.6, we show how to define sectional curvature and Ricci curvature in terms of the edge curvature. Finally, in Appendix 2, we explore methods to construct solutions on new graphs from given solutions. An interesting model occurs in the theory of random graphs which suggests how geometry could naturally emerge in scale-free networks, so prevalent in biology and social networks.

7.2 The Geometric Spectrum

We use the notation $G = (V, E)$ to denote a graph with vertex set V and edge set E. Graphs will be assumed simple (no multiple edges or loops) and undirected. We will suppose also that G is locally finite, that is the degree of each vertex is finite. For $x \in V$ define the *tangent space to G at x* to be the set of oriented edges with base point x together with the zero vector: $T_xG = \{\mathbf{xy} : y \sim x\} \cup \{\mathbf{0}\}$. Define the *tangent bundle to G* to be the union: $TG = \cup_{x\in V} T_xG$. Inclusion of the zero vector is useful when we come to discuss holomorphic mappings between graphs in Appendix 2.

A 1-*form* on G is a map $\omega : TG \to \mathbb{C}$ such that $\omega(\mathbf{xy}) = -\omega(\mathbf{yx})$ and $\omega(\mathbf{0}) = 0$. To a function $\phi : V \to \mathbb{C}$, we can naturally associate a 1-form, the *derivate* $\nabla\phi$, by setting $\nabla\phi(\mathbf{xy}) = \phi(y) - \phi(x)$ and $\nabla\phi(\mathbf{0}) = 0$. For differential calculus on a simplex, see the book by Romon [24], where the notion of 1-form coincides with that given above (Sect. IV, Section 1.1 of [24]).

For two 1-forms ω, η, define their *pointwise symmetric product at* $x \in V$ by

$$\langle \omega, \eta \rangle_x = \frac{1}{d_x} \sum_{y\sim x} \omega(\mathbf{xy})\eta(\mathbf{xy}) \,,$$

where d_x is the degree of vertex x. The above definition is the complex symmetric analogue of standard L^2 products that arises in functional analytic theory on a graph; in the latter situation it is replaced by a Hermitian product rather than a symmetric product, see Chap. 1 in this volume. Write $(\nabla\phi)^2(x)$ for the *symmetric square of the derivative* of the function $\phi : V \to \mathbb{C}$:

$$(\nabla\phi)^2(x) := \frac{1}{d_x} \sum_{y\sim x} \big(\phi(y) - \phi(x)\big)^2 \,.$$

The Tutte Laplacian (or harmonic Laplacian) on G is defined by

$$\Delta\phi(x) = \frac{1}{d_x} \sum_{y\sim x} \big(\phi(y) - \phi(x)\big) \,.$$

The choice of conventions means that the spectrum is negative and lies in the interval $[-2, 0]$.

Given a graph $G = (V, E)$ together with a real-valued function $\gamma : V \to \mathbb{R}$, we are interested in the equation:

$$\frac{\gamma(x)}{d_x}\Big(\sum_{y\sim x} \big(\phi(y) - \phi(x)\big)\Big)^2 = \sum_{y\sim x} \big(\phi(y) - \phi(x)\big)^2 \,, \tag{7.2}$$

at each vertex x, where $\phi : V \to \mathbb{C}$ is a complex-valued function. In the notation above, this has the convenient form of (7.1) of the Introduction. Solutions with

$\gamma \equiv 0$ have been called *holomorphic functions*[1] and have been used to give a description of massless fields in a combinatorial setting [4], see also Appendix 2. Note that the equations are invariant under the transformations

$$\phi \mapsto \lambda\phi + \mu \quad (\lambda \in \mathbb{C} \setminus \{0\}, \mu \in \mathbb{C}), \quad \text{and} \quad \phi \mapsto \overline{\phi}. \tag{7.3}$$

We shall consider two solutions related in this way as *equivalent*. Equation (7.2) only depends on the derivative $\nabla\phi$ and more generally can be defined for an arbitrary 1-form ω on replacing $\phi(y) - \phi(x)$ with $\omega(\mathbf{xy})$.

If in (7.2) we label the neighbors of x by $y_1, \ldots, y_d$ and write $z_\ell = \phi(y_\ell) - \phi(x)$, then if the z_ℓ are all real, an application of the Cauchy-Schwarz inequality gives

$$\sum_{\ell=1}^{d} {z_\ell}^2 = \frac{\gamma}{d}\left(\sum_{\ell=1}^{d} z_\ell\right)^2 \leq \gamma \sum_{\ell=1}^{d} {z_\ell}^2, \tag{7.4}$$

so that if the z_ℓ are not all zero, necessarily $1 \leq \gamma$. Otherwise said, if $\gamma < 1$ at least one of the z_ℓ must be complex. When defining the geometric spectrum below, we impose the condition $\gamma(x) \leq 1$ for all $x \in V$ with the inequality strict if $d_x \geq 3$. This is necessary for our invariance requirement that we discuss later.

For a given graph, we would like to know what are the admissible functions $\gamma : V \to \mathbb{R}$ for which (7.2) has a solution. Define the *geometric spectrum* of G to be the collection of equivalence classes of such functions:

$$\Sigma = \{\gamma : V \to [-\infty, 1] \subset \mathbb{R} : \exists \text{ non} - \text{const } \phi : V \to \mathbb{C} \text{ satisfying (7.2)} \\ \textit{with } \gamma(x) < 1 \text{ if } d_x \geq 3\},$$

where two functions are identified when they determine a common solution ϕ and agree on the compliment of the set $\{x \in V : \Delta\phi(x) = (\nabla\phi)^2(x) = 0\}$. The function γ may take on the value $-\infty$ at points where the Laplacian vanishes.

By a *framework* in Euclidean space, we mean a graph that is realized as a subset of Euclidean space with edges straight line segments joining the vertices. We say that it is *immersed* if all vertices are distinct and *embedded* if it is immersed and edges only intersect at end points. The framework is called *invariant* if for a particular γ, it satisfies (7.2) with ϕ the restriction to the vertices of some orthogonal projection to the complex plane *independently of any similarity transformation of the framework*.

[1] A notion of *holomorphic function* somewhat similar to this has been introduced by S. Barré and A. Zeghib [7]; however, in addition to (7.2) with $\gamma \equiv 0$, Barré and Zeghib require that ϕ be harmonic. An alternative notion of *discrete holomorphic function* in the special case of quad-graphs is given by Bobenko et al. [8].

Questions that arise are:

- For a given graph G, what is its geometric spectrum?
- Does a solution to (7.2) arise from an embedding of the graph as an invariant framework in Euclidean space?
- Even if the answer to the last question is no, can we still derive geometric quantities such as edge length and curvature from a solution?
- To what extent do such quantities depend only on γ rather than on the choice of solution ϕ?

For an arbitrary graph, the geometric spectrum is determined by a fairly complicated set of algebraic equations. For graphs of sufficiently small order, these can be solved with computer software, for example MAPLE; see Appendix 1. One can check that for the complete graph on $N+1$ vertices, with $2 \leq N \leq 5$, the geometric spectrum consists of the single value $N/(N+1)$, with corresponding invariant realization as the 1-skeleton of a regular simplex in $\mathbb{R}^N$. It is not known if this remains so for $N > 5$. At the other extreme, for a cyclic or linear graph (connected graphs with the least internal connections), the geometric spectrum arises from realizations in the plane of corresponding frameworks with all edges of equal length. Now, the geometric spectrum has continuous components with complicated branching behavior. These properties are discussed in Appendix 1.

In general, after taking into account the freedom (7.3), an equation count shows that for a graph on N vertices, (7.2) is equivalent to $2N$ real equations in $3N-4$ real parameters, together with the N inequalities: $\gamma < 1$ at each vertex. Empirical evidence indicates that the more connected the graph, the more restricted its geometric spectrum. Another question that now arises is:

- Does the geometric spectrum have any relation to the Laplace spectrum?

The spectrum of the Tutte Laplacian is reviewed in Sect. 1.6.1 of Chap. 1 of this volume, and for the physical Laplacian, see Chap. 6. Any connection is not immediately obvious, since both have different structures: for a graph on N vertices, its (Tutte) Laplacian has precisely N eigenvalues, whereas, as noted above, the geometric spectrum can range from a single element to continuous components. The structure of the geometric spectrum seems to be related to rigidity in the context of bar-and-joint frameworks, although this connection needs to be explored further. In recent work, Zelazo et al. introduce a rigidity matrix associated to a framework which they show is similar (in the mathematical sense) to a weighted graph Laplacian; the eigenvalues of the rigidity matrix are related to the rigidity of the bar-and-joint framework [29]. However, the rigidity that occurs in our situation is not quite the same. Although arising from embeddings in Euclidean space (see Sect. 7.3 below), edge length can change through different realizations. Examples occur with both the dodecahedron and icosahedron, which are one of a family of realizations of graphs consisting of their respective 1-skeletons with vertices given by the columns of the matrices (7.25) and (7.26) below.

In order to interpret γ, it is instructive to consider Eq. (7.1) for the case of a smooth hypersurface in Euclidean space, where we find an interesting connection with mean-curvature.

Theorem 7.1 ([2]) *Let M^n be a smooth hypersurface in $\mathbb{R}^{n+1}$ ($n \geq 1$) and let g denote the metric on M^n induced from the standard metric on $\mathbb{R}^n$. Let $\phi : (M^n, g) \to \mathbb{C}$ be any orthogonal projection; then*

$$(\Delta\phi)^2 = -H^2(\nabla\phi)^2\,, \tag{7.5}$$

where H is the mean curvature of M^n, and where in local coordinates, $\Delta\phi = g^{ij}(\phi_{ij} - \Gamma_{ij}^k\phi_k)$ and $(\nabla\phi)^2 = g^{ij}\phi_i\phi_j$ (summing over repeated indices).

In the case when $n = 1$, the theorem confirms the identity

$$c''(s) = \kappa(s)ic'(s)\,,$$

for a regular curve $c : I \subset \mathbb{R} \to \mathbb{C}$ parametrized with respect to arc length. It is necessary that M^n be a *hypersurface* in order to satisfy the smooth version of (7.1). For example, consider the surface in $\mathbb{R}^4$ parametrized in the form:

$$(x^1, x^2) \mapsto (x^1, x^2, x^1x^2, x^1 + x^2)\,.$$

Let $\phi : \mathbb{R}^4 \to \mathbb{C}$ be the projection $\phi(x^1, x^2, x^3, x^4) = x^1 + x^2i$. Then it is readily checked that the function γ defined by (7.1) is not even real.

Given the above theorem, we expect an invariant framework that closely coincides with a smooth hypersurface to have γ approximately equal to $-1/H^2$ modulo a scaling factor (Eq. (7.5) is not scale invariant; in order to make it so, a volume term should be added).

The study of constant elements of the geometric spectrum is particularly interesting and leads to the association of a polynomial invariant to a finite graph, which we refer to as the γ-polynomial. Its definition is given in Appendix 1, where examples are given of its construction. Further work needs to be done to understand if this invariant is related to other polynomial invariants and to what extent it distinguishes isomorphism classes of graphs.

7.3 Invariant Stars and the Lifting Problem

A star graph, or bipartite graph $K_{1,d}$, has one internal vertex connected to d external vertices; there are no other connections. A star framework in $\mathbb{R}^N$ with internal vertex located at the origin can be specified by a $(N \times d)$-matrix W whose columns are the components of the external vertices. We will refer to W as the *star matrix*. Provided the center of mass of the external vertices does not coincide with the origin, then it defines a line through the origin which we refer to as the *axis of the star*. We are interested in a particular class of star frameworks whose external vertices form what we call a configuration in a plane orthogonal to the axis of the star.

7.3.1 Configured Stars

For $N \geq 2$, a collection of points $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ in $\mathbb{R}^{N-1}$ forms a *configuration* if the $((N-1) \times d)$-matrix $U = (\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_d)$ whose columns have as components the coordinates $v_{\ell j}$ of $\mathbf{v}_\ell$ $(j = 1, \ldots, N-1;\ \ell = 1, \ldots, d)$, satisfies:

$$UU^t = \rho I_{N-1}\,, \qquad \sum_{\ell=1}^{d} \mathbf{v}_\ell = \mathbf{0}\,, \tag{7.6}$$

for some non-zero constant ρ (necessarily positive), where $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^{N-1}$ and U^t denotes the transpose of U. Necessarily $\mathrm{rank}(U) = N-1$ so that $d \geq N$. A star in $\mathbb{R}^N$ whose external vertices form a configuration in a plane not passing through the origin, is referred to as a *configured star*. If all edges are of identical length, we call the star *regular*. An *invariant* of a configured star is a quantity that is invariant by orthogonal transformation. The following lemma characterizes configured stars [3].

Lemma 7.1 *Consider a configured star in $\mathbb{R}^N$ $(N \geq 2)$ with internal vertex the origin connected to d external vertices $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$ $(d \geq N)$. Let $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ be the $(N \times d)$-matrix whose columns are the components $x_{\ell j}$ of $\mathbf{x}_\ell$ $(j = 1, \ldots, N;\ \ell = 1, \ldots, d)$. Then*

$$WW^t = \rho I_N + \sigma \mathbf{u}\mathbf{u}^t, \qquad \sum_{\ell=1}^{d} \mathbf{x}_\ell = \sqrt{d(\sigma+\rho)}\,\mathbf{u}\,, \tag{7.7}$$

where the unit vector $\mathbf{u} \in \mathbb{R}^N$ is the axis of the star, $\rho > 0$ and $\rho + \sigma > 0$. The quantities d, ρ, σ are all invariants of the star; the vector $\mathbf{u}$ is normal to the affine plane containing $\mathbf{x}_1, \ldots, \mathbf{x}_d$.

Conversely, any matrix $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ satisfying (7.7) determines a configured star with internal vertex the origin and external vertices $\mathbf{x}_1, \ldots, \mathbf{x}_d$.

Proof Consider a configured star in standard position given by (7.6). Set

$$V = \left(\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_d \\ c & c & \cdots & c \end{array}\right)$$

and let $A : \mathbb{R}^N \to \mathbb{R}^N$ be an orthogonal transformation; set $\mathbf{x}_d = A\begin{pmatrix} \mathbf{v}_d \\ c \end{pmatrix}$. Then $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d) = AV$ and

$$WW^t = AVV^tA^t = \rho I_N + \sigma(A\mathbf{e}_N)(A\mathbf{e}_N)^t\,,$$

where

$$\sigma = dc^2 - \rho\,. \tag{7.8}$$

Furthermore $\sum_{\ell=1}^{d} \mathbf{x}_\ell = dcA\mathbf{e}_N$, which gives the form (7.7) with $\mathbf{u} = A\mathbf{e}_N$. The independence of the quantities d, ρ, σ under the orthogonal transformation A is clear.

Conversely, suppose we are given an $(N \times d)$-matrix $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ satisfying (7.7). Let A be an orthogonal transformation such that $A\mathbf{u} = \mathbf{e}_N$ and let $V = AW$. Write

$$V = \left(\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_d \\ y_{1N} & y_{2N} & \cdots & y_{dN} \end{array}\right).$$

Then

$$VV^t = \rho I_N + \sigma \mathbf{e}_N \mathbf{e}_N{}^t \qquad \text{and} \qquad \sum_{\ell} \begin{pmatrix} \mathbf{v}_\ell \\ y_{\ell N} \end{pmatrix} = \sqrt{d(\sigma+\rho)}\, \mathbf{e}_N\,, \tag{7.9}$$

so that $\sum_\ell \mathbf{v}_\ell = 0$ and $\sum_\ell y_{\ell N} = \sqrt{d(\sigma+\rho)}$. Furthermore, (7.9) implies that $\sum_\ell y_{\ell N}{}^2 = \rho + \sigma$. In particular

$$d \sum_{\ell}^{d} y_{\ell N}{}^2 = \left(\sum_{\ell}^{d} y_{\ell N} \right)^2 .$$

But then equality in the inequality (7.4) implies that $y_{1N} = y_{2N} = \cdots = y_{dN} = \sqrt{(\sigma+\rho)/d}$. □

Corollary 7.1 *Let $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ define a configured star and let $\phi : \mathbb{R}^N \to \mathbb{C}$ be the orthogonal projection $\phi(y_1, y_2, \ldots, y_N) = y_1 + \mathrm{i}y_2$. Then if $z_\ell = \phi(\mathbf{x}_\ell) = x_{\ell 1} + \mathrm{i}x_{\ell 2}$, we have*

$$\frac{\sigma}{d(\sigma+\rho)} \left(\sum_{\ell=1}^{d} z_\ell \right)^2 = \sum_{\ell=1}^{d} z_\ell{}^2 ,$$

where ρ and σ are given by (7.7). In particular, with reference to Eq. (7.2), $\gamma = \sigma/(\sigma+\rho)$ is real and depends only on the star invariants.

Proof Let $\mathbf{u} = (u_1, \ldots, u_N)$ be the unit normal to the plane of the star. Then for each $j = 1, \ldots, N$, we have

$$\sum_{\ell=1}^{d} x_{\ell j} = \sqrt{d(\sigma+\rho)}\, u_j\,.$$

Thus

$$\begin{aligned} \left(\sum_{\ell=1}^{d} z_\ell \right)^2 &= \sum_{k,\ell=1}^{d} (x_{k1}x_{\ell 1} - x_{k2}x_{\ell 2} + 2\mathrm{i}x_{k1}x_{\ell 2}) \\ &= d(\sigma+\rho)(u_1{}^2 - u_2{}^2 + 2\mathrm{i}u_1u_2) = d(\sigma+\rho)(u_1 + \mathrm{i}u_2)^2\,, \end{aligned}$$

whereas

$$\sum_{\ell=1}^{d} {z_\ell}^2 = \sum_{\ell=1}^{d} ({x_{\ell 1}}^2 - {x_{\ell 2}}^2 + 2\mathrm{i}x_{\ell 1}x_{\ell 2}) = (WW^t)_{11} - (WW^t)_{22} + 2\mathrm{i}(WW^t)_{12} = \sigma(u_1 + \mathrm{i}u_2)^2 .$$

The formula now follows. □

7.3.2 *Invariant Stars*

To test whether a framework in Euclidean space is invariant, it suffices to see whether the star about each of its vertices is invariant at the internal vertex. A consequence of Corollary 7.1 is that any *configured* star is invariant at its internal vertex. The star framework about the vertex of any regular polytope is configured, so that the underlying framework of a regular polytope is invariant [3]. On the other hand, not all invariant stars are configured. For example, the star in $\mathbb{R}^3$ with $2r$ external vertices represented by the columns of the $(3 \times (2r))$-matrix

$$W = \begin{pmatrix} x_1 & x_2 & \cdots & x_r & x_1 & x_2 & \cdots & x_r \\ s_1 & s_2 & \cdots & s_r & -s_1 & -s_2 & \cdots & -s_r \\ t_1 & t_2 & \cdots & t_r & -t_1 & -t_2 & \cdots & -t_r \end{pmatrix},$$

where the vectors $\mathbf{s} = (s_1, \ldots, s_r)$ and $\mathbf{t} = (t_1, \ldots, t_r)$ are orthogonal and of the same length, is invariant, but it is only configured when $x_1 = x_2 = \cdots = x_r$. As a specific case, take W to be the star on four vertices represented by the following matrix.

$$W = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{7.10}$$

The external vertices lie on the corners of a cube as indicated in Fig. 7.2.

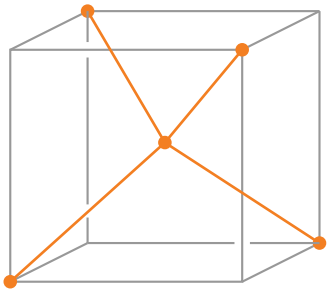


Fig. 7.2 An invariant star which is not configured

Contrary to a configured star, the external vertices don't lie in any half-space whose boundary passes through the internal vertex. When we come to discuss Gaussian curvature in Sect. 7.5, such a star will have negative curvature associated to the internal vertex. On the other hand, a configured star will always have positive curvature at the internal vertex.

We can give a characterization of all invariant stars in $\mathbb{R}^3$ with the following.

Proposition 7.1 *Consider a general star in* $\mathbb{R}^3$ *with matrix*

$$W = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ x_{31} & x_{32} & \cdots & x_{3d} \end{pmatrix},$$

where we assume that $d \geq 2$*, that the columns of* W *are non-zero and that there is at least one* $i = 1, 2, 3$ *such that* $\sum_{\ell=1}^{d} x_{i\ell}$ *is non-zero. Then* W *is invariant if and only if there exists a real number* γ *such that*

$$\frac{\gamma}{d}\left[\left(\sum_{\ell=1}^{d} x_{i\ell}\right)^2 - \left(\sum_{\ell=1}^{d} x_{j\ell}\right)^2\right] = \sum_{\ell=1}^{d} {x_{i\ell}}^2 - \sum_{\ell=1}^{d} {x_{j\ell}}^2 \tag{7.11}$$

$$\frac{\gamma}{d}\left(\sum_{\ell=1}^{d} x_{i\ell}\right)\left(\sum_{\ell=1}^{d} x_{j\ell}\right) = \sum_{\ell=1}^{d} x_{i\ell}x_{j\ell} \tag{7.12}$$

for all $i, j = 1, 2, 3,\ i \neq j$.

Proof Let $A = (a_{ij})$ be an arbitrary orthogonal transformation. Then

$$AW = \begin{pmatrix} a_{1j}x_{j1} & a_{1j}x_{j2} & \cdots & a_{1j}x_{jd} \\ a_{2j}x_{j1} & a_{2j}x_{j2} & \cdots & a_{2j}x_{jd} \\ a_{3j}x_{j1} & a_{3j}x_{j2} & \cdots & a_{3j}x_{jd} \end{pmatrix}$$

where, in the matrix, we sum over repeated indices. Now project to the complex plane via $(y_1, y_2, y_3) \mapsto y_i + \mathrm{i}y_2$, so that the end points of the star are projected to the complex numbers

$$z_\ell = \sum_{j=1}^{3} (a_{1j} + \mathrm{i}a_{2j})x_{j\ell}, \quad \ell = 1, \ldots, d\,.$$

In order that W represents an invariant star we require that

$$\frac{\gamma}{d}\left(\sum_{\ell=1}^{d} z_\ell\right)^2 = \sum_{\ell=1}^{d} {z_\ell}^2\,, \tag{7.13}$$

for some real number γ independently of A. Since A is an orthogonal matrix, we have

$$\sum_{j=1}^{3}(a_{1j}+\mathrm{i}a_{2j})^2=0 \quad \Rightarrow \quad (a_{13}+\mathrm{i}a_{23})^2=-(a_{11}+\mathrm{i}a_{21})^2-(a_{12}+\mathrm{i}a_{22})^2\,,$$

and (7.13) is equivalent to

$$\begin{aligned}\frac{\gamma}{d}\Bigg\{&(a_{11}+\mathrm{i}a_{21})^2\left(\left(\sum_{\ell=1}^{d}x_{1\ell}\right)^2-\left(\sum_{\ell=1}^{d}x_{3\ell}\right)^2\right)\\&+(a_{12}+\mathrm{i}a_{22})^2\left(\left(\sum_{\ell=1}^{d}x_{2\ell}\right)^2-\left(\sum_{\ell=1}^{d}x_{3\ell}\right)^2\right)\\&+2\sum_{j<k}(a_{1j}+\mathrm{i}a_{2j})(a_{1k}+\mathrm{i}a_{2k})\left(\sum_{\ell=1}^{d}x_{j\ell}\right)\left(\sum_{\ell=1}^{d}x_{k\ell}\right)\Bigg\}\\=\;&(a_{11}+\mathrm{i}a_{21})^2\sum_{\ell=1}^{d}({x_{1\ell}}^2-{x_{3\ell}}^2)+(a_{12}+\mathrm{i}a_{22})^2\sum_{\ell=1}^{d}({x_{2\ell}}^2-{x_{3\ell}}^2)\\&+2\sum_{j<k}(a_{1j}+\mathrm{i}a_{2j})(a_{1k}+\mathrm{i}a_{2k})\sum_{\ell=1}^{d}x_{j\ell}x_{k\ell}\end{aligned}$$

and the sufficiency of Eqs. (7.11) and (7.12) follow.

To see that the equations are necessary, it suffices to set

$$A=\begin{pmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}$$

for θ an arbitrary parameter and to compare coefficients of $\cos^2\theta, \sin^2\theta$ and $\cos\theta\sin\theta$. □

If $\sum_{\ell=1}^{d}x_{i\ell}=0$ for all $i=1,2,3$, then for γ finite, whatever the projection to the plane, the left-hand side of (7.13) vanishes. In this case, in certain circumstances, it is desirables to allow γ to take on the value $-\infty$ and to interpret (7.13) by continuity arguments, even with a non-zero right-hand side. This can occur for example, when the degree $d=2$; see Section "First Cases" in Appendix 1 and below.

In the statement of the above proposition, we do not exclude the possibility that some columns of W may coincide. In this case the star will not be embedded. Indeed, the only invariant star satisfying the hypotheses of the proposition with $d=2$ must

have identical external vertices. To see this, we perform a similarity transformation so that the matrix W has the form

$$W = \begin{pmatrix} 1 & x_{12} \\ 0 & 0 \\ 0 & x_{32} \end{pmatrix}.$$

Then the projection of the star to the plane is real and provided $x_{12} \neq -1$, by the Cauchy-Schwarz inequality (7.4), we must have $\gamma = 1$. The same conclusion arises if $x_{32} \neq 0$ and $x_{12} = -1$; to see this we permute the 1st and 3rd lines of W prior to projection. Equations (7.11) and (7.12) now imply that necessarily $x_{32} = 0$ and $x_{12} = 1$. This leaves the case when $x_{12} = -1$ and $x_{32} = 0$, i.e. column two is minus column one. As discussed in the preceding paragraph, it makes sense to interpret this as an invariant star with $\gamma = -\infty$.

Any star with just one external vertex is automatically invariant with $\gamma = 1$. On the other hand, for $d = 3$ we have the following.

Corollary 7.2 *Any invariant star in $\mathbb{R}^3$ with three distinct external vertices must be configured.*

Proof By a similarity transformation of $\mathbb{R}^3$, we may suppose that W has the form

$$W = \begin{pmatrix} 1 & x_{12} & x_{13} \\ 0 & x_{22} & x_{23} \\ c & c & c \end{pmatrix},$$

for some constant c, that is, all the external vertices lie in the plane $x_3 = c$ and further, by a rotation about the x_3 axis and a dilation, the first external vertex has the form indicated. Write $\mathbf{r}_j$ for the jth line of W and $\sum \mathbf{r}_j$ for the sum of its components. Thus (7.11) and (7.12) can be written

$$\frac{\gamma}{3}\left((\sum \mathbf{r}_j)^2 - (\sum \mathbf{r}_k)^2\right) = \mathbf{r}_j \cdot \mathbf{r}_j - \mathbf{r}_k \cdot \mathbf{r}_k \quad \text{and} \quad \frac{\gamma}{3}\left(\sum \mathbf{r}_j\right)\left(\sum \mathbf{r}_k\right) = \mathbf{r}_j \cdot \mathbf{r}_k$$

for all $j \neq k$.

If $c = 0$, then we deduce that $\frac{\gamma}{3}\left(\sum \mathbf{r}_j\right)^2 = \mathbf{r}_j \cdot \mathbf{r}_j$ for $j = 1, 2$ and from the Cauchy-Schwarz inequality, we must have $\gamma \geq 1$ which contradicts our hypothesis that $\gamma < 1$. Thus $c \neq 0$ and from the second equation, noting that $\sum \mathbf{r}_3 = 3c$, we deduce that

$$\gamma c \sum \mathbf{r}_j = c \sum \mathbf{r}_j .$$

for $j = 1, 2$. Then if $\sum \mathbf{r}_j \neq 0$ $(j = 1, 2)$ we would have $\gamma = 1$, a contradiction, so that $\sum \mathbf{r}_j = 0$ for $j = 1, 2$, in particular $1 + x_{12} + x_{13} = 0$ and $x_{22} + x_{23} = 0$.

We exclude the case $x_{22} = x_{23} = 0$ which again gives a contradiction. But now the necessary conditions

$$1 + {x_{12}}^2 + {x_{13}}^2 = {x_{22}}^2 + {x_{23}}^2 \quad \text{and} \quad x_{12}x_{22} + x_{13}x_{23} = 0\,,$$

have only $x_{12} = x_{13} = -1/2$ and $x_{22} = -x_{23} = \pm\sqrt{3}/2$ as possible solutions. Thus the external vertices are symmetrically placed as the third roots of unity in the plane $x_3 = c$ and the star is configured. □

Note that, up to dilation and rotation, the only configuration of three points in the plane is given by the third roots of unity. In particular, any star on this configuration is regular.

Examples of invariant frameworks in $\mathbb{R}^3$ with vertices of degree three are the 1-skeleta of the tetrahedron, the cube and the dodecahedron. However, the necessity that invariant stars about a vertex of degree three be configured, means that the piecing together of such stars to form an extensive invariant framework is bound to be restrictive. On the other hand, if the degree is four, then one has more flexibility. For example, if one fixes two of the external vertices then there is a certain freedom in the choice of the other two. This may make it possible to construct an invariant mesh which approximates arbitrarily closely a smooth surface. Such problems are for future investigation.

7.3.3 The Lifting Problem

Given a solution ϕ to (7.2), at each vertex x, our aim is to construct a configured star in some Euclidean space $\mathbb{R}^N$ whose external vertices project to the points $\phi(y) - \phi(x)$ $(y \sim x)$ of the complex plane. To do this, we establish a converse to Corollary 7.1. We shall refer to the problem of constructing such a star as *the lifting problem*. At a vertex of degree three with ϕ holomorphic, this is the Theorem of Axonometry of Gauss [15].

Fix a vertex x of degree d and label its neighbors $y_1, \dots, y_d$. Set $z_\ell = \phi(y_\ell) - \phi(x)$ $(\ell = 1, \dots, d)$, which we suppose not all zero. From (7.2):

$$\frac{\gamma}{d}\left(\sum_{\ell=1}^{d} z_\ell\right)^2 = \sum_{\ell=1}^{d} {z_\ell}^2 \quad (\gamma \in \mathbb{R})\,. \tag{7.14}$$

For a given N with $2 \le N \le d$, we wish to construct a configured star $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ in $\mathbb{R}^N$ with z_ℓ the orthogonal projection of $\mathbf{x}_\ell$. For convenience, write $z_\ell = x_{\ell 1} + \mathrm{i}x_{\ell 2} = \alpha_\ell + \mathrm{i}\beta_\ell$, so that

$$W = \left(\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & \cdots & \alpha_d \\ \beta_1 & \beta_2 & \cdots & \beta_d \\ x_{13} & x_{23} & \cdots & x_{d3} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1N} & x_{2N} & \cdots & x_{dN} \end{array}\right).$$

For the case $N = 2$, see Appendix 1. For $N \geq 3$, we are required to solve the system:

$$WW^t = \rho I_N + \sigma \mathbf{u}\mathbf{u}^t, \qquad \sum_{\ell=1}^{d} \mathbf{x}_\ell = \sqrt{d(\sigma+\rho)}\,\mathbf{u}\,, \tag{7.15}$$

for $x_{\ell j}$ ($\ell = 1, \ldots, d$; $j = 3, \ldots, N$), $\rho > 0$, σ such that $\rho + \sigma > 0$ and $\mathbf{u} \in \mathbb{R}^N$ unit, with $\gamma = \sigma/(\sigma+\rho)$. This is a matter of linear algebra which we now detail for the case $N = 3$. The general case is dealt with in [3].

Let $\{z_1, \ldots, z_d; \gamma\}$ be a non-trivial solution to (7.14) satisfying $\gamma < 1$. Set

$$\rho = \frac{1}{2}\sum_{\ell} z_\ell \overline{z}_\ell - \frac{\gamma}{2d}\Big(\sum_{\ell} z_\ell\Big)\Big(\sum_{\ell} \overline{z}_\ell\Big) > 0\,, \tag{7.16}$$

and

$$\sigma = \frac{\gamma\rho}{1-\gamma} \quad (\Rightarrow \sigma + \rho = \rho/(1-\gamma) > 0)\,. \tag{7.17}$$

Define

$$u_1 = \frac{1}{\sqrt{d(\sigma+\rho)}}\sum_{\ell=1}^{d}\alpha_\ell, \quad u_2 = \frac{1}{\sqrt{d(\sigma+\rho)}}\sum_{\ell=1}^{d}\beta_\ell\,; \tag{7.18}$$

and let $u_3 = \sqrt{1 - {u_1}^2 - {u_2}^2}$. Set

$$A := \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_d \\ \beta_1 & \beta_2 & \cdots & \beta_d \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \qquad X := \begin{pmatrix} x_{13} \\ x_{23} \\ \vdots \\ x_{d3} \end{pmatrix}. \tag{7.19}$$

Then (7.15) is equivalent to solving

$$AX = B := u_3 \begin{pmatrix} \sigma u_1 \\ \sigma u_2 \\ \sqrt{d(\sigma+\rho)} \end{pmatrix} \tag{7.20}$$

subject to the constraint:

$$X^t X = \rho I_{N-2} + \sigma {u_3}^2. \tag{7.21}$$

It is important to note the sign ambiguity: the equations are invariant under the simultaneous replacement of u_3 by $-u_3$ and of X by $-X$. This ambiguity represents two choices for the configured star.

When A has maximal rank 3, the system (7.20) and (7.21) has the unique solution $X = A^{+}B$, where $A^{+} = A^{t}(AA^{t})^{-1}$ (together with the sign ambiguity discussed above). If the rows of A are dependent then AA^{t} is no longer invertible and ${u_1}^2 + {u_2}^2 = 1 \Rightarrow u_3 = 0$, so we are required to solve the system $AX = 0$ with the constraint $X^{t}X = \rho$. There is now a 1-parameter family of solutions. This case occurs if and only if the complex numbers z_ℓ in (7.14) satisfy (see [3]):

$$d\sum_{\ell=1}^{d}|z_\ell|^2 + (\gamma - 2)\Big|\sum_{\ell=1}^{d} z_\ell\Big|^2 = 0\,.$$

Thus, apart from special cases, we have a lift about each vertex of a solution to (7.2) into $\mathbb{R}^3$ which is unique modulo translation along the axis of projection $\mathbb{R}^3 \to \mathbb{R}^2$ and up to the twofold ambiguity corresponding to the sign of X. This already enables certain geometric quantities to be defined in an unambiguous way, for example edge length. Furthermore, the twofold ambiguity may sometimes be removed by a requirement of global consistency, as is the case with the cube: a choice at one vertex imposes a choice of lift at neighboring vertices.

The problem of when a global lifting of a given graph exists remains relatively unexplored. An obvious geometric obstruction occurs when the lifts at neighboring vertices defines a *different* length to the connecting edge. This is particularly relevant when we try to lift into $\mathbb{R}^3$ since then, in general, edge length of each lift is unique (see, for example [3], Example 4.4). However, in general there is a smooth family of lifts into $\mathbb{R}^N$ when $N > 3$ subject to the constraint that $N \leq d$ (d = degree of the vertex), so that it may still be possible to find a global lift into a higher dimensional Euclidean space. A next step in order to extend the lifting beyond the immediate neighborhood of a vertex, will be to relax the condition that a lifted star be configured while maintaining invariance.

Examples of invariant frameworks other than the regular polytopes are given in [2, 3]. A particularly interesting example is given by a double cone on a convex planar regular polygon as illustrated in Fig. 7.3, sometimes referred to as an *n-gonal bipyramid*. In this case, there is a unique height (distance from the plane of the polygon to the apex) for which it becomes invariant. The invariant stars about the lateral vertices are only configured when the polygon is a square, which corresponds to the octahedron.

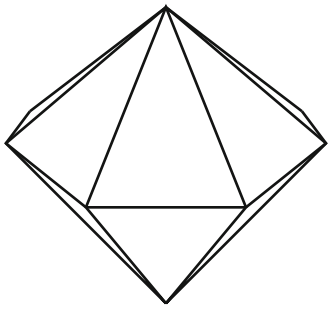


Fig. 7.3 The double cone on a regular convex polygon has a unique height for which it becomes invariant

7.4 Edge Length and the Gauss Map

Let $G = (V, E)$ be a graph and let $\gamma : V \to [-\infty, 1]$ be an element of its geometric spectrum. We are interested in quantities related to distance and curvature, defined in terms of a lift of a vertex and its neighbors to a configured star, that depend only on the geometric spectrum rather than the representative solution ϕ to (7.2). Sometimes this will be possible, sometimes not. To a given solution ϕ, suppose that the matrix A of (7.19) is of maximal rank 3. This property is clearly independent of the equivalence (7.3). The most obvious object that first arises is the axis $\mathbf{u}$ of a configured star, whose components are given by (7.18), with $u_3 = \sqrt{1 - {u_1}^2 - {u_2}^2}$. In fact, we may write

$$u_1 + \mathrm{i}u_2 = \sqrt{\frac{1-\gamma}{d\rho}} \sum_{\ell=1}^{d} z_\ell \,,$$

which, from the expression (7.16) for ρ, is scale invariant under $z_\ell \mapsto r z_\ell$ ($\forall \ell \,, r \in \mathbb{R}^*$), but not invariant under complex conjugation $z_\ell \mapsto \overline{z_\ell}$ ($\forall \ell$), which changes the sign of u_2. This ambiguity represents a global freedom in the solution of the form $\phi \mapsto \overline{\phi}$, which corresponds to the isometry $(y_1, y_2, y_3) \mapsto (y_1, -y_2, y_3)$ in the ambient Euclidean space, which will not affect subsequent quantities such as edge curvature and edge length. Note also that the axis is well-defined as $\gamma \to -\infty$.

However, there is a further sign ambiguity in the choice of u_3 which we can trace back to the two choices of lifted star. What additional information is required in order to make a unique choice for the lifted star? One way to do this is to define *orientation* on a graph.

7.4.1 Graph Orientation

In [4], a notion of orientation was considered on a regular graph of degree d, say, whereby the graph is endowed with an edge coloring of the d colors $\{1, 2, \ldots, d\}$. Thus each edge is colored in such a way that no two edges of the same color are incident at a vertex. This enables one to uniquely label the edges at each vertex to give an ordering. One could then attempt to apply a right-hand rule say, in order to make a choice of lift. However, although this can be done with the solution corresponding to the framework of the tetrahedron in a way consistent with its canonical embedding, it turns out to be impossible for the cube and the dodecahedron. In the latter examples, any edge coloring with three colors leads to at least one of the two choices of lifted stars being directed in the opposite way required. We therefore proceed to define an orientation in terms of an edge coloring together with an n-form at each vertex of degree d.

Definition 7.1 Let $G = (V, E)$ be a graph with largest vertex degree equal to M. Then an *edge coloring* of G is an association of one of the colors $\{1, 2, \ldots, M, M+1\}$ to each edge so that no two same colors are incident at any vertex. By a theorem of Vizing, any graph can be colored with either M or $M+1$ colors (see [12]). We make the convention to choose the minimum M colors when possible. Given an edge coloring of G, at each vertex, a *volume form* is an alternating mapping θ of the edges which takes on the value $+1$ or -1. Thus if x is a vertex with d incident edges $e_1, \ldots, e_d$ arranged so that $color(e_j) < color(e_k)$ for $j < k$, then $\theta_x(e_1, \ldots, e_d) = \pm 1$ with $\theta_x(e_{\sigma(1)}, \ldots e_{\sigma(d)}) = \text{sign}\,(\sigma)\theta(e_1, \ldots, e_d)$ for any permutation σ of $\{1, \ldots, d\}$. An *orientation* of G is given by an edge coloring together with a volume form at each vertex.

Suppose that $N = 3$ and that the matrix A of Eq. (7.19) is of maximal rank 3. There is now either a unique lifted configured star in the case when $u_3 = 0$, or two choices if $u_3 \neq 0$ depending on the sign chosen for $u_3 = \pm\sqrt{{u_1}^2 + {u_2}^2}$. The star matrix is now given by

$$W = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_d \\ \beta_1 & \beta_2 & \cdots & \beta_d \\ x_{13} & x_{23} & \cdots & x_{d3} \end{pmatrix},$$

where the last row is only defined up to sign. Suppose that the vertex in question has an orientation according to Definition 7.1. Without loss of generality, we can suppose that the edges are colored with the colors $\{1, 2, \ldots, d\}$, in such a way that the external vertex $\mathbf{x}_\ell$ is joined to the internal vertex by the edge with color ℓ. Suppose that the volume form satisfies $\theta(e_1, \ldots, e_d) = \varepsilon$, where $\varepsilon \in \{+1, -1\}$. Then provided the determinant of the 3×3-minor given by the first three columns of W is non-zero, we choose the sign of the third row so that

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ x_{13} & x_{23} & x_{33} \end{vmatrix} = \varepsilon\delta,$$

where $\delta > 0$. If on the other hand this determinant vanishes, then we proceed in a lexicographic ordering, to choose next the minor formed from columns 1, 2 and 4 and so on, until we encounter a non-zero determinant and apply the above rule.

Example 7.1 Consider the framework of a regular octahedron with vertices placed at the points $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$, $(0, 0, \pm 1)$. Then this can be edge-colored as indicated in Fig. 7.4 and there is a volume form which gives the lifts that correspond to the standard embedding in $\mathbb{R}^3$. However, in order to do this at the lateral vertices, we have to impose an additional condition that the star be regular. This is because at these vertices $u_3 = 0$ and we do not satisfy the conditions of the discussion above. The corresponding solution to (7.2) has $\gamma = 1/2$ and $\rho = \sigma = 2$.

Consider the vertex x_0 and define the volume form θ_{x_0} by $\theta_{x_0}(1234) = -1$ (for convenience, we write $\theta(1234)$ rather than $\theta(e_1, e_2, e_3, e_4)$). Then, with this

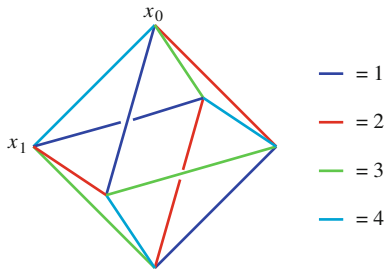


Fig. 7.4 An edge coloring for the octahedron

edge-coloring, at this vertex $z_1 = 1$, $z_2 = \mathrm{i}$, $z_3 = -1$ and $z_4 = -\mathrm{i}$. Thus $u_1 = u_2 = 0$ and $u_3 = \pm 1$. The solution to (7.20) is given by $(x_{13}, \ldots, x_{43}) = \pm(1, 1, 1, 1)$. In order to be consistent with the orientation, we must take the negative sign, to give the lifted star:

$$W = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & -1 & -1 \end{pmatrix}$$

whose sign of the determinant of the 3×3-minor given by the first three columns is negative, which coincides with the sign of $\theta_{x_0}(1234)$.

At the vertex x_1, we choose $\theta_{x_1}(1234) = +1$. Then the edge-coloring dictates that $z_1 = -1 + \mathrm{i}$, $z_2 = 1 + \mathrm{i}$, $z_3 = \mathrm{i}$, $z_4 = \mathrm{i}$, so that $u_1 = 0$, $u_2 = 1$ and $u_3 = 0$. The solution to (7.20) gives a 1-parameter family of lifted stars:

$$W = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -\frac{\cos t}{\sqrt{2}} & -\frac{\cos t}{\sqrt{2}} & \frac{\cos t}{\sqrt{2}} + \sin t & \frac{\cos t}{\sqrt{2}} - \sin t \end{pmatrix}.$$

If we now impose the condition that the lift must be a regular star, then there are just two solutions given by $t = \pi/2$ or $t = 3\pi/2$. The choice $t = 3\pi/2$ is required in order that the determinant of the 3×3-minor consisting of the first three columns be positive, to coincide with the sign of $\theta_{x_1}(1234)$. This gives the lift that coincides with the canonical embedding of the octahedron. We proceed similarly with the other vertices, defining the appropriate volume form, with the proviso that the stars at the lateral vertices be regular.

7.4.2 *Edge Curvature and Edge Length*

For a given graph $G = (V, E)$ together with a solution ϕ to (7.2), a unique choice of lifted star at each vertex now determines a unit vector $\mathbf{u}$ which gives the associated Gauss map $\mathbf{u} : V \to S^2$. If two vertices are connected by an edge, we may connect

them by the shortest geodesic arc in S^2, so realizing a copy of G in S^2. For a given edge $e = \overline{xy} \in E$, we can define the *edge-curvature* to be the length of the spherical arc joining $\mathbf{u}(x)$ to $\mathbf{u}(y)$. For example, for the standard cube, the curvature of each edge is $\arccos(1/3)$; for the octahedron, it is $\pi/2$.

Edge-length clearly depends on the choice of representative solution ϕ. A convenient way to obtain the mean length of edges incident with a particular vertex is to reverse the order of multiplication of $W = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_d)$ and W^t:

$$W^t W = \begin{pmatrix} ||\mathbf{x}_1||^2 & \langle \mathbf{x}_1, \mathbf{x}_2\rangle & \cdots & \langle \mathbf{x}_1, \mathbf{x}_d\rangle \\ \langle \mathbf{x}_1, \mathbf{x}_2\rangle & ||\mathbf{x}_2||^2 & \cdots & \langle \mathbf{x}_2, \mathbf{x}_d\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{x}_1, \mathbf{x}_d\rangle & \langle \mathbf{x}_2, \mathbf{x}_d\rangle & \cdots & ||\mathbf{x}_d||^2 \end{pmatrix},$$

where $\langle \mathbf{x}_j, \mathbf{x}_k\rangle$ denotes the standard Euclidean inner product of $\mathbf{x}_i$ and $\mathbf{x}_j$. But then

$$\begin{aligned} \sum_\ell ||\mathbf{x}_\ell||^2 &= \text{trace}\, W^t W = \text{trace}\, WW^t \\ &= N\rho + \sigma||\mathbf{u}||^2 = N\rho + \sigma\,. \end{aligned}$$

The latter quantity can be expressed in terms of γ and z_ℓ from (7.16) and the relation $\gamma = \sigma/(\sigma + \rho)$ to give the mean of the values $||\mathbf{x}_\ell||^2$:

$$\frac{1}{d}\sum_\ell ||\mathbf{x}_\ell||^2 = \frac{\big(N + (1-N)\gamma\big)}{d(1-\gamma)}\rho\,. \tag{7.22}$$

This equation expresses the mean length of the edges of a lift to a configured star in $\mathbb{R}^N$ whose external vertices $\mathbf{x}_\ell$ project to z_ℓ. This motivates our definition of edge length in a graph.

Let $G = (V, E)$ be a graph coupled to a solution $\phi : V \to \mathbb{C}$ to Eq. (7.2). For each $x \in V$, following (7.16), set

$$\rho(x) = \frac{1}{2}\left\{\sum_{y\sim x} |\phi(y) - \phi(x)|^2 - \frac{\gamma(x)}{d_x}\Big|\sum_{y\sim x}(\phi(y) - \phi(x))\Big|^2\right\}\,,$$

where d_x is the degree of G at x.

Definition 7.2 If $x \in V$ is a vertex of degree d_x such that $\gamma(x) < 1$, then we define the *median edge length at x relative to* ϕ to be the quantity $r(x) \geq 0$ whose square is given by

$$r(x)^2 = \frac{\big[N + (1-N)\gamma(x)\big]}{d_x[1-\gamma(x)]}\,\rho(x)\,.$$

If $\overline{xy} \in E$ is an edge which joins x to y such that both $\gamma(x) < 1$ and $\gamma(y) < 1$, then we define the *length of* $\overline{xy}$ *relative to* ϕ to be the mean $\ell(\overline{xy})$ of the median edge lengths at x and y:

$$\ell(\overline{xy}) = \frac{r(x) + r(y)}{2}.$$

As emphasized in the above definition, the lengths so defined are *relative* to the solution ϕ of (7.2), which is only defined up to $\phi \mapsto \lambda\phi + \mu$ for $\lambda, \mu \in \mathbb{C}$. This means that the only meaningful quantities are *relative* lengths, say $\ell(e)/\ell(f)$, for two edges $e, f \in E$. In particular, if both d and γ are constant on the graph, we may take the quantity 2ρ defined by (7.16) as a measure of median edge length at each vertex:

$$r(x)^2 = 2\rho = \sum_{y\sim x} |\phi(y) - \phi(x)|^2 - \frac{\gamma}{d}\Big|\sum_{y\sim x}(\phi(y) - \phi(x))\Big|^2.$$

If desired, we can define an absolute length by normalizing with respect to the *square* L^2*-norm of the derivative of* ϕ:

$$||\mathrm{d}\phi||^2 = \sum_{\overline{xy}\in E} |\mathrm{d}\phi(\overline{xy})|^2 = \frac{1}{2}\sum_{x,y\in V, x\sim y} |\phi(y) - \phi(x)|^2,$$

and setting

$$r_{\mathrm{abs}}(x)^2 = \frac{r(x)^2}{||\mathrm{d}\phi||^2},$$

where $r(x)$ is the median edge length at x relative to ϕ. If $e \in E$ is an edge joining x to y such that both $\gamma(x) < 1$ and $\gamma(y) < 1$, then we define the *absolute length of* e *relative to* ϕ to be the mean $\ell_{\mathrm{abs}}(e)$ of the absolute median edge lengths at x and y:

$$\ell_{\mathrm{abs}}(e) = \frac{r_{\mathrm{abs}}(x) + r_{\mathrm{abs}}(y)}{2}.$$

Then both the quantities $r_{\mathrm{abs}}(x)$ and $\ell_{\mathrm{abs}}(e)$ are independent of the freedom (7.3).

The median edge length of Definition 7.2 is defined so as to give the length of the edges of a corresponding regular star in $\mathbb{R}^N$, when such exists. In particular, if $G = (V, E)$ is the 1-skeleton of a regular polytope and $\phi : V \to \mathbb{C}$ associates to each vertex its value after an orthogonal projection, then the edge-length at each vertex coincides with the lengths of the edges of the regular polytope. More generally, we can interpret edge length at each vertex as the length of the edges of the "best fit" configured star at that vertex. The median edge length then gives the average length at two adjacent vertices.

Example 7.2 Consider the graph on five vertices sketched in Fig. 7.5, with solutions ϕ to (7.2) normalized so as to take the value 0 at the central vertex and 1 on one of

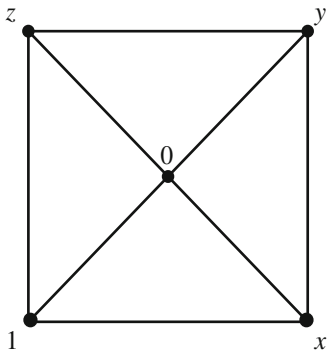


Fig. 7.5 The most general non-constant solution can be normalized to take on the value 0 at the central vertex and 1 at any one of the other vertices

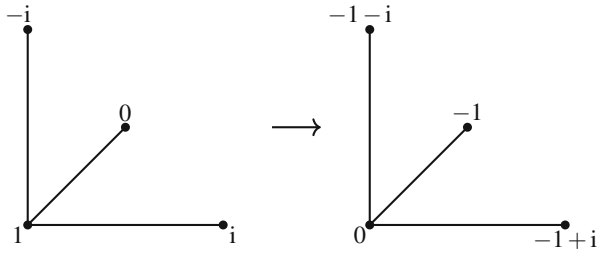


Fig. 7.6 The solution is now normalized to take on the value 0 at the internal vertex of the star

the other vertices. The symmetry of the figure means that this determines the most general non-constant solution.

There are two solutions to (7.2) with γ constant, namely:

$$\begin{aligned} &\gamma = 1/3\,; \quad x = \pm \mathrm{i}\,,\ y = -1\,,\ z = \mp \mathrm{i}\,; \\ &\gamma = 1\,; \qquad x = yz\,,\ y = \tfrac{1}{2} \pm \tfrac{\sqrt{3}}{2}\mathrm{i}\,,\ z = 3 \pm 2\sqrt{2}\,. \end{aligned}$$

We reject the latter solution, since the inequality $\gamma < 1$ is violated. Consider the solution with $\gamma = 1/3$. Let us construct the lift at the bottom left-hand vertex. First, we normalize so that the solution takes on the value 0 at this vertex as indicated in Fig. 7.6.

