BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 57, Number 1, January 2020, Pages 153–159
https://doi.org/10.1090/bull/1647
Article electronically published on September 24, 2018

Elementary applied topology, by Robert W. Ghrist, Createspace, 2014, ISBN 978-1-5028-8085-7, US$19.99

Persistence theory: From quiver representations to data analysis, by Steve Y. Oudot, Mathematical Surveys and Monographs, Vol. 209, American Mathematical Society, Providence, RI, 2015, viii+218 pp., ISBN 978-1-4704-2545-6, US$65.00

1. Leonhard Euler, the first applied topologist

In addition to contributing to almost all branches of modern mathematics, Leonhard Euler was also the first to use a topological argument—a hundred or so years before topology was formally established—in order to answer a practical question. The problem was that of the Seven Bridges of Königsberg, and Euler's solution is often regarded as the inception of both topology and graph theory. Here is the problem. During the 1700s, Königsberg was a Prussian city located in modern-day Kaliningrad, Russia, on the banks of the river Pregel. The city also included two large islands, Kneiphof and Lomse, connected to each other and to the mainland by seven bridges. The problem of the Seven Bridges of Königsberg asked whether or not it was possible to devise a walk around the city, i.e., the mainland and the two islands, which crossed each bridge exactly once.

Euler answered this question in the negative by showing that discarding unimportant details (e.g., distances and the shapes of the land masses) leads to a topological problem, whose solution reduces to evaluating a computable invariant. Specifically, whether or not the number of odd-degree nodes in a surrogate connectivity graph—where land masses correspond to nodes and bridges to edges—is either zero or two. In modern language, a path which traverses all edges in a graph exactly once is known as an *Eulerian walk*; such a walk exists if (this direction was stated by Euler and proved by Carl Hierholzer [22]) and only if (this is the implication that Euler proved [17]) the graph is connected, and the number of nodes with odd degree