The choice $n = N = 3$ is determined and from (7.16) and (7.17); we obtain $\rho = 2$ and $\sigma = 1$. From (7.18) we find that ${u_1}^2 + {u_2}^2 = 1$ so that $u_3 = 0$. Then the 3×1-matrix B vanishes and system (7.20) has general solution $\mathbf{x}_3 := (x_{13}, x_{23}, x_{33}) = (2\lambda\,, -\lambda\,, -\lambda)$. The constraint (7.21) requires that $\langle \mathbf{x}_3, \mathbf{x}_3 \rangle = \rho = 2$,

so that $\lambda = \pm 1/\sqrt{3}$. The star matrix W (whose columns give the positions of the external star vertices) is given by:

$$W = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 1 & -1 \\ \pm\frac{2}{\sqrt{3}} & \mp\frac{1}{\sqrt{3}} & \mp\frac{1}{\sqrt{3}} \end{pmatrix}.$$

We can proceed similarly with the central vertex of degree 4. Now we can choose $N = 3$ or $N = 4$. In either case, $u_1 = u_2 = 0$, $\rho = 2$ and $\sigma = 1$, so that, for $N = 3$ we must have $u_3 = \pm 1$. Then

$$A = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 0 \\ 2\sqrt{3} \end{pmatrix}$$

and the unique (minimizing) solution is given by

$$Z = A^{+}B = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}.$$

The star matrix W is given by

$$W = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{pmatrix},$$

where the last line is only defined up to sign.

The (unique) common dimension to define edge length is $N = 3$. At the central vertex the edge length is $\sqrt{7}/2$ and at any of the other vertices, it is $\sqrt{7/3}$. Thus the median edge length of the edge joining the central vertex to one of the other vertices is $\frac{(\sqrt{7}/2)+\sqrt{7/3}}{2} \sim 1.4252$, whereas the median edge length of one of the outside edges is $\sqrt{7/3} \sim 1.5275$. So, for example, the shortest path joining x to z is given by passing through the central vertex. Note that, as already remarked, the edge lengths are only defined up to a multiple and so only relative edge lengths have meaning.

7.4.3 Path Metric Space Structure and Curvature in the Sense of Alexandrov

Does the notion of edge length, either relative or absolute, that we have defined above, endow a graph with the structure of a *path metric space* (see [16])? We first of all note a triangle inequality around complete subgraphs on three vertices.

Given a function $\phi : V \to \mathbb{C}$ and a vertex $x \in V$, we say that ϕ *is constant on the star centered on* x if the restriction of ϕ to x and its neighbors $y \sim x$, is constant.

Proposition 7.2 (Local Triangle Inequality) *Let $G = (V, E)$ be a graph coupled to a solution $\phi : V \to \mathbb{C}$ to Eq. (7.2). Suppose $x, y, z \in V$ are three vertices of a complete subgraph: $x \sim y$, $y \sim z$, $z \sim x$, such that the inequality $\gamma < 1$ is satisfied at each vertex. Then the triangle inequality is satisfied:*

$$\ell(\overline{xy}) + \ell(\overline{xz}) \geq \ell(\overline{yz})\,.$$

If further ϕ is non-constant on the star centered on x, then the inequality is strict.

Proof This is an immediate consequence of the definition. Specifically,

$$\ell(\overline{xy}) + \ell(\overline{xz}) = \frac{1}{2}(r(x) + r(y)) + \frac{1}{2}(r(x) + r(z)) = \ell(\overline{yz}) + r(x) \geq \ell(\overline{yz})\,,$$

since because of the inequality $\gamma(x) < 1$, we have $r(x) \geq 0$. If further, ϕ is non-constant on the star centered on x, then $r(x) > 0$ and the inequality is strict. □

In spite of this local triangle inequality, we may encounter a difficulty in trying to endow a graph coupled to a solution ϕ to (7.2) with a metric space structure. This may arise when, for a given vertex x, the function ϕ is constant on the star centered on x, as well as on the star centered on one of its neighbors y. Then $\ell(\overline{xy}) = 0$. We can either agree to allow distinct points to have zero distance between them, and so consider rather a *pseudo-metric space structure*, or we can avoid this situation by introducing a notion of *collapsing*.

Definition 7.3 Let (G, ϕ) be a graph coupled to a solution to Eq. (7.2). Then we *collapse* G to a new graph $\widetilde{G}$ by removing all edges that connect vertices at which ϕ takes on identical values; then remove all isolated vertices.

It is clear that after collapse, if we let $\widetilde{\phi}$ denote the restriction of ϕ to $\widetilde{G}$, then $\widetilde{\phi}$ also satisfies (7.2) with $\widetilde{\gamma} = \gamma\widetilde{d}/d$ where $\widetilde{d}$ is the new degree at each vertex. Indeed, if we check at a vertex x, then if y is a neighbor at which $\phi(y) = \phi(x)$, then since only the difference $\phi(y) - \phi(x)$ occurs in (7.2), removing the edge $\overline{xy}$ only affects the degree. However, it is to be noted that collapsing may disconnect a graph.

Let (G, ϕ) be a connected graph coupled to a solution of Eq. (7.2). Suppose that for each edge $\overline{xy}$, its length $\ell(\overline{xy}) > 0$. Given a path $\overline{c} := \overline{x_0 x_1 x_2 \cdots x_n}$ joining x to y (so we have $x = x_0$, $y = x_n$ and $x_j \sim x_{j+1}$ for all $j = 0, \ldots, n-1$), we define the length $\ell(\overline{c})$ to be the sum:

$$\sum_{j=0}^{n-1} \ell(\overline{x_j x_{j+1}})\,.$$

Let X be the underlying topological space formed from the union of the vertices and the edges (where edges only intersect at a common end point). We can extend path length to include points belonging to edges in the obvious way, by identifying the segment $[0, \ell(\overline{xy})] \subset \mathbb{R}$ with $\overline{xy}$ and attributing length linearly along the interval $[0, \ell(\overline{xy})]$. Now any two points $\xi, \eta \in X$ have a well-defined distance $d(\xi, \eta)$ between them given as the length of the shortest path joining them. A geodesic segment $[\xi, \eta]$ is a path of length $d(\xi, \eta)$ joining ξ and η. Since any pair of points $\xi, \eta \in X$ can be joined by a continuous path of length $d(\xi, \eta)$, X has the structure of a geodesic space. In this setting, curvature bounds in the sense of Alexandrov arise by comparing triangles in the path metric space X with those in a 2-dimensional space form $M(K)$ of constant Gaussian curvature K; see Chap. 2 by E. Saucan in this volume. In Example 7.2, due to the simple nature of the graph, we can find an exact comparison.

If we let $a = (\sqrt{7}/2 + \sqrt{7/3})/2$ and $b = \sqrt{7/3}$, then there exists a unique radius r of the 2-sphere for which the graph can be placed as indicated sketched in Fig. 7.7, where we require two neighboring exterior vertices to be subtended by a right-angle at the central vertex and the two geodesic segments to have respective lengths a and b. A numerical calculation shows that the required radius r is approximately given by 0.9811, which yields Gauss curvature $K = 1/r^2 \sim 1.0389$, very close to 1.

For a general graph coupled to a solution ϕ of (7.2), we can only expect to obtain bounds on the curvature in the sense of Alexandrov, rather than an exact figure as in the above example. If we first normalize ϕ so as to use absolute edge length, then an interesting problem would be to express such bounds in terms of the geometric spectrum γ.

Note that the positive curvature of the above example can also be deduced by drawing a Euclidean triangle with side-lengths as indicated in Fig. 7.8.

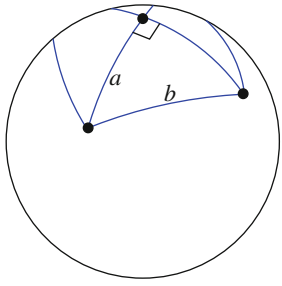


Fig. 7.7 There is a unique 2-sphere into which the graph can be embedded to realize the correct edge lengths

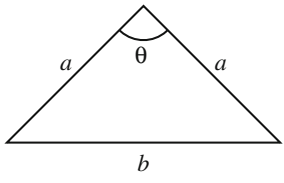


Fig. 7.8 Angle deficit at the internal vertex also implies positive curvature

The angle θ subtended by the two external vertices is calculated to be about 1.13 radians, yielding an angular deficit for the sum of the interior angles at the central vertex to be approximately $2\pi - 4 \times 1.13 \sim 2.04$ radians. Gaussian curvature in this sense, defined in terms of angular deficit is the subject of the next section.

7.5 Gaussian Curvature

We first review some classical notions of curvature associated to polytopes defined in terms of angular deficiency at a vertex, see also Sect. 6.2.1 of Chap. 6 in this volume. Even though we are mainly interested in liftings into dimension $N = 3$, we shall in the first instance consider general N, since, as we shall see, this provides a practical way to calculate γ for any regular polytope. Our approach is pragmatic in that our expression for curvature is approximate, exact formulae being difficult in general to write down.

7.5.1 The Theorem of Descartes

A *convex* polytope is by definition, the closed intersection of a finite number of half-spaces (whether this be in Euclidean space, or in spherical space). In the case when a polytope is regular (convex or not), its vertices all lie on a sphere called the *circumsphere* [11]. It is useful to use *absolute angle measure* when measuring solid angles (see [17, 26]). We will write $H^M(\Lambda)$ for the M-dimensional Hausdorff measure of a set Λ in these units. Then, in any dimension, the angle is measured as a fraction of the total angle subtended by a sphere centered at the point in question. Thus in two dimensions, a right-angle has value $1/4$, whereas in three dimensions, the angle subtended by the vertex figure of a cube has value $1/8$. Equivalently, $H^2(\Lambda) = 1/8$, where $\Lambda \subset S^2$ represents one eighth portion: $x, y, z > 0$ of the sphere $x^2 + y^2 + z^2 = 1$.

A family $\{F_1, \ldots, F_r\}$ of $(N-1)$-dimensional convex polytopes in $\mathbb{R}^N$ form an *elementary polytope* P of dimension N if: (i) for all j, k, $F_j \cap F_k$ is either empty or a face of each of F_j and F_k; (ii) $\cup_j F_j$ is an $(N-1)$-dimensional manifold.

Given an elementary polytope $P \subset \mathbb{R}^N$, denote by $\mathscr{P}$ the *face decomposition of* P; thus $\mathscr{P}$ is the collection of all (open) faces of all dimension, consisting of the vertices, edges, $\ldots$, $(N-1)$-faces. For $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{w} \in S^{N-1}$, following Ehrenborg [14], we define the quantity $R(\mathbf{x}, \mathbf{w}) := \lim_{s \to 0+} \mathbf{1}_P(\mathbf{x} + s \cdot \mathbf{w})$, where $\mathbf{1}_P$ is the characteristic function of P. Note that this takes on the value 0 or 1. Then given a face $F \in \mathscr{P}$, we have $R(\mathbf{x}, \mathbf{w}) = R(\mathbf{y}, \mathbf{w})$ for all $\mathbf{x}, \mathbf{y} \in F$; write this as $R(F, \mathbf{w})$ and define $\Lambda_F = \{\mathbf{w} \in S^{N-1} : R(F, \mathbf{w}) = 1\}$.

Let $S(\Lambda_F) = H^{N-2}(\partial\Lambda_F)$ and let $\sigma_N = H^N(S^N)$, so that in absolute angle measure, $\sigma_N = 1$. Let F be a face of P of dimension $\leq N-3$. We define the *deficiency at* F to be the quantity:

$$\delta(F) := \sigma_{N-2} - S(\Lambda_F)\,.$$

Note that if $Q \subset S^{N-1}$ is spherically convex (that is, it is the intersection of hemispheres), then $S(Q) = H^{N-2}(\partial Q)$ is proportional to the Haar measure of all the great circles which intersect Q. The following theorem generalizes a classical result of Descartes.

Theorem 7.2 ([14, 17, 26]) *Let P be an elementary polytope with face decomposition $\mathscr{P}$. Then*

$$\sum_{F\in\mathscr{P},\ \dim F\leq N-3} \varepsilon(F)\delta(F) = \sigma_{N-2}\,\varepsilon(P)\,,$$

where $\varepsilon(F)$ denotes the Euler characteristic of F given by $(-1)^k$ when F is of dimension k and $\varepsilon(P)$ is the Euler characteristic of P.

In the case when $N = 3$ and P is a convex polyhedron (now using radians for our measure), we obtain the classical theorem of Descartes:

$$\sum_{\ell} \delta(\mathbf{v}_\ell) = 4\pi\,,$$

where the sum is taken over the vertices of P. Here, the deficiency $\delta(\mathbf{v}_\ell)$ is $2\pi-$ (the sum of the internal angles at $\mathbf{v}_\ell$ of the faces which contain $\mathbf{v}_\ell$) (see also Theorem 2.1 of Chap. 6).

7.5.2 Vertex Curvature

The definition of curvature that we give below coincides with the angular deficiency discussed in the above paragraph when the solution to (7.2) arises from the underlying framework of a regular convex polyhedron. The term *vertex-curvature* is used to distinguish it from the previously defined *edge-curvature*. In the case when $N = 3$, the vertex curvature coincides with the same notion considered by M. Keller, Chap. 6, Sect. 6.1.3 in this volume for graphs embedded in a surface, provided the embedding permits a configuration about the vertex in question.

Definition 7.4 (N-dimensional Vertex-Curvature) Let (G,ϕ) be a pair consisting of a graph G coupled to a solution ϕ to *(7.2)* with $\gamma < 1$. Let y be a vertex of G and let d be the degree of G at y. Let $\{\mathbf{v}_1,\ldots,\mathbf{v}_d\}$ be the configuration in $\mathbb{R}^{N-1}$

(centered on $\mathbf{0} \in \mathbb{R}^{N-1}$) of an associated lifted configured star in $\mathbb{R}^N$ with invariant ρ given by *(7.6)*. Let

$$\mathbf{x}_\ell = \frac{1}{\sqrt{\rho + dr^2(1-\gamma)}} \begin{pmatrix} \sqrt{d(1-\gamma)}\,\mathbf{v}_\ell \\ \sqrt{\rho} \end{pmatrix} \in S^{N-1} \quad (\ell = 1, \dots, d)\,,$$

be the corresponding normalized vertices of a configured star in $\mathbb{R}^N$ centered on $\mathbf{0}$. Let Λ be the convex hull in S^{N-1} of the set $\{\mathbf{x}_1, \dots, \mathbf{x}_d\} \subset S^{N-1}$. We define the *N-dimensional vertex-curvature of* (G, ϕ) *at* y to be the deficit:

$$\delta_\gamma(y) = 1 - H^{N-2}(\partial\Lambda)\,,$$

in absolute angle measure.

Note that in the above definition, if we rescale the vertices $\mathbf{v}_\ell$ by $\mathbf{v}_\ell \mapsto \mathbf{v}_\ell^{\sim} = \lambda \mathbf{v}_\ell$, say, then $r \mapsto \widetilde{r} = \lambda r$, $\rho \mapsto \widetilde{\rho} = \lambda^2 \rho$ and both $\mathbf{x}_j$ and the curvature $\delta_\gamma(y)$ are well-defined and independent of this scaling.

Let us now consider the different dimensional curvatures that arise from Definition 7.4. In general, we would like this curvature to depend only on γ and not on the solution ϕ to (7.2). For this reason, we will consider only cases when the configuration given by $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ in the above definition, coincides with the vertices of a regular polytope in $\mathbb{R}^{N-1}$. Note that the vertices of a regular polytope always form a configuration [3].

By convention, at a vertex of degree 1, we assign the curvature $\delta = 1$. At a vertex of degree 2, the 2-curvature just measures the exterior angle in absolute angle measure. Note that in this dimension, a configured star with two external vertices corresponds to two line segments in the plane of equal length meeting at the internal vertex (see Appendix 1). If θ is the exterior angle, then from Section "First Cases" in Appendix 1,

$$\gamma = \frac{2\cos\theta}{\cos\theta - 1}\,,$$

from which one deduces the following expression for the vertex curvature.

Proposition 7.3 *Let* (G, ϕ) *be a pair consisting of a graph* G *coupled to a solution* ϕ *to* (7.2). *Let* y *be a vertex of degree* 2 *where* $\gamma \leq 1$. *Then the* 2*-dimensional vertex-curvature of* (G, ϕ) *at* y *is given by the quantity:*

$$\delta_\gamma(y) = \frac{1}{2\pi} \arccos\left(\frac{\gamma}{\gamma - 2}\right).$$

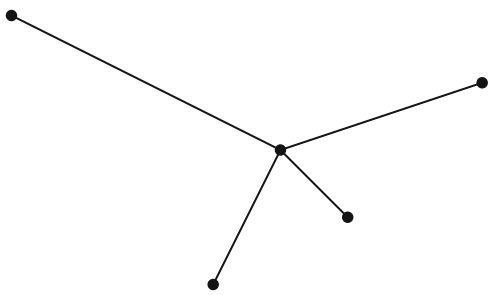
Note that $\lim_{\gamma \to 1^-} \delta_\gamma(y)$ is well-defined and equals $1/2$. For example, if G is a cyclic graph of even order $2n$ and ϕ is a function taking on alternate values at neighboring vertices, then ϕ satisfies (7.2) with $\gamma = 1$. The total curvature is then given by n. If G is a convex polygon in the plane and ϕ the corresponding position function, then the total 2-curvature is equal to 1, or in radians, to 2π, as required. We now proceed to higher dimensional curvature.

Proposition 7.4 (3-dimensional Vertex-Curvature) *Let (G, ϕ) be a pair consisting of a graph G coupled to a solution ϕ to (7.2). Let y be a vertex of G and let d be the degree of G at y. Suppose that the vertices of a lifted configured star coincide with those of a regular planar polygon. Then the 3-dimensional vertex-curvature of (G, ϕ) at y is given by the quantity:*

$$\delta_\gamma(y) = 1 - \frac{d}{2\pi} \arccos\left\{ \frac{1 + 2(1-\gamma)\cos\frac{2\pi}{d}}{3 - 2\gamma} \right\} . \tag{7.23}$$

Proof The configuration of vertices is given by $\mathbf{v}_\ell = \mathrm{e}^{2\pi \mathrm{i}\ell/d}$ $(\ell = 1, \ldots, d)$. The boundary of the convex hull Λ of the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$ in S^2 is made up of arcs of great circles of length $\alpha = \arccos(\mathbf{x}_1 \cdot \mathbf{x}_2)$. In absolute angle measure, the deficit, or 3-curvature, is given by $1 - \frac{d}{2\pi}\alpha$. Substitution of the expressions for $\mathbf{x}_\ell$ given by Definition 7.4 gives the required formula. □

Although the above proposition only applies to configured stars whose vertices coincide with a regular planar polygon, since the expression for the curvature depends only on γ, we can apply the formula to more general situations. However, we can only expect this to give an approximation of the true angular deficit. Note that $\lim_{\gamma \to 1-} \delta_\gamma(y)$ is well-defined and equals 1 and that $\lim_{\gamma \to -\infty} \delta_\gamma(y) = 0$. As a further remark, an invariant framework may contain an invariant star which is not configured, for which one would desire a negative curvature, as illustrated:



An example of an invariant star with negative curvature at the internal vertex

The challenge in future investigations will be to characterize conditions that allow for a global lifting to an invariant framework, but without the constraint that each star be configured. Thus we allow stars such as that given by (7.10) with negative curvature at the internal vertex given as $2\pi - \sum \theta_j$ where the θ_j are the angles between the incident edges. This requires an ordering of the edges in order to decide how to calculate the angles. However, since in this example, all the angles are equal, simple trigonometry gives $2\pi - 8\arcsin(\sqrt{2/3}) \sim -1.36$ for the curvature in radian measure.

For the convex polyhedra, Proposition 7.4 gives the required result. For example, the 1-skeleton of a tetrahedron in $\mathbb{R}^3$ has $d = 3$ and $\gamma = 3/4$. The 3-curvature at each vertex is given by $\delta = 1/2$, giving a total curvature of 2. To obtain the total curvature in radian measure, we multiply by 2π to give the value 4π, which confirms

the theorem of Descartes. The same proposition can also be used to calculate the values of γ for the regular convex polyhedra as given in the Introduction. For example, the dodecahedron has 20 vertices, so by the Theorem of Descartes, the deficit δ at each vertex must equal $1/10$. The degree of each vertex is $d = 3$ and substitution into (7.23) gives the required value $\gamma = \frac{3(1-\sqrt{5})}{2(3-\sqrt{5})}$.

Example 7.3 Consider the double cone on the triangle discussed at the end of Sect. 7.3. Now there is an underlying polytope and we can calculate angular deficit at each vertex in the traditional way. By the theorem of Descartes, the total angular deficit is 4π. However, let us calculate it by taking a lift to a configured star at each vertex as determined by the corresponding solution to (7.2). In the original figure, the stars at the lateral vertices are not configured, so an error will occur. We find:

$$\delta_{\text{apex}} = 2\pi - 3\arccos\frac{1}{7} \quad \text{and} \quad \delta_{\text{lat}} = 2\pi - 4\arccos\frac{5}{7},$$

to give a total curvature of

$$\delta_{\text{tot}} = 3\delta_{\text{lat}} + 2\delta_{\text{apex}} = 10\pi - 6\left(\arccos\frac{1}{7} + 2\arccos\frac{5}{7}\right) \sim 4.244 \times \pi\,.$$

The above example illustrates one of the problems in defining the curvature. The advantage of lifting to a *configured* star is that, in dimension $N = 3$, the lift is unique up to sign changes, to give a uniquely defined curvature. However, any expression of the total curvature as an invariant quantity would need to involve some approximation.

For the 4-curvature, there are some special cases to consider. We list these in the proposition below.

Proposition 7.5 (4-dimensional Vertex-Curvature) *Let (G, ϕ) be a pair consisting of a graph G coupled to a solution ϕ to (7.2). Let y be a vertex of G and let d be the degree of G at y. Suppose that $d \in \{4, 6, 12, 20\}$ and that the configuration associated to a lifted configured star coincides with the vertices of a regular 3-polytope (having d vertices). Then depending on the degree, the 4-dimensional vertex-curvature of (G, ϕ) at y is given by one of the expressions below:*

Degree	*Configuration*	*Curvature*
4	*Tetrahedron*	$2 - \frac{3}{\pi}\arccos\left(\frac{\gamma}{4-2\gamma}\right)$
6	*Octahedron*	$3 - \frac{6}{\pi}\arccos\left(\frac{1}{5-3\gamma}\right)$
12	*Icosahedron*	$6 - \frac{15}{\pi}\arccos\left(\frac{6(\sqrt{5}+1)\gamma - 11 - 7\sqrt{5}}{2[6(\sqrt{5}+3)\gamma - 23 - 7\sqrt{5}]}\right)$
20	*Dodecahedron*	$10 - \frac{15}{\pi}\arccos\left(\frac{5\gamma - 1 - 2\sqrt{5}}{2[-5\gamma + 8 - \sqrt{5}]}\right)$

Proof Given two vectors $\mathbf{u}, \mathbf{v} \in S^M(r)$ in a sphere of radius r, the arc of the great circle joining $\mathbf{u}$ to $\mathbf{v}$ is given by

$$\theta \mapsto \left(\cos\theta - \frac{(\mathbf{u}\cdot\mathbf{v})\sin\theta}{\sqrt{r^4-(\mathbf{u}\cdot\mathbf{v})^2}}\right)\mathbf{u} + \frac{r^2\sin\theta}{\sqrt{r^4-(\mathbf{u}\cdot\mathbf{v})^2}}\,\mathbf{v}$$
$$(0 \le \theta \le \arcsin\tfrac{\sqrt{r^4-(\mathbf{u}\cdot\mathbf{v})^2}}{r^2}) .$$

When $r = 1$, this is unit speed. Furthermore, the tangent vector to this arc at $\mathbf{v}$ is given by

$$\mathbf{t} = \frac{1}{\sqrt{r^4-(\mathbf{u}\cdot\mathbf{v})^2}}(-r^2\mathbf{u} + (\mathbf{u}\cdot\mathbf{v})\mathbf{v}) . \tag{7.24}$$

The area of a spherical polygon with n sides and with interior angles θ_k ($k = 1, \ldots, n$) is given by

$$KA = \sum_k \theta_k - (n-2)\pi ,$$

where K is the curvature of the sphere. We are required to calculate the spherical surface area of the boundaries of the various vertex figures in S^3. These are made up of faces lying in great 2-spheres which are either triangles, or in the case of the dodecahedron, pentagons, whose edges are arcs of great circles. In order to calculate the interior angles, we calculate the scalar product of the unit tangents to these edges at a vertex. By symmetry, any vertex will do. We calculate this for the icosahedron and the dodecahedron, the other cases being similar.

For the dodecahedron, a configuration of vertices is given by

$$\begin{pmatrix} 0 & \pm\lambda & \pm\lambda^{-1} & \pm 1 \\ \pm\lambda^{-1} & 0 & \pm\lambda & \pm 1 \\ \pm\lambda & \pm\lambda^{-1} & 0 & \pm 1 \end{pmatrix} , \tag{7.25}$$

where $\lambda = \frac{1+\sqrt{5}}{2}$ (see [11], Sect. 3.8). With the notation of Definition 7.4, three consecutive vertices around one pentagonal face are given by

$$\mathbf{x}_1 = \frac{1}{\sqrt{8+3a^2}}\begin{pmatrix} a\lambda^{-1} \\ a\lambda \\ 0 \\ \sqrt{8} \end{pmatrix}, \mathbf{x}_2 = \frac{1}{\sqrt{8+3a^2}}\begin{pmatrix} a \\ a \\ a \\ \sqrt{8} \end{pmatrix}, \mathbf{x}_3 = \frac{1}{\sqrt{8+3a^2}}\begin{pmatrix} 0 \\ a\lambda^{-1} \\ a\lambda \\ \sqrt{8} \end{pmatrix},$$

where $a = \sqrt{d(1-\gamma)}$ and $\lambda = (1+\sqrt{5})/2$. These three vertices determine a great 2-sphere in S^3 which contains the pentagonal face. With this arrangement, $\mathbf{x}_2$ is the

central vertex joined to $\mathbf{x}_1$ and $\mathbf{x}_3$ by arcs of great circles. In order to calculate the interior angle at each vertex of the pentagon, we calculate the tangent to each of these arcs at $\mathbf{x}_2$. To do this we apply (7.24). For the first arc we set $\mathbf{u} = \mathbf{x}_1$ and $\mathbf{v} = \mathbf{x}_2$, to obtain the tangent vector:

$$\mathbf{t}_1 = \frac{1}{2\sqrt{3a^2+8}\sqrt{a^2+12-4\sqrt{5}}} \begin{pmatrix} \frac{a^2}{2}(3-\sqrt{5}) + 4(3-\sqrt{5}) \\ -\frac{a^2}{2}(3+\sqrt{5}) + 4(1-\sqrt{5}) \\ \sqrt{5}a^2 + 8 \\ -a\sqrt{8}(3-\sqrt{5}) \end{pmatrix}.$$

For the second arc, we set $\mathbf{u} = \mathbf{x}_3$ and $\mathbf{v} = \mathbf{x}_2$, to obtain:

$$\mathbf{t}_2 = \frac{1}{2\sqrt{3a^2+8}\sqrt{a^2+12-4\sqrt{5}}} \begin{pmatrix} \sqrt{5}a^2 + 8 \\ \frac{a^2}{2}(3-\sqrt{5}) + 4(3-\sqrt{5}) \\ -\frac{a^2}{2}(3+\sqrt{5}) + 4(1-\sqrt{5}) \\ -a\sqrt{8}(3-\sqrt{5}) \end{pmatrix}.$$

Then

$$\mathbf{t}_1 \cdot \mathbf{t}_2 = \frac{-a^2+16-8\sqrt{5}}{2(a^2+12-4\sqrt{5})},$$

which gives the cosine of the interior angle (it is indeed the interior angle, being greater than $\pi/2$). Then the area (in absolute angle measure) of each pentagonal face is given by

$$\frac{1}{4\pi}\left\{5\arccos\left(\frac{-a^2+16-8\sqrt{5}}{2(a^2+12-4\sqrt{5})}\right) - 3\pi\right\},$$

so that the surface area of the spherical dodecahedron is given by twelve times this quantity. We then obtain the angular deficiency, or 4-curvature:

$$\delta = 10 - \frac{15}{\pi}\arccos\left(\frac{-a^2+16-8\sqrt{5}}{2(a^2+12-4\sqrt{5})}\right).$$

On substituting the value of a, we obtain the required formula.

For the icosahedron, a configuration of vertices is given by

$$\begin{pmatrix} 0 & \pm 1 & \pm\lambda \\ \pm 1 & \pm\lambda & 0 \\ \pm\lambda & 0 & \pm 1 \end{pmatrix}, \tag{7.26}$$

where $\lambda = \frac{1+\sqrt{5}}{2}$. Three vertices which form one of the triangular faces are given by:

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ \lambda \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ \lambda \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} \lambda \\ 0 \\ 1 \end{pmatrix},$$

where $\lambda = (1+\sqrt{5})/2$. Then $\rho = 2+2\lambda^2 = 5+\sqrt{5}$, $n = 12$ and $r^2 = (5+\sqrt{5})/2$. This gives the corresponding vertices in S^3 as:

$$\mathbf{x}_\ell = \frac{\sqrt{2}}{\sqrt{2+d(1-\gamma)}} \begin{pmatrix} \sqrt{\frac{d(1-\gamma)}{5+\sqrt{5}}}\,\mathbf{v}_\ell \\ 1 \end{pmatrix} \quad (\ell = 1, 2, 3)\,.$$

We then proceed as above for the dodecahedron to calculate the angular deficiency. □

Example 7.4 The 600-cell is a convex 4-dimensional regular polytope made up of 600 tetrahedral 3-polytopes. It has 120 vertices and 720 edges. Its vertex figure is a regular icosahedron. If we consider an orthogonal projection onto the complex plane and let ϕ associate the corresponding value to each vertex, then ϕ satisfies (7.2) with γ constant. We can find the value of γ as follows.

Since the edges of the 600-cell all have the same length, in the notation of the above proof, we must have the distance from the origin to $\mathbf{x}_1$, that is 1, equal to the distance between two neighbors of the vertex figure: $||\mathbf{x}_1 - \mathbf{x}_2||$. One can readily calculate:

$$||\mathbf{x}_1 - \mathbf{x}_2||^2 = \frac{8d(1-\gamma)}{(5+\sqrt{5})[2+d(1-\gamma)]}\,,$$

to obtain the negative value:

$$\gamma = \frac{5(1-2\sqrt{5})}{3}\,.$$

One can now confirm the generalization of the theorem of Descartes (Theorem 7.2).

The 600-cell has 5 tetrahedra around each of its edges, each having dihedral angle $\arccos(1/3)$. Thus the angular deficiency at each edge (in absolute angle measure) is given by:

$$\delta_e = 1 - \frac{5}{2\pi}\arccos\frac{1}{3}\,.$$

Substitution of the above value of γ into the third formula of Proposition 7.5 gives the deficit, or curvature at each vertex, as

$$\delta_v = 6 - \frac{15}{\pi}\arccos\frac{1}{3}\,.$$

On then finds that $120\delta_v - 720\delta_e = 0$, as required.

We can proceed similarly to obtain explicit formulae for higher dimensional N-curvature. This is simplified by the fact that there are just three regular polytopes in dimensions $N \geq 5$, namely the N-simplex, the N-cube and the cross-polytope, with vertex figures an $(N-1)$-simplex in the first two cases and another cross-polytope in the last case. To find the N-curvature requires the calculation of the $(N-2)$-dimensional measure of $(N-2)$-simplices in great spheres of S^{N-1}. This is a standard, but non-trivial procedure using Schläfli's differential equality [25]. See also the expository article of J. Milnor for a nice account and references [21, pp. 281–295]. The article of J. Murakami provides explicit expressions in the 3-sphere [22]. We do not attempt to derive these formulae here.

7.6 Other Curvatures

Recall from Sect. 7.4, that given an edge $e = \overline{xy} \in E$, the *edge-curvature* of e is defined to be the unique angle $\theta(e) := \arccos(\langle \mathbf{u}(x), \mathbf{u}(y) \rangle_{\mathbb{R}^N}) \in [0, \pi]$ between the axes of lifted configured stars. By analogy with Riemannian geometry, various other curvatures can now be defined. If we let $\ell(e)$ denote the length of an edge $e = \overline{xy}$ as given by Definition 7.2, and $\theta(e)$ its edge-curvature, then the radius of the best-fit circle is given by $r(e) = \ell(e)/\theta(e)$ (by *best-fit circle*, we mean the circle subtending the same arc length $\ell(e)$ for the given angle $\theta(e)$). The *normal curvature* of e is then the reciprocal $1/r(e) = \theta(e)/\ell(e)$. The *mean curvature* at a vertex x can then be defined to be the mean of the normal curvatures of the edges incident with x. Since $\ell(e)$ depends on the scaling $\phi \mapsto \lambda\phi$, this quantity also depends on the scaling; the mean curvature should be thought of as the analogue of the same notion in the smooth setting, whereby we locally embed a Riemannian manifold in a Euclidean space.

From Theorem 7.1, we already have γ related to mean curvature by the formula $\gamma = -1/H^2$. However, this latter quantity will only approximate mean curvature for an invariant framework that closely approximates a smooth surface. For example, the cube has $\gamma \equiv 0$, which is far from the mean curvature of a sphere, whereas the 600-cell has negative γ, which is more realistic. We don't know if the two notions are related, or if they are related to more classical definitions of mean curvature for polytopes such as are given in [18, 23].

Ricci curvature is one of the most natural curvatures intrinsic to a Riemannian manifold. Recall that given two unit directions $\mathbf{X}$ and $\mathbf{Y}$, the sectional curvature $\mathrm{Sec}(\mathbf{X}, \mathbf{Y})$ can be interpreted as the Gaussian curvature of a small geodesic surface generated by the plane $\mathbf{X} \wedge \mathbf{Y}$. This in turn is the product of the principal curvatures which are the extremal values of the normal curvatures. The Ricci curvature $\mathrm{Ric}(\mathbf{X}, \mathbf{X})$ is then the sum: $\sum_j \mathrm{Sec}(\mathbf{X}, \mathbf{Y}_j)$ taken over an orthonormal frame $\{\mathbf{Y}_j\}$ with each $\mathbf{Y}_j$ orthogonal to $\mathbf{X}$. This motivates the following definition.

Definition 7.5 Given a vertex $x \in V$ and two edges $e_1 = \overline{xy_1}$ and $e_2 = \overline{xy_2}$ with endpoint x, we define the *sectional curvature* $\mathrm{Sec}\,(e_1, e_2)$ *determined by* e_1 *and* e_2 to be the product:

$$\mathrm{Sec}_x\,(e_1, e_2) = \frac{\theta(e_1)}{\ell(e_1)}\frac{\theta(e_2)}{\ell(e_2)}\,,$$

where $\theta(e_j)$ $(j = 1, 2)$ are the edge-curvatures. For an edge $e = \overline{xy}$, the *Ricci curvature* $\mathrm{Ric}\,(e, e)$ is the sum

$$\mathrm{Ric}_x\,(e, e) = \ell(e)^2 \sum_{z\sim x, z\neq y} \mathrm{Sec}_x(e, \overline{xz}) = \ell(e)^2 \sum_{z\sim x, z\neq y} \frac{\theta(\overline{xy})}{\ell(\overline{xy})}\frac{\theta(\overline{xz})}{\ell(\overline{xz})}\,,$$

and the *scalar curvature at* x is given by

$$\mathrm{Scal}_x = \sum_{y\sim x} \mathrm{Ric}\,(\overline{xy}, \overline{xy})/\ell(\overline{xy})^2.$$

As in Riemannian geometry, the Ricci curvature is scale invariant. In fact, it makes complete sense to introduce the factor $\ell(e)^2$, since the Ricci curvature is bilinear in its arguments. On the other hand, both the sectional curvature and the scalar curvature will decrease as we dilate a solution to (7.2) by some scaling $\phi \mapsto \lambda\phi$ (λ constant), since that will modify edge length by the same factor λ. In Riemannian geometry, one usually applies the polarization identity to define $\mathrm{Ric}\,(\mathbf{X}, \mathbf{Y})$, however, there would seem to be no reasonable interpretation for the sum of two edges in our setting.

In [20], Romon, provides numerical calculations of various curvatures associated to the convex regular polyhedra, notably the Ricci curvature of Ollivier. In Table 7.2, we produce similar calculations for the above curvatures, together with Ollivier's Ricci curvature for comparison. Since the sectional and scalar curvatures are scale dependent, we have taken edge length equal to 1 in their definitions. Rational values are exact.

Table 7.2 Table of the various curvatures associated to the convex polyhedra with Ollivier's Ricci curvature included for comparison

Curvature	Tetrahedron	Cube	Octahedron	Icosahedron	Dodecahedron
Vertex (Gauss)	1/2 = 0.50	1/4 = 0.25	1/3 = 0.33	1/6 = 0.17	1/10 = 0.10
Edge	0.30	1/4 = 0.25	1/4 = 0.25	0.18	0.11
Sectional	0.09	1/16 = 0.63	1/16 = 0.63	0.03	0.01
Scalar	0.54	3/8 = 0.36	3/4 = 0.75	0.60	0.07
Ricci	0.18	1/8 = 0.13	3/16 = 0.19	0.12	0.02
Ollivier's Ricci	4/3 = 1.33	2/3 = 0.67	1 = 1.00	2/5 = 0.40	0 = 0.00

Appendix 1: The Geometric Spectrum, Gröbner Bases and the γ-Polynomial

To compute the geometric spectrum, even for simple graphs, is quite challenging. We consider some fundamental cases and then make a simplifying assumption in order to apply the technique of Gröbner bases. This motivates the construction of a new polynomial invariant associated to a graph.

First Cases

Consider a vertex v_0 of degree 2. Suppose that a solution ϕ to (7.2) is non-constant on a neighborhood of this vertex, thus ϕ takes a different value on at least one of its two neighbors v_1, v_2, say v_1. By the normalization (7.3), we may suppose that $\phi(v_0) = 0$ and that $\phi(v_1) = 1$. If we let $\phi(v_2) = z$ as illustrated in Fig. 7.9, then at vertex v_0, Eq. (7.2) takes the form:

$$\frac{\gamma}{2}(1+z)^2 = 1 + z^2\,.$$

Suppose that $z \neq -1$. Then the requirement that γ be real is equivalent to

$$\text{either} \quad \Im(z) = 0 \quad \text{or} \quad |z| = 1.$$

If z is real and $z \neq \pm 1$, then $\gamma > 1$, so we reject this case since γ will not lie in the geometric spectrum. Suppose that z is not real and write it in polar coordinates: $z = \mathrm{e}^{\mathrm{i}\alpha}$. Then

$$\gamma = \frac{2\cos\alpha}{1+\cos\alpha} = \frac{2\cos\theta}{\cos\theta - 1}\,,$$

where $\theta = \pi - \alpha$ is the external angle. The two limiting cases $\alpha = 0$ and $\alpha = \pi$ correspond to $\gamma = 1$ and $\gamma = -\infty$, respectively. This justifies the admissibility of the value 1 in the geometric spectrum at a vertex of degree 2 (see Sect. 7.3.2).

As a consequence, for a cyclic graph, solutions to (7.2) correspond to realizations in the plane as a polygonal framework with sides of equal length. A cyclic graph on

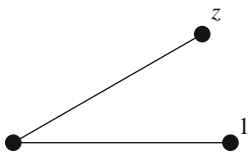


Fig. 7.9 The general invariant star at a vertex of degree 2

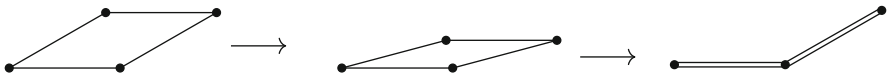


Fig. 7.10 An invariant cyclic graph of order 4 has geometric spectrum corresponding to realizations in the plane as a polygonal framework with sides of equal length

three vertices can only have one such realization (up to similarity transformation) as an equilateral triangle, so that $\gamma = 2/3$ is the only value in the geometric spectrum.

A cyclic graph on four vertices is realized as a rhombus, with the possibility of its edges collapsing onto themselves as indicated in Fig. 7.10.
In this case, the geometric spectrum has continuous components corresponding to continuous deformations of the rhombus which may be parametrized by one of the internal angles. Branching phenomena occurs as edges collapse onto themselves. As the number of vertices increases, so too does the number of variables that parametrize the geometric spectrum.

Let us turn to the other extreme, that of a complete graph. The complete graph on three vertices is cyclic and as shown above, there is just one element in its geometric spectrum, the constant value $\gamma = 2/3$. Since any invariant star in $\mathbb{R}^3$ with three external vertices is necessarily configured [3], it follows that a complete graph on four vertices which has a realization as an invariant framework in $\mathbb{R}^3$ is necessarily a tetrahedron with γ taking the constant value $3/4$. However, the results of Sect. 7.3 only guarantee a local lifting of a vertex and its neighbors in an invariant way, so in general, we don't know if such a realization exists. However, we can use a computer to solve the equations for a sufficiently small number of vertices.

Consider the complete graph on $N+1$ vertices. Label the vertices by the integers $0, 1, \ldots, N$ and set $\gamma_j = \gamma(j), z_j = \phi(j)$. After normalization (7.3), we may suppose that $z_0 = 0$ and $z_1 = 1$. Then (7.2) becomes:

$$\frac{\gamma_k}{N}\left(\sum_{j=0}^{N}(z_k - z_j)\right)^2 = \sum_{j=0}^{N}(z_k - z_j)^2 \quad (k = 0, 1, 2, \ldots, N)\,,$$

with the constraints that each γ_k be real and < 1. The software MAPLE can now solve this algebraic system at least up to $N = 5$ (the complete graph on 6 vertices), which confirms that the only element in the geometric spectrum is given by γ constant equal to $N/(N+1)$. We currently don't have a mathematical (non-computer) proof of this fact and don't know if the constancy of γ persists for $N > 5$.

The Constant Geometric Spectrum and the γ-Polynomial

From now on, to simplify matters, we consider only constant values of γ that may lie in the spectrum and attempt an algebraic geometric approach to compute this set which we refer to as the *constant geometric spectrum*.

Let $G = (V, E)$ be a connected graph. We are interested in the possible real numbers γ for which there are non-constant solutions to the equation:

$$\gamma \Delta \phi^2 = (d\phi)^2 \,. \tag{7.27}$$

Any solution is invariant by $\phi \mapsto \lambda\phi + \mu$, for complex constants $\lambda, \mu (\lambda \neq 0)$. We will no longer insist that $\gamma < 1$. Consider first how to parametrize all possible complex fields on the graph under this invariance.

Label the vertices of the graph $x_1, x_2, \ldots, x_N$ and consider a non-constant complex field ϕ that assigns the value $\phi(x_k) = z_k$ to vertex x_k. Then the space of all such fields is identified with the complex space $\mathbb{C}^N \setminus \{\mu(1, 1, \ldots, 1) : \mu \in \mathbb{C}\}$. Up to the equivalence $(z_1, \ldots, z_N) \sim (z_1 + \mu, \ldots, z_N + \mu)$, we can identify these fields with the set $\Pi \setminus \{0\}$, where Π is the linear subspace $\Pi = \{\mathbf{Z} = (z_1, \ldots, z_N) \in \mathbb{C}^N : z_1 + \cdots + z_n = 0\} \subset \mathbb{C}^N$. In effect, given any non-constant field $(z_1, \ldots, z_N)$, then $(z_1 + \mu, \ldots, z_N + \mu)$ lies in the plane $z_1 + \cdots + z_N = 0$, when we set $\mu = -\frac{1}{N}(z_1 + \cdots + z_N)$. By non-constancy, this is non-zero. Furthermore, it is clear that any two equivalent fields correspond to the same point.

Now consider the relation $\mathbf{Z} \sim \lambda\mathbf{Z}$, for $\lambda \in \mathbb{C} \setminus \{0\}$. This defines the *moduli space* of fields up to equivalence to be $\mathscr{Z} := \mathbb{C}\mathrm{P}^{N-2}$. Specifically, given a point $[z_1, \ldots, z_{N-1}] \in \mathscr{Z}$ in homogeneous coordinates, we define a representative field by $(z_1, \ldots, z_{N-1}, z_N = -\sum_{k=1}^{N-1} z_k) \in \mathbb{C}^N$. In practice, we can set a field equal to 0 and 1 on two selected vertices x_0 and x_1, respectively, and label the other vertices arbitrarily. This is only one chart and we miss those fields which coincide at these two vertices.

If we consider γ as an arbitrary *complex* parameter, then (7.27) imposes a constraint at each vertex, so we have N equations in $N - 1$ parameters. In general these are independent so that this is an overdetermined system, which may have no solutions.

For each $\ell = 2, \ldots, N$, consider the following set of N polynomials defined over the algebraically closed field $\mathbb{C}$. The variables are the values $\{z_1, \ldots, z_N\}$ of a field on G with constraints $z_1 = 0$ and $z_\ell = 1$; we suppose the degree of vertex j is d_j and that $z_{jk} \in \{z_1, \ldots, z_N\}$ $(k = 1, \ldots, d_j)$ are the values of the field on the neighbors x_{jk} of x_j. The polynomials are then defined by

$$f_j^{\ell} := \frac{\gamma}{d_j}\left(\sum_{k=1}^{d_j}(z_j - z_{jk})\right)^2 - \sum_{k=1}^{d_j}(z_j - z_{jk})^2 \qquad (z_1 = 0,\, z_\ell = 1)\,,$$

in the $N - 1$ *complex* variables $\{\gamma, z_2, z_3, \ldots, \widehat{z_\ell}, \ldots, z_N\}$. Recall some facts and terminology from commutative algebra. We are particularly interested in the techniques of Gröbner bases, for which we refer the reader to [1, 27].

For an ideal $I = < f_1, \ldots, f_N >$ in a polynomial ring $\mathbb{C}[x_1, x_2, \ldots, x_M]$, we denote by $V(I)$ the corresponding variety: $f_1 = 0$, $f_2 = 0$, $\ldots, f_N = 0$. Then I is said to be *zero-dimensional* if $V(I)$ is finite. A *Gröbner basis* for I is a basis of polynomials

which can be constructed from $f_1, \ldots, f_N$ using a particular algorithm, called the *Buchberger algorithm*. To employ this algorithm, one is required first to choose an order on monomials. We shall only be concerned with *lexicographical order* here, which means we first choose an ordering of the variables, say $x_1 > x_2 > \cdots > x_M$ and then order monomials $x^\alpha := {x_1}^{\alpha_1} \cdots {x_M}^{\alpha_M}$, $x^\beta := {x_1}^{\beta_1} \cdots {x_M}^{\beta_M}$, by $x^\alpha < x^\beta$ if and only if the first coordinate from the left for which α_i and β_i are different, satisfies $\alpha_i < \beta_i$. With respect to the monomial order, every polynomial f in I has a leading term $\mathrm{lt}(f)$ which is the product $\mathrm{lt}(f) = \mathrm{lc}(f)\mathrm{lm}(f)$ of the leading coefficient with the leading monomial.

A set of non-zero polynomials $G = \{g_1, \ldots, g_P\}$ in I is called a *Gröbner basis for I* if and only if for all $f \in I$ such that $f \neq 0$, there is a g_j in G such that $\mathrm{lm}(g_j)$ divides $\mathrm{lm}(f)$. The Gröbner basis is further called *reduced* if for all j, $\mathrm{lc}(g_j) = 1$ and g_j is reduced with respect to $G \setminus \{g_j\}$, that is, no non-zero term in g_j is divisible by any $\mathrm{lm}(g_k)$ for any $k \neq j$. A theorem of Buchberger states that every non-zero ideal has a unique reduced Gröbner basis with respect to a monomial order [9]. Gröbner bases are particularly useful for understanding the solution set of a system of polynomial equations.

Let I be an ideal in the polynomial ring $\mathbb{C}[x_1, x_2, \ldots, x_M]$ and let $G = \{g_1, \ldots, g_P\}$ be the unique reduced Gröbner basis with respect to the lexicographical ordering induced by the order $x_1 > x_2 > \cdots > x_M$. Then $V(I)$ is finite if and only if for each $j = 1, \ldots, M$, there exists a $g_k \in G$ such that $\mathrm{lm}\, g_k = x_j^{n_j}$ for some natural number n_j. As a consequence, if I is a zero-dimensional ideal, it follows that we can order $g_1, \ldots, g_P$ so that g_1 contains only the variable x_M, g_2 contains only x_M, x_{M-1} and so on. This is because the leading monomial of one element, g_1 say, of G must be a power of x_M and then no other term of g_1 can contain powers of any other variable (for such terms would be greater that any power of x_M with respect to the monomial order), and so on for successive elements $g_2, g_3, \ldots$ of G. We note also that $V(I)$ is empty if and only if $1 \in G$.

It is also the case that, with the above hypothesis, the polynomial g_1 is the *least degree univariate polynomial in x_M which belongs to I* (any zero-dimensional ideal contains such a polynomial for every variable). For if there was another univariate polynomial $p(x_M)$ with $\deg p < \deg g_1$, then $\mathrm{lm}\, p$ would divide $\mathrm{lm}\, g_1$ in a strict sense, which would contradict the fact that G is a reduced Gröbner basis. Let us now return to the case under consideration.

For each $\ell = 2, \ldots, N$, consider the ideal $I_\ell =< {f_1}^\ell, \ldots, {f_N}^\ell >$. Suppose that for each $\ell = 2, \ldots, N$ this admits a least degree univariate polynomial p_ℓ in γ. This can be constructed by first choosing a lexicographical ordering of the variables with γ the smallest and then applying an algorithm (say the Buchberger algorithm) to construct the unique reduced Gröbner basis for I_ℓ. The first element of this basis gives p_ℓ.

Definition 7.6 We define the *γ-polynomial* $p = p_G$ of the connected finite graph $G = (V, E)$ to be the least common multiple of the least degree univariate

polynomials p_ℓ ($\ell = 2, \ldots, N$) in γ associated to the Eq. (7.2) for fields $(z_1, \ldots, z_N)$ on G with $z_1 = 0$ and $z_\ell = 1$:

$$p := \operatorname{lcm}(p_2, \ldots, p_N),$$

when each p_ℓ exists.

The γ-polynomial $p(\gamma)$ is defined up to rational multiple and has rational coefficients. This is because the initial polynomials f_j^ℓ used to define p all have integer coefficients and the Buchberger algorithm then generates polynomials with rational coefficients—it involves at most division by coefficients—see [1]. Clearly p depends only on the isomorphism class of a graph and in the case when the Eq. (7.27) admit no solutions for γ constant and complex, then $p \equiv 1$. In this case we shall say that p *is trivial*. The polynomials p_ℓ and so p may still be well-defined even if the solution set of the equations is infinite (that is the corresponding ideal is no longer zero-dimensional). In fact we know of no case when they are not well-defined.

The constant geometric spectrum arises as real roots of p (the problem of establishing the discreteness of the spectrum is clearly intimately related to knowing if p is well-defined in all cases). However, not all real roots may occur in the spectrum, for in general they must also solve the other equations determined by the Gröbner basis: $g_1 = 0, \ldots, g_P = 0$. Examples below illustrate this property. We know of no two non-isomorphic connected graphs with non-trivial γ-polynomial having the same γ-polynomial.

The examples of the triangle C_3 (the cyclic graph on three vertices) and the bipartite graphs K_{23} and K_{33} are instructive. We label the vertices as shown in Fig. 7.11 and consider fields ϕ taking the values $\phi(x_j) = z_j$ at each vertex x_j.

For the triangle, there are precisely two solutions to (7.27) when we normalize so that $z_1 = 0, z_2 = 1$; specifically $z_3 = \frac{1}{2} \pm \mathrm{i}\frac{\sqrt{3}}{2}$. Then $p = p_2 = p_3 = 3\gamma - 2$ is the γ polynomial and the only constant element (in fact the only element) of the geometric spectrum is the unique root $\gamma = 2/3$.

For K_{23}, we find $p_2 = 1$ with no solution and $p_3 = \gamma^2 - 2\gamma + 1 = (\gamma - 1)^2$ with solution $z_1 = 0$, $z_3 = 1, z_2 = 0, z_5 = \lambda$ arbitrary and $z_4 = [1 + \lambda \pm \sqrt{3}(1 - \lambda)\mathrm{i}]/2$. Then $p = \gamma^2 - 2\gamma + 1$ and the constant geometric spectrum is given by $\Sigma = \{1\}$.

For K_{33}, we find $p_2 = 9\gamma^2 - 26\gamma + 17 = (\gamma - 1)(9\gamma - 17)$ and $p_4 = 9\gamma^3 - 35\gamma^2 + 43\gamma - 17 = (\gamma - 1)p_2$, so that the γ-polynomial $p = 9\gamma^3 - 35\gamma^2 + 43\gamma - 17$.

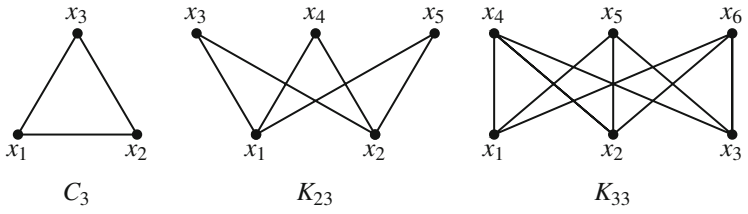


Fig. 7.11 Three graphs for which the γ-polynomial can be easily calculated by an appropriate labeling of the vertices

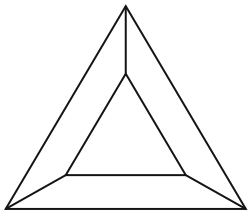


Fig. 7.12 An example of a simple graph with complicated γ-polynomial

Although this has $\gamma = 17/9$ as a root, the constant geometric spectrum $\Sigma = \{1\}$. In fact for $\gamma = 1, z_1 = 0, z_2 = 1$ we find a two complex parameter family of solutions to (7.27). The next example shows that even for simple graphs, the γ-polynomial can be quite complicated.

Consider the graph of constant degree three on six vertices whose edges form two concentric triangles as shown in Fig. 7.12.
The γ-polynomial is given by

$$\begin{aligned} &5859375\gamma^{10} - 67656250\gamma^9 + 333521875\gamma^8 - 926025000\gamma^7 \\ &\quad +1603978830\gamma^6 - 1808486028\gamma^5 + 1339655598\gamma^4 \\ &\quad -639892872\gamma^3 + 186760323\gamma^2 - 29598858\gamma + 1883007\,. \end{aligned}$$

This has eight real roots, four of which are rational: $\gamma = 3/5, 21/25, 1, 3$. There remain two conjugate complex roots. Since the γ-polynomial differs from that of the bipartite graph K_{33} calculated above, we deduce that these two graphs of constant degree three on six vertices cannot be isomorphic.

It is interesting to consider the real roots that are ≥ 1, which can be seen to arise from surjective mappings $\phi : V \to S_N := \{z_0, \ldots, z_N\} \subset \mathbb{C}$, where $z_0, \ldots, z_N$ are the images under any orthogonal projection to the complex plane of the vertices of a regular N-simplex in $\mathbb{R}^N$. This is because of the translation-invariant relation

$$(z_0 + \cdots + z_N)^2 = (N+1)({z_0}^2 + \cdots + {z_N}^2)$$

between these projections whenever $N \geq 2$ [13]. For example, if $N = 4$, we may take $S_4 = \{0, 1, \mathrm{i}, 1+\mathrm{i}\}$. When $N = 1$, we take $S_1 = \{0, 1\}$ and $(z_0+z_1)^2 = {z_0}^2+{z_1}^2$.

The root 3 corresponds to the solution $\phi : V \to S_1 = \{0, 1\}$ to (7.2) which takes the value 0 say, on the vertices of the inner triangle and 1 on the vertices of the outer triangle. The root 1 corresponds to the solution $\phi : V \to S_2 = \{0, 1, \frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}\}$, as indicated in Fig. 7.13. Thus, ϕ takes on the colors red, green and blue which are in bijective correspondence with S_2.

More generally, if we can "color" the vertices of a regular graph G of degree d with S_N in such a way that each vertex is connected by an edge to precisely one vertex of each of the other colors, then $\gamma = d/(N+1)$ will be an element of the constant geometric spectrum. Clearly we must have $d \geq N$. For the complete graph on $N+1$ vertices, we can take the coloring given by any bijection $V \to S_N$ to give $\gamma = N/(N+1)$. The above example can be generalized by taking two

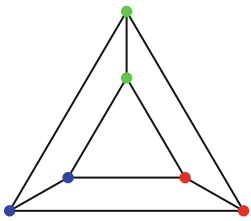


Fig. 7.13 Certain roots of the γ-polynomial correspond to graph colorings

Table 7.3 The γ-polynomial for the complete graph on $N+1$ vertices

N	$p(\gamma)$
3	$8\gamma^2 - 18\gamma + 9$
4	$5\gamma - 4$
5	$18\gamma^2 - 45\gamma + 25$
6	$7\gamma - 6$
7	$32\gamma^2 - 84\gamma + 49$

copies of the complete graph on $N+1$ vertices and connecting each vertex of one of the graphs to precisely one vertex of the other in a bijective correspondence. Now $d = N + 1$ and $\gamma = 1$ lies in the constant geometric spectrum. Relations to vertex colorings suggest potential connections between the γ-polynomial and other more well-known polynomial invariants, such as the Tutte polynomial.

Finally, let us consider the γ-polynomial of the complete graph on $N+1$ vertices. We may apply the method of Gröbner bases with an appropriate lexicographical ordering which produces the γ-polynomial as its first basis element. In fact, by the symmetry of the complete graph, it is clear that all the univariate polynomials p_ℓ ($\ell = 2, \ldots, N$) in Definition 7.6 are identical. If we denote the γ-polynomial by $p(\gamma)$, then using MAPLE, we obtain the suggestive list given in Table 7.3.

It is reasonable to conjecture that for N even, the γ-polynomial is given by $p(\gamma) = (N+1)\gamma - N$ and that for $N = 2k+1$ odd, it is given by

$$\begin{aligned} p(\gamma) &= 2(k+1)^2\gamma^2 - 3(k+1)(2k+1)\gamma + (2k+1)^2 \\ &= \Big(2(k+1)\gamma - (2k+1)\Big)\Big((k+1)\gamma - (2k+1)\Big). \end{aligned}$$

Appendix 2: New Solutions from Old

Holomorphic Mappings Between Graphs

The natural class of mappings between graphs which preserve Eq. (7.2) are the so-called *holomorphic* mappings. These were introduced for simple graphs under the name *semi-conformal* mappings by Urakawa [28], as the class of maps which preserve local harmonic functions (i.e. functions which are harmonic at a vertex). The notion was later extended to non-simple graphs by Baker and Norine [5, 6],

who used the term *holomorphic mapping*. In [4], it was shown that the holomorphic mappings are precisely the class of mappings which preserve local *holomorphic* functions, defined as solutions to (7.2) for which $\gamma \equiv 0$. The definition requires that we restrict to mappings of graphs that determine a well-defined mapping of the tangent space at each vertex, which also justifies our inclusion of the zero vector in the definition of tangent space (see Sect. 7.2).

Definition 7.7 Let $f : G = (V, E) \to H = (W, F)$ be a mapping between graphs. Then f is *holomorphic* if

(i) $x \sim y$ implies either $f(x) = f(y)$ or $f(x) \sim f(y)$;
(ii) there exists a function $\lambda : V \to \mathbb{N}$ such that for all $x \in V$ and for all $z' \sim z = f(x)$, we have

$$\lambda(x) = \lambda(x, z') = \sharp\{x' \in V : x' \sim x, f(x') = z'\},$$

is independent of the choice of z'; we set $\lambda(x) = 0$ if $f(x') = z$ for all $x' \sim x$. Call λ the *dilation of* f.

Proposition 7.6 *Let $f : G = (V, E) \to H = (W, F)$ be a holomorphic mapping between graphs of dilation $\lambda : V \to \mathbb{N}$. Suppose $\psi : W \to \mathbb{C}$ satisfies the equation*

$$\mu(\Delta\psi)^2 = (\mathrm{d}\psi)^2,$$

for some $\mu : W \to \mathbb{R}$. Then for each $x \in V$ such that $\lambda(x) \neq 0$, the function $\phi = \psi \circ f$ satisfies (7.2) at x with

$$\gamma(x) = \frac{d_x \mu(f(x))}{\lambda(x) d_{f(x)}}, \tag{7.28}$$

where d_z also denotes the degree of a vertex $z \in W$.

Proof Let $f : G = (V, E) \to H = (W, F)$ be a holomorphic mapping between graphs of dilation $\lambda : V \to \mathbb{N}$. Let $x \in V$ and set $z = f(x)$. Then

$$\frac{\mu(z)}{d_z}\left(\sum_{z' \sim z}(\psi(z') - \psi(z))\right)^2 = \sum_{z' \sim z}(\psi(z') - \psi(z))^2.$$

Since f is holomorphic

$$\sum_{x' \sim x}[(\psi \circ f)(x') - (\psi \circ f)(x)] = \lambda(x) \sum_{z' \sim z}(\psi(z') - \psi(z)).$$

Suppose that $\lambda(x) \neq 0$. Then

$$\begin{aligned}\sum_{x' \sim x}[(\psi \circ f)(x') - (\psi \circ f)(x)]^2 &= \lambda(x) \sum_{z' \sim z}(\psi(z') - \psi(z))^2 \\ &= \frac{\lambda(x)\mu(z)}{d_z}\left(\sum_{z' \sim z}(\psi(z') - \psi(z))\right)^2 \\ &= \frac{\mu(z)}{\lambda(x)d_z}\left(\sum_{x' \sim x}[(\psi \circ f)(x') - (\psi \circ f)(x)]\right)^2,\end{aligned}$$

from which the formula follows. If on the other hand $\lambda(x) = 0$, then $f(x') = f(x)$ for all $x' \sim x$ and both sides of (7.2) vanish. □

Given a holomorphic mapping between graphs, the above proposition shows how an element of the geometric spectrum on the co-domain determines one on the domain. A simple example of a holomorphic mapping between planar graphs is given in Fig. 7.14. In this example, the outer "wheel" of the domain graph is mapped cyclically onto the outer wheel of the image, covering it twice, while the central vertex of the domain is mapped onto the central vertex of the image.

The dilation at the central vertex is given by $\lambda = 2$, whereas it equals 1 at the other vertices. With reference to Example 7.2, we see that the constant value $\gamma \equiv 1/3$ is also an element of the geometric spectrum of the domain graph, the doubling of the degree at the central vertex being exactly compensated for by the doubling of the dilation in the formula (7.28).

Applying Normalization to Construct New Solutions

Another method to construct new solutions to (7.2) is to exploit the freedom to normalize a solution ϕ by (7.3). Whilst not exhaustive, we describe some examples.

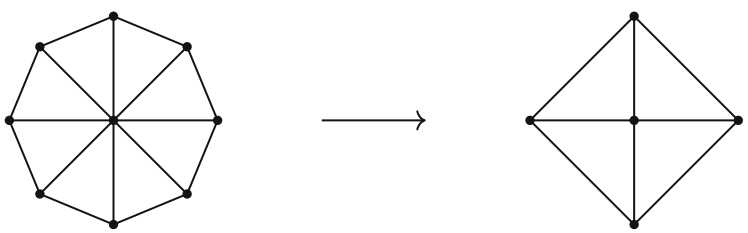


Fig. 7.14 An example of a holomorphic mapping between graphs: the outer wheel covers its image twice, while the central vertex is preserved

Fig. 7.15 Solutions to equation (7.2) on two graphs can be normalized at two vertices in such a way that the graphs can be connected by edges to give a new solution

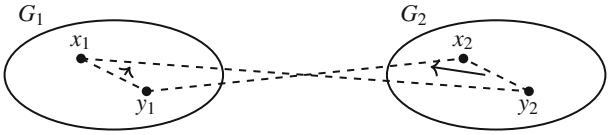


Fig. 7.16 Solutions to equation (7.2) on two graphs can be normalized at two vertices to allow edge rotations which connect the two graphs, giving a new solution

The process of collapsing was described at the end of Sect. 7.4, whereby edges that connect vertices on which a solution to (7.2) takes on the same value can be removed. This operation can be reversed as follows. Given a non-constant solution ϕ_1 to (7.2) on a graph $G_1 = (V_1, E_1)$, then for two vertices $x_1, y_1 \in V_1$ where $\phi_1(x_1) \neq \phi_1(y_1)$, we can apply the normalization (7.3) and suppose that $\phi_1(x_1) = 0$ and $\phi_1(y_1) = 1$. Similarly for a non-constant solution ϕ_2 to (7.2) on a graph $G_2 = (V_2, E_2)$, we may find two vertices x_2, y_2 and normalize so that $\phi_2(x_2) = 0$ and $\phi_2(y_2) = 1$. As illustrated in Fig. 7.15, define a new graph G whose vertex set $V = V_1 \cup V_2$ and whose edge set $E = E_1 \cup E_2 \cup \{\overline{x_1x_2}, \overline{y_1y_2}\}$.

Note that we are not obliged to add both edges, and indeed we can connect any vertices on which ϕ_1 and ϕ_2 take on the same value. We then define $\phi : V \to \mathbb{C}$ by $\phi(x) = \phi_1(x)$ if $x \in V_1$ or $\phi(x) = \phi_2(x)$ if $x \in V_2$. Clearly ϕ satisfies (7.2), but with γ modified to take into account the fact that the degrees at x_1, y_1, x_2, y_2 have increased by one.

As a variant, if both x_1 and y_1 and x_2 and y_2 are connected by edges in G_1 and G_2, respectively, then we can remove these edges and replace them by $\overline{x_1y_2}$ and $\overline{y_1x_2}$. Thus the new graph $G = (V, E)$ (illustrated in Fig. 7.16) has $V = V_1 \cup V_2$ and $E = (E_1 \setminus \{\overline{x_1y_1}\}) \cup (E_2 \setminus \{\overline{x_2y_2}\}) \cup \{\overline{x_1y_2}, \overline{y_1x_2}\}$. As before we can define a solution to (7.2) to be the restriction of ϕ_1 to V_1 and ϕ_2 to V_2, but now the degrees are preserved, so γ is unchanged.

The Preferential Attachment Model

In random graph theory, an important generative model is the preferential attachment scheme which proceeds as follows. For the current graph G, add a new vertex y and add an edge $\overline{xy}$ from y by randomly and independently choosing x in proportion to the degree of x in G. As demonstrated rigorously by Chung and Lu, as the order of the graph approaches infinity, this model generates the scale-free graphs that are

so prevalent in biology and social networks [10]. Geometry may be seen to emerge from this process by exploiting our construction of invariant stars.

Consider an invariant (not necessarily configured) star in $\mathbb{R}^N$ with internal vertex located at the origin and with d external vertices. Let $\mathbf{x} \in \mathbb{R}^N$ be the center of mass of the external vertices. Suppose that $\mathbf{x} \neq \mathbf{0}$. Let $b > 0$ denote the distance of the center of mass from the origin along the axis of the star.

Lemma 7.2 ([2]) *The addition of a new external vertex at any point other than $-db$ along the axis of the star produces a new invariant star. Furthermore, if γ denotes the invariant of the original star and $x \in \mathbb{R}$ is the position along the axis of the star, then the new star invariant is given by*

$$\widetilde{\gamma} = \frac{(d+1)(x^2 + db^2\gamma)}{(x+db)^2} \,. \tag{7.29}$$

Proof Without loss of generality, we may suppose that the center of mass of the star lies along the y^N-axis. In particular, if $\mathbf{v}_1, \ldots, \mathbf{v}_d$ denote the external vertices, then

$$\sum_{\ell=1}^{d} \mathbf{v}_\ell = db\mathbf{e}_N \,.$$

We now add a new vertex at the point $x\mathbf{e}_N$, for some $x \in \mathbb{R}$. Thus the new star matrix is given by

$$(\mathbf{v}_1 | \cdots | \mathbf{v}_d | x\mathbf{e}_N) \ .$$

Let $A = (a_{jk})$ be an arbitrary orthogonal transformation of $\mathbb{R}^N$ and let $\phi : \mathbb{R}^N \to \mathbb{C}$ be the projection $\phi(y^1, \ldots, y^N) = y^1 + \mathrm{i}y^2$. Set $z_\ell = \phi \circ A(\mathbf{v}_\ell)$ for $\ell = 1, \ldots, d$ and $z_{d+1} = \phi \circ A(x\mathbf{e}_N)$. Then

$$z_\ell = \sum_{j=1}^{N} (a_{1j} + \mathrm{i}a_{2j}) v_\ell{}^j \,, \qquad z_{d+1} = x(a_{1N} + \mathrm{i}a_{2N}) \,.$$

Furthermore, $\sum_{\ell=1}^{d} z_\ell = db(a_{1N} + \mathrm{i}a_{2N})$, so that

$$\sum_{\ell=1}^{d} {z_\ell}^2 = \frac{\gamma}{d} \left(\sum_{\ell=1}^{d} z_\ell \right)^2 = db^2\gamma(a_{1N} + \mathrm{i}a_{2N})^2 \,,$$

where γ is the invariant of the original star. We require that there is a real number $\widetilde{\gamma}$ such that

$$\widetilde{\gamma} \left(z_{d+1} + \sum_{\ell=1}^{d} z_\ell \right)^2 = (d+1) \left({z_{d+1}}^2 + \sum_{\ell=1}^{d} {z_\ell}^2 \right) \,.$$

But this is uniquely given by (7.29). □

We note that as x approaches $-db$, then $|\widetilde{\gamma}|$ becomes arbitrary large. Indeed, when $x = -db$, then the Laplacian of ϕ vanishes at the internal vertex of the new star, so that $\widetilde{\gamma}$ is not well-defined in this case.

Suppose we are given a solution to (7.2) on a graph $G = (V, E)$. For a given $x \in V$, let $\{y_1, \ldots, y_d\}$ be its neighbors. Then by the lifting property described in Sect. 7.3, there is an invariant star K in $\mathbb{R}^3$ (or $\mathbb{R}^2$ if the degree of x is two) whose external vertices project to the values $\phi(y_1), \ldots, \phi(y_d)$ and whose internal vertex projects to $\phi(x)$. Introduce a new vertex y and form the graph $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ for which $\widetilde{V} = V \cup \{y\}$ and $\widetilde{E} = E \cup \{\overline{xy}\}$. Define $\widetilde{\phi} : \widetilde{V} \to \mathbb{C}$ by $\widetilde{\phi}(u) = \phi(u)$ for $u \in V$ and $\widetilde{\phi}(y) = z$, where z is the projection of any point along the axis K. Then $\widetilde{\phi}$ satisfies (7.2) on $\widetilde{G}$ with γ modified according to the above lemma (note that any vertex of degree one always satisfies (7.2) with $\gamma = 1$). The choice of distance along the axis of the star may depend on some other parameter of the model, for example, so as to best uniformize γ.

Acknowledgements The author would like to express his thanks to Pascal Romon and the referee whose insightful comments have helped to improve this work.

References

1. Adams, W.W., Loustaunau, P.: An introduction to Groebner bases. Graduate Studies in Mathematics, vol. 3. American Mathematical Society, Providence (1996)
2. Baird, P.: Constant mean curvature polytopes and hypersurfaces via projections. Differ. Geom. Appl. **33**, 199–212 (2014)
3. Baird, P.: An invariance property for frameworks in Euclidean space. Linear Algebra Appl. **440**, 243–265 (2014)
4. Baird, P., Wehbe, M.: Twistor theory on a finite graph. Commun. Math. Phys. **304**(2), 499–511 (2011)
5. Baker, M., Norine, S.: Riemann-Roch and Abel-Jacobi theory on a finite graph. Adv. Math. **215**(2), 766–788 (2007)
6. Baker, M., Norine, S.: Harmonic morphisms and hyperelliptic curves. Int. Math. Res. Not. **2009**, 2914–2955 (2009)
7. Barré, S., Zeghib, A.: Real and discrete holomorphy: introduction to an algebraic approach. J. Math. Pures Appl. **87**, 495–513 (2007)
8. Bobenko, A., Mercat, C., Sursis, Y.B.: Linear and nonlinear theories of discrete analytic functions. Integrable structure and isomonodromic green's function. J. Reine Angew. Math. **583**, 117–161 (2005)
9. Buchberger, B.: Ein algorithmus zum auffinden der basiselemente des resklassenringes nach einem nulldimensionalen polynomideal. Ph.D. thesis, Inst. University of Innsbruck, Innsbruck, Austria (1965)
10. Chung, F., Lu, L.: Complex Graphs and Networks. CBMS Regional Conference Series in Mathematics, vol. 107. American Mathematical Society, Providence (2004)
11. Coxeter, H.S.M.: Regular Polytopes, 3rd edn. Dover Publications, New York (1973)
12. Diestel, R.: Graph Theory, 3rd edn. Graduate Texts in Mathematics. Springer, Berlin (2005)
13. Eastwood, M.G., Penrose, R.: Drawing with complex numbers. Math. Intelligencer **22**, 8–13 (2000)
14. Ehrenborg, R.: The Perles-Shephard identity for non-convex polytopes. Technical Report, University of Kentucky (2007). http://citeseerx.ist.psu.edu/

15. Gauss, C.F.: Werke, Zweiter Band. Königlichen Gesellschaft der Wissenschaften, Göttingen (1876)
16. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces. Modern Birkhäuser Classics, 3rd edn. Birkhäuser, Basel (2007)
17. Grünbaum, B., Shephard, G.C.: Descartes' theorem in n dimensions. Enseign. Math (2) **37**, 11–15 (1991)
18. Hsu, L., Kusner, R., Sullivan, J.: Minimizing the squared mean curvature integral for surfaces in space forms. Exp. Math. **1**(3), 191–207 (1992)
19. Jost, J., Liu, S.: Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs. Discret. Comput. Geom. **51**, 300–322 (2014)
20. 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
21. Milnor, J.: The Schläfli differential equality. John Milnor Collected Papers, vol. 1, Geometry. Publish or Perish, Inc., Houston, TX (1994)
22. Murakami, J.: Volume formulas for a spherical tetrahedron. Proc. Am. Math. Soc. **140**(9), 3289–3295 (2012)
23. Polthier, K.: Polyhedral surfaces of constant mean curvature. Tech. rep., Freie Universitat Berlin (2002). URL http://page.mi.fu-berlin.de/polthier/articles/habil/polthier-habil2002.pdf
24. Romon, P.: Introduction à la géométrie différentielle discrète. Editions-Ellipses, Paris (2013)
25. Schläfli, L.: On the multiple integral $\int^n dx\,dy\cdots dz$, whose limits are $p_1 = a_1x + b_1y + \cdots + h_1z > 0, p_2 > 0, \ldots, p_n > 0$, and $x^2 + y^2 + \cdots z^2 < 1$. Q. J. Math. **2**, 269–301 (1858)
26. Shephard, G.C.: Angle deficiences of convex polytopes. J. Lond. Math. Soc. **43**, 325–336 (1968)
27. Sturmfels, B.: What is a Groebner basis? Not. AMS **52**(11), 2–3 (2005)
28. Urakawa, H.: A discrete analogue of the harmonic morphism and Green kernel comparison theorems. Glasgow Math. J. **42**, 319–334 (2000)
29. Zelazo, D., Franchi, A., Bülthoff, H.H., Giordano, P.R.: Decentralized rigidity maintenance control with range measurements for multi-robot systems. arXiv:1309.0535v3[cs.SY] (2014)

Chapter 8
Discrete Minimal Surfaces of Koebe Type

Alexander I. Bobenko, Ulrike Bücking, and Stefan Sechelmann

Abstract There is an increasing interest to find suitable discrete analogs for known geometric notions and shapes like minimal surfaces. In this article, we consider parametrized surfaces which lead to quadrilateral meshes. In particular, we choose a parametrization where the second fundamental form is diagonal. In addition to the discrete surface we consider a line congruence at the vertices which can be interpreted as a discrete Gauss map. This easily leads to parallel offset meshes. Comparing the areas of two such parallel planar quadrilaterals can then be used to define discrete mean and Gaussian curvature analogously as in the smooth case. This approach leads to a simple notion of discrete minimal surfaces which contains several known definitions as special cases. We especially focus on discrete minimal surfaces whose discrete Gauss map is given by a Koebe polyhedron, i.e. a polyhedral surface with edges tangent to the unit sphere. This case is closely connected to the theory of S-isothermic discrete minimal surfaces. We remind the construction scheme and present analogs for several known smooth minimal surface.

8.1 Introduction

Minimal surfaces have been studied for a long time, but still contain unsolved aspects. Stimulated by the relevance for computer graphic and visualization, there is actually an increasing interest to find suitable discrete analogs for known geometric notions and shapes like minimal surfaces. Usually, one can suggest various discretizations which have quite different properties. See [6, 15, 17, 18, 24, 27, 38] for a choice of examples from the huge variety of different approaches for discrete minimal surfaces. We consider this class of surfaces in the context of discrete differential geometry. The goal of this theory is to find a discretization which inherits as many essential properties of the smooth theory as possible.

A.I. Bobenko • U. Bücking (✉) • S. Sechelmann
Technische Universität Berlin, Institut für Mathematik, Sekr. MA 8-3, Straße des 17. Juni 136, 10623 Berlin, Germany
e-mail: bobenko@math.tu-berlin.de; scheerer@math.tu-berlin.de; stefan@sechel.de

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_8

Our approach is based on the discretization of parametrized surfaces which lead to quadrilateral meshes. In particular, we choose a parametrization where the second fundamental form is diagonal. Such *conjugate nets* exist (locally) for general surfaces in $\mathbb{R}^3$ and lead to discrete parametrized surfaces built from planar quadrilateral faces, which are particular *polyhedral surfaces*. In addition to the discrete surface we consider a line congruence at the vertices which can be interpreted as a discrete Gauss map. This easily leads to parallel offset meshes, see Fig. 8.3 below. Comparing the areas of two such parallel planar quadrilaterals can then be used to define discrete mean and Gaussian curvature analogously as in the smooth case. This approach leads to a simple notion of discrete minimal surfaces which contains several known definitions as special cases.

We especially focus on discrete minimal surfaces whose discrete Gauss map is given by a Koebe polyhedron, i.e. a polyhedral surface with edges tangent to the unit sphere. This case is closely connected to the theory of S-isothermic discrete minimal surfaces as studied in [7]. We remind the construction scheme which, in particular, allows to build an analog for (a finite part of) any given smooth minimal surface.

S-isothermic discrete minimal surfaces can be constructed using suitable orthogonal circle patterns. These are configurations of circles for a given planar graph G such that to each vertex of G there corresponds a circle. If two vertices are connected by an edge in G then the corresponding circles intersect orthogonally. The circles corresponding to the vertices of a face of G intersect in one point. Furthermore, for any circle corresponding to an interior vertex the disks filling the neighboring circles have mutually disjoint interiors. For the construction of such circle patterns we use a variational principle from [4, 37]. This has advantages for the explicit calculation of examples and has been used to create our various examples of Sect. 8.4.

An important connection to the smooth theory is demonstrated by the convergence of a large class of S-isothermic discrete minimal surfaces to their smooth analogons. The main ideas are explained in Sect. 8.5; for a detailed version and a proof see [12].

8.2 Discrete Minimal Surfaces

Every smooth immersed surface in $\mathbb{R}^3$ can (locally) be written as a parametrization, that is as a smooth mapping

$$\mathscr{F} : \mathbb{R}^2 \supset D \to \mathbb{R}^3, \quad (x, y) \mapsto \mathscr{F}(x, y),$$

such that in any point the partial derivatives $\partial_x \mathscr{F}$ and $\partial_y \mathscr{F}$ are linearly independent. We focus on special parametrizations where the second fundamental form is diagonal, that is $\partial_{xy} \mathscr{F} \in \operatorname{span}(\partial_x \mathscr{F}, \partial_y \mathscr{F})$, where ∂_{xy} denotes the mixed partial derivative. Such parametrizations are also called *conjugate nets*. This condition can be interpreted as the planarity of infinitesimal quadrilaterals with vertices $\mathscr{F}(x, y)$, $\mathscr{F}(x+\varepsilon, y)$, $\mathscr{F}(x+\varepsilon, y+\varepsilon)$, $\mathscr{F}(x, y+\varepsilon)$. As a discrete analog, it is therefore natural

to consider discrete parametrized surfaces built from planar quadrilateral faces. The corresponding discrete parameter space will be a special cell decomposition.

Recall that a *cell decomposition of a two-dimensional manifold* possibly with boundary is a graph embedded in the manifold such that the complement of the graph is (topologically) a disjoint union of open disks. Thus, a cell decomposition decomposes a two-dimensional manifold into vertices, edges, and faces. The sets of vertices and edges will be denoted by V and E respectively. The cell decomposition is called *regular*, if it has no loops (edges with the same vertex on both ends), and if the boundary of a face contains an edge or vertex at most once. A regular cell decomposition is called *strongly regular* if two edges have at most one vertex in common at their boundaries, and if two faces have at most one edge or one vertex in common at their boundaries.

Definition 8.1 A strongly regular cell decomposition $\mathscr{D}$ of a two-dimensional oriented manifold possibly with boundary is called a *quad-graph*, if all 2-cells (faces) of $\mathscr{D}$ are embedded, carry the same orientation, and are quadrilaterals, that is there are exactly four edges incident to each face. See Fig. 8.1 for an example.

A mapping $F : \mathscr{D} \to \mathbb{R}^3$ from a quad-graph $\mathscr{D}$ is called a *quadrilateral surface* if each face of $\mathscr{D}$ is mapped by F to a planar quadrilateral in $\mathbb{R}^3$, see Fig. 8.2.

Given a smooth immersed surface $\mathscr{F} : \mathbb{R}^2 \supset D \to \mathbb{R}^3$ there exists in every point a normal direction, i.e. a line orthogonal to the surface. In fact, there is even (locally) a smooth map $\mathscr{N} : D \to \mathbb{S}^2$, called Gauss map, such that $\mathscr{N}(p)$ is orthogonal to the surface at $\mathscr{F}(p)$. As a discrete analog of normals to a surface we introduce lines at the vertices building a line congruence:

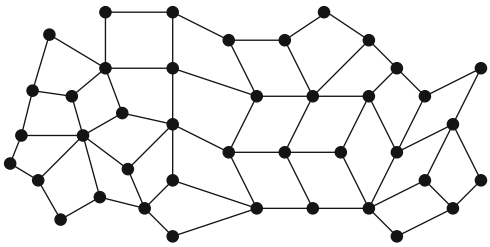


Fig. 8.1 Example of a quad-graph

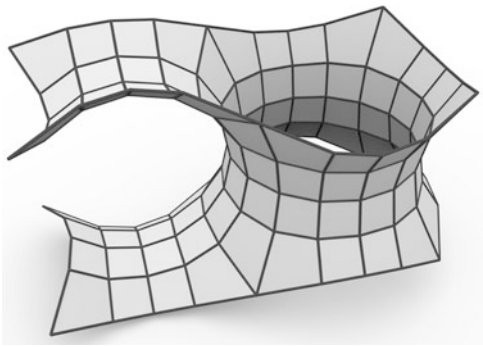


Fig. 8.2 Example of a quadrilateral surface

Definition 8.2 Let $\mathbb{L}^3$ be the space of affine lines in $\mathbb{R}^3$ and denote by $\mathbb{L}^3(p)$ the subset of affine lines in $\mathbb{R}^3$ passing through the point $p \in \mathbb{R}^3$. A *line congruence* for a quadrilateral surface $F : \mathscr{D} \to \mathbb{R}^3$ is a mapping $\ell : V(\mathscr{D}) \to \mathbb{L}^3$ which assigns to every vertex $v \in V(\mathscr{D})$ an affine line $\ell(v) \in \mathbb{L}^3(F(v))$ such that for neighboring vertices v_i, v_j the corresponding lines $\ell(v_i)$ and $\ell(v_j)$ are coplanar, i.e. intersect or are parallel.

Remark 8.1 Line congruences belong to projective geometry. Therefore, instead of affine lines in $\mathbb{L}^3$, it is natural to consider the space of lines in $\mathbb{RP}^3$.

Given a quadrilateral surface $F : \mathscr{D} \to \mathbb{R}^3$ with a line congruence ℓ note that for every edge $e = [v_1, v_2] \in E(\mathscr{D})$ joining the vertices v_1, v_2 the lines $\ell(v_1)$ and $\ell(v_2)$ and the edge of the surface $F([v_1, v_2])$ are coplanar. Therefore one can easily construct parallel surfaces to F, see Fig. 8.3 (left).

Conversely, given two such parallel quadrilateral surfaces F_1 and F_2, we can easily obtain a line congruence by adding lines through corresponding points. Furthermore, the vectors joining the points corresponding to the same vertex v can be interpreted as a *generalized discrete Gauss map* $N(v) = F_2(v) - F_1(v)$. Now consider these vectors for one face of the quadrilateral surface F_1 and translate the vectors within that face such that they all start at one vertex, see Fig. 8.3 (right). Their endpoints still lie in the corresponding plane of the parallel surface F_2. Thus moving all vectors $N(v)$ such that they all start at the origin we again obtain a quadrilateral surface $N : \mathscr{D} \to \mathbb{R}^3$ parallel to F_1 and F_2.

These reasonings show that for a given quadrilateral surface F we can identify a line congruence ℓ with a one parameter family λN for $\lambda \neq 0$ of associated generalized discrete Gauss maps by $\ell(v) = \{F(v) + \mu N(v) \ : \ \mu \in \mathbb{R}\}$ for every vertex v.

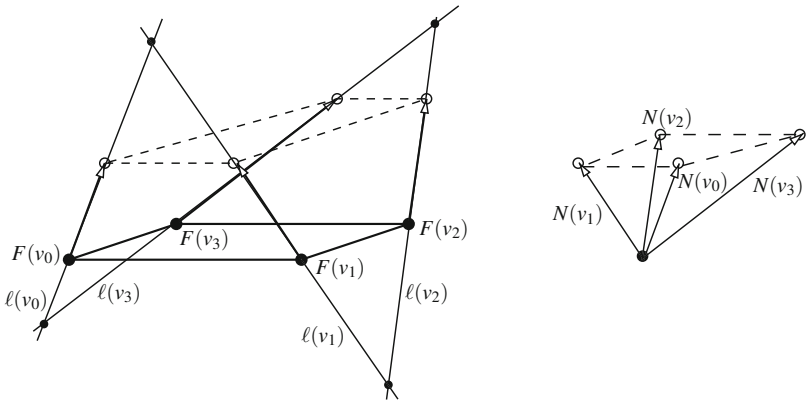


Fig. 8.3 *Left:* planar face of a quadrilateral surface F with line congruence ℓ and a corresponding parallel face (*dashed lines*), *Right:* corresponding face of the associated generalized discrete Gauss map N

Definition 8.3 A quadrilateral surface $F : \mathscr{D} \to \mathbb{R}^3$ with a generalized discrete Gauss map $N : \mathscr{D} \to \mathbb{R}^3$, i.e. a parallel quadrilateral surface N, is called a *line congruence net* (F, N).

8.2.1 Discrete Curvatures

Given a smooth surface $\mathscr{F}$ with its Gauss map $\mathscr{N}$, a parallel surface $\mathscr{F}_t$ can (locally) be defined by the formula $\mathscr{F}_t = \mathscr{F} + t\mathscr{N}$ for sufficiently small $t \in \mathbb{R}$. The classical *Steiner formula* states that the infinitesimal area of the parallel surface $\mathscr{F}_t$ is the quadratic polynomial

$$dA(\mathscr{F}_t) = (1 - 2tH + t^2K)dA(\mathscr{F}). \tag{8.1}$$

Here $dA(\mathscr{F})$ denotes the infinitesimal area of the surface $\mathscr{F}$ and H and K are its mean and Gaussian curvature respectively. See also Chap. 4, Sect. 4.3.3, of this volume.

For quadrilateral surfaces, parallel offset surfaces or equivalently a generalized discrete Gauss map can be used in an analogous way to define Gaussian and mean curvature.

First we note that the oriented area functional $A(P)$ of a planar polygon P is a quadratic form. The corresponding symmetric bilinear form $A(P, Q)$ is called *mixed area*. It will be important for the following theory and may be obtained as follow.

Consider two planar polygons $P = (\mathbf{p}_0, \ldots, \mathbf{p}_{k-1})$ and $Q = (\mathbf{q}_0, \ldots, \mathbf{q}_{k-1})$ with vertices $\mathbf{p}_i$ and $\mathbf{q}_i$ respectively and whose corresponding edges are parallel. Then $A(P, Q) = \frac{1}{2}(A(P+Q) - A(P) - A(Q))$ where $P+Q = (\mathbf{p}_0 + \mathbf{q}_0, \ldots, \mathbf{p}_{k-1} + \mathbf{q}_{k-1})$ is the polygon obtained by pointwise addition. For further use, we express the mixed area using the determinant. Let $\mathbf{n}$ be the unit normal vector of the parallel planes containing the polygons P and Q. Then

$$A(P, Q) = \frac{1}{4}\sum_{j=0}^{k-1}[\det(\mathbf{p}_j, \mathbf{q}_{j+1}, \mathbf{n}) + \det(\mathbf{q}_j, \mathbf{p}_{j+1}, \mathbf{n})]. \tag{8.2}$$

Given a line congruence net (F, N), consider one quadrilateral face f with area $A(f) \neq 0$. Denote by n the corresponding face of the surface N. Then the one parameter family of parallel offset faces is $f_t = f + tn$ for $t \in \mathbb{R}$. The area of f_t is

$$A(f_t) = A(f + tN, f + tN) = \left(1 + 2t\frac{A(f, n)}{A(f)} + t^2\frac{A(n)}{A(f)}\right)A(f). \tag{8.3}$$

Comparing this equation with the corresponding formula for smooth surfaces (8.1) leads to the following curvature definitions on faces, see [8, 28].

Definition 8.4 Let (F, N) be a line congruence net such that the faces of F are not degenerate, i.e. $A(f) \neq 0$. Then the *discrete Gaussian curvature* K and the *discrete mean curvature* H are defined by

$$K = \frac{A(n)}{A(f)} \qquad \text{and} \qquad H = -\frac{A(f,n)}{A(f)}. \tag{8.4}$$

Here f denotes a quadrilateral face of F and n is the corresponding face of N.

Note that the Gaussian curvature is defined as the quotient of the areas of the Gauss image and of the original surface analogously as in the smooth case. Moreover, as for a quadrilateral surface with a line congruence the generalized Gauss map is only defined up to a common factor, the curvatures at faces are also defined up to multiplication by a common constant.

The principle curvatures κ_1, κ_2 are in the smooth case defined by the formulas $H = (\kappa_1 + \kappa_2)/2$ and $K = \kappa_1\kappa_2$ as the zeros of the quadratic polynomial $(1 - 2tH + t^2K) = (1 - t\kappa_1)(1 - t\kappa_2)$.

Discrete principle curvatures can be defined in an analogous way in the following case. Let (F, N) be a line congruence net such that all faces of F are non-degenerate and all quadrilaterals of F are convex (or equivalently all quadrilaterals of N are convex), that is the vertices of every face are on the boundary of their convex hull. Then the *discrete principle curvatures* κ_1, κ_2 on the faces are the zeros of the quadratic polynomial $(1 - 2tH + t^2K) = (1 - t\kappa_1)(1 - t\kappa_2)$ where H and K are the discrete mean and the discrete Gaussian curvature respectively. Note that in this case $H^2 - K \geq 0$ and Minkowski's first inequality $A(f,n)^2 - A(f)A(n) \geq 0$ applies, see [32].

Using these definitions, one easily obtains special classes of quadrilateral surfaces. This article focuses on discrete minimal surfaces:

Definition 8.5 A line congruence net (F, N) is called a *discrete minimal surface* if $H = 0$ for all faces.

8.2.2 Characterization of Discrete Minimal Surfaces

Lemma 8.1 *A line congruence net (F, N) is a discrete minimal surface if and only if the mixed areas vanish, $A(f, n) = 0$, for all quadrilaterals f of F and the corresponding quadrilaterals n of N.*

Thus, we first study planar quadrilaterals P, Q with parallel edges whose mixed area vanishes, $A(P, Q) = 0$, and which are called *dual*.

Lemma 8.2 *For every planar quadrilateral there exists a dual one. For non-zero quadrilaterals the dual is unique up to scaling and translation.*

Furthermore two planar quadrilaterals $P = (\mathbf{p}_0, \ldots, \mathbf{p}_3)$ and $Q = (\mathbf{q}_0, \ldots, \mathbf{q}_3)$ with parallel corresponding edges are dual to each other if and only if their diagonals are antiparallel, that is $\mathbf{p}_2 - \mathbf{p}_0 \parallel \mathbf{q}_3 - \mathbf{q}_1$ and $\mathbf{p}_3 - \mathbf{p}_1 \parallel \mathbf{q}_2 - \mathbf{q}_0$.

Proof In the following, we will only consider quadrilaterals with non-zero edges. The degenerate cases can easily be treated using projective geometry.

We first show the characterization of dual quadrilaterals by their diagonals. Using Eq. (8.2) and the multilinearity of the determinant we obtain

$$4A(P,Q) = \det(\mathbf{p}_0 - \mathbf{p}_2, \mathbf{q}_1 - \mathbf{q}_3, \mathbf{n}) + \det(\mathbf{p}_1 - \mathbf{p}_3, \mathbf{q}_2 - \mathbf{q}_0, \mathbf{n}).$$

As corresponding edges $\mathbf{p}_{i+1} - \mathbf{p}_i$ and $\mathbf{q}_{i+1} - \mathbf{q}_i$ (for indices (mod 4)) are parallel, we further deduce

$$\det(\mathbf{p}_0 - \mathbf{p}_2, \mathbf{q}_1 - \mathbf{q}_3, \mathbf{n}) = \det(\mathbf{p}_1 - \mathbf{p}_3, \mathbf{q}_2 - \mathbf{q}_0, \mathbf{n}),$$

which proves the claim.

The conditions of antiparallel diagonals for dual quadrilaterals show the uniqueness claim, as Q is already fixed by specifying $\mathbf{q}_0 \in \mathbb{R}^3$ and a common scaling factor $\lambda \neq 0$.

Existence of a quadrilateral dual to a non-zero quadrilateral P can be seen as follows. Let $\mathbf{m}$ be the intersection point of the diagonals of P and define the two diagonal directions

$$e_1 = \frac{\mathbf{p}_0 - \mathbf{m}}{\|\mathbf{p}_0 - \mathbf{m}\|} = -\frac{\mathbf{p}_2 - \mathbf{m}}{\|\mathbf{p}_2 - \mathbf{m}\|} \quad \text{and} \quad e_2 = \frac{\mathbf{p}_1 - \mathbf{m}}{\|\mathbf{p}_1 - \mathbf{m}\|} = -\frac{\mathbf{p}_3 - \mathbf{m}}{\|\mathbf{p}_3 - \mathbf{m}\|}.$$

Set

$$Q^* = (\mathbf{q}_0^*, \dots, \mathbf{q}_3^*) := \left(-\frac{e_2}{\|\mathbf{p}_0 - \mathbf{m}\|}, -\frac{e_1}{\|\mathbf{p}_1 - \mathbf{m}\|}, \frac{e_2}{\|\mathbf{p}_2 - \mathbf{m}\|}, \frac{e_1}{\|\mathbf{p}_3 - \mathbf{m}\|} \right).$$

Then Q^* is a planar quadrilateral whose diagonals are antiparallel to those of P by construction. Furthermore, the edges of Q^* are parallel to the corresponding edges of P. We just show this for one edge as the other cases are analogous.

$$\mathbf{q}_1^* - \mathbf{q}_0^* = -\frac{e_1}{\|\mathbf{p}_1 - \mathbf{m}\|} + \frac{e_2}{\|\mathbf{p}_0 - \mathbf{m}\|} = \frac{\mathbf{p}_1 - \mathbf{p}_0}{\|\mathbf{p}_0 - \mathbf{m}\| \|\mathbf{p}_1 - \mathbf{m}\|}$$

Although dual quadrilaterals always exist locally, global dualization requires additional properties in order to result again in a surface.

Definition 8.6 A quadrilateral surface $F : \mathscr{D} \to \mathbb{R}^3$ is called a *discrete Koenigs net* if it admits a dual net $F^* : \mathscr{D} \to \mathbb{R}^3$, i.e. F^* is a quadrilateral surface such that corresponding quadrilaterals of F and F^* are dual.

There are further characterizations of Koenigs nets, in particular in projective geometry, see for example [5, Chap. 2.3] for more details.

Theorem 8.1 *Let F be a discrete Koenigs net. Then the line congruence net (F, N) with $N = F^*$ is a discrete minimal surface.*

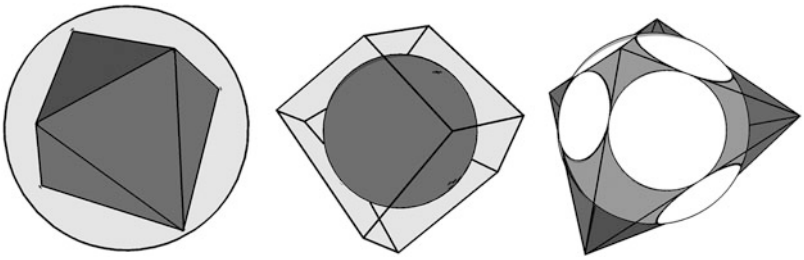


Fig. 8.4 Three possible types of spherical polyhedra

Proof By Lemma 8.1 we know that $H = 0$ is equivalent to $A(f, n) = 0$ for all corresponding faces f of F and n of N. This holds by the definition of the dual surface $N = F^*$.

For smooth minimal surfaces $\mathscr{F}$ the characterization $\mathscr{N} = \mathscr{F}^*$ for the Gauss map $\mathscr{N}$ is due to Christoffel [14], for a modern treatment see for example [20].

In order to obtain a discrete Gauss map N which is related to the unit sphere $\mathbb{S}^2$ as for its smooth analog, one may consider spherical polyhedra. There are three natural types of such polyhedra (see Fig. 8.4) with

- vertices on $\mathbb{S}^2$:
  In this case $\|N(v)\| = 1$ and for all parallel surfaces $F_t = F + tN$ the distance between corresponding vertices is constant.
  Furthermore, we obtain a *circular net* $N : \mathscr{D} \to \mathbb{S}^2$, that is for every quadrilateral exists a circle through its vertices. The dual net $F = N^*$ is then also a circular net, see [5] for a proof and more details.
- faces tangent to $\mathbb{S}^2$:
  In particular, for every vertex of the polyhedron the incident faces are tangent to a cone of revolution touching $\mathbb{S}^2$. Thus we have a *conical net* and this property is conserved by dualization, see [1, 9, 28] for more details.
- edges tangent to $\mathbb{S}^2$, called *Koebe polyhedra*:
  In this case every face of N intersects $\mathbb{S}^2$ in a circle which touches the boundary edges of the face from inside.

Note that from circular nets in $\mathbb{S}^2$ one obtains discrete minimal surfaces which have been defined in [2]. Moreover, circular nets belong to Moebius geometry whereas conical nets are part of Laguerre geometry. Some examples of Koebe polyhedra can also be found in the new project DGD GALLERY[1] for storage and publication of digital research data (see also [21]).

In the following, we will only consider Koebe polyhedra further. These are closely connected to the theory of circle patterns. Above, we have already identified a family of touching circles on $\mathbb{S}^2$ corresponding to the faces of the Koebe polyhedron. There is another family of circle corresponding to the vertices. Consider

[1] http://gallery.discretization.de.

every vertex of the Koebe polyhedron as the apex of a cone which touches $\mathbb{S}^2$. The intersections of the cones with $\mathbb{S}^2$ are circles which touch at points where the edges of the Koebe polyhedron touch $\mathbb{S}^2$. Moreover, the circles corresponding to faces intersect the circles corresponding to vertices orthogonally at these touching points.

Conversely, to an orthogonal circle pattern in $\mathbb{S}^2$ we can associated a Koebe polyhedron by choosing one family of touching circles as above and adding the corresponding cones. Their apices are the vertices of the polyhedron which are by construction connected by edges touching $\mathbb{S}^2$. Thus to an orthogonal circle pattern there are associated two corresponding Koebe polyhedra.

Furthermore, instead of cones, we may consider the spheres with centers at the vertices of a Koebe polyhedron which intersect $\mathbb{S}^2$ orthogonally. This also leads to the same family of circles as for the touching cones. Also, Koebe polyhedra are thus discrete S-isothermic surface which were originally introduced in [3] and will be explained in the following.

Definition 8.7 Let $\mathscr{D}$ be a bipartite quad-graph, that is the vertices are colored white and black such that every edge is incident to one white and one black vertex. $\mathscr{D}$ is called a *S-quad-graph* if all interior black vertices have degree 4 and if the white vertices can be labelled $\textcircled{c}$ and $\textcircled{s}$ in such a way that each quadrilateral has one white vertex labelled $\textcircled{c}$ and one white vertex labelled $\textcircled{s}$. Furthermore, the $\textcircled{c}$-labelled white vertices have degree 4 and the $\textcircled{s}$-labelled white vertices have even degree.

Let $\mathscr{D}$ be an S-quad-graph and let $V_b(\mathscr{D})$ be the set of all black vertices. A *discrete S-isothermic surface* is a map $F_b : V_b(\mathscr{D}) \to \mathbb{R}^3$ with the following properties:

(i) If $v_1, \dots, v_4 \in V_b(\mathscr{D})$ are the neighbors of a $\textcircled{c}$-labelled vertex in cyclic order, then $F_b(v_1), \dots, F_b(v_4)$ lie on a circle in $\mathbb{R}^3$ in the same cyclic order. This defines a map from the $\textcircled{c}$-labelled white vertices to the set of circles in $\mathbb{R}^3$.

(ii) If $v_1, \dots, v_{2m} \in V_b(\mathscr{D})$ are the neighbors of a $\textcircled{s}$-labelled vertex in cyclic order, then $F_b(v_1), \dots, F_b(v_{2m})$ lie on a sphere in $\mathbb{R}^3$. This defines a map from the $\textcircled{s}$-labelled white vertices to the set of spheres in $\mathbb{R}^3$.

(iii) If v_c and v_s are the $\textcircled{c}$-labelled and the $\textcircled{s}$-labelled vertices of a quadrilateral of $\mathscr{D}$, the circle corresponding to v_c intersects the sphere corresponding to v_s orthogonally.

Given an S-quad-graph $\mathscr{D}$, we can construct an associated graph $\mathscr{D}_{\textcircled{s}}$ by taking as vertices all $\textcircled{s}$-labelled white vertices. Two such vertices are connected by an edge in $\mathscr{D}_{\textcircled{s}}$ if they are incident to the same black vertex in $\mathscr{D}$. If $F_b : V_b(\mathscr{D}) \to \mathbb{R}^3$ is a discrete S-isothermic surface, the *central extension* of F_b is the discrete quadrilateral surface $F : V(\mathscr{D}_{\textcircled{s}}) \to \mathbb{R}^3$ defined by

$$F(v) = \text{the center of the sphere corresponding to } v.$$

By abuse of notation, we also call F a discrete S-isothermic surface.

Theorem 8.2 *Every discrete S-isothermic surface is a Koenigs net (i.e. dualizable).*

Proof As the vertices of the quad-graph $\mathscr{D}_{\textcircled{S}}$ have even degree, the edges can be labelled $+$ and $-$ such that opposite edges of every quadrilateral carry the same label. Given the centers of spheres c and the radii r of a discrete S-isothermic surface F, define new centers of spheres c^* and radii r^* by

$$r^*(v) = \frac{1}{r(v)} \qquad \text{and} \qquad c^*(v_1) - c^*(v_0) = \pm \frac{c(v_1) - c(v_0)}{r(v_1) r(v_0)}, \tag{8.5}$$

where the sign $\pm$ is given by the label of the edge $[v_1, v_0] \in E(\mathscr{D}_{\textcircled{S}})$.

We first show that this definition leads indeed to a quadrilateral surface F^*. In particular, we need to show that for any quadrilateral of F we get a new corresponding quadrilateral of F^* by (8.5). Let c_0, c_1, c_2, c_3 and r_0, r_1, r_2, r_3 be the centers of spheres and the corresponding radii of a quadrilateral of F enumerated in counterclockwise orientation. By assumption there is a circle inscribed in this quadrilateral with center m and radius R. It touches the boundary edges at the points $p_{01}, p_{12}, p_{23}, p_{30}$ where $p_{ij} = c_i + r_i \frac{c_j - c_i}{r_i + r_j}$ are images of black vertices. See Fig. 8.5. Note that the quadrilaterals $P_i = (c_i, p_{i,i+1}, m, p_{i-1,i})$, where all indices are taken (mod 4), are orthogonal rhombi. Now scale P_i by $1/(Rr_i)$ and reflect it in the edge labelled $+$. The new orthogonal rhombi P_i^* then have two edges of length $1/R$ and another two of length $r_i^* = 1/r_i$. By suitable translations these orthogonal rhombi P_i^* may be combined into a new quadrilateral face with an inscribed circle with radius $1/R$. The corresponding vertices $c_0^*, \dots, c_3^*$ of this quadrilateral then satisfy (8.5) by construction and $\|c_{i+1}^* - c_i^*\| = r_{i+1}^* + r_i^*$ for the radii $r_0^*, \dots, r_3^*$.

Therefore it remains to show that corresponding quadrilaterals of F and F^* are dual. Using the above notations, it is by Lemma 8.2 sufficient to show that

$$c_1^* - c_3^* \parallel c_2 - c_0 \iff c_1^* - c_0^* + c_0^* - c_3^* \parallel c_2 - c_0,$$

where $\parallel$ means that the vectors are parallel. By the definition of c^* this condition is equivalent to

$$\frac{c_1 - c_0}{r_0 r_1} + \frac{c_3 - c_0}{r_3 r_0} \parallel c_2 - c_0. \tag{8.6}$$

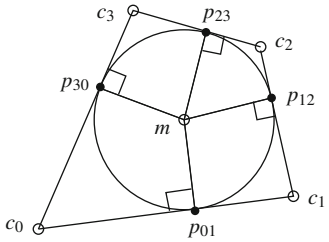


Fig. 8.5 A quadrilateral face of an S-isometric surface

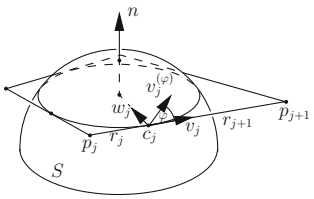


Fig. 8.6 The vector $v_j^{(\varphi)}$ is obtained by rotating v_j in the tangent plane to the sphere S at c_j

As F^* is a quadrilateral surface, we have with the definition of c^*

$$\frac{1}{r_0}\left(\frac{c_1-c_0}{r_1}+\frac{c_3-c_0}{r_3}\right)=\frac{1}{r_2}\left(\frac{c_2-c_3}{r_1}+\frac{c_2-c_1}{r_3}\right)$$
$$\Longleftrightarrow\quad \left(\frac{1}{r_1}+\frac{1}{r_3}\right)(c_2-c_0)=\left(\frac{r_2}{r_0}+1\right)\left(\frac{c_1-c_0}{r_1}+\frac{c_3-c_0}{r_3}\right).$$

This immediately shows that condition (8.6) holds.

Definition 8.8 An *S-isothermic discrete minimal surface* (or discrete minimal surface of Koebe type) is an S-isothermic discrete surface $F : V(\mathscr{D}_{\textcircled{S}}) \to \mathbb{R}^3$ whose dual surface is a Koebe polyhedron.

The *associated family* F_φ of an S-isothermic discrete minimal surface F_0 consists of the one-parameter family of discrete surfaces that are obtained by the following construction. Before dualizing the Koebe polyhedron, rotate each edge by an equal angle φ in the plane which is tangent to the unit sphere in the point where the edge touches the unit sphere, see Fig. 8.6.

This definition implies that the discrete associated family keeps essential properties of the smooth associated family. In particular, the surfaces are isometric and have the same Gauss map.

Theorem 8.3 ([7, Theorem 8]) *The discrete surfaces F_φ of the associated family of an S-isothermic discrete minimal surface F_0 consist of touching spheres. The radii of the spheres do not depend on φ.*

In the generic case, when the S-quad-graph is a part of the regular square grid, there are also circles through the points of contact. The normals of these circles do not depend on φ.

8.3 Construction of Koebe Polyhedra and Discrete Minimal Surfaces

Definition 8.8 leads to the following *construction scheme* for discrete minimal analogs of known smooth minimal surfaces, see also [7].

Step 1: Consider a smooth minimal surface together with its conformal curvature line parameterization. Map the curvature lines to the unit sphere by the

Gauss map to obtain a qualitative picture. The goal is to *understand the combinatorics of the curvature lines*.
From the combinatorial picture of the curvature lines we choose finitely many curvature lines to obtain a finite cell decomposition of the unit sphere $\mathbb{S}^2$ (or of a part or of a branched covering of $\mathbb{S}^2$) with quadrilateral cells. The choice of the number of curvature lines corresponds to the level of refinement (and possibly to the choice of different length parameters). Generically, all vertices have degree 4. Exceptional vertices correspond to umbilic or singular points, ends or boundary points of the minimal surface. So the cell decomposition (or a suitable subdivision) leads to an S-quad-graph which provides a *combinatorial conformal parameterization*.

Step 2: Given the combinatorics from step 1, *construct an orthogonal spherical circle pattern* where circles correspond to white vertices of the S-quad-graph. If two vertices are incident to the same face the corresponding two circles intersect orthogonally. Circles corresponding to vertices which are not incident to the same face, but to the same black vertex touch. The interiors of the disks filling such touching circles are disjoint. At boundary vertices we use information about the smooth minimal surface to specify angles. Ends have also to be taken care of but we do not consider this case further.

Step 3: From the circle pattern, *construct the Koebe polyhedron*. Choose half of the circles, that is those corresponding to white vertices labelled Ⓒ or those labelled Ⓢ. The two choices lead to different discrete surfaces close to each other. Take these circles and build the spheres which intersect $\mathbb{S}^2$ orthogonally in these circles. Then build the Koebe polyhedron by joining the apices corresponding to touching circles.

Step 4: *Dualize the Koebe polyhedron* to arrive at the desired discrete minimal surface.

Note that we obtain the geometry (discrete minimal surface) from the combinatorics of the curvature lines (and eventually some boundary conditions). Therefore the first two steps are most important and may require some care. In Sect. 8.3.2, we present some more details on how to find the combinatorics of the curvature lines and the boundary conditions for the concrete examples presented in Sect. 8.4. Also note that the dualization may be impossible for the whole Koebe polyhedron. Then we first cut the polyhedron along suitable edges.

8.3.1 Construction of Koebe Polyhedra and Spherical Circle Patterns

Given the combinatorics, there remains the task to construct the corresponding spherical circle pattern and the Koebe polyhedron. The existence of Koebe polyhedra was first studied by Koebe [23]. There are several generalizations of Koebe's theorem, for example

Theorem 8.4 *Every polytopal cell decomposition of the sphere can be realized by a polyhedron with edges tangent to the sphere. This realization is unique up to projective transformations which fix the sphere.*

A proof of this theorem has been given in [11], see also [4, 31, 34] for generalizations. Bobenko and Springborn [4] contains a variational proof. A suitable stereographic projection of the orthogonal circle pattern on $\mathbb{S}^2$ which corresponds to the sought-after polyhedron is shown to be the critical point of a strictly convex functional. This implies existence and uniqueness. By its explicit formula the functional is also easy to compute and therefore particularly suitable for explicit constructions. Furthermore, an adaption of the functional can be used to compute planar circle patterns with given angles at boundary vertices corresponding to Neumann boundary conditions.

As orthogonal circle patterns are crucial for our construction, we specify this notion.

Definition 8.9 Let $\mathscr{D}$ be a bipartite quad-graph. Let G be the associated graph constructed from all white vertices of $\mathscr{D}$, that is $V(G) = V_w(\mathscr{D})$. Two vertices of G are connected by an edge if they are incident to the same face in $\mathscr{D}$.

An *orthogonal planar circle pattern for* $\mathscr{D}$ *or* G is a configuration of circles in the complex plane $\mathbb{C}$, such that to each white vertex of $V_w(\mathscr{D}) = V(G)$ there corresponds a circle (all with the same orientation). If two white vertices are incident to the same face in $\mathscr{D}$, i.e. are incident in G, the corresponding circles intersect orthogonally. Furthermore, if two white vertices are incident to the same black vertex of $V_b(\mathscr{D})$, but are not incident to the same face in $\mathscr{D}$, the corresponding circles touch and have disjoint interiors.

Remark 8.2 Note that an orthogonal planar circle pattern for $\mathscr{D}$ exists only if all black vertices have degree 4. For the graph G this means that all faces are quadrilaterals.

The notion of orthogonal circle patterns may easily be generalized, for example to spherical or hyperbolic geometry. Accordingly, a *spherical orthogonal circle pattern for* $\mathscr{D}$ *or* G is a configuration of circles on the sphere $\mathbb{S}^2$ intersecting orthogonally corresponding to the combinatorics of $\mathscr{D}$ or G respectively analogously as for an orthogonal planar circle pattern. In other terms, spherical orthogonal circle patterns for $\mathscr{D}$ or G and orthogonal planar circle patterns for $\mathscr{D}$ or G are related by a suitable orthogonal projections.

Theorem 8.5 *Let $\mathscr{D}$ be a strongly regular cell decomposition of a compact oriented surface with or without boundary which is a bipartite quad-graph. Let G be the associated graph built from all white vertices. Let $\Phi \in (0,\infty)^V$ be a function on the set $V = V(G)$ of vertices of G which correspond to the white vertices of $\mathscr{D}$ such that $\Phi(v) = 2\pi$ for interior vertices $v \in V(G)$. Φ prescribes for boundary*

vertices of G the Neumann boundary conditions. A planar orthogonal circle pattern corresponding to these data exists if and only if the following condition is satisfied:

If $V' \subseteq V$ is any nonempty set of vertices and $E' \subseteq E = E(G)$ is the set of all edges which are incident with any vertex $v \in V'$, then

$$\sum_{v \in V'} \Phi(v) \leq \sum_{e \in E'} \pi,$$

where equality holds if and only if $V' = V$.

If it exists, the orthogonal circle pattern is unique up to similarities.

See [4] for a proof of a more general statement.

Theorem 8.5 is useful for the application of Step 2 of our construction scheme above as we usually do not obtain in Step 1 a cell decomposition of the whole sphere but only of a part. In order to construct the Koebe polyhedron, the main task is to suitably specify the corresponding boundary conditions after stereographic projection.

However, for the examples presented in Sect. 8.4 the boundary conditions are simplest for the *spherical* circle patterns. Therefore, we have actually used a method developed by Springborn for calculating the spherical circle patterns directly by a variational principle, see [7, Sect. 8] or [37] for more details. The solutions of the spherical circle pattern problem with given boundary angles are in one-to-one correspondence with the critical points of another functional. Since this functional is not convex and has negative index at least one, the critical points cannot be obtained by simply minimizing the functional. In order to numerically compute the spherical circle pattern, a convenient reduced functional is used instead. Existence and uniqueness of a solution are not yet proven. Nevertheless, this method has proven to be amazingly powerful, in particular to produce the spherical circle patterns for the examples in Sect. 8.4. More details concerning the implementation can be found in [36].

8.3.2 Construction of S-Isothermic Discrete Minimal Surfaces with Special Boundary Conditions

In the following we explain some details of how to perform the first step (finding the combinatorial parametrization and boundary conditions) of the general constructions scheme for *S*-isothermic discrete minimal surfaces. We restrict ourselves to examples with boundary conditions specified below (Schwarzian chains).

8.3.2.1 Boundary Conditions and Reduction of Symmetries

We consider the family of smooth minimal surfaces which are bounded by a closed curve consisting of finitely many arcs of positive length each of which

either (i) *lies within a plane which intersects the surface orthogonally.* Then this boundary arc is a curvature line and the surface may be continued smoothly across the plane by reflection in this plane.

or (ii) *lies on a straight line.* Then this boundary arc is an asymptotic line and the surface may be continued smoothly across this straight line by 180°-rotation about it.

In each case the image of the boundary arc under the Gauss map is (a part of) a great circle on the unit sphere. The implications in (i) and (ii) are well-known properties of smooth minimal surfaces (Schwarz's reflection principles); see for example [16].

Since the boundary conditions are translated into angle conditions for boundary circles, an S-isothermic discrete minimal surface may also be continued by reflection in the boundary planes or by 180°-rotation about straight boundary lines. In particular, the boundary circles of the spherical circle pattern intersect the corresponding great circle orthogonally.

Thus in order to simplify the construction of a discrete analog, we only consider a *fundamental piece* of the smooth (and the discrete) minimal surface. This piece of the surface is bounded by planar curvature lines and/or straight asymptotic lines (like the whole surface itself) and the whole surface is obtained from the fundamental piece by successive reflection/rotation in its boundary planes/lines and the new obtained boundary planes/lines. See Fig. 8.7 for an example of a fundamental piece. Furthermore, there is no strictly smaller piece with this property.

8.3.2.2 Combinatorics of Curvature Lines

Given a fundamental piece of a smooth minimal surface, first determine

- a combinatorial picture of the image of the boundary arcs under the Gauss map (which are parts of great circles on $\mathbb{S}^2$),
- the angles between different boundary arcs on $\mathbb{S}^2$,
- and eventually the lengths of the boundary arcs on $\mathbb{S}^2$ if the angles do not uniquely determine the boundary curve on $\mathbb{S}^2$ (up to rotations of the sphere). This case is more difficult; see Remark 8.3.

Second, from the smooth curvature line parametrization we deduce the following conditions.

(1) Umbilic and singular points are taken from the smooth minimal surface, but only their combinatorial locations matter. The smooth surface also determines the number of curvature lines meeting at these points. For interior umbilic or

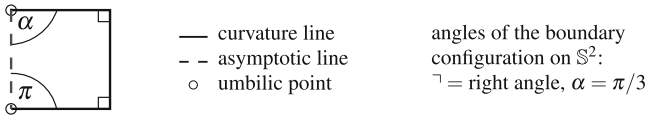


Fig. 8.7 A combinatorial picture of the boundary conditions on $\mathbb{S}^2$ of a fundamental piece of Schwarz's H surface

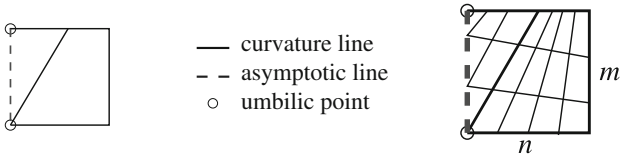


Fig. 8.8 A combinatorial conformal parametrization of a fundamental domain of Schwarz's H surface

singular points, we additionally have to take care how many times the interior region is covered by the Gauss map. This case does not occur in any of the examples presented in Sect. 8.4, so we assume for simplicity that all umbilic or singular points lie on the boundary.

(2) At a regular intersection point of a boundary curvature line and a boundary asymptotic line possibly additional curvature lines are taken from the smooth surface (depending on the intersection angle).

(3) The curvature lines of the combinatorial parametrization divide the domain into combinatorial squares. The only exceptions are combinatorial triangles formed by two curvature lines and by an asymptotic line on the boundary.

Hence, first determine all special points from conditions (1) and (2) and continue the additional curvature line(s) meeting at these points. In this way, the combinatorial domain is divided into finitely many subdomains such that condition (3) holds and the discrete curvature lines correspond to or approximate the smooth (infinite) curvature line pattern; see Fig. 8.8 (left).

Given this coarse subdivision the parametrization of the subdomains is obvious. The only additional conditions occur at common boundaries of two subdomains, where the number of crossing curvature lines has to be equal on both sides. The remaining free integer parameters of the discrete minimal surface correspond to smooth parameters of the continuous minimal surface such as scaling or quotients of lengths (for example of the boundary curves).

Remark 8.3 By construction, the free parameters of the combinatorial curvature line parametrization take integer values and the number of (quadrilateral) subdomains is finite. Also, all combinatorial curvature lines are closed modulo the boundary. In general these properties do not hold for curvature lines of smooth minimal surfaces. Furthermore, there may be dependencies between the numbers of curvature lines of different types which have to be approximated by the choice of the free integer parameters of the combinatorial curvature line parametrization. These three aspects affect the different appearances of the smooth minimal surfaces and its discrete minimal analog.

In the special case that the smooth curvature lines are closed modulo the boundary and the combinatorial curvature line parametrization has only one free parameter (corresponding to overall scaling), the discrete minimal analog will have exactly the same boundary planes/straight lines (up to translation and scaling). In this way we can construct some triply-periodic S-isothermic discrete minimal analogs to triply-periodic minimal surfaces.

8.4 Examples of Discrete Minimal Surfaces

In the following, the construction of discrete minimal surfaces explained in the beginning of Sect. 8.3 and in Sect. 8.3.2 is applied to some examples. In each case we present the boundary conditions and deduce suitable reduced conditions, a combinatorial picture of the curvature and/or the asymptotic lines and the image of the reduced boundary conditions under the Gauss map. If this image is uniquely determined by the intersection angles between arcs of great circles (up to rotation of the sphere), then the corresponding spherical circle pattern will also be unique (due to the simple combinatorics). This is the case in Sects. 8.4.1–8.4.6, for symmetric quadrilaterals and for the cubic frames considered in Sect. 8.4.7.3. Reflection in straight boundary lines and/or boundary planes results in triply periodic discrete minimal surfaces analogously as for smooth surfaces. This is the case in Sects. 8.4.3–8.4.6 without generalizations and in Sects. 8.4.1, 8.4.2 and 8.4.7.3 for special cases. The figures often show one cubical unit cell of the periodic lattice.

Pictures of the corresponding smooth minimal surfaces can be found in textbooks like [16, 26], in [22], on Brakke's web-page [10], or in some of the original treatises cited below.

Notation
For the combinatorial pictures of the curvatures lines we use the following notation.

——	curvature line	Angles of the boundary configuration on $\mathbb{S}^2$:	
– –	asymptotic line	⌝	right angle
○	umbilic point	$\alpha, \beta, \gamma, \delta, \eta$	given angles
n,m,k,l	integer parameters of curvature or asymptotic lines		

8.4.1 Gergonne's Surface

Gergonne's surface traces back to Gergonne [19], who posed the first geometric problem leading to minimal surfaces with free boundaries in 1816. A correct solution was only found by H.A. Schwarz in 1872; see [35, pp. 126–148]. Figure 8.9 shows two discrete minimal analogues.

Boundary conditions: Given a cuboid take two opposite faces as boundary faces and non-collinear diagonals of two other opposite faces, as in Fig. 8.10a. Then the two axes of 180°-rotation symmetry (see Fig. 8.10b) will lie on the minimal surface and cut it into four congruent fundamental pieces bounded by three straight asymptotic lines and one planar curvature line; see Fig. 8.10c. Its images under the Gauss map are spherical triangles with angles $\frac{\pi}{2}$, $\frac{\pi}{2}$, and α.

A combinatorial picture of the asymptotic lines or of the curvature lines is shown in Fig. 8.10d. The two parameters correspond to the free choice of two length

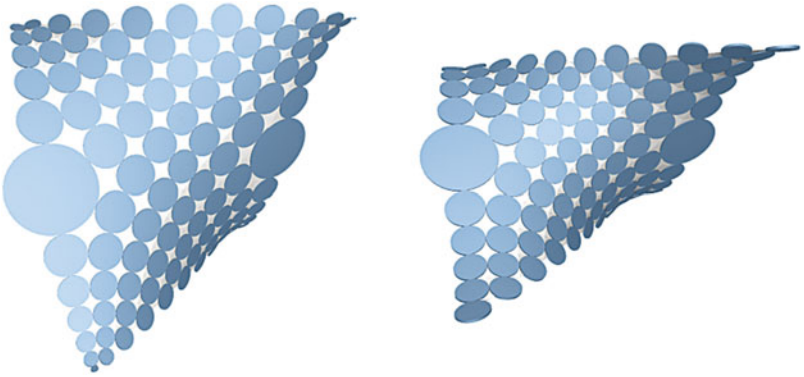


Fig. 8.9 Discrete Gergonne's surface with $\alpha = \frac{\pi}{6}$ (*left*, case 3 of Fig. 8.10d with $m = 1, n = 3$) and $\alpha = \frac{\pi}{4}$ (*right*, case 2 of Fig. 8.10d with $n = 4$)

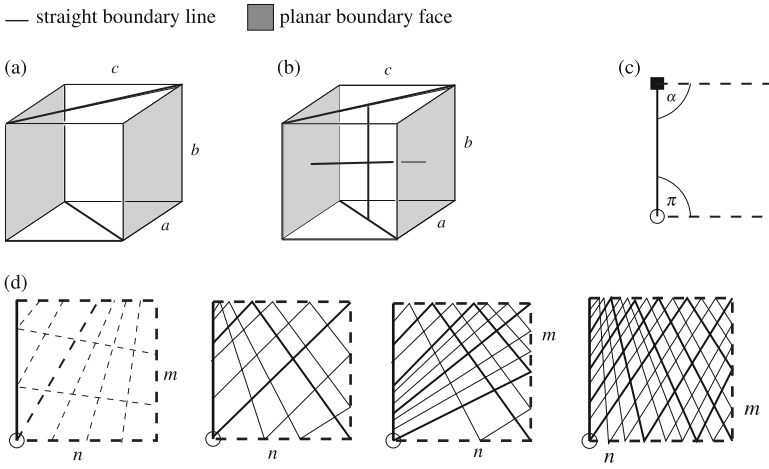


Fig. 8.10 Gergonne's surface: boundary conditions and combinatorial conformal parametrizations. (**a**) Boundary conditions. (**b**) Symmetry axes. (**c**) Reduced boundary conditions. (**d**) Different types of combinatorial conformal parametrizations

parameters of the cuboid, given the angle $(\frac{\pi}{2} - \alpha)$ between the diagonal and the planar boundary face. If $\alpha = \frac{\pi}{4}$, the minimal surface [see Fig. 8.9 (right)] can be continued by reflection and rotation in the boundary faces/edges to result in a triply periodic (discrete) minimal surface.

8.4.2 Schwarz's CLP Surface

Schwarz's CLP surface is one of the triply periodic minimal surfaces already constructed by H.A. Schwarz, see [35, vol. 1, pp. 92–125]. The name CLP is due to Schoen [33]. He considered the labyrinth formed by the periodic surface and associated the name to properties of the underlying spatial lattice.

Boundary conditions: Consider a frame of two pasted rectangles (with edge lengths a, b, c) which enclose an angle $\alpha \in (0, \pi)$, as in Fig. 8.11 (left). Then there is a plane of reflection symmetry orthogonal to the sides with length a and a corresponding planar curvature line. If $b = c$ there is another plane of reflection symmetry through both corners with angle α and yet another curvature line. This curvature line persists if $b \neq c$. The image of a fundamental domain for the generic case $b \neq c$ under the Gauss map is a spherical triangle with angles $\frac{\pi}{2}$, $\frac{\pi}{2}$, and $(\pi - \alpha)$.

The combinatorics of curvature lines of one fourth of the surface are depicted in Fig. 8.11 (right). The two integer parameters correspond to the two boundary lengths of the first rectangle. Thus for the whole frame there are three parameters corresponding to the three lengths a, b, c. If $\alpha = \frac{\pi}{2}$ the surface [see Fig. 8.12 (right)]

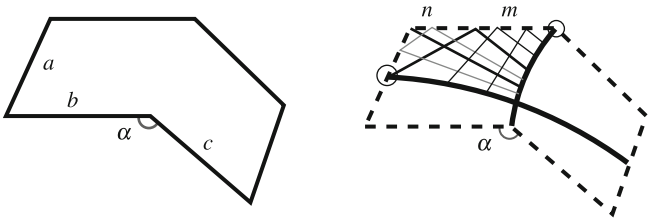


Fig. 8.11 A generalization of Schwarz's CLP surface: boundary frame (*left*) and combinatorics of curvature lines (*right*)

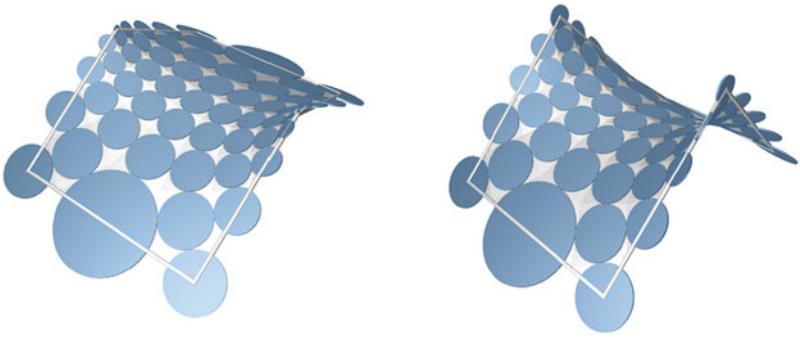


Fig. 8.12 A generalization of Schwarz's CLP surface (*left*, $\alpha = \frac{2\pi}{3}$, $m = 5, n = 1$) and Schwarz's CLP surface (*right*, $\alpha = \frac{\pi}{2}$, $m = 5, n = 1$)

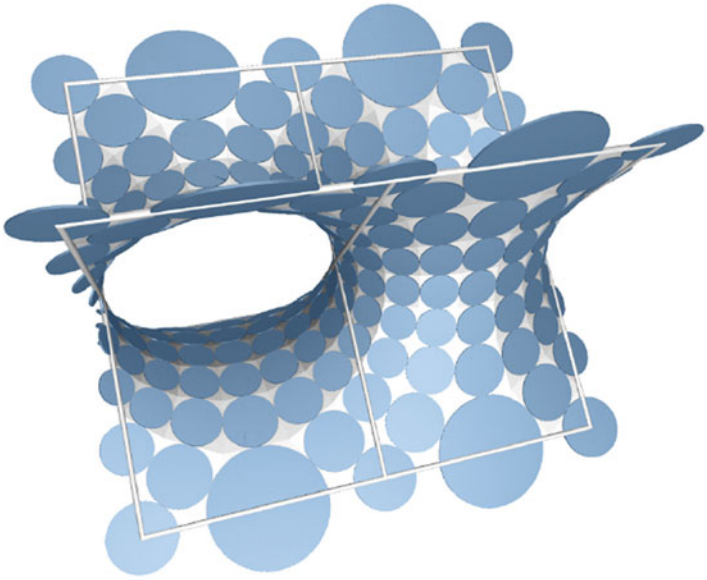


Fig. 8.13 Schwarz's CLP surface

can be continued across its boundaries to build a triply periodic discrete minimal surface. Figure 8.13 shows a cubical unit cell.

8.4.3 Schwarz's D Surface

Schwarz's D surface is another triply periodic minimal surfaces already constructed by H.A. Schwarz, see [35, vol. 1, pp. 92–125]. The surface was named D by A. Schoen because its labyrinth graphs are 4-connected 'diamond' networks.

Boundary conditions: From the edges of a cuboid take a closed boundary frame as in Fig. 8.14 (left). The three straight lines of 180°-rotation symmetry lying within the minimal surface and one of the three planes of reflection symmetry orthogonal to the boundary frame which give planar curvature lines are depicted in Fig. 8.14 (left). Therefore a fundamental piece is a triangle bounded by two straight asymptotic lines and one planar curvature line. Its image under the Gauss map is a spherical triangle with angles $\frac{\pi}{2}$ (between the two asymptotic lines) and $\frac{\pi}{4}$ and $\frac{\pi}{3}$.

The simple combinatorics of the curvature lines is shown in Fig. 8.14 (right). The only parameter corresponds to overall scaling or refinement. The minimal surface

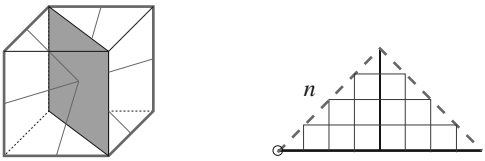


Fig. 8.14 Schwarz's D surface: boundary frame with symmetries (*left*) and combinatorics of curvature lines of a fundamental piece (*right*)

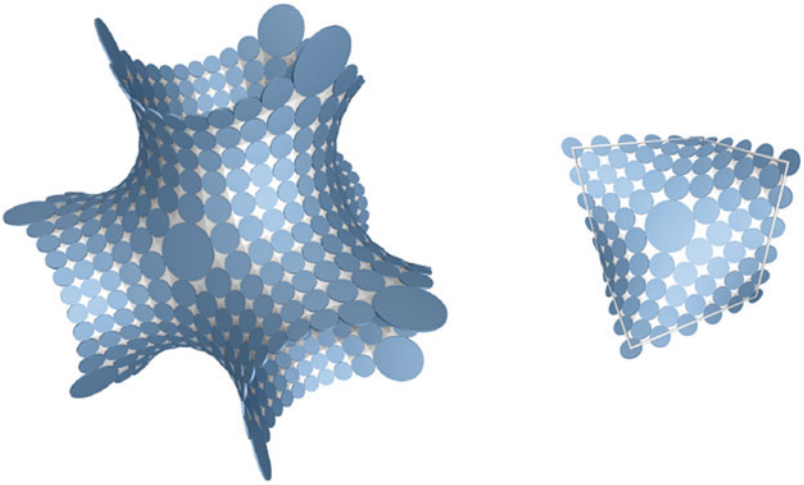


Fig. 8.15 Schwarz's D surface ($n = 4$)

constructed by the boundary frame can be continued across the boundary to result in a triply periodic minimal surface, see Fig. 8.15.

8.4.4 Neovius's Surface

Schwarz [35] began to consider minimal surfaces bounded by two straight lines and an orthogonal plane. His student E.R. Neovius continued and deepened this study and found another triply periodic surface, see [25]. This surface has the same symmetry group as Schwarz's P surface and was named C(P) by A. Schoen.

Boundary conditions: One unit cell of the lattice of Neovius's surface is basically a cubical cell with one central chamber and necks out of the middle of each edge of the cube; see Fig. 8.16. By symmetry, it is sufficient to consider one eighth of the unit cell. All the faces of this cuboid are boundary planes and we have the same three planes of reflection symmetry and three lines of 180°-rotation symmetry as

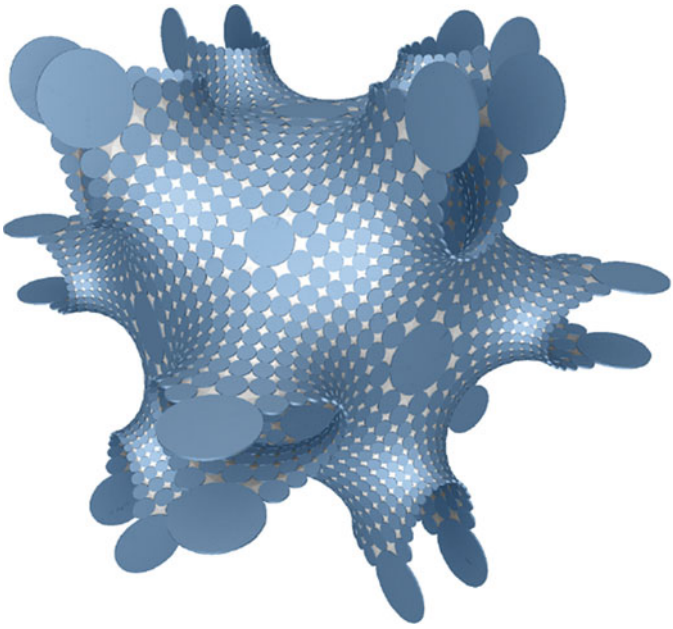


Fig. 8.16 Neovius's surface ($n = 15$)

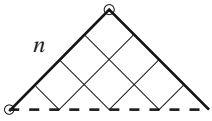


Fig. 8.17 Combinatorics of curvature lines of a fundamental piece of Neovius surface

for Schwarz's D surface; see Fig. 8.14 (left). Therefore a fundamental piece is a triangle bounded by one straight asymptotic lines and two planar curvature lines. Its image under the Gauss map is a spherical triangle with angles $\frac{3\pi}{4}$ (between the two curvature lines) and $\frac{\pi}{4}$ and $\frac{\pi}{3}$.

The combinatorics of the curvature lines is again very simple, see Fig. 8.17.

8.4.5 Schwarz's H Surface

Schwarz's H surface is another triply periodic minimal surface which was already known to H.A. Schwarz, see [35, vol. 1,pp. 92–125].

Boundary conditions: Take the edges of two parallel copies of an equilateral triangle as boundary frame for a minimal surface spanned in between them, as in Fig. 8.18. Then there are one plane of reflection symmetry parallel to the triangles and three other orthogonal planes as symmetry group of the equilateral triangle. Thus a fundamental piece is bounded by three planar curvature lines and one straight

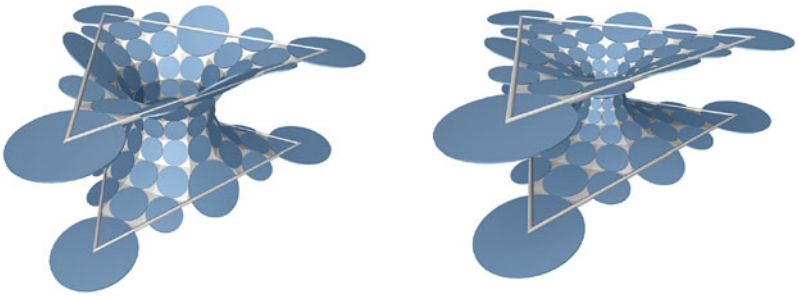


Fig. 8.18 Schwarz's H surfaces with different heights (*left*: $m = 3, n = 3$, *right*: $m = 3, n = 6$)

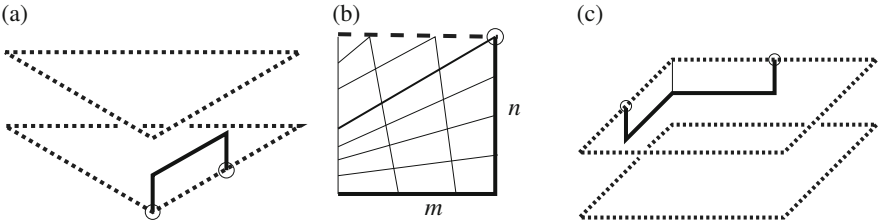


Fig. 8.19 Schwarz's H surface and (a generalization of) Schoen's I-6 surface. (**a**) Boundary frame of Schwarz's H surface with a fundamental piece. (**b**) Combinatorics of the curvature lines of (one half of) a fundamental piece. (**c**) Boundary frame of Schoen's I-6 surface with a fundamental piece

asymptotic line; see Fig. 8.19a. Its image under the Gauss map is a spherical triangle with angles $\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{3}$; see Fig. 8.7.

The combinatorics of the curvature lines is shown in Fig. 8.8 or 8.19b. The two parameters correspond to the side length of the equilateral triangle and to their distance.

8.4.6 Schoen's I-6 Surface and Generalizations

About 1970, the physicist and crystallographer A. Schoen discovered many triply periodic minimal surfaces. His reports [33] were a bit sketchy, but Karcher [22] established the existence of all of Schoen's surfaces.

Boundary conditions: Similarly as for Schwarz's H surface, take the edges of two parallel copies of a rectangle as boundary frame for a minimal surface between them. There are one plane of reflection symmetry parallel to the rectangles and two other orthogonal planes as symmetry group of the rectangle (four in case of a

square). Thus a fundamental piece is generically bounded by three planar curvature lines and one straight asymptotic line; see Fig. 8.19c. Its image under the Gauss map is a spherical triangle with angles $\frac{\pi}{2}$, $\frac{\pi}{2}$, and $\frac{\pi}{4}$. There is one additional curvature line splitting the piece into two parts which can easily be found be reflection symmetry in the case of squares.

For both parts the combinatorics of curvature lines are the same as for the fundamental piece of Schwarz's H surface; see Fig. 8.19b or Fig. 8.8. The three remaining parameters correspond to the side lengths of the two rectangles and to their distance.

In an analogous way to the construction of Schwarz's H surface and Schoen's I-6 surface, we may consider all regular symmetric planar polygons with sides of equal length or to non-regular planar polygons with additional symmetry planes. The fundamental piece is always combinatorially the same as for Schwarz's H surface. The only difference is the angle $\alpha = \frac{\pi}{n}$ for an n-gon. These minimal surfaces may also be understood as an approximation to a piece of the catenoid; see Fig. 8.21 (left). The construction may easily be generalized to further similar examples; see Fig. 8.20 (right) and Fig. 8.21 (right).

8.4.7 Polygonal Boundary Frames

As a special class of Plateau's problems, all non-planar, simple, unknotted polygons can be considered as boundary conditions. The main task is to determine the corresponding discrete combinatorics of the conformal parameter lines. The number of different special cases increases with the number of boundary segments, so we confine ourselves to quadrilaterals, pentagons, and a cubical boundary frame as one more complicated example. Since the boundary frame consists only of straight

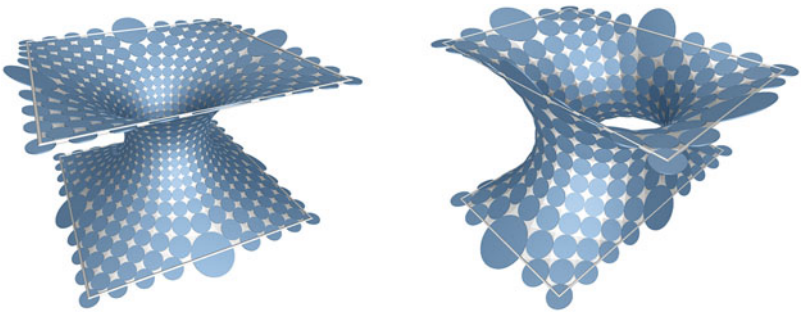


Fig. 8.20 Schoen's I-6 surface (*left*, $m = 7, n = 14$) and a generalization (*right*)

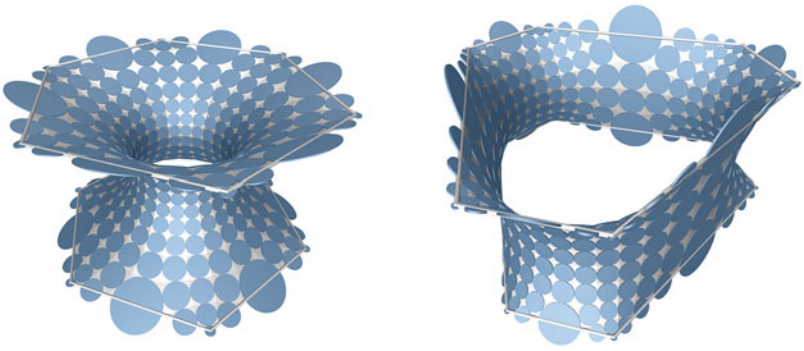


Fig. 8.21 An approximation of a piece of a catenoid using a polygonal boundary frame (*left*, $n = 10, m = 3$) and a generalization (*right*, $m_1 = 3, n_1 = 6, m_2 = 7, n_2 = 2$)

asymptotic lines, we use the conjugate minimal surface from the associated family for the construction, see Definition 8.8.

8.4.7.1 Quadrilateral

The minimal surface spanned by a quadrilateral with equal side lengths and equal angles of $\frac{\pi}{3}$ was the first known solution in the class of Plateau's problems. In 1865, H.A. Schwarz found the explicit solution [35, pp. 6–125]. About the same time, B. Riemann independently solved this problem [30, pp. 326–329]. His paper [29] appeared posthumously in 1867. At the same time H. A. Schwarz sent his prize-essay to the Berlin Academy. Later on, Plateau's problem was tackled for other polygonal boundaries, see for example the historical remarks in [16, 26].

Given a non-planar quadrilateral of straight boundary lines as in Fig. 8.22, the combinatorics of asymptotic lines is easily found; see Fig. 8.23b (left). Then the corresponding curvature lines may also be determined, as in Fig. 8.23b (right). The two parameters correspond to a global scaling and the to a ratio of lengths of the boundary segments. As explained in Remark 8.3, in general we do not obtain exactly a given quadrilateral boundary frame but an approximation which converges for a suitable choice of parameter values.

In the case of a non-planar symmetric quadrilateral the combinatorics of curvature lines are obvious; see Fig. 8.23a. The only parameter corresponds to a global scaling or refinement. Thus we can obtain any symmetric quadrilateral boundary frame as the exact boundary of a discrete minimal surface, see Fig. 8.22.

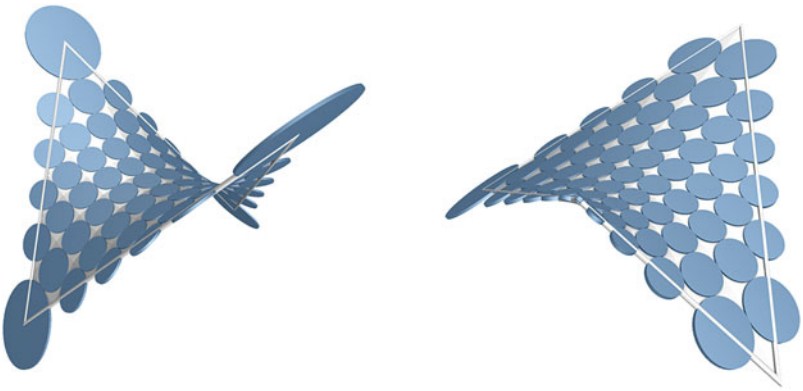


Fig. 8.22 Two examples of discrete quadrilateral boundary frames: a symmetric version (*left* with $\alpha = \beta = \gamma = \pi/6$, $n = 9$) and a general version (*right*)

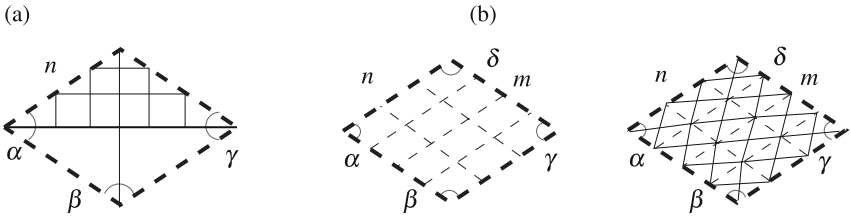


Fig. 8.23 Boundary conditions and combinatorial parametrizations for symmetric (**a**) and general (**b**) quadrilaterals

8.4.7.2 Pentagon

In the symmetric case the boundary configuration allows a plane of mirror symmetry containing one boundary vertex and cutting the opposite edge. The remaining reduced domain is similar to the boundary conditions of Gergonne's surface and thus has the same combinatorics of asymptotic lines; see Fig. 8.24a.

The general case is more difficult. In principle, there are two possibilities for the combinatorial position of an additional umbilic point; see Fig. 8.24b. The integer parameters correspond to lengths of boundary segments. See Fig. 8.25 for two examples.

8.4.7.3 A Cubical Frame

As a last example consider a more complicated polygonal boundary frame as illustrated in Fig. 8.27a. By construction, the minimal surface spanned by this frame has two planes of reflection symmetry and yet two planar curvature lines.

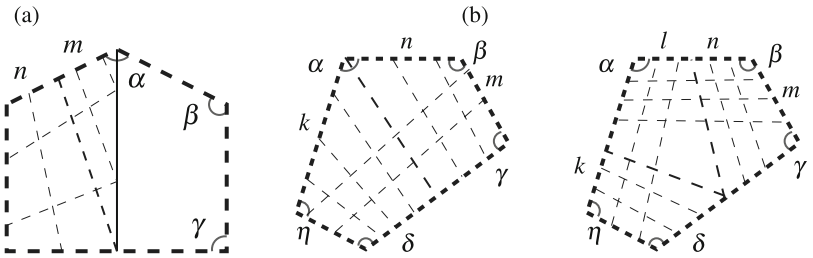


Fig. 8.24 Boundary conditions and combinatorics of asymptotic lines for symmetric (**a**) and general (**b**) pentagons

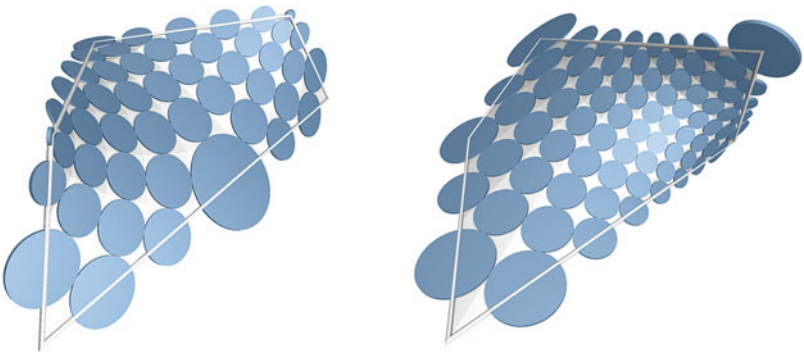


Fig. 8.25 Two examples of discrete pentagons: a symmetric boundary frame (*left*, $m = 4, n = 4$) and a general boundary frame (*right*)

A fundamental piece is therefore bounded by three straight asymptotic lines and two curvature line. Its image under the Gauss map is a spherical triangle with three right angles.

The possible combinatorics of curvature lines are depicted in Fig. 8.27b. The three integer parameters correspond to the three edge lengths of the cuboid. A minimal surface constructed by this boundary conditions is shown in Fig. 8.26 (left) and can be continued across the boundary to result in a triply periodic minimal surface. Of course, this example may be generalized to related problems. Figure 8.26 (right) shows an example of such a generalization.

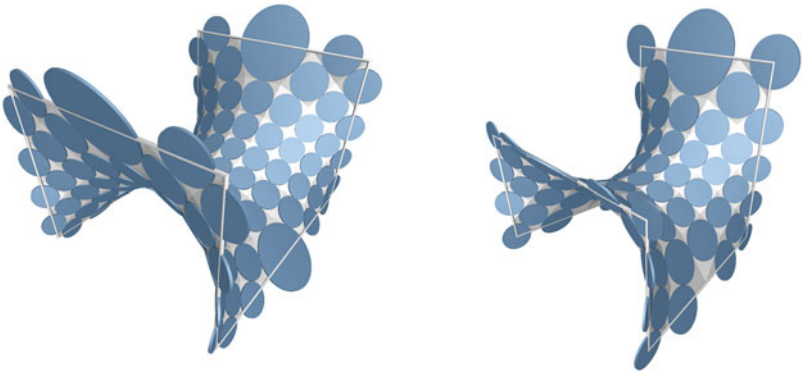


Fig. 8.26 An example of a discrete minimal surface with cubical boundary frame (*left*, first case of Fig. 8.27b with $k = 2, n = 4$) and a generalization (*right*)

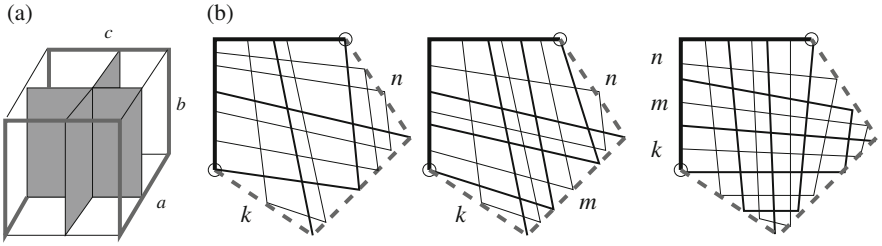


Fig. 8.27 A general cubical boundary frame. (**a**) A cubic boundary frame and its symmetry planes. (**b**) Different possible types for the combinatorics of curvature lines of a fundamental piece

8.5 Weierstrass Representation and Convergence of Discrete Minimal Surfaces

The convergence of the *S*-isothermic discrete minimal surfaces can easily be deduced from the convergence of the orthogonal circle patterns, which are used for the construction of the associated Koebe polyhedron, to the corresponding smooth Gauss map. More precisely, we prove the convergence of the stereographic projections of the spherical orthogonal circle patterns to the corresponding complex Gauss map. Then the convergence of the corresponding *S*-isothermic discrete minimal surfaces is obtained using a suitable discrete Weierstrass representation.

8.5.1 Discrete Weierstrass Representation

The classical Weierstrass representation relates a conformal mapping of the complex plane, i.e. a bijective holomorphic mapping, to a smooth minimal surface. In the discrete case, we take suitable orthogonal circle patterns with the same combinatorics as discrete conformal mappings. This leads to an analogous discrete Weierstrass representation formula.

Theorem 8.6 (Discrete Weierstrass Representation; [7, Theorem 6]) *Let $\mathscr{D}$ be an S-quad-graph. Let $\mathscr{C}$ be a planar orthogonal circle pattern for $\mathscr{D}$. Denote the centers and intersection points of this circle pattern by $c_{\mathscr{C}}$ and $p_{\mathscr{C}}$ respectively. Then an S-isothermic discrete minimal surface $F : V(\mathscr{D}_{\textcircled{s}}) \to \mathbb{R}^3$ is given by the following formula:*

Let $v_1, v_2 \in V(\mathscr{D}_{\textcircled{s}})$ be two vertices, which correspond to touching circles of the pattern. Let $y \in V_b(\mathscr{D})$ be the black vertex between v_1 and v_2, which corresponds to the point of contact. Then the centers $F(v_1)$ and $F(v_2)$ of the corresponding touching spheres of the S-isothermic discrete minimal surface F satisfy

$$F(v_2) - F(v_1) = \pm \Re \left(\frac{r(v_2) + r(v_1)}{1 + |p|^2} \frac{\overline{c_{\mathscr{C}}(v_2) - c_{\mathscr{C}}(v_1)}}{|c_{\mathscr{C}}(v_2) - c_{\mathscr{C}}(v_1)|} \begin{pmatrix} 1 - p^2 \\ i(1 + p^2) \\ 2p \end{pmatrix} \right), \tag{8.7}$$

where $p = p_{\mathscr{C}}(y)$ and the radii $r(v_j)$ of the spheres are

$$r(v_j) = \left| \frac{1 + |c_{\mathscr{C}}(x_j)|^2 - |c_{\mathscr{C}}(x_j)p|^2}{2|c_{\mathscr{C}}(x_j) - p|} \right|. \tag{8.8}$$

The sign on the right-hand side of equation (8.7) depends on the label of the edge connecting v_1 with v_2.

8.5.2 Convergence of Discrete Minimal Surfaces

The discrete Weierstrass representation and its smooth analogon are the basis of our convergence proof. In fact, note that C^∞-convergence of the orthogonal circle patterns to a conformal mapping implies the C^∞-convergence of the associated discrete S-isothermic minimal surfaces (suitably scaled and translated) to the corresponding smooth minimal surface.

We restrict ourselves to the class of (discrete) minimal surfaces which are bounded, homeomorphic to a disk, and have a boundary curve consisting of planar curvature lines or straight asymptotic lines. As a simple example, consider a part of Schwarz's D surface.

First, we fix some notation. Let *SGD* be the regular square grid cell decomposition of the complex plane, that is the vertices are $V(SGD) = \mathbb{Z} + i\mathbb{Z}$ and the edges are given by pairs of vertices $[z, z']$ with $z, z' \in V(SGD)$ and $|z - z'| = 1$. Bicolor the vertices $V(SGD)$ such that the black vertices are $V_b(SGD) = \{n + im \in \mathbb{Z} + i\mathbb{Z} : n + m = 1 \pmod 2\}$ and the white vertices are $V_w(SGD) = \{n + im \in \mathbb{Z} + i\mathbb{Z} : n + m = 0 \pmod 2\}$.

Adding circles of radius 1 with centers at the vertices $V_b(SGD)$ leads to an isoradial circle pattern filling the whole complex plane, where circles intersect only orthogonally or touch.

Definition 8.10 Consider a cell decomposition of a two-dimensional manifold which is isomorphic to a subset of *SGD* and let G be its bicolored graph. An *SG-circle pattern* with the combinatorics of G is a configuration of circles, such that to each white vertex of G there corresponds a circle (all with the same orientation). If two vertices are incident to the same face then the corresponding circles intersect orthogonally. Furthermore, if two vertices are incident to the same black vertex of G but are not incident to the same face, the corresponding circles touch and have disjoint interiors.

Let $\mathscr{D}$ be simply connected bounded region in the complex plane whose boundary is a convex kite with straight edges, or a stereographic projection of a convex spherical kite whose edges are parts of great circles and lying strictly within one half-sphere. Let $\mathscr{R} = \{x+iy : x, y \in [0, 1]\}$ be the closed unit square in $\mathbb{R}^2 \cong \mathbb{C}$.

For $n \in \mathbb{N}$, denote by SG_n the cell decomposition *SGD* scaled by the factor $1/(2n) > 0$. Denote the subdecomposition corresponding to all vertices of SG_n contained in $\mathscr{R}$ by $SG_n^{\mathscr{R}}$. By abuse of notation, we will not distinguish between the abstract cell decomposition $SG_n^{\mathscr{R}}$ and its embedding into $\mathscr{R}$. The isoradial orthogonal circle pattern of all circles with the same radius $1/(2n)$ and centers in the white vertices $V_w(SG_n^{\mathscr{R}})$ is denoted by $\mathscr{C}_n$. The four vertices $a + ib \in V_w(SG_n^{\mathscr{R}})$ with $a, b \in \{0, 1\}$ at the corners of $\mathscr{R}$ will be referred to as *corner points* of $\mathscr{R}$.

Assume that for each $n \in \mathbb{N}$ there is an *SG*-circle pattern $\mathscr{C}_n^{\mathscr{D}}$ with the combinatorics of $SG_n^{\mathscr{R}}$ such that all boundary circles intersect the boundary $\partial\mathscr{D}$ orthogonally and the circles corresponding to corner points of $SG_n^{\mathscr{R}}$ intersect the two corresponding boundary arcs of $\partial\mathscr{D}$ orthogonally. In the case of a convex kite (i.e. straight edges), the existence of $\mathscr{C}_n^{\mathscr{D}}$ is guaranteed by Theorem 8.5. The general question on existence for the spherical case is still open, see Sect. 8.3 and [7] for further remarks and special cases. Nevertheless, existence can be proven for some special cases, for example for a symmetric spherical quadrilateral with angles $\alpha = 2\pi/3 = \gamma$, $\beta = \pi/2 = \delta$. This choice of angles leads to four fundamental pieces of Schwarz's D surface bounded by four planar curvature lines, see Sect. 8.4.3.

Given the cell decomposition $SG_n^{\mathscr{R}}$, we can easily obtain a triangulation by adding edges between white vertices which are incident to the same face. Given the circle patterns $\mathscr{C}_n$ and $\mathscr{C}_n^{\mathscr{D}}$, we denote the embeddings of the above triangulations by $T_n \subset \mathscr{R}$ and $T_n^{\mathscr{D}} \subset \mathscr{D}_n$ respectively. Here $\mathscr{D}_n$ is the convex hull of the centers and the intersection points of $\mathscr{C}_n^{\mathscr{D}}$. Note that all triangles of T_n and $T_n^{\mathscr{D}}$ are right-angled.

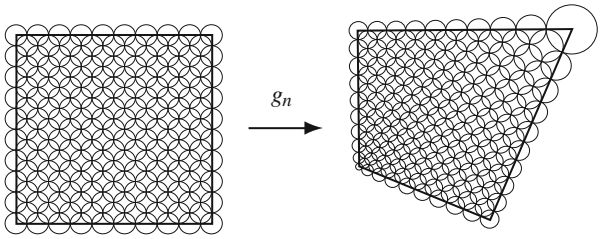


Fig. 8.28 An example of two corresponding orthogonal circle patterns filling a square and a kite-shaped convex quadrilateral respectively

Now we construct a homeomorphism corresponding to the circle patterns which approximates a given conformal mapping from $\mathscr{R}$ onto $\mathscr{D}$.

Fix a conformal homeomorphism from $\mathscr{R}$ onto $\mathscr{D}$ which maps the corner points of $\mathscr{R}$ to the corner points of $\mathscr{D}$. Let g^{corn} be the restriction of g to the corner points of $\mathscr{R}$. Denote by g_n^{corn} the corresponding bijective mapping which maps the corner points of $\mathscr{R}$ to the corresponding centers of circles of $\mathscr{C}_n^{\mathscr{D}}$. As approximating mapping we define $g_n : \mathscr{R} \to \mathscr{D}_n$ to be the simplicial map determined by the correspondence of vertices and edges of T_n and $T_n^{\mathscr{D}}$ which agrees with g_n^{corn} at the corner points. Recall that a simplicial map for two triangulations as T_n and $T_n^{\mathscr{D}}$ maps vertices to corresponding vertices and is linear on every triangular face. An example of $\mathscr{C}_n$ and $\mathscr{C}_n^{\mathscr{D}}$—and hence g_n—is given in Fig. 8.28.

For abbreviation we denote $\mathscr{R}^* = \mathscr{R} \setminus \{\text{corner points}\}$. Compact sets $K \subset \mathscr{R}^*$ will always mean compact subsets of $\mathbb{C}$ with respect to the standard metric which are contained in $\mathscr{R}^*$. Furthermore, we define the radius function $r_n : V_w(SG_n^{\mathscr{R}}) \to (0, \infty)$ of $\mathscr{C}_n^{\mathscr{D}}$ which assigns to every white vertex the radius of the corresponding circle of $\mathscr{C}_n^{\mathscr{D}}$. Now we can state our convergence result:

Theorem 8.7 *The mappings g_n converge to the unique conformal homeomorphism $g : \mathscr{R} \to \mathscr{D}$ which coincides with g^{corn} at the corner points. The convergence is uniform on $\mathscr{R}$ and in C^∞ on compact subsets of $\mathscr{R}^*$. Furthermore, the quotients of corresponding radii, $2nr_n$, converge to $|g'|$ in C^∞ uniformly on compact subsets of $\mathscr{R}^*$.*

A detailed version and a generalization of Theorem 8.7 as well as a proof can be found in [12, Chap. 7] or in [13].

Remark 8.4 If the angle of $\mathscr{D}$ at a corner point is $\pi/2$ then the convergence is also in C^∞ on compact sets of $\mathscr{R}$ including the corresponding corner point. Instead of $\mathscr{R}$, we may consider the rotated region $\mathscr{R}' = e^{i\pi/4}\mathscr{R}$ and the part of SG_n contained in $\mathscr{R}'$. Note that the corresponding combinatorics differ from $SG_n^{\mathscr{R}}$. Theorem 8.7 also holds in this case.

Theorem 8.8 *Let $\mathscr{D}$ be the projection of a symmetric spherical quadrilateral $Q \subset \mathbb{S}^2$ bounded by parts of great circles which is entirely contained in one half of the sphere. Assume that for all $n \in \mathbb{N}$ orthogonal circle patterns $\mathscr{C}_n^{\mathscr{D}}$ with the*

combinatorics of $SG_n^{\mathscr{R}}$ exist such that all boundary circles intersect the boundary $\partial\mathscr{D}$ orthogonally and the circles corresponding to corner points of $SG_n^{\mathscr{R}}$ intersect two corresponding boundary lines of $\partial\mathscr{D}$ orthogonally. For $n \in \mathbb{N}$, define g_n as in Theorem 8.7.

Let F_n be the discrete minimal surface obtained from the stereographic projection of $\mathscr{C}_n^{\mathscr{D}}$ by dualization and scaled as in Theorem 8.6. Then the scaled S-isothermic discrete minimal surfaces M_n corresponding to $\frac{1}{n^2}F_n$ converge uniformly for the images of compact subsets of $\mathscr{R}^*$ to a smooth minimal surface M. The stereographic projection of the Gauss map, determining M up to translation and scaling, is given by the conformal map $g : \mathscr{R} \to \mathscr{D}$.

The proof is a straight forward application of the Weierstrass Representation Theorem 8.6, see also the proof of [7, Theorem 9].

Acknowledgements The authors were supported by the DFG Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics".

References

1. Bobenko, A.I., Hoffmann, T.: S-conical CMC surfaces. Towards a unified theory of discrete surfaces with constant mean curvature. In: Bobenko, A.I. (ed.) Advances in Discrete Differential Geometry, pp. 287–308. Springer, Berlin (2016)
2. Bobenko, A.I., Pinkall, U.: Discrete isothermic surfaces. J. Reine Angew. Math. **475**, 187–208 (1996)
3. Bobenko, A.I., Pinkall, U.: Discretization of surfaces and integrable systems. In: Bobenko, A.I., Seiler, R. (eds.) Discrete Integrable Geometry and Physics, pp. 3–58. Clarendon Press, Oxford (1999)
4. Bobenko, A.I., Springborn, B.A.: Variational principles for circle patterns and Koebe's theorem. Trans. Am. Math. Soc. **356**, 659–689 (2004)
5. Bobenko, A.I., Suris, Y.B.: Discrete Differential Geometry. Integrable Structure. Graduate Studies in Mathematics, vol. 98. American Mathematical Society, Providence, RI (2008). doi:http://dx.doi.org/10.1090/gsm/098
6. Bobenko, A.I., Suris, J.B.: Discrete Koenigs nets and discrete isothermic surfaces. Int. Math. Res. Not. **11**, 1976–2012 (2009)
7. Bobenko, A.I., Hoffmann, T., Springborn, B.A.: Minimal surfaces from circle patterns: Geometry from combinatorics. Ann. Math. **164**(1), 231–264 (2006)
8. Bobenko, A.I., Pottmann, H., Wallner, J.: A curvature theory for discrete surfaces based on mesh parallelity. Math. Ann. **348**, 1–24 (2010)
9. Bobenko, A., Hoffmann, T., König, B., Sechelmann, S.: S-conical minimal surfaces. Towards a unified theory of discrete minimal surfaces (2015). http://www.discretization.de/en/publications/. Preprint
10. Brakke, K.A.: Triply periodic minimal surfaces. Internet (cited 2017 Jan 10). http://www.susqu.edu/facstaff/b/brakke/evolver/examples/periodic/periodic.html
11. Brightwell, G.R., Scheinerman, E.R.: Representations of planar graphs. SIAM J. Discrete Math. **6**(2), 214–229 (1993)
12. Bücking, U.: Approximation of conformal mappings by circle patterns and discrete minimal surfaces. Ph.D. thesis, Technische Universität Berlin (2007). Published online at http://opus.kobv.de/tuberlin/volltexte/2008/1764/

13. Bücking, U.: Approximation of conformal mappings by circle patterns. Geom. Dedicata **137**, 163–197 (2008)
14. Christoffel, E.B.: Über einige allgemeine Eigenschaften der Minimumsflächen. J. Reine Angew. Math. **67**, 218–228 (1867)
15. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Studies in Mathematics and its Applications. North-Holland, Amsterdam (1978)
16. Dierkes, U., Hildebrandt, S., Küster, A., Wohlrab, O.: Minimal Surfaces I. Springer, Berlin (1992)
17. Dziuk, G.: An algorithm for evolutionary surfaces. Numer. Math. **58**, 603–611 (1991)
18. Dziuk, G., Hutchinson, J.E.: The discrete Plateau problem: algorithm and numerics. Math. Comput. **68**, 1–23 (1999)
19. Gergonne, J.D.: Questions proposées/résolues. Ann. Math. Pure Appl. **7**, 68, 99–100, 156, 143–147 (1816)
20. Hertrich-Jeromin, U.: Introduction to Möbius Differential Geometry. London Mathematical Society Lecture Note Series, vol. 300. Cambridge University Press, Cambridge (2003)
21. Joswig, M., Mehner, M., Sechelmann, S., Techter, J., Bobenko, A.I.: DGD Gallery: Storage, sharing, and publication of digital research data. In: Bobenko, A.I. (ed.) Advances in Discrete Differential Geometry, pp. 421–439. Springer, Berlin (2016)
22. Karcher, H.: The triply periodic minimal surfaces of Alan Schoen and their constant mean curvature companions. Manuscr. Math. **64**, 291–357 (1989)
23. Koebe, P.: Kontaktprobleme der konformen Abbildung. Abh. Sächs. Akad. Wiss. Leipzig Math.-Natur. Kl. **88**, 141–164 (1936)
24. Müller, C., Wallner, J.: Oriented mixed area and discrete minimal surfaces. Discrete Comput. Geom. **43**, 303–320 (2010)
25. Neovius, E.R.: Bestimmung zweier spezieller periodischer Minimalflächen, auf welchen unendlich viele gerade Linien und unendlich viele ebene geodätische Linien liegen. J. C. Frenckell & Sohn, Helsingfors (1883)
26. Nitsche, J.C.C.: Lectures on Minimal Surfaces, vol. 1. Cambridge University Press, Cambridge (1989)
27. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Exp. Math. **2**, 15–36 (1993)
28. Pottmann, H., Liu, Y., Wallner, J., Bobenko, A.I., Wang, W.: Geometry of multi-layer freeform structures for architecture. ACM Trans. Graph. **26**(3), 65 (2007). In: Proceedings of SIGGRAPH
29. Riemann, B.: Über die Fläche vom kleinsten Inhalt bei gegebener Begrenzung. Abh. Königl. Ges. d. Wiss. Göttingen, Mathem. Cl. **13**, 3–52 (1867). K. Hattendorff, edit.
30. Riemann, B.: Gesammelte Mathematische Werke - Nachträge. Teubner, Leipzig (1876/1902)
31. Rivin, I.: A characterization of ideal polyhedra in hyperbolic 3-space. Ann. Math. **143**, 51–70 (1996)
32. Schneider, R.: Convex bodies: the Brunn-Minkowski Theory. Cambridge University Press, Cambridge (1993)
33. Schoen, A.H.: Infinite periodic minimal surfaces without self-intersections. Technical Report D-5541, NASA (1970). Technical Note
34. Schramm, O.: How to cage an egg. Invent. Math. **107**(3), 543–560 (1992)
35. Schwarz, H.A.: Gesammelte Mathematische Abhandlungen, vol. 1. Springer, Berlin (1890)
36. Sechelmann, S.: Discrete minimal surfaces, Koebe polyhedra, and Alexandrov's theorem. variational principles, algorithms, and implementation. Diploma thesis, Technische Universität Berlin (2007). www.sechel.de
37. Springborn, B.A.: Variational principles for circle patterns. Ph.D. thesis, Technische Universität Berlin (2003). Published online at http://opus.kobv.de/tuberlin/volltexte/2003/668/
38. Tsuchiya, T.: Discrete solution of the Plateau problem and its convergence. Math. Comput. **49**, 157–165 (1987)

Chapter 9
Robust and Convergent Curvature and Normal Estimators with Digital Integral Invariants

Jacques-Olivier Lachaud, David Coeurjolly, and Jérémy Levallois

Abstract We present, in details, a generic tool to estimate differential geometric quantities on digital shapes, which are subsets of $\mathbb{Z}^d$. This tool, called *digital integral invariant*, simply places a ball at the point of interest, and then examines the intersection of this ball with input data to infer local geometric information. Just counting the number of input points within the intersection provides curvature estimation in 2D and mean curvature estimation in 3D. The covariance matrix of the same point set allows to recover principal curvatures, principal directions and normal direction estimates in 3D. We show the multigrid convergence of all these estimators, which means that their estimations tend toward the exact geometric quantities on—smooth enough—Euclidean shapes digitized with finer and finer gridsteps. During the course of the chapter, we establish several multigrid convergence results of digital volume and moments estimators in arbitrary dimensions. Afterwards, we show how to set automatically the radius parameter while keeping multigrid convergence properties. Our estimators are then demonstrated to be accurate in practice, with extensive comparisons with state-of-the-art methods. To conclude the exposition, we discuss their robustness to perturbations and noise in input data and we show how such estimators can detect features using scale-space arguments.

J.-O. Lachaud (✉)
CNRS, LAMA, UMR 5127, Université Savoie Mont Blanc, F-73376 Le Bourget-du-lac, France

CNRS, LJK, UMR 5224, Université Grenoble-Alpes, F-38041 Saint-Martin-d'Héres, France
e-mail: jacques-olivier.lachaud@univ-savoie.fr

D. Coeurjolly
CNRS, Université de Lyon, LIRIS, UMR 5205, Université de Lyon, F-69621 Lyon, France
e-mail: david.coeurjolly@liris.cnrs.fr

J. Levallois
CNRS, LAMA, UMR 5127, Université Savoie Mont Blanc, F-73376 Le Bourget-du-lac, France

CNRS, INSA-Lyon, LIRIS, UMR 5205, Université de Lyon, F-69621 Lyon, France
e-mail: jeremy.levallois@liris.cnrs.fr

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9_9

9.1 Curvature Estimation on Discrete Data

Context and Objectives In many shape processing applications, the estimation of differential quantities on the shape boundary is usually an important step. Their correct estimation makes easier further processing, like quantitative evaluation, feature detection, shape matching or visualization. This paper focuses on estimating the curvature tensor on the boundary of digital shapes. Such digital structures are subsets of the 3-dimensional digital space $\mathbb{Z}^3$ and come generally from the digitization of some Euclidean shape. Of course, the curvature tensor estimation should be as close as possible to the curvature tensor of the underlying Euclidean shape before digitization. Digital data form a special case of discrete data with specific properties: (1) digital data cannot sample the boundary of the Euclidean shape (i.e. they do not lie on the shape boundary), (2) digital data are distributed around the true sample according to arithmetic noise, which looks rather uniform over a range $[-h, h]$ from a statistical point of view, where h is the digitization grid step. Another way of stating these characteristics is to say that the Hausdorff distance between the Euclidean shape and its digitization is some $O(h)$. Of course, the quality of the estimation should be improved as the digitization step gets finer and finer. This property is called the *multigrid convergence* [12, 29]. It is similar in spirit with the *stability* property in geometry processing: given a continuous shape and a specific sampling of its boundary, the estimated measure should converge to the Euclidean one when the sampling become denser (e.g. [2, 45]).

Curvature Estimation on Meshes Digital data being discrete in nature, it is interesting to look at the curvature estimation techniques on triangulated meshes. In computer graphics and geometry processing, there exists a vast family of techniques to estimate either the mean or Gaussian curvatures, or sometimes the full curvature tensor. Most of them are local (i.e. limited to a 1-ring or 2-ring of neighbors) but exhibit correct results for nice meshes. They generally fall into three categories: fitting, discrete methods, curvature tensor estimation. We may refer to [56] and [24] for comprehensive evaluations, and Desbrun et al. [20] or Bobenko and Suris [5] for an entirely discrete theory. Most of them do not have theoretical convergence guarantees even without noise on the mesh. We may quote [47] and [54] as approaches trying to tackle perturbation through averaging.

For Gaussian curvature estimated with Gauss-Bonnet approach (angle defect), Xu [58] provides a stability theorem for triangulated mesh whose vertices lie on the underlying smooth manifold, with valence 6 and parallelogram condition (each 1-ring of neighbors is projected as a parallelogram onto a plane). Assuming a sampling with density δ, he provides an additional convergence property whenever the sampling is perturbated by some $O(\delta^\alpha)$, but $\alpha > 2$ (inadequate for discrete data). Note that if the triangulated mesh does not satisfy these requirements, such estimation does not converge.

The integral measures of curvatures, based on normal cycle theory [16, 17] is another notable approach for estimating curvature information on a triangulated mesh. The authors exhibit some convergence results for triangulated meshes with

vertices lying on the underlying smooth Euclidean shape boundary. In this case, if the mesh has Hausdorff distance to shape boundary below ϵ, convergence is obtained with speed/error $O(\epsilon)$ under some hypotheses.

Finally, in geometry processing, interesting mathematical tools have been developed to design differential estimators on smooth surfaces based on integral invariants [48, 49]. They consist in moving a kernel along the shape surface and in computing integrals on the intersection between the shape and the kernel. The authors have demonstrated that some integral quantities provide interesting curvature information when the kernel size tends to zero. They also achieve stability depending on the kernel radius and on ϵ, for instance in the case of a mesh sampling. Our new estimators rely on the same ideas.

Curvature Estimation on Point Clouds When having only discrete data (i.e. a point cloud), the most natural way to approach curvature(s) is to fit a polynomial surface of degree two at least. Perhaps the most representative of these techniques is the osculating jets of Cazals and Pouget [10]. The authors provide $O(\delta^2)$ convergence results when the data is a surface sampling, assuming δ is the density of points. There is no theoretical result in presence of noise, although the least-square fitting of osculating jets is very robust to noise in practice.

Another family of techniques exploits the Voronoi diagram [1, 44, 45]. The idea behind these approaches is, instead of fitting the tangent space, to estimate at best the orthogonal space. The convolved covariance measure introduced by Mérigot et al. [45] is particularly appealing since this measure achieves robustness even for arbitrary compact sets, essentially in $O(\sqrt{\epsilon})$. It is, in some sense, an integral measure of the covariance matrix of the normal cone around the point of interest. However, convergence of curvature(s) is subject to several parameters r and R which contribute contradictorily to the error. In practice, this approach gives results comparable to osculating jets for curvatures.

Recently, several authors have developed new interesting approaches for estimating the normal vector field on noisy point clouds, even in the presence of sharp features [6, 41, 59]. Furthermore, Boulch and Marlet [6] gives probabilistic convergence results. Although they cannot be used "as is" for curvature computation, they could be used in parallel with curvature estimation techniques to locate sharp features in a first pass, and to limit curvature estimations to smooth zones.

Following integral invariants approaches proposed in [48, 49], Digne and Morel [21] propose several differential estimators on point clouds. These estimators (normal vector field, curvature tensor...) also consider a spherical integration kernel and covariance matrices are constructed from point clouds or oriented point clouds (*i.e.* points equipped with normal vector). For a large set of estimators, the authors provide convergence results (as the spherical kernel radius tends to zero) in the continuous case. Our estimators on digital surfaces follow similar principles making explicit convergence speed.

Last, a very recent approach coming from geometric measure theory uses the theory of varifolds to design stable mean curvature estimations [7, 8]. This theory is generic enough to include smooth manifolds, discrete meshes, point clouds,

and digital data into the same framework. For geometric estimation, this approach requires to have both position and normal approximation. If both are convergent (in some sense), then the regularized first variation of the varifold measure converges toward the mean curvature. The speed of convergence of this approach as well as its accuracy in practice remain to be explored.

Curvature Estimation on Digital Contours and Surfaces In digital geometry, we usually consider multigrid convergence as an essential criterion [12]. Hence, in dimension 2, parameter free convergence results have been obtained for length [11] and normal vector estimation [57]. Based either on binomial convolution principles [22, 42], or polynomial fitting [50], convergence results can also be obtained for higher order derivatives of digital curves. Algorithms are parametrized by the size of the convolution or fitting kernel support and convergence theorems hold when such support size is an increasing function of the grid resolution and some shape characteristics.

For curvature estimation along 2D curves, multigrid convergence of parameter-free estimators is still challenging, although accurate experimental results have been obtained with maximal digital circular arcs [53] and with global optimization [28]. In 3D digital space, several empirical methods exist for estimating curvatures, but none achieves multigrid convergence (e.g. see [23, 38]). In [14], we recently presented a digital estimator for mean curvature for 2D and 3D digital objects, which achieves multigrid convergence in $O(h^{\frac{1}{3}})$.

Desirable Properties for Digital Curvature Estimators Our objective is to design a curvature tensor estimator for digital data such that: (1) it is provably multigrid convergent, (2) it is accurate in practice, (3) it is computable in an exact manner, (4) it can be efficiently computed either locally or globally (evaluation at a single surface point or extraction of the curvature tensor field), (5) it is robust to further perturbations (like bad digitization around the boundary, outliers).

Contributions and Outline of the Chapter We achieve such a goal by adapting the integral invariant tool originally designed for smooth manifolds and triangulated meshes [48, 49]. We begin in Sect. 9.2 by giving some background on digitization processes with a few useful lemma, and by recalling the multigrid convergence property as well as integral invariants in the continuous setting. In order to define a sound digital version of such object, it appears clearly that digital moments (of up to 2nd order) must replace continuous moments. This is why we study digital moments in Sect. 9.3 and we establish several results related to volume and moments approximation. Then, Sect. 9.4 shows how curvature in 2D and mean curvature in 3D can be approximated with digital integral invariants. This approach relies solely on simple volume estimates. We are then ready to address in Sect. 9.5 the more involved issue of estimating principal curvatures and directions, as well as the normal vector. This is done by building a local digital covariance matrix. Section 9.6 provides a method to automatically set the radius parameter of integral invariants so as to achieve multigrid convergence. Section 9.8 presents a comprehensive evaluation of the practical accuracy of our estimators. Their

robustness to perturbation on input data is then illustrated on numerous examples. Last, such estimators are shown to be useful for feature detection on digital surfaces. Singularities (edges) on shapes are detected by examining the behavior of curvature estimators in scale-space.

9.2 Background: Multigrid Convergence and Integral Invariants

9.2.1 Digital Space and Digitizations Operators

Since we are interested in evaluating both theoretically and experimentally the behavior of several differential estimators on digital object boundaries, we first have to formalize links between Euclidean objects and digital ones with the help of a digitization process. Let us consider a family $\mathbb{X}$ of compact subsets of $\mathbb{R}^d$ whose smoothness requirements will be specified later. The digital space is defined as the set of points of $\mathbb{R}^d$ with integer coordinates, naturally denoted by $\mathbb{Z}^d$. We denote $\mathbb{G}_h(X)$ the *Gauss digitization* of X in a $d-$dimensional grid of grid step h:

$$\mathbb{G}_h(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, (h \cdot \mathbf{z}) \in X \right\} , \tag{9.1}$$

where $h \cdot \mathbf{z}$ is the uniform scaling of $\mathbf{z}$ by factor h. If $\mathbf{z} \in \mathbb{Z}^d$, then $Q\mathbf{z}$ denotes the (closed) unit d-dimensional cube of $\mathbb{R}^d$ centered on $\mathbf{z}$ and aligned with the axes of $\mathbb{Z}^d$. We further define $Q_h(\mathbf{z}) := h \cdot Q\mathbf{z}$, so-called *h-cube*, as the scaled version of $Q\mathbf{z}$ (*i.e.* $Q_h(\mathbf{z})$ is a d-dimensional cube centered at $h \cdot \mathbf{z}$ with edge length h). In addition to the Gauss digitization operator, we consider the *inner Jordan digitization* $\mathbb{J}^-{}_h(X)$ and *outer Jordan digitization* $\mathbb{J}^+{}_h(X)$ at step h of a shape $X \in \mathbb{X}$ as follows (see Fig. 9.1):

$$\mathbb{J}^-{}_h(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, Q_h(\mathbf{z}) \subset X \right\} , \tag{9.2}$$

$$\mathbb{J}^+{}_h(X) := \left\{ \mathbf{z} \in \mathbb{Z}^d, Q_h(\mathbf{z}) \cap X \neq \emptyset \right\} . \tag{9.3}$$

Given a digital set $Z \subset \mathbb{Z}^d$, the *body* of Z is the embedding of Z into $\mathbb{R}^d$, denoted by $[\cdot]_h$ and defined as:

$$[Z]_h := \bigcup_{\mathbf{z} \in Z} Q_h(\mathbf{z}) . \tag{9.4}$$

Let us now formalize relationships between Gauss, Jordan digitizations and the Euclidean shape X. We call *Jordan strip* the digitization $\mathbb{J}^0{}_h(X) := \mathbb{J}^+{}_h(X) \setminus \mathbb{J}^-{}_h(X)$. Its body is clearly a union of h-cubes, with a thickness of at least one h-cube. First, it is straightforward to check that:

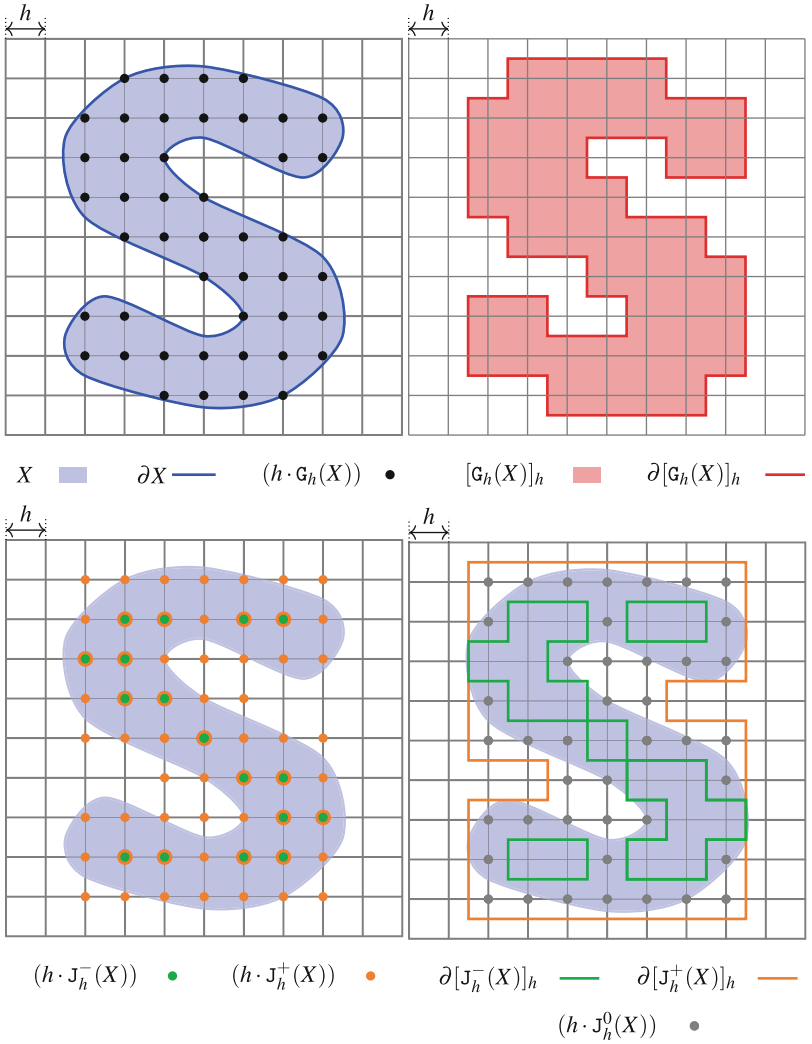


Fig. 9.1 Illustration of the digitization models and notations

Lemma 1 $\mathbb{J}^-{}_h(X) \subset \mathbb{G}_h(X) \subset \mathbb{J}^+{}_h(X)$ *and* $[\mathbb{J}^-{}_h(X)]_h \subset X \subset [\mathbb{J}^+{}_h(X)]_h$.

From these objects, it is natural to consider the relations of ∂X, the topological boundary of X, with the digitized boundaries $\partial[\mathbb{G}_h(X)]_h$, $\partial[\mathbb{J}^-{}_h(X)]_h$ or $\partial[\mathbb{J}^+{}_h(X)]_h$. We have:

Lemma 2 $[\mathbb{J}^0{}_h(X)]_h = [\mathbb{J}^+{}_h(X)]_h \setminus \mathrm{Int}([\mathbb{J}^-{}_h(X)]_h)$ *and* $\partial X \subset [\mathbb{J}^0{}_h(X)]_h$.

Proof The first equality is straightforward. For the inclusion $\partial X \subset [\jmath^0{}_h(X)]_h$, on the one hand we have by Lemma 1 that $[\jmath^-{}_h(X)]_h \subset X$, hence $\mathrm{Int}([\jmath^-{}_h(X)]_h) \subset \mathrm{Int}(X)$, so $\partial X \cap \mathrm{Int}([\jmath^-{}_h(X)]_h) = \emptyset$. On the other hand we have by the same lemma that $X \subset [\jmath^+{}_h(X)]_h$. By compactness of X, we have $\partial X \subset X$ and thus $\partial X \subset [\jmath^+{}_h(X)]_h$. Putting these two facts together concludes. □

The ϵ-*offset* of a shape X is the set of points at distance lower or equal to ϵ from X. It is denoted by X^ϵ. Furthermore, the *medial axis* $\mathrm{MA}(\partial X)$ of ∂X is the subset of $\mathbb{R}^d$ whose points have more than one closest point to ∂X. The *reach* reach(X) of X is the infimum of the distance between ∂X and its medial axis. Shapes with positive reach have principal curvatures bounded by $\pm 1/\mathrm{reach}(X)$. They have a C^2 smooth boundary almost everywhere. The boundary of the Gauss digitization of X is close to the surface ∂X for smooth shapes:

Theorem 1 ([36]) *Let X be a compact domain of $\mathbb{R}^d$ such that the reach of ∂X is greater than ρ. Then, for any digitization step $0 < h < \frac{2\rho}{\sqrt{d}}$, the Hausdorff distance between sets ∂X and $\partial[\mathbb{G}_h(X)]_h$ is less than $\sqrt{d}h/2$. Hence*

$$\partial[\mathbb{G}_h(X)]_h \subset (\partial X)^{\frac{\sqrt{d}}{2}h}. \tag{9.5}$$

This is also true for Jordan digitizations, with lesser hypotheses on X, but with a larger bound:

Lemma 3 *Let X be a compact domain of $\mathbb{R}^d$. Jordan digitizations of X are close to the boundary of X in the Hausdorff sense:*

$$\partial[\jmath^-{}_h(X)]_h \subset (\partial X)^{\sqrt{d}h}, \tag{9.6}$$

$$\partial[\jmath^+{}_h(X)]_h \subset (\partial X)^{\sqrt{d}h}, \tag{9.7}$$

$$[\jmath^0{}_h(X)]_h \subset (\partial X)^{\sqrt{d}h}. \tag{9.8}$$

Proof First of all, one can check that $\partial[\jmath^0{}_h(X)]_h = \partial[\jmath^-{}_h(X)]_h \cup \partial[\jmath^+{}_h(X)]_h$. It follows that proving the last inclusion implies the first two ones.

Now, let $\mathbf{x} \in [\jmath^0{}_h(X)]_h$. There exists some h-cube $Q_h(\mathbf{z})$ that contains $\mathbf{x}$, such that (i) $Q_h(\mathbf{z}) \cap X \neq \emptyset$ since $\mathbf{z} \in \jmath^+{}_h(X)$, and (ii) $Q_h(\mathbf{z}) \not\subset X$ since $\mathbf{z} \notin \jmath^-{}_h(X)$. From (i), there exists $\mathbf{y}_1 \in Q_h(\mathbf{z}) \cap X$. From (ii), there exists $\mathbf{y}_2 \in Q_h(\mathbf{z})$ and $\mathbf{y}_2 \notin X$. The straight segment $[\mathbf{y}_1\mathbf{y}_2]$ joining $\mathbf{y}_1$ to $\mathbf{y}_2$ is an arc that goes from X to the complementary of X. By compactness of X, there exists some $\mathbf{y}_3 \in [\mathbf{y}_1\mathbf{y}_2]$ with $\mathbf{y}_3 \in \partial X$. By convexity of $Q_h(\mathbf{z})$, $\mathbf{y}_1 \in Q_h(\mathbf{z})$ and $\mathbf{y}_2 \in Q_h(\mathbf{z})$ implies $[\mathbf{y}_1\mathbf{y}_2] \subset Q_h(\mathbf{z})$. It follows that $\mathbf{y}_3$ belongs also to $Q_h(\mathbf{z})$. We have just found a point in the same h-cube as $\mathbf{x}$, which lies in ∂X. Since two points in some h-cube may not be further away than $\sqrt{d}h$, $\mathbf{x}$ is no further away than this distance from ∂X. □

The following result will also be useful.

Lemma 4 (Proof of Lemma 10 [36]) *Let X be a compact domain of $\mathbb{R}^d$ such that the reach of ∂X is greater than ρ, and let ϵ be some value smaller or equal to ρ. Then*

$$\mathrm{Vol}(\partial X^{\epsilon}) \leq 2^{d+1}\mathrm{Area}(\partial X)\epsilon. \tag{9.9}$$

9.2.2 Multigrid Convergence of Global and Local Geometric Estimators

As discussed in various previous works (see for instance [12] for a survey), the idea of multigrid convergence is that when we define a quantity estimator on the digitization of some shape $X \subset \mathbb{R}^d$, we check if the estimated quantity converges (theoretically and/or experimentally) to the associated one on X when h tends to zero. More formally,

Definition 1 (Multigrid Convergence for Local Geometric Quantities) Given a digitization process $\mathtt{D}$, a local discrete geometric estimator $\hat{E}$ of some geometric quantity E is *multigrid convergent* for the family of shapes $\mathbb{X}$ if and only if, for any $X \in \mathbb{X}$, there exists a grid step $h_X > 0$ such that the estimate $\hat{E}(\mathtt{D}_h(X), \hat{\mathbf{x}}, h)$ is defined for all $\hat{\mathbf{x}} \in \partial[\mathtt{D}_h(X)]_h$ with $0 < h < h_X$, and for any $\mathbf{x} \in \partial X$,

$$\forall \hat{\mathbf{x}} \in \partial[\mathtt{D}_h(X)]_h \text{ with } \|\hat{\mathbf{x}}-\mathbf{x}\|_\infty \leq h, |\hat{E}(\mathtt{D}_h(X), \hat{\mathbf{x}}, h)-E(X,\mathbf{x})| \leq \tau_{X,\mathbf{x}}(h), \tag{9.10}$$

where $\tau_{X,\mathbf{x}} : \mathbb{R}^+ \setminus \{0\} \to \mathbb{R}^+$ has null limit at 0. This function defines the speed of convergence of $\hat{E}$ toward E at point $\mathbf{x}$ of X. The convergence is *uniform* for X when every $\tau_{X,\mathbf{x}}$ is bounded from above by a function τ_X independent of $\mathbf{x} \in \partial X$ with null limit at 0.

Note that when a geometrical quantity is global (*e.g.* area or volume), we do not need an explicit mapping between ∂X and $\partial[\mathtt{D}_h(X)]_h$, and Definition 1 can be rephrased to define the simpler *multigrid convergence of global geometric quantities* [12].

Instead of estimating the geometric quantity for all $\hat{\mathbf{x}} \in \partial[\mathtt{G}_h(X)]_h$, classical local discrete estimators estimate the quantity at cells of the cellular boundary of a digital set, otherwise said at elements of the interpixel representation of the digital set boundary (pointels, linels or surfels). We usually consider a canonical Euclidean embedding of $k-$cells into $\mathbb{R}^d$ (1-cells are mapped into unitary Euclidean segments, 2-cells into unit squares…), scaled by the factor h. Furthermore the estimated quantity $\hat{E}(\mathtt{D}_h(X), \hat{\mathbf{x}}, h)$ is constant for all $\hat{\mathbf{x}}$ belonging to the embedding of a boundary $k-$cell.

9.2.3 Integral Invariants in the Continuous Setting

In geometry processing, integral invariants have been widely investigated to define estimators of differential quantities (see [48, 49] for a complete overview). For short, the main idea is to move a kernel on points $\mathbf{x} \in \partial X$ and to compute integrals on the intersection between X and the kernel. Even if different kernels (e.g., Euclidean ball, Euclidean sphere) and different integration functions can be considered, we focus here on volumetric integral invariants defined as follows:

Definition 2 Given $X \in \mathbb{X}$ and a radius $R \in \mathbb{R}^{+*}$, the volumetric integral $V_R(\mathbf{x})$ at $\mathbf{x} \in \partial X$ is given by (see Fig. 9.3)

$$V_R(\mathbf{x}) := \int_{B_R(\mathbf{x})} \chi(\mathbf{p}) d\mathbf{p}\,, \tag{9.11}$$

where $B_R(\mathbf{x})$ is the Euclidean ball with radius R and center $\mathbf{x}$ and $\chi(\cdot)$ the characteristic function of X. In dimension 2, we simply denote $A_R(\mathbf{x})$ such quantity. Several authors have detailed connections between $V_R(\mathbf{x})$ and curvature (resp. mean curvature) at $\mathbf{x}$ for shapes in $\mathbb{R}^2$ (resp. $\mathbb{R}^3$) [9, 48, 49].

Lemma 5 ([49]) *For a sufficiently smooth shape X in $\mathbb{R}^2$, $\mathbf{x} \in \partial X$, we have*

$$A_R(\mathbf{x}) = \frac{\pi}{2} R^2 - \frac{\kappa(X, \mathbf{x})}{3} R^3 + O(R^4)\,, \tag{9.12}$$

where $\kappa(X, \mathbf{x})$ is the curvature of ∂X at $\mathbf{x}$. For a sufficiently smooth shape X in $\mathbb{R}^3$ and $\mathbf{x} \in \partial X$, we have

$$V_R(\mathbf{x}) = \frac{2\pi}{3} R^3 - \frac{\pi H(X, \mathbf{x})}{4} R^4 + O(R^5)\,, \tag{9.13}$$

where $H(X, \mathbf{x})$ is the mean curvature of ∂X at $\mathbf{x}$.

Such results are obtained by Taylor expansion at $\mathbf{x}$ of the surface ∂X approximated by a parametric function $y = f(x)$ in 2D and $z = f(x, y)$ in 3D. From Eqs. (9.12) and (9.13) and with a fixed radius R, one can derive local estimators $\tilde{\kappa}^R$ and $\tilde{H}^R$ respectively:

$$\tilde{\kappa}^R(X, \mathbf{x}) := \frac{3\pi}{2R} - \frac{3A_R(\mathbf{x})}{R^3}, \quad \tilde{H}^R(X, \mathbf{x}) := \frac{8}{3R} - \frac{4V_R(\mathbf{x})}{\pi R^4}\,. \tag{9.14}$$

In this way, as R tends to zero, both estimated values converge to expected ones (respectively κ and H). More formally:

$$\tilde{\kappa}^R(X, \mathbf{x}) = \kappa(X, \mathbf{x}) + O(R), \quad \tilde{H}^R(X, \mathbf{x}) = H(X, \mathbf{x}) + O(R)\,. \tag{9.15}$$

Similarly, directional information such as principal curvatures and thus Gaussian curvature can be retrieved from integral computations. Indeed, instead of computing the measure of $B_R(\mathbf{x}) \cap X$ as in Definition 2, we consider its covariance matrix. Given a non-empty subset $Y \subset \mathbb{R}^d$, the *covariance matrix* of Y is given by

$$\mathscr{V}(Y) := \int_Y (\mathbf{p} - \overline{Y})(\mathbf{p} - \overline{Y})^T d\mathbf{p} = \int_Y \mathbf{p}\mathbf{p}^T d\mathbf{p} - \mathrm{Vol}(Y)\overline{Y}\overline{Y}^T, \tag{9.16}$$

where $\overline{Y}$ is the *centroid* of Y and $\mathrm{Vol}(Y)$ its volume. For non negative integers p, q and s, we recall the definition of *pqs-moments* $m^{pqs}(Y)$ of Y:

$$m^{pqs}(Y) := \iiint_Y x^p y^q z^s dxdydz\,. \tag{9.17}$$

Note that the volume $\mathrm{Vol}(Y)$ is the 0-moment $m^{000}(Y)$, and that the centroid $\overline{Y}$ is the vector of 1-moments normalized by the 0-moment, i.e. $(m^{100}(Y), m^{010}(Y), m^{001}(Y))^T / m^{000}(Y)$. For simplicity, let us denote by A the Euclidean set $B_R(\mathbf{x}) \cap X$. The covariance matrix of A is then rewritten as[1]:

$$\begin{aligned}\mathscr{V}(A) = &\begin{bmatrix} m^{200}(A) & m^{110}(A) & m^{101}(A) \\ m^{110}(A) & m^{020}(A) & m^{011}(A) \\ m^{101}(A) & m^{011}(A) & m^{002}(A) \end{bmatrix} \\ &- \frac{1}{m^{000}(A)} \begin{bmatrix} m^{100}(A) \\ m^{010}(A) \\ m^{001}(A) \end{bmatrix} \otimes \begin{bmatrix} m^{100}(A) \\ m^{010}(A) \\ m^{001}(A) \end{bmatrix}^T . \end{aligned} \tag{9.18}$$

In [48], authors have demonstrated that eigenvalues and eigenvectors of $\mathscr{V}(A)$ provide principal curvature and principal direction information:

Lemma 6 ([48], Theorem 2) *Given a shape $X \in \mathbb{X}$, the eigenvalues λ_1, λ_2, λ_3 of $\mathscr{V}(A)$, where $A := B_R(\mathbf{x}) \cap X$ and $\mathbf{x} \in \partial X$, $\lambda_1 \geq \lambda_2 \geq \lambda_3$, have the following Taylor expansion:*

$$\lambda_1 = \frac{2\pi}{15} R^5 - \frac{\pi}{48}(3\kappa_1(X,\mathbf{x}) + \kappa_2(X,\mathbf{x}))R^6 + O(R^7)\,, \tag{9.19}$$

$$\lambda_2 = \frac{2\pi}{15} R^5 - \frac{\pi}{48}(\kappa_1(X,\mathbf{x}) + 3\kappa_2(X,\mathbf{x}))R^6 + O(R^7)\,, \tag{9.20}$$

$$\lambda_3 = \frac{19\pi}{480} R^5 - \frac{9\pi}{512}(\kappa_1(X,\mathbf{x}) + \kappa_2(X,\mathbf{x}))R^6 + O(R^7)\,, \tag{9.21}$$

where $\kappa_1(X,\mathbf{x})$ and $\kappa_2(X,\mathbf{x})$ denote the principal curvatures of ∂X at $\mathbf{x}$.[2]

[1] $\otimes$ denotes the usual tensor product in vector spaces.

[2] There is a typographic error in the λ_1 expression in [48].

Hence, similarly to Eq. (9.14), one can define local estimators $\tilde{\kappa}_1^R$, $\tilde{\kappa}_2^R$ and finally the Gaussian curvature $\tilde{K}_R := \tilde{\kappa}_1^R \cdot \tilde{\kappa}_2^R$ as functions of $\{\lambda_i\}_{1,2,3}$ and R. From Lemma 6, all these estimators approach expected quantities when R tends to 0.

When dealing with digital shapes $\mathrm{D}_h(X)$, implementation of these estimators becomes straightforward: choose a radius R, center a Euclidean (or digital) ball at chosen points of $\partial[\mathrm{D}_h(X)]_h$ (e.g. centroids of linels or surfels), compute the quantities (area, volume, covariance matrix) and finally estimate curvature information $\tilde{\kappa}$, $\tilde{H}$, $\tilde{\kappa}_1$, $\tilde{\kappa}_2$ or $\tilde{K}$.

However, several issues are hidden in this approach: What are meaningful values for R according to the shape size and geometry? Do points of $\partial[\mathrm{D}_h(X)]_h$ converge to points $\mathbf{x} \in \partial X$ for which Lemmas 5 and 6 are valid? Does counting the number of pixels (resp. voxels) converge to $A_R(\mathbf{x})$ (resp. $V_R(\mathbf{x})$)? Does the digital covariance matrix converges to the expected one? The rest of the chapter addresses all these questions.

9.3 Digital Moments

Integral invariants rely on the precise estimation of the volume and covariance matrix of specific Euclidean subsets. These quantities can be expressed as functions of zeroth, first and second order moments of Euclidean subsets. It is thus of critical importance to estimate properly moments in the digital world in order to use integral invariants for approximating curvatures of digital shapes.

Since the approximation of moments is directly related to area and volume estimation, but also to integral approximation, there exists a vast literature on this topic. It is known since Gauss and Dirichlet that counting the number of integer points within a convex set provides an order one approximation of its total area (in 2D) / volume (in dD). In fact, much better bounds are achievable for 2D shapes if the boundary is strictly C^3-convex [26]. This holds also in higher dimensions [25, 33, 46]. The estimation of moments of digital sets was tackled in a series of papers of Klette and Žunić [30–32]. To sum up their results, they give error upper bounds that are similar to Huxley's bound for arbitrary moments of 2D shapes.

However, we will not use these bounds for several reasons: (i) they are valid for (strictly) smooth convex shapes while integral invariants may involve non smooth and non convex shapes, (ii) the bounds are given as big "O" and some shape geometry is hidden in the constant. We prefer possibly weaker bounds but we want them to be explicit and valid for more general shapes.

9.3.1 Moments and Digital Moments

Let X be some compact domain of $\mathbb{R}^d$. The *$p_1 \cdots p_d$-moment of X* is defined as

$$m^{p_1 \cdots p_d}(X) := \int \cdots \int_X {x_1}^{p_1} \cdots {x_d}^{p_d} dx_1 \ldots dx_d. \tag{9.22}$$

The $0\cdots0$-moment of X is the volume of X (denoted $\mathrm{Vol}(X)$). For any subset Z of $\mathbb{Z}^d$, the $p_1\cdots p_d$-*digital moment of* Z *at step* h is defined as

$$\hat{m}_h^{p_1\cdots p_d}(Z) := h^{d+p_1+\cdots+p_d} \sum_{(z_1,\ldots,z_d)\in Z} {z_1}^{p_1}\cdots {z_d}^{p_d}. \tag{9.23}$$

The $0\cdots0$-digital moment of Z is the digital volume of Z (denoted by $\widehat{\mathrm{Area}}(Z,h)$ when $d=2$ and by $\widehat{\mathrm{Vol}}(Z,h)$ when $d\geq 3$).

In the sequel, points and vectors of $\mathbb{R}^d$ and $\mathbb{Z}^d$ are written in bold, and for some point or vector $\mathbf{z}\in\mathbb{Z}^d$, its coordinates or components are written with subscripts as $\mathbf{z}=(z_1,\ldots,z_d)$. We wish to bound the error between moments of X and digital moments of the digitization of X as some function of the digitization gridstep h. We thus give a particular attention to moments and digital moments of h-cubes. The following equalities are easily obtained by simple integration.

Lemma 7 *Let* $\mathbf{z}\in\mathbb{Z}^d$. *Point* $\mathbf{z}$ *is the Gauss digitization of* h*-cube* $Q_h(\mathbf{z})$*, but also its inner or outer Jordan digitization. First orders moments and digital moments of* h*-cubes follow*

$$m^{0\cdots0}(Q_h(\mathbf{z})) = h^d \qquad \hat{m}_h^{0\cdots0}(\{\mathbf{z}\}) = h^d \tag{9.24}$$

$$m^{10\cdots0}(Q_h(\mathbf{z})) = h^{d+1}z_1 \qquad \hat{m}_h^{10\cdots0}(\{\mathbf{z}\}) = h^{d+1}z_1 \tag{9.25}$$

$$m^{110\cdots0}(Q_h(\mathbf{z})) = h^{d+2}z_1z_2 \qquad \hat{m}_h^{110\cdots0}(\{\mathbf{z}\}) = h^{d+2}z_1z_2 \tag{9.26}$$

$$m^{20\cdots0}(Q_h(\mathbf{z})) = h^{d+2}\left(z_1^2+\frac{h^2}{12}\right) \qquad \hat{m}_h^{20\cdots0}(\{\mathbf{z}\}) = h^{d+2}z_1^2 \tag{9.27}$$

Discrepancies between digital and continuous moments appear for moments $p_1\cdots p_d$, when one of the p_i is greater or equal to 2.

9.3.2 General Results for Volume Estimation Errors

The theorem below shows that the error between the volume of a shape X and the naive volume estimation on its digitization by simple enumeration is smaller than the volume of the offset of ∂X with distance $\sqrt{d}h$.

Theorem 2 *Let* X *be a compact domain of* $\mathbb{R}^d$. *Let* $\mathtt{D}$ *be any digitization process such that* $\mathtt{J}^-{}_h(X)\subset\mathtt{D}_h(X)\subset\mathtt{J}^+{}_h(X)$. *Digital and continuous volumes are related as follows:*

$$\left|\mathrm{Vol}(X)-\widehat{\mathrm{Vol}}(\mathtt{D}_h(X),h)\right| \leq \mathrm{Vol}(\partial X^{\sqrt{d}h}). \tag{9.28}$$

Proof First of all, 0-order continuous and digital moments of h-cubes coincide, so we have the following equality:

$$\begin{aligned}\widehat{\mathrm{Vol}}(\mathbb{D}_h(X),h) &= \sum_{\mathbf{z}\in\mathbb{D}_h(X)} \hat{m}_h^{0\cdots0}(\{\mathbf{z}\}) \\ &= \sum_{\mathbf{z}\in\mathbb{D}_h(X)} m^{0\cdots0}(Q_h(\mathbf{z})) \qquad \text{(Lemma 7)} \\ &= \mathrm{Vol}([\mathbb{D}_h(X)]_h)\,. \end{aligned} \tag{9.29}$$

Then, by denoting $A\Delta B$ the symmetric difference between sets A and B, we bound the volume difference as

$$|\mathrm{Vol}(X) - \mathrm{Vol}([\mathbb{D}_h(X)]_h)| \le \mathrm{Vol}(X\Delta[\mathbb{D}_h(X)]_h)\,. \tag{9.30}$$

Indeed, given two sets A and B, we have

$$\begin{aligned}|\,\mathrm{Vol}(A) - \mathrm{Vol}(B)| &= |\,\mathrm{Vol}(A\setminus B) - \mathrm{Vol}(B\setminus A)| &(9.31)\\ &\le |\,\mathrm{Vol}(A\setminus B) + \mathrm{Vol}(B\setminus A)| &(9.32)\\ &\le \mathrm{Vol}(A\Delta B)\,. &(9.33)\end{aligned}$$

Now, for any sets A, B, Y_1, Y_2 with $A\subset Y_1\subset B$ and $A\subset Y_2\subset B$, we have $Y_1\Delta Y_2\subset B\setminus A$. This follows from $Y_1\Delta Y_2 = (Y_1\cup Y_2)\setminus(Y_1\cap Y_2)$. Then, obviously $(Y_1\cup Y_2)\subset B$ and $A\subset(Y_1\cap Y_2)$.

Now, by Lemma 1, we have that $[\mathbb{J}^-{}_h(X)]_h \subset X \subset [\mathbb{J}^+{}_h(X)]_h$. But by hypothesis, $\mathbb{J}^-{}_h(X)\subset\mathbb{D}_h(X)\subset\mathbb{J}^+{}_h(X)$, so we also have $[\mathbb{J}^-{}_h(X)]_h\subset[\mathbb{D}_h(X)]_h\subset[\mathbb{J}^+{}_h(X)]_h$. We may thus apply the preceding property setting $A{:=}[\mathbb{J}^-{}_h(X)]_h$, $B{:=}[\mathbb{J}^+{}_h(X)]_h$, $Y_1{:=}X$ and $Y_2{:=}[\mathbb{D}_h(X)]_h$. We get

$$X\Delta[\mathbb{D}_h(X)]_h \subset [\mathbb{J}^+{}_h(X)]_h \setminus [\mathbb{J}^-{}_h(X)]_h\,. \tag{9.34}$$

Putting preceding relations together gives

$$\begin{aligned}\left|\mathrm{Vol}(X) - \widehat{\mathrm{Vol}}(\mathbb{D}_h(X),h)\right| &\le \mathrm{Vol}([\mathbb{J}^+{}_h(X)]_h \setminus [\mathbb{J}^-{}_h(X)]_h) \quad \text{(using Eq.(9.29), (9.30) and (9.34))}\\ &\le \mathrm{Vol}([\mathbb{J}^0{}_h(X)]_h) \quad \text{(definition of Jordan strip)}\\ &\le \mathrm{Vol}(\partial X^{\sqrt{d}h})\,. \quad \text{(Lemma 3)} \end{aligned} \tag{9.35}$$

This concludes. □

It would be tempting to extend the preceding result to arbitrary moments. However, some moments are not non-negative measures and we cannot use directly the preceding argument. Hence, we postpone results on moment estimation to a later section.

Lemma 4 in conjunction with the preceding theorem allows us to relate this error to the shape area, for smooth enough shapes.

Corollary 1 *Let X be a compact domain of $\mathbb{R}^d$ such that the reach of ∂X is greater than ρ, and h is smaller than $\frac{\rho}{\sqrt{d}}$. Let D be any digitization process such that $\mathrm{J}^-_h(X) \subset \mathrm{D}_h(X) \subset \mathrm{J}^+_h(X)$. Digital and continuous volumes are related as follows:*

$$\left|\mathrm{Vol}(X) - \widehat{\mathrm{Vol}}(\mathrm{D}_h(X), h)\right| \leq 2^{d+1}\sqrt{d}\mathrm{Area}(\partial X)h\,. \tag{9.36}$$

9.3.3 Volume Approximation Within a Ball of Radius R

Integral invariants rely on moment estimation along the boundary of a shape X within a ball $B_R(\mathbf{x})$ of given radius R, for $\mathbf{x} \in \partial X$. Most results on volume and moments estimation are valid for smooth enough shapes, which is not the case of $X \cap B_R(\mathbf{x})$. We thus establish results for intersection of smooth shapes. The following lemma is required.

Lemma 8 *Let A, B be compact domains of $\mathbb{R}^d$ and ϵ some positive number. Then* $(\partial(A \cap B))^\epsilon = ((\partial A) \cap B)^\epsilon \cup (A \cap (\partial B))^\epsilon$.

Proof Figure 9.2 illustrates this lemma. It suffices to show that $\partial(A \cap B) = ((\partial A) \cap B) \cup (A \cap (\partial B))$. For $\boxed{\subset}$, this comes from the facts that $\partial(A \cap B) \subset A \cap B$ and $\partial(A \cap B) \subset (\partial A) \cup (\partial B)$. For $\boxed{\supset}$, let $\mathbf{y}$ be a point of $(\partial A) \cap B$. If $\mathrm{Int}(A)$ is the interior of A, we have that $\partial A = A \setminus \mathrm{Int}(A)$ since A is compact. It follows that $\mathbf{y} \in (A \setminus \mathrm{Int}(A)) \cap B$, then $\mathbf{y} \in (A \cap B) \setminus (\mathrm{Int}(A) \cap B)$. But it holds that $\mathrm{Int}(A \cap B) \subset (\mathrm{Int}(A) \cap B)$. Hence $(A \cap B) \setminus \mathrm{Int}(A \cap B) \supset (A \cap B) \setminus (\mathrm{Int}(A) \cap B)$. Noticing that $(A \cap B) \setminus \mathrm{Int}(A \cap B) = \partial(A \cap B)$, we conclude that $\mathbf{y} \in \partial(A \cap B)$. The case $\mathbf{y} \in A \cap \partial B$ is similar. □

We prove below that the volume of the $\sqrt{d}h$-offset of the boundary of a compact domain X intersected with a ball a radius R can be upper bounded by a constant

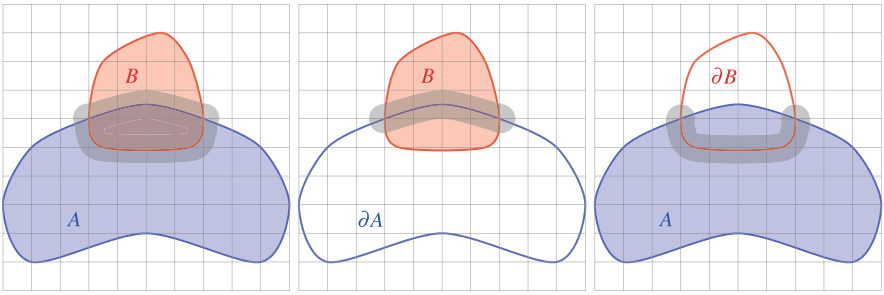


Fig. 9.2 Illustration of Lemma 8

times h. Furthermore, this bound does not depend on the geometry of X, as long as the radius R is smaller than half the reach of X.

Theorem 3 *Let X be a compact domain of $\mathbb{R}^d$ such that the reach of ∂X is greater than ρ. Let $\mathbf{x} \in \mathbb{R}^d$. Let R (the radius of the ball) and h (the gridstep) be some positive numbers such that $h \leq \frac{R}{\sqrt{2d}}$ and $2R \leq \rho$.*

$$\mathrm{Vol}\left((\partial(X \cap B_R(\mathbf{x})))^{\sqrt{d}h}\right) \leq K_1(d) R^{d-1} h\,, \tag{9.37}$$

with $K_1(d) := 2\sqrt{d}\left(d + \left(1 + \frac{1}{\sqrt{2}}\right)^d\right) V_d + \frac{2\sqrt{6}}{3}\left(\frac{4+\sqrt{2}}{2}\right)^d V_{d-1}$. As a corollary, this volume is upper bounded by $68Rh$ when $d = 2$, and upper bounded by $154R^2h$ for $d = 3$.

Proof Lemma 8 shows that

$$\mathrm{Vol}\left((\partial(X \cap B_R(\mathbf{x})))^{\sqrt{d}h}\right) \leq \mathrm{Vol}\left((\partial X \cap B_R(\mathbf{x}))^{\sqrt{d}h}\right) + \mathrm{Vol}\left((X \cap \partial B_R(\mathbf{x}))^{\sqrt{d}h}\right). \tag{9.38}$$

The rightmost term is not hard to bound above. Letting V_d be the volume of the d-dimensional ball of radius 1, we proceed as follows:

$$\begin{aligned}
&\mathrm{Vol}\left((X \cap \partial B_R(\mathbf{x}))^{\sqrt{d}h}\right) \\
&\leq \mathrm{Vol}\left((\partial B_R(\mathbf{x}))^{\sqrt{d}h}\right) \\
&= \mathrm{Vol}\left(B_{R+\sqrt{d}h}(\mathbf{x})) - \mathrm{Vol}(B_{R-\sqrt{d}h}(\mathbf{x})\right) \\
&= V_d\left[(R + \sqrt{d}h)^d - (R - \sqrt{d}h)^d\right] \\
&= 2V_d\left[\binom{d}{1} R^{d-1}\sqrt{d}h + \sum_{k=1}^{\lfloor \frac{d-1}{2} \rfloor} \binom{d}{2k+1} R^{d-(2k+1)} \left(\sqrt{d}h\right)^{2k+1}\right] \\
&\leq 2V_d\left[d\sqrt{d}R^{d-1}h + R^{d-1}\sqrt{d}h \sum_{k=1}^{\lfloor \frac{d-1}{2} \rfloor} \binom{d}{2k+1} \frac{1}{\sqrt{2}^{d-(2k+1)}}\right] \text{ (since } \sqrt{d}h \leq R/\sqrt{2}) \\
&\leq 2V_d\left[d\sqrt{d}R^{d-1}h + \left(1 + \frac{1}{\sqrt{2}}\right)^d \sqrt{d}R^{d-1}h\right] \quad \left(\text{since } \sum_{i=0}^{d} \binom{d}{i} x^{d-i} = (1+x)^d\right) \\
&\leq 2V_d\left(d + \left(1 + \frac{1}{\sqrt{2}}\right)^d\right)\sqrt{d}R^{d-1}h\,.
\end{aligned} \tag{9.39}$$

The other term is harder to bound. The idea is to define a kind of cylinder of thickness $2\sqrt{d}h$ that includes the set $(\partial X \cap B_R(\mathbf{x}))^{\sqrt{d}h}$, and then we bound its volume.

We define the set $P{:=}\partial X \cap B_{R+\sqrt{2dh}}(\mathbf{x})$. It is a $d-1$-manifold which is homeomorphic to a $d-1$-disk since $R+\sqrt{2dh} \leq \rho$. We choose thus a local parameterization $\mathbf{z} : \mathscr{U} \to P$ sending any $\mathbf{u} \in \mathscr{U}$ into P, where $\mathscr{U} \subset \mathbb{R}^{d-1}$. Let $\mathbf{n}$ be the function that associates to each point of ∂X its outward unit normal vector. Let also $\mathbf{N}$ be the function which assigns to each $\mathbf{u} \in \mathscr{U}$ the vector $\mathbf{N}(\mathbf{u}) = \mathbf{n}(\mathbf{z}(\mathbf{u}))$, i.e. the unit vector normal to ∂X at $\mathbf{z}(\mathbf{u})$. By definition of the reach, the map $\mathbf{Z}^t : \mathscr{U} \to \mathbb{R}^d$ such that $\mathbf{Z}^t(\mathbf{u}) = \mathbf{z}(\mathbf{u}) + t\mathbf{N}(\mathbf{u})$ is injective for $-\rho \leq t \leq \rho$. Let S denote the shape operator of P, defined for surfaces by $\mathrm{S}\mathbf{v}{:=} -\nabla_{\mathbf{v}}\mathbf{N}$ for any tangent vector $\mathbf{v}$ of P. It is straightforward to check that $\mathrm{S}\frac{\partial \mathbf{z}}{\partial u_i} = -\frac{\partial \mathbf{N}}{\partial u_i}$.

Looking now at partial derivatives of $\mathbf{Z}^t$, it follows that $\frac{\partial \mathbf{Z}^t}{\partial u_i} = (\mathrm{Id} - t\mathrm{S})\frac{\partial \mathbf{z}}{\partial u_i}$. The deformation of an area element of $\mathbf{z}$ onto $\mathbf{Z}^t$ is thus given by $\det(\mathrm{Id} - t\mathrm{S})$ and we may write:

$$\begin{aligned}\mathrm{Area}(\mathbf{Z}^t(\mathscr{U})) &= \int \cdots \int_{\mathscr{U}} \left\| \frac{\partial \mathbf{Z}^t}{\partial u_1} \wedge \cdots \wedge \frac{\partial \mathbf{Z}^t}{\partial u_{d-1}} \right\| du_1 \cdots du_{d-1}, \\ &= \int \cdots \int_{\mathscr{U}} \det(\mathrm{Id} - t\mathrm{S}) \left\| \frac{\partial \mathbf{z}}{\partial u_1} \wedge \cdots \wedge \frac{\partial \mathbf{z}}{\partial u_{d-1}} \right\| du_1 \cdots du_{d-1}.\end{aligned} \tag{9.40}$$

Now, since ∂X has positive reach, function $\mathbf{N}$ is differentiable almost everywhere (a.e.). Furthermore, the normal vector variation is a.e. bounded by $1/\rho$. It follows that $\det(\mathrm{Id} - t\mathrm{S}) \leq (1 + |t|/\rho)^{d-1}$ almost everywhere. Injecting this inequality into Eq.(9.40) gives

$$\mathrm{Area}(\mathbf{Z}^t(\mathscr{U})) \leq (1 + |t|/\rho)^{d-1}\mathrm{Area}(P). \tag{9.41}$$

For any $x \in \mathbb{R}$, $0 < x \leq \rho$, the cylindric shape $C(x){:=}\{\mathbf{Z}^t(\mathbf{u}), \mathbf{u} \in \mathscr{U}, t \in \mathbb{R}, -x \leq t \leq x\}$ has a volume that is bounded as follows:

$$\begin{aligned}\mathrm{Vol}(C(x)) &= \int_{-x}^{x} \mathrm{Area}(\mathbf{Z}^t(\mathscr{U}))dt \\ &\leq 2\mathrm{Area}(P) \int_0^x (1 + t/\rho)^{d-1} dt \\ &= 2\mathrm{Area}(P)\frac{\rho}{d}\left((1 + x/\rho)^d - 1\right).\end{aligned} \tag{9.42}$$

We look more precisely at the volume of $C(\sqrt{dh})$. After some simple computations, we get:

$$\begin{aligned}\mathrm{Vol}(C(\sqrt{dh})) &\leq 2\mathrm{Area}(P)\frac{\rho}{d}\frac{\sqrt{dh}}{\rho}\left(\sum_{i=1}^{d}\binom{d}{i}\left(\frac{\sqrt{dh}}{\rho}\right)^{i-1}\right), \\ &\leq 4\left(\frac{4+\sqrt{2}}{4}\right)^d \frac{\sqrt{d}}{d}\mathrm{Area}(P)h,\end{aligned} \tag{9.43}$$

since $\sqrt{d}h \leq \frac{R}{\sqrt{2}} \leq \frac{\sqrt{2}\rho}{4}$ and using the binomial expansion of $(1+\frac{\sqrt{2}}{4})^d$.

It remains to show that (i) $(\partial X \cap B_R(\mathbf{x}))^{\sqrt{d}h} \subset C(\sqrt{d}h)$ and (ii) to estimate Area(P).

(i) Let $\mathbf{y} \in (\partial X \cap B_R(\mathbf{x}))^{\sqrt{d}h}$. There is some $\mathbf{y}' \in \partial X \cap B_R(\mathbf{x})$ with $\|\mathbf{y}-\mathbf{y}'\| \leq \sqrt{d}h$. As $\mathbf{y}$ is within the reach of ∂X, since $\mathfrak{d}(y, \partial X) \leq \sqrt{d}h \leq \rho/2$, there is only one closest point $\mathbf{y}''$ to $\mathbf{y}$ on ∂X. Let t be the distance $\|\mathbf{y}-\mathbf{y}''\|$. It follows that $t \leq \|\mathbf{y}-\mathbf{y}'\| \leq \sqrt{d}h$ since $\mathbf{y}'$ belongs also to ∂X. And we may write $\mathbf{y} = \mathbf{y}'' + t\mathbf{n}(\mathbf{y}'')$.

If $\mathbf{y}' = \mathbf{y}''$, then $\mathbf{y}'' \in P$ and $\mathbf{y}$ belongs to $C(\sqrt{d}h)$ since $t \leq \sqrt{d}h$.

Otherwise, we prove that $\|\mathbf{y}'-\mathbf{y}''\| \leq \sqrt{2d}h$. Without loss of generality, assume $\mathbf{y}$ is outside of X. Since $\mathbf{y}'' \in \partial X$ and X has reach greater than ρ, there is an outside osculating ball of radius ρ at $\mathbf{y}''$ [36], which contains no point of X except $\mathbf{y}''$. It contains of course $\mathbf{y}$. We denote $\mathbf{p}$ the orthogonal projection of $\mathbf{y}'$ onto the straight line passing through $\mathbf{y}''$ and pointing along the normal direction to ∂X. This straight line goes through $\mathbf{y}$ and through the center $\mathbf{c}$ of this osculating ball. We may also write $\mathbf{p} = \mathbf{y}'' + p\mathbf{n}(\mathbf{y}'')$, where p is the signed distance between $\mathbf{y}''$ and P along the normal direction. We use now Pythagoras' theorem to get

$$\|\mathbf{y}'-\mathbf{y}''\|^2 = p^2 + \|\mathbf{y}'-\mathbf{p}\|^2, \qquad \|\mathbf{y}'-\mathbf{p}\|^2 + (t-p)^2 = \|\mathbf{y}-\mathbf{y}'\|^2,$$

which implies

$$\|\mathbf{y}'-\mathbf{y}''\|^2 = \|\mathbf{y}-\mathbf{y}'\|^2 + t(2p-t). \tag{9.44}$$

Using first the fact that $\mathbf{c} = \mathbf{y}'' + \rho\mathbf{n}(\mathbf{y}'')$ and point $\mathbf{y}'$ is outside the osculating ball of radius ρ, and second the fact that $\|\mathbf{y}-\mathbf{y}'\| \leq \sqrt{d}h$, we obtain

$$\begin{aligned} &\|\mathbf{y}'-\mathbf{p}\|^2 + (\rho-p)^2 \geq \rho^2 \ \text{ and } \ \|\mathbf{y}'-\mathbf{p}\|^2 + (t-p)^2 \leq dh^2 \\ &\qquad \Rightarrow dh^2 \geq t^2 + 2p(\rho - t) \\ &\qquad \Rightarrow p \leq \frac{dh^2 - t^2}{\rho} \qquad\qquad (\text{since } t \leq \sqrt{d}h \leq \rho/2). \end{aligned}$$

We now inject the last inequality into Eq.(9.44) to bound $\|\mathbf{y}'-\mathbf{y}''\|$:

$$\begin{aligned} \|\mathbf{y}'-\mathbf{y}''\|^2 &\leq (\|\mathbf{y}-\mathbf{y}'\|^2 - t^2) + 2\frac{t}{\rho}(dh^2 - t^2) \\ &\leq 2dh^2. \qquad\qquad (\text{since } t/\rho < 1/2 \text{ and } \|\mathbf{y}-\mathbf{y}'\| \leq \sqrt{d}h) \end{aligned}$$

It follows that $\|\mathbf{y}''-\mathbf{x}\| \leq \|\mathbf{y}''-\mathbf{y}'\| + \|\mathbf{y}'-\mathbf{x}\| \leq \sqrt{2d}h + R$, and since $t \leq \sqrt{d}h$ we conclude that $\mathbf{y} \in C(\sqrt{d}h)$.

(ii) To estimate Area(P), we go back to its parametric definition as Area($\mathbf{z}(\mathscr{U})$). Since the maximal distance of an element of P to $\mathbf{x}$ is $R + \sqrt{2d}h$ and is smaller than the reach, it follows that P projects injectively orthogonally on the tangent plane to ∂X at $\mathbf{x}$. We may thus choose $\mathscr{U}$ to be the orthogonal projection of P onto this tangent plane. It follows that we can define $\mathbf{z}$ as $\forall \mathbf{u} \in \mathscr{U}, \mathbf{z}(\mathbf{u}) := (u_1, \ldots, u_{d-1}, f(\mathbf{u}))$, where f is a height function. The area of P is then:

$$\begin{aligned} \text{Area}(P) &= \int \cdots \int_{\mathscr{U}} \left\| \frac{\partial \mathbf{z}}{\partial u_1} \wedge \cdots \wedge \frac{\partial \mathbf{z}}{\partial u_{d-1}} \right\| du_1 \cdots du_{d-1}. \\ &= \int \cdots \int_{\mathscr{U}} \sqrt{1 + \left(\frac{\partial f}{\partial u_1}\right)^2 + \cdots + \left(\frac{\partial f}{\partial u_{d-1}}\right)^2} du_1 \cdots du_{d-1}. \end{aligned} \tag{9.45}$$

Let $\Delta := R + \sqrt{2d}h$. Now, on the boundary of a shape with reach greater than ρ, the angle variation of the normal vector cannot exceed the angle variation on the sphere of radius ρ. It follows that we can bound above this angle variation by measuring the angle variation β between the pole and a point at distance Δ of the pole onto this sphere. One easily checks that $\sin\beta = \frac{\Delta}{2\rho}$. Since $\Delta = R + \sqrt{2d}h \leq \rho$, we get $\beta \leq \frac{\pi}{6}$. It follows immediately that $\left|\frac{\partial f}{\partial u_i}\right|$ cannot exceed the tangent of β, that is $\left|\frac{\partial f}{\partial u_i}\right| \leq \tan\beta \leq \frac{\sqrt{3}}{3}$. We may now bound above all partial derivatives in Eq.(9.45) to get:

$$\begin{aligned} \text{Area}(P) &\leq \int \cdots \int_{\mathscr{U}} \sqrt{1 + \frac{1}{3} + \ldots + \frac{1}{3}} du_1 \cdots du_{d-1}. \\ &\leq \sqrt{\frac{d+2}{3}} \text{Area}(\mathscr{U}) \\ &\leq \sqrt{\frac{d+2}{3}} V_{d-1}(2R)^{d-1}, \end{aligned} \tag{9.46}$$

since $\mathscr{U}$ is included in the disk of radius Δ centered on $\mathbf{x}$ and $\Delta \leq 2R$.

We are now in position to conclude the proof.

$$\begin{aligned} \text{Vol}\left((\partial X \cap B_R(\mathbf{x}))^{\sqrt{d}h}\right) &\leq \text{Vol}(C(\sqrt{d}h)) && \text{(using (i))} \\ &\leq 4\left(\frac{4+\sqrt{2}}{4}\right)^d \frac{\sqrt{d}}{d} \text{Area}(P)h && \text{(with Eq.(9.43))} \\ &\leq \frac{2\sqrt{6}}{3}\left(\frac{4+\sqrt{2}}{2}\right)^d V_{d-1} R^{d-1} h && \text{(with Eq.(9.46)).} \end{aligned}$$

In the last line, we also use the fact that for $d \geq 2$ we have $\sqrt{1+2/d} \leq \sqrt{2}$. Putting all together gives the result. □

9.3.4 Errors on Volume and Moment Estimation Within a Ball of Radius R

Theorem 3 gives us the main key to bound above the error in volume estimation, and more generally moments estimation, within a ball around the boundary of a compact domain X. We summarize in the following theorem our results on moments estimation.

Theorem 4 *Let X be a compact domain of $\mathbb{R}^d$ such that the reach of ∂X is greater than ρ. Let D be any digitization process such that $\mathrm{J}^-{}_h(X) \subset \mathrm{D}_h(X) \subset \mathrm{J}^+{}_h(X)$. Let $\mathbf{x}$ be any point of $\mathbb{R}^d$. Let R (the radius of the ball) and h (the gridstep) be some positive numbers such that $h \leq \frac{R}{\sqrt{2d}}$ and $2R \leq \rho$. Let $(p_i)_{i=1\ldots d}$ be the integers defining the moment exponents, with $0 \leq p_i \leq 2$, and let $\sigma := p_1 + \cdots + p_d$, with $\sigma \leq 2$. Then digital moments within a ball are multigrid convergent toward continuous moments as follows*

$$\begin{aligned}\left|m^{p_1\cdots p_d}(X \cap B_R(\mathbf{x})) - \hat{m}_h^{p_1\cdots p_d}(\mathrm{D}_h(X \cap B_R(\mathbf{x})))\right| \\ \leq K_1(d) R^{d-1} (\|\mathbf{x}\|_\infty + 2R)^\sigma \, h + \frac{h^4}{12} V_d R^d,\end{aligned} \tag{9.47}$$

where V_d is the volume of the unit d-dimensional ball. Furthermore, the term in h^4 is only present when one p_i is equal to 2.

Proof The difficult part lies in the fact that moments may not be positive. To simplify notations, let $Y := X \cap B_R(\mathbf{x})$. We must split the error in moment estimation into three sets, corresponding to parts of Y lying in $[\mathrm{J}^-{}_h(Y)]_h$, in $[\mathrm{D}_h(Y)]_h \setminus [\mathrm{J}^-{}_h(Y)]_h$ and in $[\mathrm{J}^+{}_h(Y)]_h \setminus [\mathrm{D}_h(Y)]_h$.

$$\begin{aligned}\left|m^{p_1\cdots p_d}(Y) - \hat{m}_h^{p_1\cdots p_d}(\mathrm{D}_h(Y))\right| \leq & \sum_{\mathbf{z} \in \mathrm{J}^+{}_h(Y) \setminus \mathrm{D}_h(Y)} \left|m^{p_1\cdots p_d}(Y \cap Q_h(\mathbf{z}))\right| \\ & + \sum_{\mathbf{z} \in \mathrm{D}_h(Y) \setminus \mathrm{J}^-{}_h(Y)} \left|m^{p_1\cdots p_d}(Y \cap Q_h(\mathbf{z})) - \hat{m}_h^{p_1\cdots p_d}(\{\mathbf{z}\})\right| \\ & + \sum_{\mathbf{z} \in \mathrm{J}^-{}_h(Y)} \left|m^{p_1\cdots p_d}(Q_h(\mathbf{z})) - \hat{m}_h^{p_1\cdots p_d}(\{\mathbf{z}\})\right|\end{aligned} \tag{9.48}$$

The first term of this sum does not have any digital moment contribution since it lies outside the digitization of X. The third term does not require to intersect Y with the h-cube, since we are within the inner Jordan digitization.

1. We look at the third term. Lemma 7 tells us that it is equal to zero as long as $0 \leq p_i \leq 1$ for $1 \leq i \leq d$. Otherwise, if one p_i is equal to 2, it is then straightforward to bound it as follows:

$$\begin{aligned}\sum_{\mathbf{z}\in \mathbb{J}^-_h(Y)} \left| m^{p_1\cdots p_d}(Q_h(\mathbf{z})) - \hat{m}_h^{p_1\cdots p_d}(\{\mathbf{z}\})\right| &= \sum_{\mathbf{z}\in \mathbb{J}^-_h(Y)} \left|\frac{h^{d+4}}{12}\right| \qquad \text{(Lemma 7)}\\ &= \frac{h^4}{12}\,\mathrm{Vol}([\mathbb{J}^-_h(Y)]_h)\\ &\leq \frac{h^4}{12}\,\mathrm{Vol}(Y)\,, \qquad (9.49)\end{aligned}$$

since the inner Jordan digitization of Y lies inside Y. It is clear then that $\mathrm{Vol}(Y) = \mathrm{Vol}(X \cap B_R(\mathbf{x})) \leq \mathrm{Vol}(B_R(\mathbf{x})) = V_d R^d$.

2. We now consider the second term. For some $\mathbf{z} \in \mathbb{D}_h(Y) \setminus \mathbb{J}^-_h(Y)$, we reason on the sign of each component z_i for i in 1 to d. We notice that, in the h-cube $Q_h(\mathbf{z})$, the component x_i has the same sign as z_i except when $z_i = 0$. Let $\epsilon_i = 1$ when $z_i \geq 0$ and $\epsilon_i = -1$ otherwise. We can thus eliminate the signs and rewrite the difference below as:

$$\begin{aligned}&\left|m^{p_1\cdots p_d}(Y \cap Q_h(\mathbf{z})) - \hat{m}_h^{p_1\cdots p_d}(\{\mathbf{z}\})\right|\\ &= \left|\int\cdots\int_{Y\cap Q_h(\mathbf{z})} {x_1}^{p_1}\cdots{x_d}^{p_d}dx_1\ldots dx_d - h^{d+\sigma}{z_1}^{p_1}\cdots{z_d}^{p_d}\right|\\ &= \left|\int\cdots\int_{Y\cap Q_h(\mathbf{z})} (\epsilon_1 x_1)^{p_1}\cdots(\epsilon_d x_d)^{p_d}dx_1\ldots dx_d - h^{d+\sigma}(\epsilon_1 z_1)^{p_1}\cdots(\epsilon_d z_d)^{p_d}\right|\\ &\leq \left|\int\cdots\int_{Y\cap Q_h(\mathbf{z})} |\epsilon_1 x_1|^{p_1}\cdots|\epsilon_d x_d|^{p_d}dx_1\ldots dx_d - h^d(\epsilon_1 h z_1)^{p_1}\cdots(\epsilon_d h z_d)^{p_d}\right|\end{aligned} \qquad (9.50)$$

The integral on the left is necessarily non-negative while the term on the right is non-negative and is subtracted to it. It follows that this error can be upper bounded by the maximum of both terms.

$$\begin{aligned}&\left|m^{p_1\cdots p_d}(Y \cap Q_h(\mathbf{z})) - \hat{m}_h^{p_1\cdots p_d}(\{\mathbf{z}\})\right|\\ &= \max\left(\int\cdots\int_{Y\cap Q_h(\mathbf{z})} |\epsilon_1 x_1|^{p_1}\cdots|\epsilon_d x_d|^{p_d}dx_1\ldots dx_d, h^d(\epsilon_1 h z_1)^{p_1}\cdots(\epsilon_d h z_d)^{p_d}\right)\end{aligned} \qquad (9.51)$$

Since points involved in the left integral are all in $B_R(\mathbf{x})$, it follows that the left term is easily bounded by $(\|\mathbf{x}\|_\infty + R)^\sigma \operatorname{Vol}(Y \cap Q_h(\mathbf{z}))$. Since the point $\mathbf{z}$ is in $\mathtt{D}_h(Y) \setminus \mathtt{J}^-{}_h(Y)$, we have $\mathbf{z} \in \mathtt{J}^+{}_h(Y) \setminus \operatorname{Int}(\mathtt{J}^-{}_h(Y))$ and Lemma 3, equation Eq.(9.8), concludes that $(h\mathbf{z}) \in \partial Y^{\sqrt{d}h}$. it follows that $\|h\mathbf{z}\|_\infty \leq \|\mathbf{x}\|_\infty + R + \sqrt{d}h \leq \|\mathbf{x}\|_\infty + 2R$, since $h < R/\sqrt{2d}$. The second term is thus bounded by $(\|\mathbf{x}\|_\infty + 2R)^\sigma \operatorname{Vol}(Q_h(\mathbf{z}))$. Since $\operatorname{Vol}(Y \cap Q_h(\mathbf{z})) \leq \operatorname{Vol}(Q_h(\mathbf{z}))$, we obtain

$$\sum_{\mathbf{z} \in \mathtt{D}_h(Y) \setminus \mathtt{J}^-{}_h(Y)} \left| m^{p_1 \cdots p_d}(Y \cap Q_h(\mathbf{z})) - \hat{m}_h^{p_1 \cdots p_d}(\{\mathbf{z}\}) \right| \leq (\|\mathbf{x}\|_\infty + 2R)^\sigma \operatorname{Vol}([\mathtt{D}_h(Y) \setminus \mathtt{J}^-{}_h(Y)]_h). \tag{9.52}$$

3. We finally look at the first term of the sum, which is easier to bound:

$$\sum_{\mathbf{z} \in \mathtt{J}^+{}_h(Y) \setminus \mathtt{D}_h(Y)} |m^{p_1 \cdots p_d}(Y \cap Q_h(\mathbf{z}))| \leq (\|\mathbf{x}\|_\infty + R)^\sigma \sum_{\mathbf{z} \in \mathtt{J}^+{}_h(Y) \setminus \mathtt{D}_h(Y)} \operatorname{Vol}(Y \cap Q_h(\mathbf{z})). \leq (\|\mathbf{x}\|_\infty + R)^\sigma \operatorname{Vol}([\mathtt{J}^+{}_h(Y) \setminus \mathtt{D}_h(Y)]). \tag{9.53}$$

Since $\mathtt{D}_h(Y) \setminus \mathtt{J}^-{}_h(Y)$ and $\mathtt{J}^+{}_h(Y) \setminus \mathtt{D}_h(Y)$ are disjoint, we may add inequalities Eqs. (9.52) and (9.53) as

$$\begin{aligned}
&\sum_{\mathbf{z} \in \mathtt{D}_h(Y) \setminus \mathtt{J}^-{}_h(Y)} \left| m^{p_1 \cdots p_d}(Y \cap Q_h(\mathbf{z})) - \hat{m}_h^{p_1 \cdots p_d}(\{z\}) \right| \\
&\quad + \sum_{\mathbf{z} \in \mathtt{J}^+{}_h(Y) \setminus \mathtt{D}_h(Y)} |m^{p_1 \cdots p_d}(Y \cap Q_h(\mathbf{z}))| \\
&\leq (\|\mathbf{x}\|_\infty + 2R)^\sigma \operatorname{Vol}([\mathtt{J}^+{}_h(Y) \setminus \mathtt{J}^-{}_h(Y)]_h) \\
&\leq (\|\mathbf{x}\|_\infty + 2R)^\sigma \operatorname{Vol}\left((\partial Y)^{\sqrt{d}h}\right) \qquad \text{(Lemma 3, Eq.(9.8), page 299)} \\
&\leq (\|\mathbf{x}\|_\infty + 2R)^\sigma K_1(d) R^{d-1} h \qquad \text{(Theorem 3)}
\end{aligned}$$

This concludes by simple addition of the bound on the third term, i.e. Eq.(9.49). □

9.3.5 Conclusion

In this section we have determined links between continuous and digital moments of order up to 2 in arbitrary dimension. We have established several approximation error bounds for arbitrary compact domains. Furthermore, we have been able to estimate the approximation error on moments of intersections of a smooth shape

with a ball of radius R and center $\mathbf{x}$. Our error bound only depends on the dimension of the space, the radius R of the ball and the norm $\|\mathbf{x}\|_\infty$, and scales linearly with the digitization gridstep h. It is worthy to note that our results apply for arbitrary digitization processes, as long as they contain the inner Jordan digitization and are included in the outer Jordan digitization. In particular, it includes the Gauss digitization. We will use these results in the next sections to show the multigrid convergence of curvature estimators based on digital integral invariants.

9.4 Multigrid Convergence of Mean Curvature in 2D and 3D

We show in the section that the local mean curvature on the boundary of a digital shape can be approximated simply by intersecting the digital shape with a ball around the point of interest and counting the number of digital points within. This is related to integral invariants results [49] (recalled in Lemma 5, page 301), which requires only the computation of the volume of the shape intersected with a ball.

9.4.1 Definition of Mean Curvature Estimators

We begin by defining a 2D digital curvature estimator and a 3D digital mean curvature estimator, whose computation requires only the enumeration of digital points within a neighborhood around the point of interest.

Definition 3 (2D Integral Digital Curvature Estimator) For any positive radius R, we define the *2D integral digital curvature estimator* $\hat{\kappa}^R$ of a digital shape $Z \subset \mathbb{Z}^2$ at any point $\mathbf{x} \in \mathbb{R}^2$ and for a grid step $h > 0$ as:

$$\forall 0 < h < R, \quad \hat{\kappa}^R(Z, \mathbf{x}, h) := \frac{3\pi}{2R} - \frac{3\widehat{\text{Area}}(B_{R/h}(\mathbf{x}/h) \cap Z, h)}{R^3}. \tag{9.54}$$

Definition 4 (3D Integral Digital Mean Curvature Estimator) For any positive radius R, we define the *3D integral mean digital curvature estimator* $\hat{H}^R$ of a digital shape $Z \subset \mathbb{Z}^3$ at any point $\mathbf{x} \in \mathbb{R}^3$ and for a grid step $h > 0$ as:

$$\forall 0 < h < R, \quad \hat{H}^R(Z, \mathbf{x}, h) := \frac{8}{3R} - \frac{4\widehat{\text{Vol}}(B_{R/h}(\mathbf{x}/h) \cap Z, h)}{\pi R^4}. \tag{9.55}$$

As one can see on Fig. 9.3, these estimators place a ball of Euclidean radius R around the point of interest $\mathbf{x}$ and count the number of digital points of Z, scaled back in $(h\mathbb{Z})^d$, within this ball. A simple linear formula is then applied on this number to get a curvature approximation.

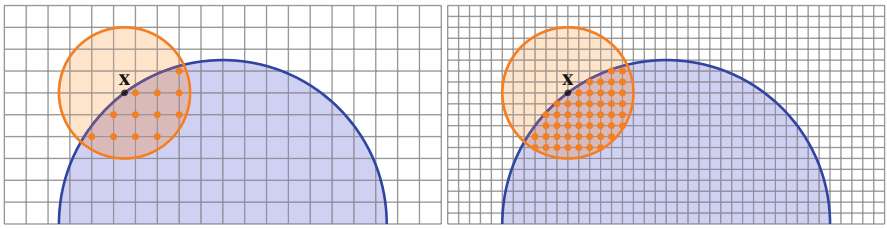


Fig. 9.3 Illustration of 2D integral digital curvature estimator $\hat{\kappa}^R$. The shape X is a disk of radius 7.5. To the left, the digitization gridstep h is 1, while h is 0.5 to the right. We choose a ball of radius $R = 3$ and we wish to estimate the curvature at some arbitrary point $\mathbf{x}$. We count the number of digital points within the *orange ball*, centered at $\mathbf{x}$ and of radius 3. To the left we count 12 points. Hence for $h = 1$, the estimated curvature $\hat{\kappa}^3(\mathbb{G}_1(X), \mathbf{x}, 1)$ is $3\pi/(2\times 3) - 3\times 12\times 1^2/3^3 \approx 0.237$. To the right we count 50 points. Hence for $h = 0.5$, the estimated curvature $\hat{\kappa}^3(\mathbb{G}_{0.5}(X), \mathbf{x}, 0.5)$ is $3\pi/(2\times 3) - 3\times 51\times 0.5^2/3^3 \approx 0.154$. Ground truth curvature is $1/7.5 \approx 0.133$

The 2D example displayed on Fig. 9.3 indicates that the finer the digitization step, the better the approximation. This is indeed true, at least for shapes with smooth enough boundary.

9.4.2 Convergence at Points of ∂X (Weak Formulation)

In this section, we show that the curvature estimator $\hat{\kappa}^R$ (resp. the mean curvature estimator $\hat{H}^R$) converges to the expected curvature (resp. mean curvature) for points $\mathbf{x}$ belonging to the boundary of a compact shape X, as long as X has positive reach. A preliminary version of this theorem that requires X to be convex with C^3-smooth boundary has been published in [14].

Theorem 5 (Convergence of Curvature Estimator $\hat{\kappa}^R$ Along ∂X) *Let X be a compact domain of $\mathbb{R}^2$ such that its boundary ∂X is C^3-smooth and has reach greater than ρ.*

Then the curvature estimator $\hat{\kappa}^R$ at any point $\mathbf{x}$ of ∂X is multigrid convergent to the curvature $\kappa(X, \mathbf{x})$ of X at point $\mathbf{x}$ for the Gauss digitization process, with convergence speed at least $O(h^{\frac{1}{3}})$ when $R = \Theta(h^{\frac{1}{3}})$ and $R < \rho/2$. More precisely, we have

$$\forall 0 < h \leq \frac{R}{2}, \quad \left|\hat{\kappa}^R(\mathbb{G}_h(X), \mathbf{x}, h) - \kappa(X, \mathbf{x})\right| \leq O\left(h^{\frac{1}{3}}\right). \tag{9.56}$$

Proof Note that the boundary is smooth enough so that integral invariants do have a Taylor expansion (and Lemma 5 applies). Furthermore, the domain being compact, it has necessarily a positive reach. Since the Gauss digitization lies in the Jordan

strip (Lemma 1, page 298), we are in the hypotheses of Theorem 4. We bound the curvature approximation error as follows:

$$
\begin{aligned}
&|\hat{\kappa}^R(\mathbb{G}_h(X), \mathbf{x}, h) - \kappa(X, \mathbf{x})| \\
&\quad \leq \left| \hat{\kappa}^R(\mathbb{G}_h(X), \mathbf{x}, h) - \tilde{\kappa}^R(X, \mathbf{x}) \right| + O(R) && \text{(from Eq.(9.15), page 301)} \\
&\quad = \left| \frac{3A_R(\mathbf{x})}{R^3} - \frac{3\widehat{\mathrm{Area}}(B_{R/h}(\mathbf{x}/h) \cap \mathbb{G}_h(X), h)}{R^3} \right| + O(R) && \text{(from Eq.(9.54) and (9.14))} \\
&\quad = \frac{3}{R^3} \left| m^{00}(X \cap B_R(\mathbf{x})) - \hat{m}_h^{00}(\mathbb{G}_h(X \cap B_R(\mathbf{x}))) \right| + O(R)\,. && (9.57)
\end{aligned}
$$

The last equality follows from the definitions of A_R and $\widehat{\mathrm{Area}}$ expressed as 00-moments, and also from the fact $\mathbb{G}_h(X \cap B_R(\mathbf{x})) = B_{R/h}(\mathbf{x}/h) \cap \mathbb{G}_h(X)$ (this is easily checked for Gauss digitization). Theorem 4 then implies

$$
|\hat{\kappa}^R(\mathbb{G}_h(X), \mathbf{x}, h) - \kappa(X, \mathbf{x})| \leq \frac{3K_1(2)}{R^2} h + O(R)\,. \tag{9.58}
$$

The error is the sum of two terms, in which R has an opposite effect. The right term requires a radius R tending to zero, while the left term is minimized by a large radius. If R is chosen as some function kh^α, where k is a constant, then the asymptotically minimizing error is for $\alpha = \frac{1}{3}$. □

We have a similar result for the digital mean curvature estimator on 3D shapes.

Theorem 6 (Convergence of Mean Curvature Estimator $\hat{H}^R$ Along ∂X) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X is C^3-smooth and has reach greater than ρ.*

Then the mean curvature estimator $\hat{H}^R$ at any point $\mathbf{x}$ of ∂X is multigrid convergent to the mean curvature $H(X, \mathbf{x})$ of X at point $\mathbf{x}$ for the Gauss digitization process, with convergence speed at least $O(h^{\frac{1}{3}})$ when $R = \Theta(h^{\frac{1}{3}})$ and $R < \rho/2$. More precisely, we have

$$
\forall 0 < h \leq \frac{R}{\sqrt{6}}, \quad \left| \hat{H}^R(\mathbb{G}_h(X), \mathbf{x}, h) - H(X, \mathbf{x}) \right| \leq O\left(h^{\frac{1}{3}}\right). \tag{9.59}
$$

Proof The proof follows exactly the same steps as the proof of Theorem 5, since on the one hand integral invariants also have a Taylor expansion in 3D and on the other hand Theorem 4 applies in arbitrary dimension. We bound the curvature approximation error as follows:

$$
\begin{aligned}
&|\hat{H}^R(\mathbb{G}_h(X), \mathbf{x}, h) - H(X, \mathbf{x})| \\
&\quad \leq \left| \hat{H}^R(\mathbb{G}_h(X), \mathbf{x}, h) - \tilde{H}^R(X, \mathbf{x}) \right| + O(R) && \text{(from Eq.(9.15), page 301)}
\end{aligned}
$$

$$= \left| \frac{4V_R(\mathbf{x})}{\pi R^4} - \frac{4\widehat{\mathrm{Vol}}(B_{R/h}(\mathbf{x}/h) \cap \mathbb{G}_h(X), h)}{\pi R^4} \right| + O(R) \qquad \text{(from Eq.(9.55) and (9.14))}$$

$$= \frac{4}{\pi R^4} \left| m^{000}(X \cap B_R(\mathbf{x})) - \hat{m}_h^{000}(\mathbb{G}_h(X \cap B_R(\mathbf{x}))) \right| + O(R)\,. \tag{9.60}$$

The last equality follows from the definitions of V_R and $\widehat{\mathrm{Vol}}$ expressed as 000-moments, and also from the fact $\mathbb{G}_h(X \cap B_R(\mathbf{x})) = B_{R/h}(\mathbf{x}/h) \cap \mathbb{G}_h(X)$. Theorem 4 then induces

$$\left| \hat{H}^R(\mathbb{G}_h(X), \mathbf{x}, h) - H(X, \mathbf{x}) \right| \leq \frac{4K_1(3)}{\pi R^2} h + O(R)\,. \tag{9.61}$$

As in 2D, the error is the sum of two terms, in which R has an opposite effect. We also obtain that the asymptotically minimizing error is for $R = kh^{\frac{1}{3}}$, with k some constant. □

9.4.3 Multigrid Convergence for Smooth Enough Shapes

Theorems 5 and 6 are not multigrid convergence theorems, since convergence results are only valid on points of ∂X. However, the exact location of ∂X is generally unknown, and only approximate digital data is available. To achieve multigrid convergent theorems, we have to take into account the possible error in the position at which the curvature is estimated.

We therefore examine the perturbation of the moments when they are evaluated at a shifted position $\mathbf{x} + \mathbf{t}$.

Lemma 9 *For any point* $\mathbf{x}$*, a positive number* R*, and any vector* $\mathbf{t}$ *with norm* $t := \|\mathbf{t}\|_2 \leq R$*, we have*

$$|\mathrm{Vol}(B_R(\mathbf{x} + \mathbf{t})) - \mathrm{Vol}(B_R(\mathbf{x}))| \leq \frac{3^d}{2^d} V_d R^{d-1}\, t\,. \tag{9.62}$$

Proof We simply bound this difference of volumes by the volume of the difference of two balls with same center, as illustrated on Fig. 9.4. More precisely, we have:

$$\begin{aligned} |\mathrm{Vol}(B_R(\mathbf{x} + \mathbf{t})) - \mathrm{Vol}(B_R(\mathbf{x}))| &= |\mathrm{Vol}(B_R(\mathbf{x} + \mathbf{t}) \Delta B_R(\mathbf{x}))| \\ &\leq \left| \mathrm{Vol}\left(B_{R+\frac{t}{2}}(\mathbf{q}) \setminus B_{R-\frac{t}{2}}(\mathbf{q}) \right) \right|, \end{aligned}$$

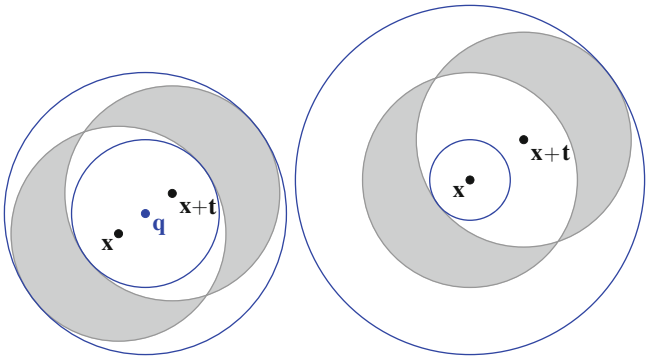


Fig. 9.4 *Left*: illustration of Lemma 9 for bounding volume. The symmetric difference of two balls is included in the shell with thickness equal to the distance between the centerpoints of these two balls and centered at their midpoint. *Right*: illustration of Lemma 11 for bounding moments. The symmetric difference of two balls is included in the shell with thickness equals to twice the distance between the centerpoints of these two balls and centered on one of the ball

where $\mathbf{q}$ is the midpoint between $\mathbf{x}$ and $\mathbf{x}+\mathbf{t}$. Since V_d is the volume of the unit d-dimensional ball, it follows that:

$$\begin{aligned} |\mathrm{Vol}(B_R(\mathbf{x}+\mathbf{t})) - \mathrm{Vol}(B_R(\mathbf{x}))| &\leq V_d \left((R+\frac{t}{2})^d - (R-\frac{t}{2})^d \right) \\ &\leq \frac{3^d}{2^d} V_d R^{d-1}\, t\,. \end{aligned}$$

□

We may now prove the uniform multigrid convergence of both the 2D curvature estimator $\hat{\kappa}^R$ and the 3D mean curvature estimator $\hat{H}^R$ towards respectively the curvature κ and the mean curvature H for the Gauss digitization process and smooth enough shapes.

Theorem 7 (Multigrid Convergence of 2D Curvature Estimator $\hat{\kappa}^R$) *Let X be a compact domain of $\mathbb{R}^2$ such that its boundary ∂X is C^3-smooth and has reach greater than ρ.*

Then the curvature estimator $\hat{\kappa}^R$ is multigrid convergent to the curvature κ for the Gauss digitization process on such shapes X, with convergence speed at least $O(h^{\frac{1}{3}})$ when $R = \Theta(h^{\frac{1}{3}})$ and $R < \rho/2$. More precisely, we have

$$\begin{aligned} \forall 0 < h \leq \frac{R}{2},\ \forall \mathbf{x} \in \partial X,\ \ \forall \hat{\mathbf{x}} \in \partial[\mathsf{G}_h(X)]_h \text{ with } \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h, \\ \left| \hat{\kappa}^R(\mathsf{G}_h(X), \hat{\mathbf{x}}, h) - \kappa(X, \mathbf{x}) \right| \leq O\left(h^{\frac{1}{3}}\right)\,. \end{aligned} \tag{9.63}$$

More precisely the bound is no greater than $(\frac{27}{4}\pi\sqrt{2} + 3K_1(2))R^{-2}h + O(R)$.

Proof Note that from Theorem 1 (page 299), we have that any point of ∂X is close to a point of $\partial[\mathbb{G}_h(X)]_h$ (distance less than $h/\sqrt{d}$) and the converse is also true. We proceed similarly as in Theorem 5, using Eq.(9.15) to get:

$$\begin{aligned}
&|\hat{\kappa}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \kappa(X, \mathbf{x})| \\
&\quad \leq \left|\hat{\kappa}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \tilde{\kappa}^R(X, \hat{\mathbf{x}})\right| + \left|\tilde{\kappa}^R(X, \hat{\mathbf{x}}) - \tilde{\kappa}^R(X, \mathbf{x})\right| + O(R).
\end{aligned}$$

The first error term is bounded by $\frac{3K_1(2)}{R^2}h$ (see proof of Theorem 5). We take care of the second error term, letting $\mathbf{t} := \hat{\mathbf{x}} - \mathbf{x}$ and $t := \|\mathbf{t}\|_2$:

$$\begin{aligned}
\left|\tilde{\kappa}^R(X, \hat{\mathbf{x}}) - \tilde{\kappa}^R(X, \mathbf{x})\right| &= \frac{3}{R^3}\left|A_R(\hat{\mathbf{x}}) - A_R(\mathbf{x})\right| \\
&= \frac{3}{R^3}\left|\mathrm{Vol}(B_R(\mathbf{x}+\mathbf{t}) \cap X) - \mathrm{Vol}(B_R(\mathbf{x}) \cap X)\right| \\
&\leq \frac{3}{R^3}\left|\mathrm{Vol}(B_R(\mathbf{x}+\mathbf{t})) - \mathrm{Vol}(B_R(\mathbf{x}))\right| \\
&\leq \frac{27}{4}\pi R^{-2}\, t\,. \quad \text{(using Lemma 9)}
\end{aligned}$$

The last bound is valid since $t = \|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq \sqrt{2}h$ by the fact that $\|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h$. By hypothesis, $h \leq R/2$, so $t < R$.

We have just shown $\left|\tilde{\kappa}^R(X, \hat{\mathbf{x}}) - \tilde{\kappa}^R(X, \mathbf{x})\right| \leq 6\pi\sqrt{2}R^{-2}h$, which concludes. □

We have a similar result for the 3D mean curvature estimator, whose proof follows the same arguments.

Theorem 8 (Multigrid Convergence of 3D Mean Curvature Estimator $\hat{H}^R$) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X is C^3-smooth and has reach greater than ρ.*

Then the 3D mean curvature estimator $\hat{H}^R$ is multigrid convergent to the mean curvature H for the Gauss digitization process on such shapes X, with convergence speed at least $O(h^{\frac{1}{3}})$ when $R = \Theta(h^{\frac{1}{3}})$ and $R < \rho/2$. More precisely, we have

$$\forall 0 < h \leq \frac{R}{\sqrt{6}},\ \forall \mathbf{x} \in \partial X,\ \forall \hat{\mathbf{x}} \in \partial[\mathbb{G}_h(X)]_h \text{ with } \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h,$$
$$\left|\hat{H}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - H(X, \mathbf{x})\right| \leq O\left(h^{\frac{1}{3}}\right). \tag{9.64}$$

More precisely the bound is no greater than $(18\sqrt{3} + 4K_1(3)/\pi)R^{-2}h + O(R)$.

9.5 Multigrid Convergence of Curvature Tensor in 3D

We show in this section how we can approach the whole curvature tensor by simply intersecting a ball with the digital input shape. However, we do not simply count the number of digital points within but we rather compute the digital covariance matrix. The digital covariance matrix is shown to be close to the covariance matrix, which is known to contain curvature information. Although most of the results presented here could be naturally extended to arbitrary dimension d, we prefer to expose them in 3D.

9.5.1 Digital Covariance Matrix and Digital Principal Curvature Estimators

Similarly to the covariance matrix, the digital covariance matrix is defined through zeroth, first and second order digital moments.

Definition 5 (Digital Covariance Matrix) For any digital subset $Z \subset \mathbb{Z}^3$, its *digital covariance matrix* $\mathcal{V}_h$ at step h is

$$\mathcal{V}_h(Z) := \begin{bmatrix} \hat{m}_h^{200}(Z) & \hat{m}_h^{110}(Z) & \hat{m}_h^{101}(Z) \\ \hat{m}_h^{110}(Z) & \hat{m}_h^{020}(Z) & \hat{m}_h^{011}(Z) \\ \hat{m}_h^{101}(Z) & \hat{m}_h^{011}(Z) & \hat{m}_h^{002}(Z) \end{bmatrix} - \frac{1}{\hat{m}_h^{000}(Z)} \begin{bmatrix} \hat{m}_h^{100}(Z) \\ \hat{m}_h^{010}(Z) \\ \hat{m}_h^{001}(Z) \end{bmatrix} \otimes \begin{bmatrix} \hat{m}_h^{100}(Z) \\ \hat{m}_h^{010}(Z) \\ \hat{m}_h^{001}(Z) \end{bmatrix}^T . \tag{9.65}$$

Following the truncated Taylor expansion of Lemma 6, we define estimators of principal curvatures from the diagonalization of the digital covariance matrix.

Definition 6 Let $Z \subset \mathbb{Z}^3$ be a digital shape and $h > 0$ a grid step. For $R \geq h$, we define the *integral principal curvature estimators* $\hat{\kappa}_1^R$ *and* $\hat{\kappa}_2^R$ of Z at point $\mathbf{y} \in \mathbb{R}^3$ and step h their respective *integral principal direction estimators* $\hat{\mathbf{w}}_1^R$ *and* $\hat{\mathbf{w}}_2^R$, and the *integral normal estimator* $\hat{\mathbf{n}}^R$ as

$$\hat{\kappa}_1^R(Z, \mathbf{y}, h) := \frac{6}{\pi R^6}(\hat{\lambda}_2 - 3\hat{\lambda}_1) + \frac{8}{5R}, \qquad \hat{\mathbf{w}}_1^R(Z, \mathbf{y}, h) := \hat{\nu}_1 \tag{9.66}$$

$$\hat{\kappa}_2^R(Z, \mathbf{y}, h) := \frac{6}{\pi R^6}(\hat{\lambda}_1 - 3\hat{\lambda}_2) + \frac{8}{5R}, \qquad \hat{\mathbf{w}}_2^R(Z, \mathbf{y}, h) := \hat{\nu}_2 \tag{9.67}$$

$$\hat{\mathbf{n}}^R(Z, \mathbf{y}, h) := \hat{\nu}_3 , \tag{9.68}$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ are the eigenvalues of $\mathcal{V}_h(B_{R/h}(\mathbf{y}/h) \cap Z)$, and $\hat{\nu}_1, \hat{\nu}_2, \hat{\nu}_3$ are their corresponding eigenvectors.

In the following sections, we show that such estimators are indeed multigrid convergent to their respective geometric quantity. Proofs rely on the convergence of digital moments, on the stability of moments with respect to small displacements, and on matrix perturbation theory. Most of the work presented here has been published in [15].

9.5.2 Some Properties of Covariance Matrix and Moments

Before proving the multigrid convergence, we need to establish a few preliminary lemmas. First of all, we notice easily that covariance matrices are translation invariant.

Lemma 10 *Translation invariance for covariance matrices:*

- $\forall Y \subset \mathbb{R}^3$, $\forall \mathbf{v} \in \mathbb{R}^3$, $\mathscr{V}(Y + \mathbf{v}) = \mathscr{V}(Y)$.
- $\forall Z \subset \mathbb{Z}^3$, $\forall \mathbf{v} \in \mathbb{Z}^3$, $\forall h > 0$, $\mathscr{V}_h(Z + \mathbf{v}) = \mathscr{V}_h(Z)$.

Then, we must examine how moments are perturbated by a positioning error of the ball.

Lemma 11 *For any point* $\mathbf{x}$*, a positive number* R*, and any vector* $\mathbf{t}$ *with norm* $t := \|\mathbf{t}\|_2 \leq R$*, we have*

$$\left|m^{000}(B_R(\mathbf{x}+\mathbf{t})) - m^{000}(B_R(\mathbf{x}))\right| \leq O(tR^2), \tag{9.69}$$

$$\left|m^{100}(B_R(\mathbf{x}+\mathbf{t})) - m^{100}(B_R(\mathbf{x}))\right| \leq O(tR^3) + O(\|\mathbf{x}\|_\infty tR^2), \tag{9.70}$$

$$\left|m^{200}(B_R(\mathbf{x}+\mathbf{t})) - m^{200}(B_R(\mathbf{x}))\right| \leq O(tR^4) + O(\|\mathbf{x}\|_\infty tR^3) + O(\|\mathbf{x}\|_\infty^2 tR^2). \tag{9.71}$$

Other moments with same order have the same respective bounds. Furthermore, these relations remain true if both $B_R(\mathbf{x}+\mathbf{t})$ *and* $B_R(\mathbf{x})$ *are intersected with the same set* X*.*

Proof We notice first that Eq.(9.69) has already been established in Lemma 9. For higher order moments pqs, we will use the following fact:

$$\emptyset \neq Y_1 \subset Y_2 \subset \mathbb{R}^3 \Rightarrow \sup_{Y \subset Y_1} |m^{pqs}(Y)| \leq \sup_{Y \subset Y_2} |m^{pqs}(Y)| \ . \tag{9.72}$$

As in Lemma 9, we notice that the difference of balls is included in the difference of two balls with same center. This is illustrated in Fig. 9.4, right.

$$\begin{aligned} |m^{pqs}(B_R(\mathbf{x}+\mathbf{t})) - m^{pqs}(B_R(\mathbf{x}))| &\leq \sup_{Y \subset B_R(\mathbf{x}+\mathbf{t}) \Delta B_R(\mathbf{x})} |m^{pqs}(Y)| \ , \\ &\leq \sup_{Y \subset B_{R+t}(\mathbf{x}) \setminus B_{R-t}(\mathbf{x})} |m^{pqs}(Y)| \ . \end{aligned}$$

Hereafter we denote $S_{R,t}(\mathbf{x}) := B_{R+t}(\mathbf{x}) \setminus B_{R-t}(\mathbf{x})$.

For first order moments, we translate the shape to the origin, then we use the previous result plus the fact that the centered 100-moment is maximized for the x-positive half-ball[3]:

$$\begin{aligned}\sup_{Y \subset S_{R,t}(\mathbf{x})} |m^{100}(Y)| &\leq \sup_{Y \subset S_{R,t}(\mathbf{0})} |m^{100}(Y)| + |x_1| \cdot m^{000}(Y) \\ &= m^{100}(B^+_{R+t}(\mathbf{0}) - B^+_{R-t}(\mathbf{0})) + O(|x_1| t R^2) \\ &= 2\pi(R^3 t + R t^3) + O((\|\mathbf{x}\|_\infty) t R^2) \\ &= O(t R^3) + O(\|\mathbf{x}\|_\infty t R^2). \end{aligned} \tag{9.73}$$

For second order moments, we translate the shape to the origin, then we use the two previous results plus the fact that the 200-moment is maximized for the ball:

$$\begin{aligned}\sup_{Y \subset S_{R,t}(\mathbf{x})} |m^{200}(Y)| &\leq \sup_{Y \subset S_{R,t}(\mathbf{0})} |m^{200}(Y)| + 2|x_1| \cdot |m^{100}(Y)| + x_1^2 \cdot m^{000}(Y) \\ &= m^{200}(S_{R,t}(\mathbf{0})) + \|\mathbf{x}\|_\infty (O(t R^3) + O(\|\mathbf{x}\|_\infty t R^2)) + \|\mathbf{x}\|_\infty^2 O(t R^2) \\ &= O(t R^4) + O(\|\mathbf{x}\|_\infty t R^3) + O(\|\mathbf{x}\|_\infty^2 t R^2) \end{aligned} \tag{9.74}$$

Other moments are proved similarly. □

9.5.3 Multigrid Convergence of Digital Covariance Matrix

With Lemma 10 and Lemma 11, we can prove the multigrid convergence of the digital covariance matrix. Theorem 9 establishes its simple convergence (covariance matrices are computed at the same point $\mathbf{y}$) and then Theorem 10 establishes its multigrid convergence.

Theorem 9 (Convergence of Digital Covariance Matrix) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X has reach greater than ρ.*

Then for any grid step h and radius R with $0 < h \leq R/\sqrt{6} < \rho/(2\sqrt{3})$, for arbitrary $\mathbf{y} \in \mathbb{R}^3$ with $d(\mathbf{y}, \partial X) \leq h$, we have:

$$\left\| \mathscr{V}_h(B_{R/h}(\mathbf{y}/h) \cap \mathbb{G}_h(X)) - \mathscr{V}(B_R(\mathbf{y}) \cap X) \right\| \leq O(R^4 h).$$

The constants hidden in the big O do not depend on the shape size or geometry. $\|\cdot\|$ denotes the spectral norm on matrices.

[3] $B^+_R(\mathbf{x})$ denotes the x-positive half ball of center $\mathbf{x}$ and radius R. Remember that x_i is the $i-$th component of $\mathbf{x}$.

Proof To simplify expressions, we set $A{:=}B_R(\mathbf{y}) \cap X$, $A_h{:=}B_{R/h}(\mathbf{y}/h) \cap \mathbb{G}_h(X) = \mathbb{G}_h(B_R(\mathbf{y}) \cap X)$ (the last equality is valid for the Gauss digitization). We begin by translating the sets A and A_h towards the origin w.r.t. $\mathbf{y}$. We must use a vector that takes into account the digitization, hence we shift A_h by the vector $\left[\frac{\mathbf{y}}{h}\right]$, the integer vector closest to $\frac{\mathbf{y}}{h}$, and we shift A with the vector $h\left[\frac{\mathbf{y}}{h}\right]$. We further set $\tilde{A}_h{:=}A_h - \left[\frac{\mathbf{y}}{h}\right]$ and $\tilde{A}{:=}A - h\left[\frac{\mathbf{y}}{h}\right]$. With these definitions and the translation invariance of digital covariance matrixes (Lemma 10), we get

$$\begin{aligned}\mathscr{V}_h(B_{R/h}(\mathbf{y}/h) \cap \mathbb{G}_h(X)) &= \mathscr{V}_h(A_h) = \mathscr{V}_h\left(\tilde{A}_h + \left[\frac{\mathbf{y}}{h}\right]\right) = \mathscr{V}_h(\tilde{A}_h) \\ &= \mathscr{V}_h(\mathbb{G}_h(\tilde{A})). \end{aligned} \tag{9.75}$$

The last equality comes from the fact that $\tilde{A}_h = \mathbb{G}_h(A) - \left[\frac{\mathbf{y}}{h}\right] = \mathbb{G}_h(A - h\left[\frac{\mathbf{y}}{h}\right]) = \mathbb{G}_h(\tilde{A})$. We use also the translation invariance of the (continuous) covariance matrix to get

$$\mathscr{V}(B_R(\mathbf{x}) \cap X) = \mathscr{V}(A) = \mathscr{V}\left(\tilde{A} + h\left[\frac{\mathbf{y}}{h}\right]\right) = \mathscr{V}(\tilde{A}). \tag{9.76}$$

With Eq.(9.75) and Eq.(9.76), we rewrite the covariance estimation error as:

$$\|\mathscr{V}_h(B_{R/h}(\mathbf{y}/h) \cap \mathbb{G}_h(X)) - \mathscr{V}(B_R(\mathbf{x}) \cap X)\| = \|\mathscr{V}_h(\mathbb{G}_h(\tilde{A})) - \mathscr{V}(\tilde{A})\|. \tag{9.77}$$

We take a closer look at $\mathscr{V}_h(\mathbb{G}_h(\tilde{A}))$:

$$\mathscr{V}_h(\mathbb{G}_h(\tilde{A})) = \begin{bmatrix} \hat{m}_h^{200}(\tilde{A}_h) & & \\ & \ddots & \end{bmatrix} - \frac{1}{\hat{m}_h^{000}(\tilde{A}_h)} \begin{bmatrix} \hat{m}_h^{100}(\tilde{A}_h) \\ \vdots \end{bmatrix} \otimes \begin{bmatrix} \hat{m}_h^{100}(\tilde{A}_h) \\ \vdots \end{bmatrix}^T.$$

We may rewrite the factor in front of the tensorial product using Theorem 4:

$$\frac{1}{\hat{m}_h^{000}(\tilde{A}_h)} = \frac{1}{m^{000}(\tilde{A})} + \left(\frac{1}{m^{000}(\tilde{A})}\right)^2 O(R^2 h).$$

Since $\tilde{A}$ is some translation of $X \cap B_R(\mathbf{y})$, it has the same volume. This volume is some $\Theta(R^3)$ since $\mathbf{y}$ is at most at distance h of ∂X and $h \leq R/\sqrt{6}$. We get:

$$\frac{1}{\hat{m}_h^{000}(\tilde{A}_h)} = \frac{1}{m^{000}(\tilde{A})} + O(R^{-4}h). \tag{9.78}$$

We bound above the terms in the tensorial product. We write $\mathbf{t}{:=}\mathbf{y} - h\left[\frac{\mathbf{y}}{h}\right]$, which is the center of the ball of $\tilde{A}$. As an example, we examine its top left term and use again Theorem 4:

$$\begin{aligned}\hat{m}_h^{100}(\tilde{A}_h) \times \hat{m}_h^{100}(\tilde{A}_h) &= \left(m^{100}(\tilde{A}) + O\left(R^2\left(\|\mathbf{t}\|_\infty + 2R\right)h\right)\right)^2 \\ &= \left(m^{100}(\tilde{A})\right)^2 + O(R^7 h) + O(R^6 h^2). \end{aligned} \tag{9.79}$$

The last equality comes from the fact that $|m^{100}(\tilde{A})| \leq m^{100}(B^+_{R+\sqrt{3}h}(\mathbf{0})) \leq 2\pi R^4$ and that $\|\mathbf{t}\|_\infty$ is smaller than h which is smaller than a constant times R.

Other terms of the tensorial product are upper bounded similarly. We look now at the upper left term of the whole covariance matrix difference:

$$
\begin{aligned}
&\left|\left(\mathscr{V}_h(\mathbb{G}_h(\tilde{A})) - \mathscr{V}(\tilde{A})\right)_{1,1}\right| \\
&= \left|\hat{m}_h^{200}(\tilde{A}_h) - \frac{1}{\hat{m}_h^{000}(\tilde{A}_h)}\left(\hat{m}_h^{100}(\tilde{A}_h)\right)^2 - m^{200}(\tilde{A}) + \frac{1}{m^{000}(\tilde{A})}\left(m^{100}(\tilde{A})\right)^2\right|. \\
&\leq \left|\hat{m}_h^{200}(\tilde{A}_h) - m^{200}(\tilde{A})\right| \qquad \text{(using Eq.(9.78) and Eq.(9.79))} \qquad (9.80) \\
&\quad + \left|\left(\frac{1}{m^{000}(\tilde{A})} + O(R^{-4}h)\right)\left(\left(m^{100}(\tilde{A})\right)^2 + O(R^7 h) + O(R^6 h^2)\right) - \frac{\left(m^{100}(\tilde{A})\right)^2}{m^{000}(\tilde{A})}\right| \\
&\leq O(R^2(\|\mathbf{t}\|_\infty + 2R)^2)h + O(R^3 h^4) \qquad \text{(Theorem 4)} \\
&\quad + O(R^4 h) + O(R^3 h^2) + O(R^2 h^3). \qquad \text{(using } m^{100}(\tilde{A}) = O(R^4)) \qquad (9.81)
\end{aligned}
$$

Noticing that from $\|\mathbf{t}\|_\infty \leq h$ and $h \leq R/\sqrt{6}$, we conclude for this term. All other terms of the matrix are bounded above in a similar way. □

Theorem 10 (Multigrid Convergence of Digital Covariance Matrix) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X has reach greater than ρ.*

Then the digital covariance matrix is multigrid convergent toward the covariance matrix for the Gauss digitization on such shapes X for any radius $R < \frac{\rho}{2}$. More precisely, we have

$$
\begin{aligned}
&\forall 0 < h < \frac{R}{\sqrt{6}},\ \ \forall \mathbf{x} \in \partial X,\ \ \forall \hat{\mathbf{x}} \in \partial[\mathbb{G}_h(X)]_h \text{ with } \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h, \\
&\qquad \left\|\mathscr{V}_h(B_{R/h}(\hat{\mathbf{x}}/h) \cap \mathbb{G}_h(X)) - \mathscr{V}(B_R(\mathbf{x}) \cap X)\right\| \leq O(R^4 h)\,. \qquad (9.82)
\end{aligned}
$$

The constants hidden in the big O do not depend on the shape size or geometry. $\|\cdot\|$ denotes the spectral norm on matrices.

Proof We split the difference of matrices into two parts:

$$
\begin{aligned}
&\|\mathscr{V}_h(B_{R/h}(\hat{\mathbf{x}}/h) \cap \mathbb{G}_h(X)) - \mathscr{V}(B_R(\mathbf{x}) \cap X)\| \\
&\quad \leq \|\mathscr{V}_h(B_{R/h}(\hat{\mathbf{x}}/h) \cap \mathbb{G}_h(X)) - \mathscr{V}(B_R(\hat{\mathbf{x}}) \cap X)\| \\
&\qquad + \|\mathscr{V}(B_R(\hat{\mathbf{x}}) \cap X) - \mathscr{V}(B_R(\mathbf{x}) \cap X)\|.
\end{aligned}
$$

The first term is directly bounded by $O(R^4 h)$ with Theorem 9, since point $\hat{\mathbf{x}}$ is at distance less than $h/\sqrt{3}$ from ∂X according to Theorem 1. It remains to bound the second term. Denoting $\mathbf{t} := \hat{\mathbf{x}} - \mathbf{x}$, $t := \|\mathbf{t}\|_2$, and $X' := X - \mathbf{x}$, we simply shift everything to the origin using the invariance by translation of the covariance matrix:

$$\|\mathscr{V}(B_R(\hat{\mathbf{x}}) \cap X) - \mathscr{V}(B_R(\mathbf{x}) \cap X)\| = \|\mathscr{V}(B_R(\mathbf{t}) \cap X') - \mathscr{V}(B_R(\mathbf{0}) \cap X')\|. \tag{9.83}$$

We will apply Lemma 11 for the different moments occurring in the covariance matrix $\mathscr{V}$. We denote by $Y_\mathbf{t}$ the set $B_R(\mathbf{t}) \cap X'$ and by $Y_\mathbf{0}$ the set $B_R(\mathbf{0}) \cap X'$.

$$\begin{aligned} \|\mathscr{V}(Y_\mathbf{t}) - \mathscr{V}(Y_\mathbf{0})\| = \Bigg\| & \begin{bmatrix} m^{200}(Y_\mathbf{t}) - m^{200}(Y_\mathbf{0}) & \\ & \ddots \end{bmatrix} \\ & - \frac{1}{m^{000}(Y_\mathbf{t})} \begin{bmatrix} m^{100}(Y_\mathbf{t}) \\ \vdots \end{bmatrix} \otimes \begin{bmatrix} m^{100}(Y_\mathbf{t}) \\ \vdots \end{bmatrix}^T \\ & + \frac{1}{m^{000}(Y_\mathbf{0})} \begin{bmatrix} m^{100}(Y_\mathbf{0}) \\ \vdots \end{bmatrix} \otimes \begin{bmatrix} m^{100}(Y_\mathbf{0}) \\ \vdots \end{bmatrix}^T \Bigg\|. \end{aligned} \tag{9.84}$$

Matrix $\mathscr{V}(Y_\mathbf{t}) - \mathscr{V}(Y_\mathbf{0})$ contains differences of geometrical moments of order two (*e.g.* $m^{200}(Y_\mathbf{t}) - m^{200}(Y_\mathbf{0})$) and quantities in the form of $\Delta := \frac{m^{100}(Y_\mathbf{t})^2}{m^{000}(Y_\mathbf{t})} - \frac{m^{100}(Y_\mathbf{0})^2}{m^{000}(Y_\mathbf{0})}$ (component $(1,1)$ in $\mathscr{V}(Y_\mathbf{t}) - \mathscr{V}(Y_\mathbf{0})$ matrix). From Lemma 11, every error on second-order moments is in $O(tR^4)$. To bound Δ quantities, we first observe that $|m^{000}(Y_\mathbf{t}) - m^{000}(Y_\mathbf{0})| = \pi R^2 (t + O(t^2) + O(tR^2))$ using Theorem 7 in [49]. Hence, we get

$$\begin{aligned} \Delta &= \frac{m^{100}(Y_\mathbf{t})^2}{m^{000}(Y_\mathbf{0}) + O(tR^2)} - \frac{m^{100}(Y_\mathbf{0})^2}{m^{000}(Y_\mathbf{0})} \\ &= O(tR^2) \frac{m^{100}(Y_\mathbf{t})^2}{m^{000}(Y_\mathbf{0})^2} + \frac{m^{100}(Y_\mathbf{t})^2 - m^{100}(Y_\mathbf{0})^2}{m^{000}(Y_\mathbf{0})}. \end{aligned}$$

Since $\frac{a}{b+O(x)} = \frac{a}{b} + \frac{a}{b^2} O(x)$, $a^2 - b^2 = (a-b)(a+b)$, and using again Lemma 11, we have

$$\begin{aligned} \Delta &= O(tR^4) + (m^{100}(Y_\mathbf{t}) + m^{100}(Y_\mathbf{0})) \frac{m^{100}(Y_\mathbf{t}) - m^{100}(Y_\mathbf{0})}{m^{000}(Y_\mathbf{0})} \\ &= O(tR^4) + (O(tR^3) + O(R^4)) \frac{m^{100}(Y_\mathbf{t}) - m^{100}(Y_\mathbf{0})}{m^{000}(Y_\mathbf{0})} \end{aligned}$$

Since the boundary of $X' = X - \mathbf{x}$ goes through point $\mathbf{0}$ and is smooth, the volume of $Y_\mathbf{0} = B_R(\mathbf{0}) \cap X'$ is some $\Theta(R^3)$. Then, Lemma 11, Eq.(9.70), bounds the error

of the 100-moments. Last, we have $t < R$ since $t < \sqrt{3}h$ and $h < R/\sqrt{6}$. It follows

$$\Delta = O(tR^4) + (O(tR^3) + O(R^4))\frac{O(tR^3)}{\Theta(R^3)} = O(tR^4).$$

The same bound is found for all terms of the matrix. Putting everything together gives the result. □

9.5.4 *Useful Results of Matrix Perturbation Theory*

We have shown above that the digital covariance matrix tends toward the continuous covariance matrix, if we compute it on the intersection of the shape with a ball. We have defined curvature estimators from the eigenvalues and eigenvectors of the digital covariance matrix. To show their multigrid convergence, it is thus necessary to understand how an error on a matrix can perturb its eigendecomposition. This is why we present first useful results from matrix perturbation theory [3, 4, 19, 55] before establishing the multigrid convergence of our estimators.

Let M and M' be two symmetric matrices, we want to quantify the difference between the eigenvalues and eigenvectors of M and the eigenvalues and eigenvectors of M' as functions of norms of $M - M'$. For instance, if M' is the covariance matrix of a noisy data and M the noise-free covariance matrix, matrix perturbation results would allow us to bound eigenvalues defect (and thus principal curvature defect as described in Lemma 6).

Let $(\lambda, \boldsymbol{v})$ be an eigenpair of M, *i.e.*, λ is an eigenvalue of M and $\boldsymbol{v}$ its associated eigenvector. The eigengap $\delta_\lambda(M)$ associated to an eigenvalue λ is the minimum distance between λ and any distinct eigenvalue of M. Note that the eigengap vanished if λ has multiplicity greater than 1. We also consider the operator norm (or 2−norm) of a matrix:

$$\|M\|_{op} := \sup_{\|\mathbf{x}\|_2=1} \|M\mathbf{x}\|_2 \,, \tag{9.85}$$

please note that if $m_{max} := \max_{m_{ij} \in M} |m_{ij}|$, then $m_{max} \leq \|M\|_{op} \leq \sqrt{nm} \cdot m_{max}$ for $n \times m$ matrix M.

The main theorem we use is due to Davis-Kahan [19] presented in a simplified version due to [43]:

Theorem 11 (Davis-Kahan) *Let M and M' be two symmetric matrices, λ an eigenvalue of M and $\delta := \delta_\lambda(M)$. Then for every eigenpair $(\lambda, \boldsymbol{v})$ of M, there exists an eigenpair $(\lambda', \boldsymbol{v}')$ of M' such that:*

$$|\lambda - \lambda'| \leq \sqrt{2}\|M - M'\|_{op} \quad \textit{and} \quad \|\boldsymbol{v} - \boldsymbol{v}'\| \leq \frac{2\|M - M'\|_{op}}{\delta} \,, \tag{9.86}$$

provided that $\|M - M'\|_{op} < \delta\sqrt{2}$.

Concerning eigenvalues, stronger results can be used to *pair* eigenvalues of M to eigenvalues of M'. For instance, sorting the values by increasing order and using the rank to pair values leads to an equivalent result on the eigenvalue difference. We mention here the Lidskii-Weyl inequality, which bounds eigenvalues defect without constraint on the eigengap:

Theorem 12 (Lidskii-Weyl Inequality) *If $\lambda_i(M)$ denotes the ordered eigenvalues of some symmetric matrix M and $\lambda_i(M')$ the ordered eigenvalues of some symmetric matrix M', then* $\max_i |\lambda_i(M) - \lambda_i(M')| \leq \sqrt{2}\|M - M'\|_{op}$.

Another consequence of Theorem 11 is that the eigengap between eigenvalues plays an important role in eigenvector deviations. For example, if M is a covariance matrix of a perfectly symmetric shape around the Oz axis, one eigenvalue would have multiplicity 2. Any infinitely small perturbations of the shape breaking the symmetry would lead to large defects of the eigenvectors associated to the eigenvalues with infinitely small eigengaps (in other words, the two eigenvectors could rotate around the third one).

In the our context where eigenvalues/eigenvectors are used to estimate principal curvature and principal directions, it means that on some flat or spherical surface patches, we can expect convergence of the principal curvature values but not the direction (in the sense of the norm of the vector deviation).

9.5.5 Multigrid Convergence of Integral Principal Curvature Estimators

The multigrid convergence of (local) digital covariance matrix and stability results of matrix perturbation theory induce multigrid convergence for integral principal curvature estimators:

Theorem 13 (Multigrid Convergence of Integral Principal Curvature Estimators $\hat{\kappa}_1^R$ and $\hat{\kappa}_2^R$) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X has reach greater than ρ and has C^3-continuity.*

Then, for the Gauss digitization process, the integral principal curvature estimators are multigrid convergent toward the principal curvatures on such shapes X for well-chosen gridsteps h and radius R. More precisely, setting $R = kh^{\frac{1}{3}}$ with k an arbitrary positive constant, we have

$$\forall h \in \mathbb{R},\ 0 < h < \min\left(\left(\frac{\rho}{2k}\right)^3,\ \left(\frac{k}{\sqrt{6}}\right)^{\frac{3}{2}}\right),$$

$$\forall \mathbf{x} \in \partial X, \forall \hat{\mathbf{x}} \in \partial[\mathbb{G}_h(X)]_h \text{ with } \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h,$$

$$\|\hat{\kappa}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \kappa_1(X, \mathbf{x})\| \leq O(h^{\frac{1}{3}}), \tag{9.87}$$

$$\|\hat{\kappa}_2^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \kappa_2(X, \mathbf{x})\| \leq O(h^{\frac{1}{3}}). \tag{9.88}$$

Proof First of all, requirements on h imply that $h < \frac{R}{\sqrt{6}}$ and $R < \frac{\rho}{2}$. We are thus in the hypotheses of Theorem 10, which indicates that the digital covariance matrix is multigrid convergent to the covariance matrix. We denote $M' := \mathcal{V}_h(B_{R/h}(\hat{\mathbf{x}}/h) \cap \mathbb{G}_h(X))$ and $M := \mathcal{V}(B_R(\mathbf{x}) \cap X)$. It follows from this convergence property that:

$$\|M' - M\|_{op} \leq O(R^4 h).$$

Since both M and M' are symmetric, it follows from matrix perturbation theory (Theorem 12) that eigenvalues of M and M' are close. Recalling that $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of the covariance matrix M and $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ are the eigenvalues of the digital covariance matrix M', it follows that $\forall i \in \{1,2,3\}, |\hat{\lambda}_i - \lambda_i| < O(R^4 h)$. For instance, for $\hat{\kappa}_1^R$, we write:

$$\begin{aligned}\hat{\kappa}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) &= \frac{6}{\pi R^6}(\hat{\lambda}_2 - 3\hat{\lambda}_1) + \frac{8}{5R} \\ &= \frac{6}{\pi R^6}\left(\lambda_2 - 3\lambda_1 + O(R^4 h)\right) + \frac{8}{5R}.\end{aligned}$$

With the hypothesis of C^3-continuity, we can substitute the truncated Taylor expansion of Lemma 6 into the latter equation. After some simple calculations, we get:

$$\hat{\kappa}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) = \kappa_1(X, \mathbf{x}) + O(R) + O(h/R^2). \tag{9.89}$$

Setting $R = kh^\alpha$, we optimize the value α to minimize all errors. The optimal value is $\alpha = \frac{1}{3}$ and the bound follows. The reasoning is strictly similar for $\hat{\kappa}_2^R$. □

As noticed in [15], slightly better bounds could be obtained if the shape X has a C^3-boundary with strictly positive curvature and if the distance between $\mathbf{x}$ and $\hat{\mathbf{x}}$ could be reduced. The formulation of this theorem in [15] (Theorem 6) requires that the shape X can be decomposed into a finite number of monotonous pieces. Here, this is unnecessary since we have rewritten convergence theorems of digital moments for shapes with the sole property of having positive reach (see Sect. 9.3).

As indicated by Theorem 11, convergence of principal curvature directions is more tricky. The problem lies near places of the surface where there are umbilical points, i.e. places where principal curvatures are equal. Otherwise said, at these points, two eigenvalues coincide and their eigenvectors span a plane. We can nevertheless state:

Theorem 14 (Multigrid Convergence of Integral Principal Direction Estimators $\hat{\mathbf{w}}_1^R$ and $\hat{\mathbf{w}}_2^R$ and of Integral Normal Estimator $\hat{\mathbf{n}}^R$) *Let X be a compact domain of $\mathbb{R}^3$ such that its boundary ∂X has reach greater than ρ and has C^3-continuity.*

Then, for the Gauss digitization process, the integral principal direction estimators are multigrid convergent toward the principal directions $\mathbf{w}_1$ and $\mathbf{w}_2$ on such shapes X for well-chosen gridsteps h and radius R, provided the principal curvatures are distinct. The integral normal vector estimator is also multigrid convergent

toward the normal direction $\mathbf{n}$*. More precisely, setting* $R = kh^{\frac{1}{3}}$ *with* k *an arbitrary positive constant, we have*

$$\exists h_X \in \mathbb{R}^+,\ \forall h \in \mathbb{R},\ 0 < h < h_X,$$

$$\forall \mathbf{x} \in \partial X, \forall \hat{\mathbf{x}} \in \partial[\mathbb{G}_h(X)]_h \text{ with } \|\hat{\mathbf{x}} - \mathbf{x}\|_\infty \leq h,$$

$$\|\hat{\mathbf{w}}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \mathbf{w}_1(X, \mathbf{x})\| \leq \frac{1}{|\kappa_1(X, \mathbf{x}) - \kappa_2(X, \mathbf{x})|} O(h^{\frac{1}{3}}), \tag{9.90}$$

$$\|\hat{\mathbf{w}}_2^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \mathbf{w}_2(X, \mathbf{x})\| \leq \frac{1}{|\kappa_1(X, \mathbf{x}) - \kappa_2(X, \mathbf{x})|} O(h^{\frac{1}{3}}), \tag{9.91}$$

$$\|\hat{\mathbf{n}}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \mathbf{n}(X, \mathbf{x})\| \leq O(h^{\frac{2}{3}}). \tag{9.92}$$

NB: Involved vectors are oriented to point to the same side.

Proof As stated in Theorem 11, eigenvectors of two similar symmetric matrices are close provided the corresponding eigengap is not null. We take the same notations as in the proof of Theorem 13 and we recall that $\|M' - M\|_{op} \leq O(R^4 h)$. We denote (λ_1, ν_1), (λ_2, ν_2) and (λ_3, ν_3) the decreasing eigenpairs of M and $(\hat{\lambda}_1, \hat{\nu}_1)$, $(\hat{\lambda}_2, \hat{\nu}_2)$ and $(\hat{\lambda}_3, \hat{\nu}_3)$ the decreasing eigenpairs of M'. Since no confusion may arise, we write κ_1 for $\kappa_1(X, \mathbf{x})$ and κ_2 for $\kappa_2(X, \mathbf{x})$. We start by expressing the eigengaps $\delta_i(M)$ using Lemma 6:

$$\begin{aligned}
\delta_1(M) &:= \lambda_1 - \lambda_2 = \frac{\pi}{24}(\kappa_2 - \kappa_1)R^6 + O(R^7),\\
\delta_3(M) &:= \lambda_2 - \lambda_3 = \frac{3\pi}{32}R^5 - \frac{\pi}{1536}(69\kappa_2 + 5\kappa_1)R^6 + O(R^7),\\
\delta_2(M) &:= \min(\delta_1(M), \delta_3(M)).
\end{aligned}$$

As R approaches zero as h tends toward zero, it is clear that there exists some $h_X > 0$ such that $\delta_1(M) \leq \delta_3(M)$. Hence we suppose hereafter that $\delta_2(M) = \delta_1(M)$.

If $\kappa_1 \neq \kappa_2$, it follows from Theorem 11 that

$$\begin{aligned}
\|\hat{\nu}_1 - \nu_1\| &\leq \frac{2}{\delta_1(M)}\|M - M'\|_{op}\\
&\leq \frac{48}{\pi(\kappa_2 - \kappa_1)} O\left(\frac{h}{R^2}\right) + O\left(\frac{h}{R}\right).
\end{aligned} \tag{9.93}$$

By Definition 6, $\hat{\nu}_1 = \hat{\mathbf{w}}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h)$. Second, Theorem 3 of [48] tells that the eigenvectors of M are close to the directions of the principal curvatures and of the normal. More precisely, we have:

$$\angle(\nu_1, \mathbf{w}_1(X, \mathbf{x})) = O\left(\frac{R}{\kappa_2 - \kappa_1}\right),$$
$$\angle(\nu_2, \mathbf{w}_2(X, \mathbf{x})) = O\left(\frac{R}{\kappa_2 - \kappa_1}\right),$$
$$\angle(\nu_3, \mathbf{n}(X, \mathbf{x})) = O\left(R^2\right).$$

Third, for unit vectors $\mathbf{u}$ and $\mathbf{v}$, $\|\mathbf{u} - \mathbf{v}\| = 2 \sin \frac{\alpha}{2}$ where α is the angle between $\mathbf{u}$ and $\mathbf{v}$. If $0 \leq \alpha \leq \frac{\pi}{2}$, it is thus straightforward to check that $\frac{\alpha}{\sqrt{2}} \leq \sin \alpha \leq \alpha$ and $\frac{\alpha}{\sqrt{2}} \leq \|\mathbf{u} - \mathbf{v}\| \leq \alpha$. Putting these three facts together with Eq.(9.93) gives:

$$\begin{aligned}
\left\|\hat{\mathbf{w}}_1^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \mathbf{w}_1(X, \mathbf{x})\right\| &\leq \|\hat{\nu}_1 - \nu_1\| + \|\nu_1 - \mathbf{w}_1(X, \mathbf{x})\| \\
&\leq \frac{1}{|\kappa_1 - \kappa_2|} O\left(\frac{h}{R^2}\right) + O\left(\frac{R}{\kappa_2 - \kappa_1}\right), \\
&\leq \frac{1}{|\kappa_1 - \kappa_2|} O\left(h^{\frac{1}{3}}\right),
\end{aligned}$$

since $R = kh^{\frac{1}{3}}$. The proof for the second principal direction is strictly similar, since $\delta_2(M) = \delta_1(M)$. For the normal estimator, we use the relation on $\delta_3(M)$ to get:

$$\begin{aligned}
\|\hat{\nu}_3 - \nu_3\| &\leq \frac{2}{\delta_3(M)} \|M - M'\|_{op} \\
&\leq \frac{48}{3\pi} O\left(\frac{h}{R}\right) + O(h).
\end{aligned}$$

Since $\hat{\nu}_3 = \hat{\mathbf{n}}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h)$ by definition and with the same reasoning as above, we derive:

$$\begin{aligned}
\|\hat{\mathbf{n}}^R(\mathbb{G}_h(X), \hat{\mathbf{x}}, h) - \mathbf{n}(X, \mathbf{x})\| &\leq \|\hat{\nu}_3 - \nu_3\| + \|\nu_3 - \mathbf{n}(X, \mathbf{x})\| \\
&\leq O\left(\frac{h}{R}\right) + O(R^2) = O\left(h^{\frac{2}{3}}\right),
\end{aligned}$$

since $R = kh^{\frac{1}{3}}$. □

To conclude this section, we have shown that digital integral invariants provide convergent estimates of curvatures values and directions in 3D. Furthermore, they provide also a 3D normal estimator to digital surfaces, whose worst case convergence speed is very fast compared to the literature (see [12] for a survey). Last, everything has been presented for the 2D and 3D case, but most of the results remain valid in arbitrary dimension. This is due to the fact that we have shown the

convergence of digital moments in arbitrary dimension, and other stability results come from matrix perturbation theory, which are also valid in arbitrary dimension. Only the Taylor expansion of integral invariant in general dimension is missing. We will show in the next sections that these estimators compare well with other estimators in practice.

9.6 Parameter-Free Digital Curvature Estimation

In the multigrid convergence framework, the notion of scale is explicit with parameter h. For instance, convergence results can be obtained for any radius integration kernel R in $O(h^{\frac{1}{3}})$. In some practical situations, we have to analyze a given digital object $Z \subset \mathbb{Z}^d$ without any information about its Euclidean embedding. As a consequence, setting suitable values for R can be challenging. Note that beside integral invariant estimators described in the previous sections, most existing curvature estimators are also parametrized by such kernel size or window size parameter.

In this section, we propose techniques to automatically set such radius parameter for a given digital object Z in dimension 2 and 3 with the property that if Z comes from a multigrid digitization process, then we remain in the convergence theorem hypothesis. For short, the key point here is to rely on a geometrical information extracted from the digital contour, here maximal digital straight segments, from which multigrid parameter h can be retrieved. Let us first formalize this property.

9.6.1 Asymptotic Laws of Straight Segments Along the Boundary of Digitized Shapes and Scale Determination

First of all, we remind the definition of digital straight segment [29, 51].

Definition 7 (Standard Line and Digital Straight Segment) The set of points $(x, y) \in \mathbb{Z}^2$ satisfying $\mu \leq ax - by < \mu + |a| + |b|$, with a, b and μ integer numbers, is called the *standard digital line* with slope a/b and shift μ. Any connected subset of pixels of a standard digital line is a *digital straight segment* (DSS for short).

On a digital set $Z \subset \mathbb{Z}^2$, we denote by $Bd(Z)$ its (cellular) topological boundary composed of pointels (0-cell) and linels (1-cell). Following the discussion in Sect. 9.2.3, the canonical embedding of $Bd(Z)$ into $\mathbb{R}^2$ coincides with $\partial[Z]_h$. On $Bd(Z)$, we can define maximal segments and maximal segment pencils:

Definition 8 (Maximal Segment and Maximal Segment Pencil [37]) The pointels composing the digital boundary $Bd(Z)$ form a 4-connected contour. They can thus be numbered consecutively as $(\mathbf{p}_i)_{i=0\ldots n-1}$. A sequence of pointels $(\mathbf{p}_i, \ldots, \mathbf{p}_j)$, indices taken modulo n, is a *maximal segment* on ∂Z iff $\{\mathbf{p}_i, \ldots, \mathbf{p}_j\}$ is a DSS, while neither $\{\mathbf{p}_{i-1}, \mathbf{p}_i, \ldots, \mathbf{p}_j\}$ nor $\{\mathbf{p}_i, \ldots, \mathbf{p}_j, \mathbf{p}_{j+1}\}$ are DSS. At a given pointel

$\mathbf{p} \in Bd(Z)$, the *pencil of maximal segment* at $\mathbf{p}$ is the set of maximal segments on $Bd(Z)$ containing $\mathbf{p}$.

In many situations, maximal segments and maximal segment pencils play a very important role in multigrid digital contour geometry processing [34, 57]. For the purpose of this paper, let us focus on the asymptotic properties of lengths of maximal segment:

Lemma 12 (Asymptotic Laws of Maximal Segments [34, 57]) *Let X be some convex shape of $\mathbb{R}^2$, with at least C^3-boundary and non-null bounded curvature. The discrete length of maximal segments in $Bd(Z)$ for $Z = \mathbb{G}_h(X)$ follows:*

- *the shortest is lower bounded by $\Omega(h^{-\frac{1}{3}})$;*
- *the longest is upper bounded by $O(h^{-\frac{1}{2}})$;*
- *their average length, denoted $L_D(Z)$, is such that:*

$$\Theta\left(h^{-\frac{1}{3}}\right) \leq L_D(Z) \leq \Theta\left(h^{-\frac{1}{3}} \log\left(\frac{1}{h}\right)\right). \tag{9.94}$$

We now have the key ingredient for the scale inference of Z (if Z is the digitization of a C^3 Euclidean shape): If we compute the average length of all maximal segments in $Bd(Z)$, Eq. (9.94) allows us to retrieve the scale parameter h. We can now introduce our parameter free curvature estimator in 2D:

Definition 9 Given $Z \subset \mathbb{Z}^2$, the *parameter-free digital curvature estimator* $\hat{\kappa}^*$ at a pointel $\mathbf{p} \in Bd(Z)$ is defined as:

$$\hat{\kappa}^*(Z, \mathbf{p}) := \frac{3\pi}{2\rho(Z)} - \frac{3A(Z, \mathbf{p})}{\rho(Z)^3}, \tag{9.95}$$

where $\rho(Z) = L_D^2(Z)$ and $A(Z, \mathbf{p}) = \mathrm{Card}(B_{\rho(Z)}(\mathbf{p}) \cap Z)$.
To rephrase the definition, we first compute the average discrete length of all maximal segments on $Bd(Z)$. Then ρ is the square of this length. The estimation $\hat{\kappa}^*(Z, \mathbf{p})$ is a function of the number of digital points in Z intersected with the ball of radius ρ centered at $\mathbf{p}$. This definition mimics continuous definition in Eq. (9.14) and is a scale-independent version of Definitions 3 and 4 (page 314).

9.6.2 Parameter-Free Digital Curvature Estimators

In dimension 2, $\hat{\kappa}^*(Z, \mathbf{p})$ is parameter-free and only related to a digital object Z. However, if Z is the digitization of a continuous shape X, *i.e.* if $Z = \mathbb{G}_h(X)$, then multigrid convergence is obtained:

Theorem 15 (Multigrid Convergence of Curvature Estimator $\hat{\kappa}^*$ [39]) *Let X be some convex shape of $\mathbb{R}^2$, with at least C^3-boundary and non null bounded*

curvature. Let $Z = \mathbb{G}_h(X)$. *Then, there exist a positive constant* h_0, *for any* $0 < h \leq h_0$, *we have,* $\forall \mathbf{x} \in \partial X$ *and* $\forall \mathbf{p} \in Bd(Z)$

$$\|h\mathbf{p} - \mathbf{x}\|_\infty \leq h \Rightarrow \left| \frac{1}{h}\hat{\kappa}^*(Z, \mathbf{p}) - \kappa(X, \mathbf{x}) \right| \leq O\left(h^{\frac{1}{3}} \log^2\left(\frac{1}{h}\right)\right) . \tag{9.96}$$

Note that $\mathbf{p} \in Bd(Z)$ implies $h\mathbf{p} \in \partial[\mathbb{G}_h(X)]_h$. The parameter-free curvature is rescaled by h in order to compare comparable shapes. To sketch the proof, Eq. (9.94) implies that $h\rho(Z)$ is in $O\left(h^{\frac{1}{3}} \log^2\left(\frac{1}{h}\right)\right)$. Theorem 7 (page 318) only requires the kernel radius R to be in $O(h^{\frac{1}{3}})$ to achieve multigrid convergence. In [39], we show that considering $R = O\left(h^{\frac{1}{3}} \log^2\left(\frac{1}{h}\right)\right)$ does not change the convergence property.

9.6.3 *Local Parameter-Free Digital Curvature Estimator and 3D Case*

Several extensions can be considered. The first one is a local definition of $\hat{\kappa}^*(Z, \mathbf{p})$. Indeed, Definition 9 imposes the same radius parameter for each point of $Bd(Z)$ (average of all maximal segment lengths). We may be interested in a local version to obtain adaptive estimations. For instance, a local version is easily defined by considering, at each point p, the average length of DSS in its maximal segment pencil, denoted $\rho(Z, \mathbf{p})$. Hence, we define:

$$\hat{\kappa}_l^*(Z, \mathbf{p}) := \frac{3\pi}{2\rho(Z, \mathbf{p})} - \frac{3A'(Z, \mathbf{p})}{\rho(Z, \mathbf{p})^3} , \tag{9.97}$$

where $A'(Z, \mathbf{p}) = \text{Card}(B_{\rho(Z,\mathbf{p})}(\mathbf{p}) \cap Z)$.

Doing so, we cannot prove anymore the multigrid convergence of $\frac{1}{h}\hat{\kappa}_l^*(Z, \mathbf{p})$ since in the maximal pencil, we may have pathological DSS with too long length (those in $O(h^{-\frac{1}{2}})$ in Lemma 12). However, experimental evaluation shows good convergence properties and we observe that $\hat{\kappa}_l^*(Z, \mathbf{p})$ outperforms $\hat{\kappa}^*(Z, \mathbf{p})$.

A second extension considers the 3D case with $Z \subset \mathbb{Z}^3$. In this case, we first want to construct a mean curvature estimator setting the radius kernel R to geometrical characteristics, denoted ρ', of $Bd(Z)$. A first issue is that in dimension 3, even if digital planes can be defined, no result similar to Lemma 12 exists. In [39], we have presented a slice based approach: intersecting Z with planes aligned with grid axis defines a set of 2D digital curves on which maximal DSS can be computed and thus average maximal DSS length can be obtained. Similarly, at a given surfel s, two specific 2D curves can be defined and thus local parameter-free estimator can be defined considering average DSS length in the two maximal segment pencils containing the projection of s. As in the local 2D case, some pathological configurations may occur preventing us to have a complete convergence proof of such estimators (see details in [39]). However, experimental convergence can be observed. The following section details these results.

9.7 Experimental Evaluation

In this section, all presented estimators are evaluated experimentally and compared with state-of-the-art techniques.

9.7.1 Implementation Details

Integral invariant estimators are based on spherical kernels with radius given by $R = kh^\alpha$ as described in theorem statements. Then, the digital object boundary is tracked and the kernel is centered on each surface elements. For 2D and 3D mean curvature estimators, the volumetric integral of the intersection between the kernel and the object is computed; for 3D principal curvature estimators, the covariance matrix of this intersection is computed and then eigenvalues and eigenvectors are deduced from it by diagonalization.

A brute-force implementation would lead to a computational cost of $O((R/h)^d)$ per surface element (i.e. the size of the kernel at grid step h). However, all quantities are additive and we can take advantage of the digital surface structure to speed up this algorithm considerably: if we consider a surface tracker for which surface elements are processed by proximity (the current surface element is a neighbor of the previous one through a translation vector $\boldsymbol{\delta}$), the area/volume estimation can be done incrementally:

$$\begin{aligned}\widehat{\text{Area}}\,(\mathbb{G}_h(X) \cap B_R(\mathbf{x}+\boldsymbol{\delta}), h) &= \widehat{\text{Area}}\,(\mathbb{G}_h(X) \cap B_R(\mathbf{x}), h)\\ &+ \widehat{\text{Area}}\,(\mathbb{G}_h(X) \cap (B_R(\mathbf{x}+\boldsymbol{\delta}) \setminus B_R(\mathbf{x})), h)\\ &- \widehat{\text{Area}}\,(\mathbb{G}_h(X) \cap (B_R(\mathbf{x}) \setminus B_R(\mathbf{x}+\boldsymbol{\delta})), h)\ .\end{aligned}$$

Similarly we have for moments:

$$\begin{aligned}\hat{m}_{p,q,s}\,(\mathbb{G}_h(X) \cap B_R(\mathbf{x}+\boldsymbol{\delta}), h) &= \hat{m}_{p,q,s}\,(\mathbb{G}_h(X) \cap B_R(\mathbf{x}), h)\\ &+ \hat{m}_{p,q,s}\,(\mathbb{G}_h(X) \cap (B_R(\mathbf{x}+\boldsymbol{\delta}) \setminus B_R(\mathbf{x})), h)\\ &- \hat{m}_{p,q,s}\,(\mathbb{G}_h(X) \cap (B_R(\mathbf{x}) \setminus B_R(\mathbf{x}+\boldsymbol{\delta})), h)\ .\end{aligned}$$

Then, if we precompute all kernels $\mathbb{G}_h(B_R(0 \pm \boldsymbol{\delta}) \setminus B_R(0))$ for some $\boldsymbol{\delta}$ displacements (based on surface element umbrella configurations, 8 in 2D and 26 in 3D for $\|\boldsymbol{\delta}\|_\infty = h$), the computational cost per surface element can be reduced to $O((R/h)^{d-1})$. Finally, in the ideal case of a Hamiltonian traversal of the surface, only the first surfel has to be computed using kernel $B_R(\hat{\mathbf{x}})$ and every subsequent neighboring surfel is processed using sub-kernels $\mathbb{G}_h(B_R(0 \pm \boldsymbol{\delta}) \setminus B_R(0))$.

To perform precise performance evaluation in both the multigrid framework and with respect to the state of the art, we need a family of Euclidean shapes $\mathbb{X}$ on which the estimated quantity is known. Table 9.1 and Fig. 9.5 present continuous

Table 9.1 Equations, parameters and domains of Euclidean shapes considered in the experimental evaluation ($t \in [0, 2\pi]$ for parametric curves)

Shape	Equation (parametric in 2D, implicit in 3D)	Parameters	Domain	k_{min}	k_{max}
Ellipse	$(x(t), y(t)) = (\rho(t) \cdot \cos(t), \rho(t) \cdot \sin(t))$ with $\rho(t) = \frac{b}{\sqrt{1-(a^2-b^2)/a^2 \cdot \cos(t+\phi)}}$	$(a, b) = (20, 7)$	$[-20, 20]^2$	0.018	0.408
Flower	$(x(t), y(t)) = (\rho(t) \cdot \cos(t), \rho(t) \cdot \sin(t))$ with $\rho(t) = r_1 + r_2 \cdot \cos(p \cdot t)$	$(r_1, r_2, p) = (20, 7, 6)$	$[-20, 20]^2$	-1.414	0.383
AccFlower	$(x(t), y(t)) = (\rho(t) \cdot \cos(t), \rho(t) \cdot \sin(t))$ with $\rho(t) = r_1 + r_2 \cdot \cos(p \cdot t^3)$	$(r_1, r_2, p) = (20, 5, 3)$	$[-20, 20]^2$	-10.45	3.1482
Sphere	$x^2 + y^2 + z^2 - a^2 = 0$	$a = 9$	$[-10, 10]^3$	0.111	0.111
Rounded cube	$x^4 + y^4 + z^4 - a^4 = 0$	$a = 9$	$[-10, 10]^3$	0	0.282
Goursat's surface	$ax^4 + ay^4 + az^4 + bx^2 + by^2 + bz^2 + c = 0$	$(a, b, c) = (0.03, -2, -8)$	$[-10, 10]^3$	-0.15	0.453

Please refer to Fig. 9.5 for illustrations

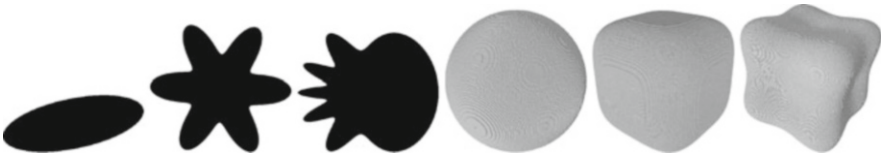


Fig. 9.5 Illustrations of 2D and 3D shapes considered in the experimental evaluation (please refer to Table 9.1 for equations and parameters): ellipse (**a**), flower (**b**), accelerated flower (**c**), sphere (**d**), rounded cube (**e**) and Goursat's surface (**f**)

shapes considered in this analysis. Please note that, for parameter-free curvature estimators, only convex ones match with theorem statements. However, we can still experimentally evaluate the behavior of estimators when considering shapes that do not satisfy all theorem hypotheses.

All curvature estimators have been implemented in DGTAL.[4] DGTAL is an opensource library devoted to digital geometry tools. Beside proposing curvature integral invariant based estimators, this library offers mathematical shapes with known curvature values and Gauss digitization of such objects on a grid with gridstep h. Furthermore, many existing curvature estimators from the literature are also available, making easy comparative evaluations.

[4]The DGTAL library is available at http://dgtal.org.

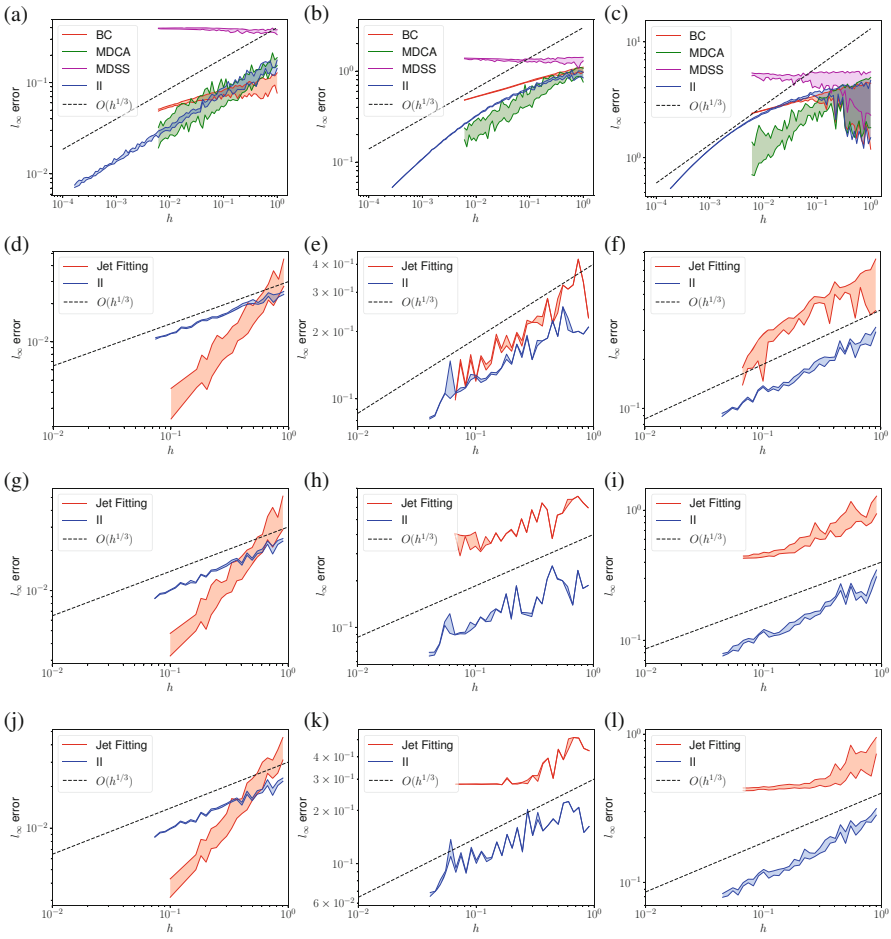


Fig. 9.6 Multigrid analysis of the estimation error with l_∞ norm in 2D and 3D. (**a**) Ellipse-$\hat{\kappa}^R$-l_∞. (**b**) Flower-$\hat{\kappa}^R$-l_∞. (**c**) AccFlower-$\hat{\kappa}^R$-l_∞. (**d**) Sphere-$\hat{H}^R$-l_∞. (**e**) RoundedCube-$\hat{H}^R$-l_∞. (**f**) Goursat-$\hat{H}^R$-l_∞. (**g**) Sphere-$\hat{\kappa}_R^1$-l_∞. (**h**) RoundedCube-$\hat{\kappa}_R^1$-l_∞. (**i**) Goursat-$\hat{\kappa}_R^1$-l_∞. (**j**) Sphere-$\hat{\kappa}_R^2$-l_∞. (**k**) RoundedCube-$\hat{\kappa}_R^2$-l_∞. (**l**) Goursat-$\hat{\kappa}_R^2$-l_∞

9.7.2 Multigrid Convergence Analysis

We have first checked experimentally that the α parameter (for the ball radius $R = kh^\alpha$) should indeed be set around $\frac{1}{3}$ to get multigrid convergence. This was empirically observed in [15], where a complete discussion on convergence behavior for different α values is detailed. In all following experiments, we have set α to $\frac{1}{3}$.

In Figs. 9.6 and 9.7, we compare our digital integral invariant estimators (II) to state-of-the-art methods for respectively the l_∞ and l_2 error norms. In dimension 2,

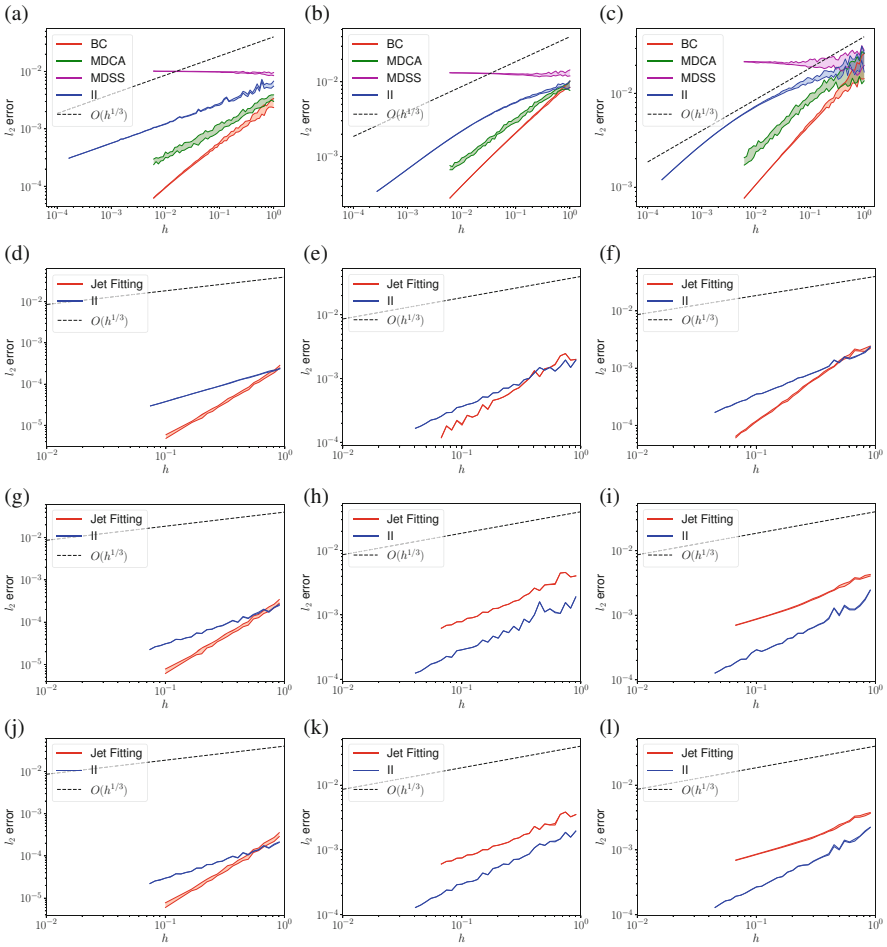


Fig. 9.7 Multigrid analysis of the estimation error with l_2 norm in 2D and 3D. (**a**) Ellipse-$\hat{\kappa}^R$-l_2. (**b**) Flower-$\hat{\kappa}^R$-l_2. (**c**) AccFlower-$\hat{\kappa}^R$-l_2. (**d**) Sphere-$\hat{H}^R$-l_2. (**e**) RoundedCube-$\hat{H}^R$-l_2. (**f**) Goursat-$\hat{H}^R$-l_2. (**g**) Sphere-$\hat{\kappa}^1_R$-l_2. (**h**) RoundedCube-$\hat{\kappa}^1_R$-l_2. (**i**) Goursat-$\hat{\kappa}^1_R$-l_2. (**j**) Sphere-$\hat{\kappa}^2_R$-l_2. (**k**) RoundedCube-$\hat{\kappa}^2_R$-l_2. (**l**) Goursat-$\hat{\kappa}^2_R$-l_2

other curvature estimators are: curvature from *Most-centered Maximal Segment* with length information (MDSS) [13, 57], curvature from *Most-centered Digital Circular Arc* (MDCA) [53] and *Binomial based convolution* (BC) [22]. In dimension 3, we have considered the curvature estimation from polynomial surface approximation (Jet Fitting) [10]. For the latter, we have also chosen a window size in $kh^{\frac{1}{3}}$. In Fig. 9.8, we detail timings in logscale for various estimators on the flower shape in 2D and the rounded cube in 3D. As expected, approaches based on object recognition in dimension 2 (MDSS and MDCA) provide faster computations. We

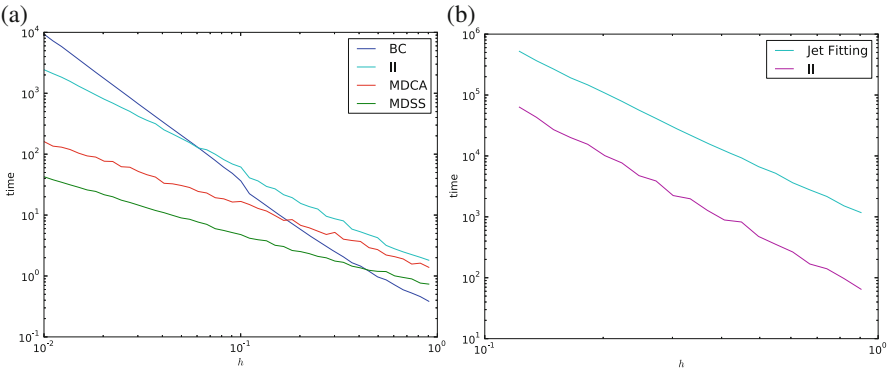


Fig. 9.8 Timings in milliseconds for 2D estimators on a flower (**a**) and 3D estimators on a rounded cube (**b**). Results have been obtained on a Intel Xeon 2.27 GHz desktop machine

also observe that II is a bit slower but has an asymptotic behavior much more favorable that BC. In dimension 3 (Fig. 9.8b), we observe that Jet fitting and II behaviors are similar and that II is 10 times faster than our implementation of Jet fitting.[5]

Finally, Fig. 9.9 illustrates various curvature maps on binary shapes.

9.7.3 Parameter-Free Estimators

As described in Sect. 9.6, the ball radius can be deduced by the contour/surface geometry in order to design a parameter-free estimator. More precisely, in dimension 2, setting the radius to the square of the average length of all maximal DSS allows us to keep a multigrid convergence property (estimator $\hat{\kappa}^*(Z,p)$ in Theorem 15). Furthermore, this ball radius can be defined locally (estimator $\hat{\kappa}_l^*(Z,p)$) to capture local behavior of the contour and thus reduce the l_2 error. The main drawback of the locally adapted estimator is that since radius may change for each surfel, we cannot use anymore the incremental propagation of area/volume or moments as described in Sect. 9.7.1. We can define an intermediate estimator based on a quantification of all local ball radii along the contour using for instance a k-means approach ($\hat{\kappa}^*_{K=5}$ denotes the local curvature estimator defined from a quantification of ball radii into five classes). Figures 9.10 and 9.12 present convergence results in dimension 2 and 3 respectively. As expected, we observe the multigrid convergence of parameter-free estimators for both l_∞ and l_2 error metrics. Furthermore, we can observe that local approaches (and local approaches

[5]For an implementation, we refer to the CGAL library available at http://www.cgal.org.

Fig. 9.9 Integral invariant based differential estimators on a digital surface (256^3 OctaFlower shape). *From left to right*, mean curvature, Gaussian curvature, first principal direction, second principal direction and normal vector field (zooms in third row)

based on a quantification) induce smaller l_2 errors. This means that better estimation can be obtained if we locally adapt the ball radius to the curve/surface geometry. Represented in scale-scale, Fig. 9.11 shows the scale that has been selected for a 2D `flower` shape (Fig. 9.12). Finally, in Fig. 9.13, we show the multiresolution behavior of $\hat{\kappa}^*(Z, p)$. Indeed, since the parameter depends only on the object geometry, we can obtain consistent curvature estimation whichever is the object resolution.

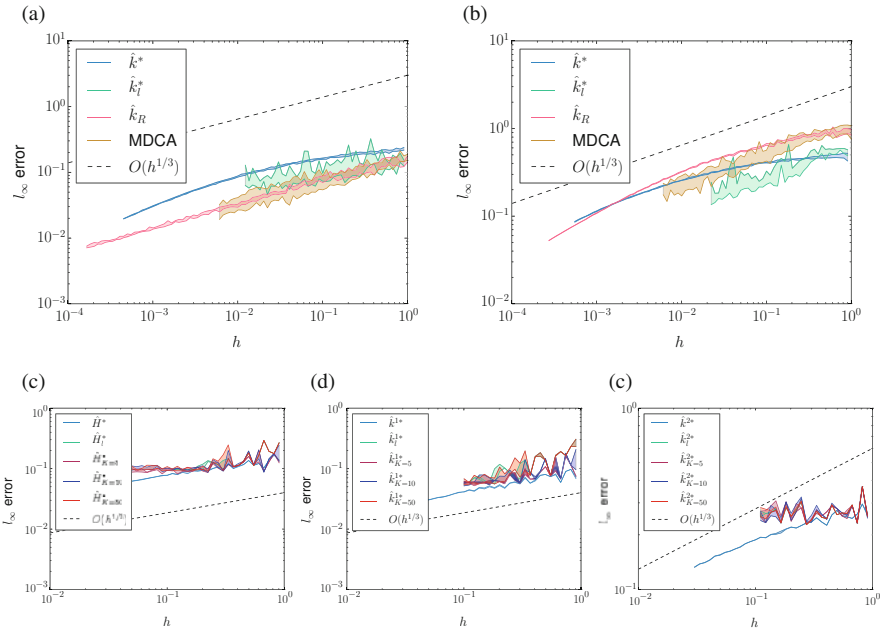


Fig. 9.10 Comparison in log-space of parameter-free curvature l_∞ errors in dimension 2 (*first row*), and parameter-free mean and principal curvatures (*second row*) l_∞ errors on a multigrid ellipsoid. (**a**) Ellipse-l_∞. (**b**) Flower-l_∞. (**c**) Ellipsoid-l_∞. (**d**) Ellipsoid-l_∞. (**e**) Ellipsoid-l_∞

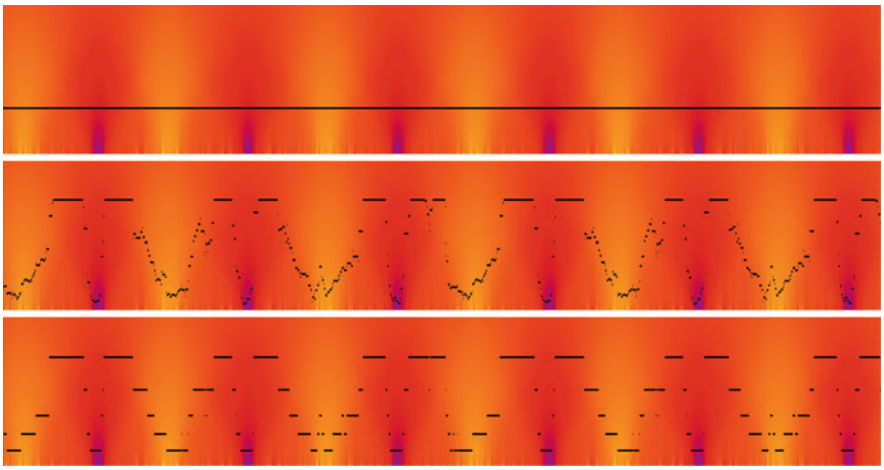


Fig. 9.11 Curvature scale-space analysis of a flower: x-axis is the curvilinear abscissa, y-axis is the kernel radius, curvature values are mapped between the *blue* (lowest curvature) and the *yellow color* (highest curvature). In black are drawn the radius $\rho(Z)$ for global estimator $\hat{\kappa}^*$ (*first row*), radii $\rho(Z,p)$ for local estimator $\hat{\kappa}_l^*$ (*second row*), and radii $\rho(Z,p)$ after K-mean clustering for local estimator $\hat{\kappa}_{K=5}^*$. (*last row*)

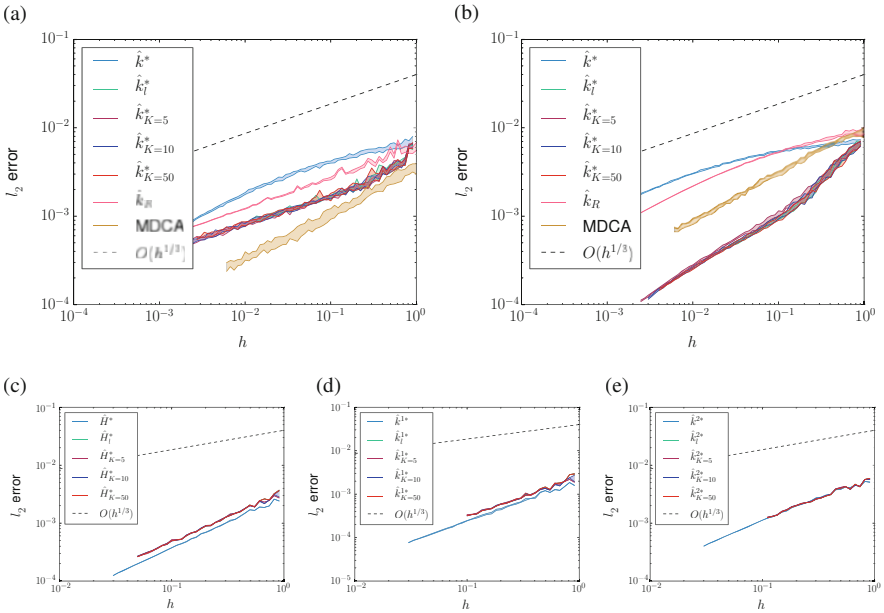


Fig. 9.12 Comparison in log-space of parameter-free curvature l_2 errors in dimension 2 (*first row*), and parameter-free mean and principal curvatures (*second row*) l_2 errors on a multigrid ellipsoid. (**a**) Ellipse-l_2. (**b**) Flower-l_2. (**c**) Ellipsoid-l_2. (**d**) Ellipsoid-l_2. (**e**) Ellipsoid-l_2

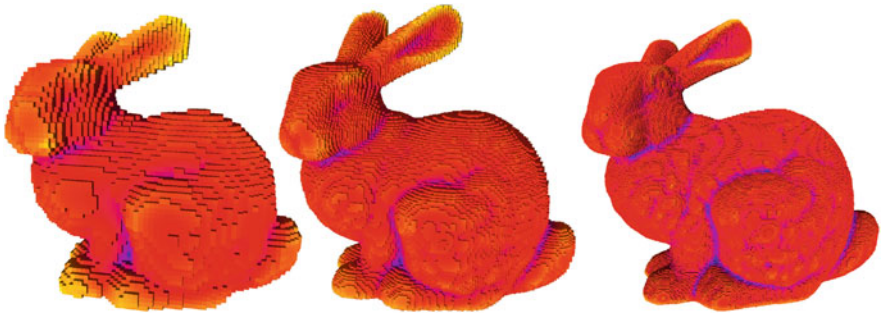


Fig. 9.13 (*Left*) Mean curvature mapped on "bunny" at different resolution using $\hat{H}_l^*$ (*yellow color is the highest curvature, blue the lowest*)

9.8 Discussion, Applications and Further Works

9.8.1 Robustness to Noise

Thanks to the volumetric integration principle of Integral Invariant estimators, robustness to noise and outliers can be expected. Intuitively, if the noise is modeled

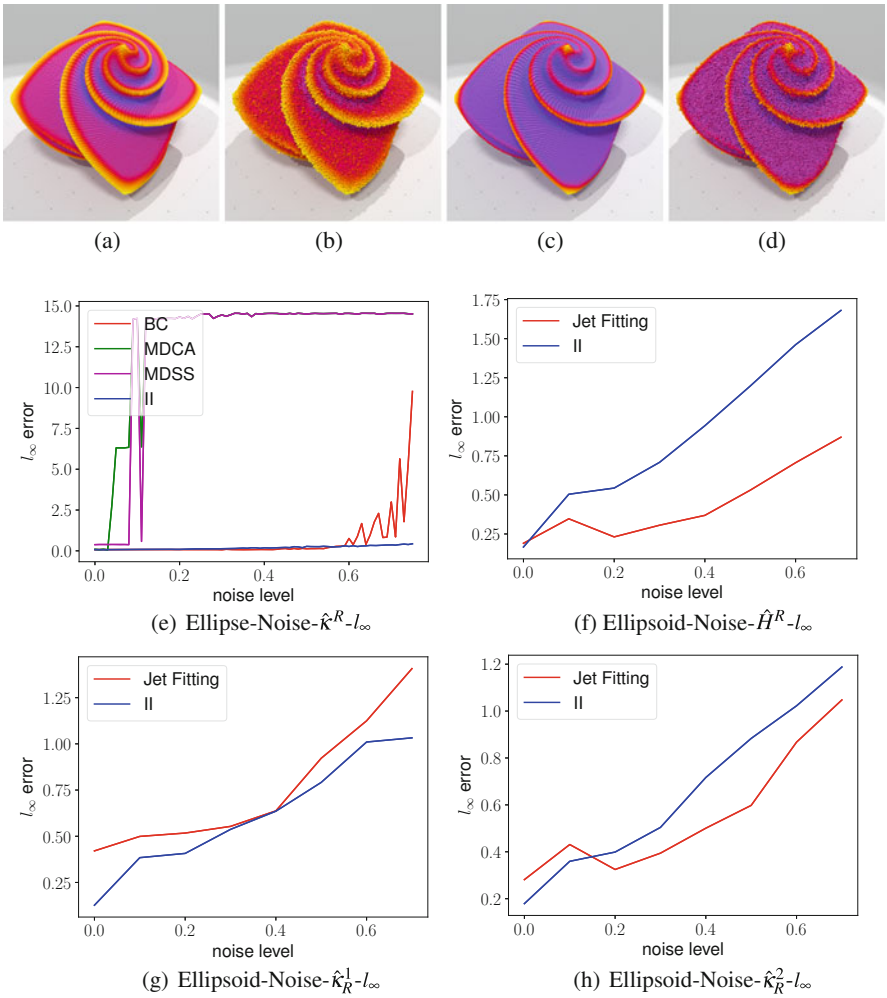


Fig. 9.14 Robustness of II estimator on the OCTAFLOWER shape: mean, (**a**)–(**b**), and Gaussian curvature, (**c**)–(**d**), estimation on a noise-free and noisy surface (noise level=0.5). Comparison of 2D estimators on different level of noise on an ellipse (**e**). Comparison of 3D estimators for mean (**f**) and principal curvatures (**g**) and (**h**) on different level of noise on an ellipsoid

as a zero-mean noise around the object surface or if outliers are of measure zero, geometrical moments would be stable.

Experimentally, robustness in dimension 2 and 3 can be observed in Fig. 9.14. Our noise model consists in swapping the grid point value at p with probability defined by a power law $\beta^{1+dt(p)}$ for some user-specified $\beta \in [0,1]$ ($dt(p)$ corresponds to the distance of p to the boundary of the original digital shape). Such noise model, so-called KANUNGO noise [27], is particularly well-adapted to evaluate the stability of digital geometry algorithms. As expected, integral invariants

approaches provide robust estimation as the noise parameter increases. Since Jet Fitting considers a principal component analysis of the point set, robustness to noise can also be observed for this approach. In dimension 2, methods which rely on the geometry of the digital contour (MDCA, MDSS) are highly perturbated, even for limited noise parameters.

9.8.2 *Feature Detection with Multiscale Approach*

As discussed in previous sections, integration radius should be set properly to have relevant curvature estimations. The radius should be set either by the user or automatically thanks to the parameter-free approach. In this section, we use the radius as a scale-space parameter to detect features on digital surfaces. First of all, as described in Eqs. (9.12) and (9.13), area and volume at a point $\mathbf{x}$ on a smooth manifold ∂X are related to curvature in 2D and mean curvature in 3D. If we consider now a piece-wise smooth surface X and a point $\mathbf{x}$ lying on a C^1-discontinuity of ∂X, Pottmann et al. [48, 49] have shown that area and volume are related to left/right side curvatures but also solid angle of normal vectors at $\mathbf{x}$ (see Fig. 9.15):

$$A_R(\mathbf{x}) = \frac{\alpha_0 R^2}{2} - \frac{\kappa_- + \kappa_+}{6} R^3 + O(R^4)\,, \tag{9.98}$$

$$V_R(\mathbf{x}) = \frac{2\alpha_0 R^3}{3} - \frac{\pi(H_- + H_+)}{8} R^4 + O(R^5)\,. \tag{9.99}$$

If we define now the following quantities:

$$G_{X,\mathbf{x}}(R) := \frac{3\pi}{2R} - \frac{3A_R(\mathbf{x})}{R^3}, \quad \mathscr{G}_{X,\mathbf{x}}(R) := \frac{8}{3R} - \frac{4V_R(\mathbf{x})}{\pi R^4}\,. \tag{9.100}$$

At a smooth C^3 point $\mathbf{x}$, these quantities converge to $\kappa(X, \mathbf{x})$ and $H(c, \mathbf{x})$ respectively as R tends to zero (see Sect. 9.2.3, Eq. (9.15)). On the contrary, at singular point $\mathbf{x}$,

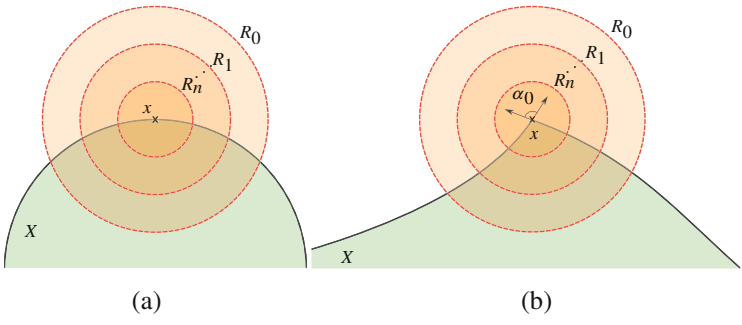


Fig. 9.15 Scale-space analysis on a spherical shape (**a**) and a shape with a singularity (**b**)

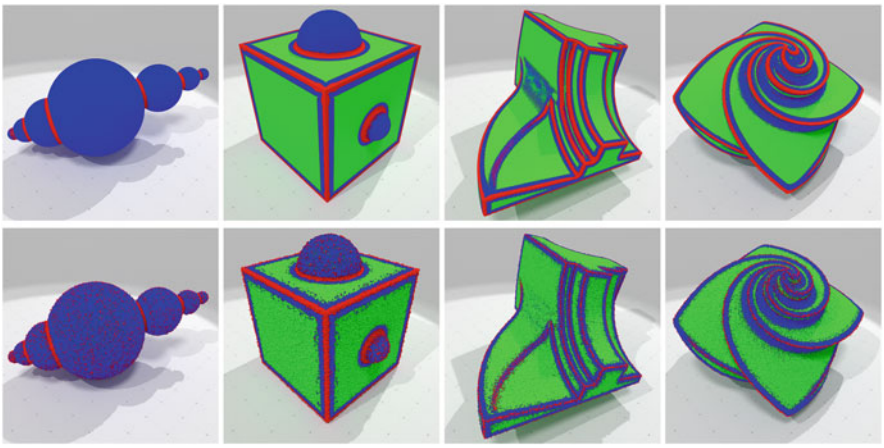


Fig. 9.16 Evaluation of feature detection on perfectly digitized and noisy shapes. SpheresUnion: $400 \times 200 \times 200$ voxels, CubeSphere: 200^3 voxels, Fandisk: 512^3 voxels, OctaFlower: 512^3 voxels. The range of radii used for all classifications is given by $r_{min} = 5$ and $r_{max} = 25$

we have:

$$G_{X,\mathbf{x}}(R) = \frac{3}{2}\frac{1}{R}(\pi - \alpha_0) + \frac{\kappa_- + \kappa_+}{6} + O(R)\,, \tag{9.101}$$

$$\mathscr{G}_{X,\mathbf{x}}(R) = \frac{8}{3}\frac{1}{R}(1 - \frac{\alpha_0}{\pi}) + \frac{H_- + H_+}{2} + O(R)\,. \tag{9.102}$$

In other words, as R tends to zero, these quantities have a dominant $\frac{1}{R}$ term. We do not go further into details, please refer to [40] for a complete state-of-the-art discussion and mathematical insights but feature detector on digital surfaces can be defined looking to the behavior of $G_{X,\mathbf{x}}(R)$ and $\mathscr{G}_{X,\mathbf{x}}(R)$ for a given range of radii: if the quantities remain constant, we classify $\mathbf{x}$ as belonging to a smooth part of the object. If the quantities follow $\Theta(R^{-1})$ speed, we classify $\mathbf{x}$ as belonging to an edge. In Fig. 9.16 we present some classification results into Edge (red), Flat (green) and Smooth (blue) classes on both noise-free and noisy data for the same set of parameters.

9.8.3 Current Limitations and Possible Research Directions

We have presented a whole set of estimators based on integral invariant principle that can be applied to analyze the local geometry of digital shapes: mean and principal curvatures, normal and principal directions. We have shown that they achieve multigrid convergence for some classes of shapes et we have made explicit their convergence speed. Furthermore, these estimators compare favorably to other methods in practice, even in the presence of noise.

These estimators suffer of course of some limitations. A first limitation is that these estimators require to know an approximation of the volumetric shape X. An approximation of the boundary of X is not a usable input for these estimators. This is a strong limitation when analyzing clouds of points. Fortunately, in digital geometry, most input shapes are approximations of X.

Secondly, these estimators are multigrid convergent for a whole range of radii and, even in the case of the parameter-free curvature estimator, it is not obvious what is the optimal value for the k parameter. We know pretty clearly the bound on errors related to moment estimations, but the error bound on the Taylor expansion is not explicit. Certainly, the parameter k depends in some ways on maximal curvature derivatives, but this result remains to be shown.

Thirdly, these estimators are rather costly to compute for "big" digital shapes (about 12s for a 256^3 object with $R = 5$, about 58s with $R = 10$). However, this has to be tempered by the fact that this is the price of being robust to noise. Indeed, even if the object is highly perturbated along its boundary, the volumetric integrals make the estimators very robust.

Many research directions could be carried on from this point. A first one would be to use other kernels than the simple ball. A good candidate is the Gaussian kernel, because computations could be conducted in the Fourier domain. Hence, computations will be almost linear in the size of the object, whatever the kernel radius. Furthermore, multiscale analysis would be much less costly to compute. It remains to show their multigrid convergence.

Another direction is related to Voronoi based approaches, especially to Voronoi Covariance Measure methods [18, 44, 45]. They are also based on covariance analysis, not of the shape but of the tangent cone. Such approaches are thus complementary and should be combined. Besides, looking at image processing methods, VCM approaches share similarities with image curvature estimators based on the image structure tensor [35, 52], and could also benefit from Fourier analysis.

Further information on digital integral invariant estimators can be found in [14, 15, 39, 40].

Acknowledgements This work has been mainly funded by DIGITALSNOW ANR-11-BS02-009, KIDICO ANR-2010-BLAN-0205 and PRIMES LABEX ANR-11-LABX-0063/ANR-11-IDEX-0007 research grants.

References

1. Alliez, P., Cohen-Steiner, D., Tong, Y., Desbrun, M.: Voronoi-based variational reconstruction of unoriented point sets. In: Proceedings of the Eurographics Symposium on Geometry Processing, vol. 7, pp. 39–48 (2007)
2. Amenta, N., Bern, M., Kamvysselis, M.: A new Voronoi-based surface reconstruction algorithm. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, pp. 415–421. ACM (1998)

3. Bauer, F., Fike, C.: Norms and exclusion theorems. Numer. Math. **2**(1), 137–141 (1960). http://dx.doi.org/10.1007/BF01386217
4. Bhatia, R.: Matrix Analysis, vol. 169. Springer, New York (1997)
5. Bobenko, A.I., Suris, Y.B.: Discrete Differential Geometry. Integrable Structure.Graduate Studies in Mathematics, vol. 98. AMS, Providence (2008). http://dx.doi.org/10.1090/gsm/098
6. Boulch, A., Marlet, R.: Fast and robust normal estimation for point clouds with sharp features. Comput. Graph. Forum **31**(5), 1765–1774 (2012)
7. Buet, B.: Approximation de surfaces par des varifolds discrets: représentation, courbure, rectifiabilité. Ph.D. Thesis, Université Claude Bernard-Lyon I, France (2014)
8. Buet, B., Leonardi, G.P., Masnou, S.: Discrete varifolds: A unified framework for discrete approximations of surfaces and mean curvature. In: Aujol, J.F., et al. (eds.) Proceedings of the Third International Conference on Scale Space and Variational Methods in Computer Vision. LNCS, vol. 9087, pp. 513–524. Springer (2015)
9. Bullard, J.W., Garboczi, E.J., Carter, W.C., Fullet, E.R.: Numerical methods for computing interfacial mean curvature. Comput. Mater. Sci. **4**, 103–116 (1995)
10. Cazals, F., Pouget, M.: Estimating differential quantities using polynomial fitting of osculating jets. Comput. Aided Geom, Des. **22**(2), 121–146 (2005)
11. Coeurjolly, D., Klette, R.: A comparative evaluation of length estimators of digital curves. IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 252–258 (2004)
12. Coeurjolly, D., Lachaud, J.O., Roussillon, T.: Multigrid convergence of discrete geometric estimators. In: Digital Geometry Algorithms, Theoretical Foundations and Applications of Computational Imaging. LNCVB, vol. 2, pp. 395–424. Springer, New York (2012)
13. Coeurjolly, D., Miguet, S., Tougne, L.: Discrete curvature based on osculating circle estimation. In: 4th International Workshop on Visual Form, *Lecture Notes in Computer Science*, vol. 2059, pp. 303–312 (2001)
14. Coeurjolly, D., Lachaud, J.O., Levallois, J.: Integral based curvature estimators in digital geometry. In: Discrete Geometry for Computer Imagery. LNCS, vol. 7749, pp. 215–227. Springer, New York (2013)
15. Coeurjolly, D., Lachaud, J.O., Levallois, J.: Multigrid convergent principal curvature estimators in digital geometry. Comput. Vis. Image Underst. **129**, 27–41 (2014)
16. Cohen-Steiner, D., Morvan, J.M.: Restricted Delaunay triangulations and normal cycle. In: Proceedings of the Nineteenth Annual Symposium on Computational Geometry, SCG'03, pp. 312–321. ACM, New York (2003). doi:http://doi.acm.org/10.1145/777792.777839
17. Cohen-Steiner, D., Morvan, J.M.: Second fundamental measure of geometric sets and local approximation of curvatures. J. Differ. Geom. **74**(3), 363–394 (2006)
18. Cuel, L., Lachaud, J.O., Thibert, B.: Voronoi-based geometry estimator for 3d digital surfaces. In: Discrete Geometry for Computer Imagery, pp. 134–149. Springer, New York (2014)
19. Davis, C.: The rotation of eigenvectors by a perturbation. J. Math. Anal. Appl. **6**(2), 159–173 (1963). doi:http://dx.doi.org/10.1016/0022-247X(63)90001-5. http://www.sciencedirect.com/science/article/pii/0022247X63900015
20. Desbrun, M., Hirani, A.N., Leok, M., Marsden, J.E.: Discrete exterior calculus (2005). Preprint, arXiv:math/0508341, Version 2: arXiv:math/0508341v2
21. Digne, J., Morel, J.M.: Numerical analysis of differential operators on raw point clouds. Numer. Math. **127**(2), 255–289 (2014). doi:10.1007/s00211-013-0584-y. https://hal.archives-ouvertes.fr/hal-01135993
22. Esbelin, H.A., Malgouyres, R., Cartade, C.: Convergence of binomial-based derivative estimation for 2 noisy discretized curves. Theor. Comput. Sci. **412**(36), 4805–4813 (2011)
23. Fourey, S., Malgouyres, R.: Normals and curvature estimation for digital surfaces based on convolutions. In: Discrete Geometry for Computer Imagery. LNCS, pp. 287–298. Springer, New York (2008)
24. Gatzke, T.D., Grimm, C.M.: Estimating curvature on triangular meshes. Int. J. Shape Model. **12**(01), 1–28 (2006). http://www.worldscientific.com/doi/abs/10.1142/S0218654306000810
25. Guo, J.: On lattice points in large convex bodies (2010). arXiv:1010.4923v2

26. Huxley, M.N.: Exponential sums and lattice points. Proc. Lond. Math. Soc. **60**, 471–502 (1990)
27. Kanungo, T.: Document degradation models and a methodology for degradation model validation. Ph.D. Thesis, University of Washington (1996)
28. Kerautret, B., Lachaud, J.O.: Curvature estimation along noisy digital contours by approximate global optimization. Pattern Recogn. **42**(10), 2265–2278 (2009). doi:10.1016/j.patcog.2008.11.013
29. Klette, R., Rosenfeld, A.: Digital Geometry: Geometric Methods for Digital Picture Analysis. Series in Computer Graphics and Geometric Modelin. Morgan Kaufmann, San Francisco (2004)
30. Klette, R., Žunić, J.: Digital approximation of moments of convex regions. Graph. Models Image Process. **61**(5), 274–298 (1999)
31. Klette, R., Žunić, J.: Multigrid convergence of calculated features in image analysis. J. Math. Imaging Vision **13**, 173–191 (2000)
32. Klette, R., Žunić, J.: On discrete moments of unbounded order. In: Proceedings of the Discrete Geometry for Computer Imagery (DGCI'2006), pp. 367–378. Springer (2006)
33. Krätzel, E., Nowak, W.G.: Lattice points in large convex bodies. Monatshefte Math. **112**, 61–72 (1991). doi:10.1007/BF01321717
34. Lachaud, J.O.: Espaces non-Euclidiens et analyse d'image : modèles déformables riemanniens et discrets, topologie et géométrie discrète. Habilitation à diriger des recherches, Université Bordeaux 1, Talence, France (2006)
35. Lachaud, J.O., Taton, B.: Deformable model with a complexity independent from image resolution. Comput. Vis. Image Underst. **99**(3), 453–475 (2005). http://www.lama.univ-savoie.fr/~lachaud/Publications/LACHAUD-JO/publications.html#Lachaud05b
36. Lachaud, J.O., Thibert, B.: Properties of Gauss digitized sets and digital surface integration. Technical Report, hal-01070289, Université de Savoie (2014)
37. Lachaud, J.O., Vialard, A., de Vieilleville, F.: Fast, accurate and convergent tangent estimation on digital contours. Image Vision Comput. **25**(10), 1572–1587 (2007). http://www.lama.univ-savoie.fr/~lachaud/Publications/LACHAUD-JO/publications.html#Lachaud07a
38. Lenoir, A.: Fast estimation of mean curvature on the surface of a 3d discrete object. In: Ahronovitz, E., Fiorio, C. (eds.) Proceedings of the Discrete Geometry for Computer Imagery (DGCI'97). Lecture Notes in Computer Science, vol. 1347, pp. 175–186. Springer, Berlin/Heidelberg (1997). http://dx.doi.org/10.1007/BFb0024839
39. Levallois, J., Coeurjolly, D., Lachaud, J.O.: Parameter-free and Multigrid Convergent Digital Curvature Estimators. In: Barcucci, A.F.E., Rinaldi, S. (eds.) 18th International Conference on Discrete Geometry for Computer Imagery (DGCI 2014). Lecture Notes in Computer Science. Springer, New York (2014). http://liris.cnrs.fr/publis/?id=6703
40. Levallois, J., Coeurjolly, D., Lachaud, J.O.: Scale-space feature extraction on digital surfaces. Comput. Graph. **51**, 177–189 (2015)
41. Li, B., Schnabel, R., Klein, R., Cheng, Z., Dang, G., Jin, S.: Robust normal estimation for point clouds with sharp features. Comput. Graph. **34**(2), 94–106 (2010). http://www.sciencedirect.com/science/article/pii/S009784931000021X
42. Malgouyres, R., Brunet, F., Fourey, S.: Binomial convolutions and derivatives estimation from noisy discretizations. In: Discrete Geometry for Computer Imagery. LNCS, vol. 4992, pp. 370–379. Springer, New York (2008)
43. Mérigot, Q.: Geometric structure detection in point clouds. Theses, Université Nice Sophia Antipolis (2009). https://tel.archives-ouvertes.fr/tel-00443038
44. Mérigot, Q., Ovsjanikov, M., Guibas, L.: Robust Voronoi-based curvature and feature estimation. In: 2009 SIAM/ACM Joint Conference on Geometric and Physical Modeling, SPM'09, pp. 1–12. ACM, New York, NY, USA (2009). http://doi.acm.org/10.1145/1629255.1629257
45. Mérigot, Q., Ovsjanikov, M., Guibas, L.: Voronoi-based curvature and feature estimation from point clouds. IEEE Trans. Vis. Comput. Graph. **17**(6), 743–756 (2011)
46. Müller, W.: Lattice points in large convex bodies. Monatshefte Math. **128**, 315–330 (1999)
47. Page, D.L., Sun, Y., Koschan, A.F., Paik, J., Abidi, M.A.: Normal vector voting: Crease detection and curvature estimation on large, noisy meshes. Graph. Models **64**(3–4), 199–229 (2002)

48. Pottmann, H., Wallner, J., Yang, Y., Lai, Y., Hu, S.: Principal curvatures from the integral invariant viewpoint. Comput. Aided Geom. Des. **24**(8–9), 428–442 (2007). doi:10.1016/j.cagd.2007.07.004
49. Pottmann, H., Wallner, J., Huang, Q., Yang, Y.: Integral invariants for robust geometry processing. Comput. Aided Geom. Des. **26**(1), 37–60 (2009). doi:10.1016/j.cagd.2008.01.002
50. Provot, L., Gérard, Y.: Estimation of the derivatives of a digital function with a convergent bounded error. In: Discrete Geometry for Computer Imagery. LNCS, pp. 284–295. Springer, New York (2011)
51. Reveillès, J.P.: Géométrie discrète, calcul en nombres entiers et algorithmique. Thèse d'état, Université Louis Pasteur, Strasbourg, France (1991). In French
52. Rieger, B., van Vliet, L.J.: Curvature of *n*-dimensional space curves in grey-value images. IEEE Trans. Image Process. **11**(7), 738–745 (2002)
53. Roussillon, T., Lachaud, J.O.: Accurate curvature estimation along digital contours with maximal digital circular arcs. In: Combinatorial Image Analysis, vol. 6636, pp. 43–55. Springer, New York (2011)
54. Rusinkiewicz, S.: Estimating curvatures and their derivatives on triangle meshes. In: Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004, pp. 486–493 (2004)
55. Stewart, G.W., Sun, J.G.: Matrix Perturbation Theory. Academic, New York (1990)
56. Surazhsky, T., Magid, E., Soldea, O., Elber, G., Rivlin, E.: A comparison of Gaussian and mean curvatures estimation methods on triangular meshes. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), vol. 1, pp. 1021–1026 (2003)
57. de Vieilleville, F., Lachaud, J.O., Feschet, F.: Maximal digital straight segments and convergence of discrete geometric estimators. J. Math. Image Vision **27**(2), 471–502 (2007)
58. Xu, G.: Convergence analysis of a discretization scheme for Gaussian curvature over triangular surfaces. Comput. Aided Geom. Des. **23**(2), 193–207 (2006). http://www.sciencedirect.com/science/article/pii/S0167839605000865
59. Zhang, J., Cao, J., Liu, X., Wang, J., Liu, J., Shi, X.: Point cloud normal estimation via low-rank subspace clustering. Comput. Graph. **37**(6), 697–706 (2013). http://www.sciencedirect.com/science/article/pii/S0097849313000824

Index

L. Najman, P. Romon (eds.), *Modern Approaches to Discrete Curvature*,
Lecture Notes in Mathematics 2184, DOI 10.1007/978-3-319-58002-9

LECTURE NOTES IN MATHEMATICS Springer

Editors in Chief: J.-M. Morel, B. Teissier;

Editorial Policy

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications – quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome.

   Manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They may be based on specialised lecture courses. Furthermore, the manuscripts should provide sufficient motivation, examples and applications. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this "lecture notes" character. For similar reasons it is unusual for doctoral theses to be accepted for the Lecture Notes series, though habilitation theses may be appropriate.

2. Besides monographs, multi-author manuscripts resulting from SUMMER SCHOOLS or similar INTENSIVE COURSES are welcome, provided their objective was held to present an active mathematical topic to an audience at the beginning or intermediate graduate level (a list of participants should be provided).

   The resulting manuscript should not be just a collection of course notes, but should require advance planning and coordination among the main lecturers. The subject matter should dictate the structure of the book. This structure should be motivated and explained in a scientific introduction, and the notation, references, index and formulation of results should be, if possible, unified by the editors. Each contribution should have an abstract and an introduction referring to the other contributions. In other words, more preparatory work must go into a multi-authored volume than simply assembling a disparate collection of papers, communicated at the event.

3. Manuscripts should be submitted either online at www.editorialmanager.com/lnm to Springer's mathematics editorial in Heidelberg, or electronically to one of the series editors. Authors should be aware that incomplete or insufficiently close-to-final manuscripts almost always result in longer refereeing times and nevertheless unclear referees' recommendations, making further refereeing of a final draft necessary. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter, a bibliography and several sample chapters. Parallel submission of a manuscript to another publisher while under consideration for LNM is not acceptable and can lead to rejection.

4. In general, **monographs** will be sent out to at least 2 external referees for evaluation.

   A final decision to publish can be made only on the basis of the complete manuscript, however a refereeing process leading to a preliminary decision can be based on a pre-final or incomplete manuscript.

   Volume Editors of **multi-author works** are expected to arrange for the refereeing, to the usual scientific standards, of the individual contributions. If the resulting reports can be

forwarded to the LNM Editorial Board, this is very helpful. If no reports are forwarded or if other questions remain unclear in respect of homogeneity etc, the series editors may wish to consult external referees for an overall evaluation of the volume.

5. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include
   - a table of contents;
   - an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   - a subject index: as a rule this is genuinely helpful for the reader.
   - For evaluation purposes, manuscripts should be submitted as pdf files.

6. Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online. After acceptance of the manuscript authors will be asked to prepare the final LaTeX source files (see LaTeX templates online: https://www.springer.com/gb/authors-editors/book-authors-editors/manuscriptpreparation/5636) plus the corresponding pdf- or zipped ps-file. The LaTeX source files are essential for producing the full-text online version of the book, see http://link.springer.com/bookseries/304 for the existing online volumes of LNM). The technical production of a Lecture Notes volume takes approximately 12 weeks. Additional instructions, if necessary, are available on request from lnm@springer.com.

7. Authors receive a total of 30 free copies of their volume and free access to their book on SpringerLink, but no royalties. They are entitled to a discount of 33.3 % on the price of Springer books purchased for their personal use, if ordering directly from Springer.

8. Commitment to publish is made by a *Publishing Agreement*; contributing authors of multiauthor books are requested to sign a *Consent to Publish form*. Springer-Verlag registers the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: a brief written (or e-mail) request for formal permission is sufficient.

Addresses:
Professor Jean-Michel Morel, CMLA, École Normale Supérieure de Cachan, France
E-mail: moreljeanmichel@gmail.com

Professor Bernard Teissier, Equipe Géométrie et Dynamique,
Institut de Mathématiques de Jussieu – Paris Rive Gauche, Paris, France
E-mail: bernard.teissier@imj-prg.fr

Springer: Ute McCrory, Mathematics, Heidelberg, Germany,
E-mail: lnm@springer.com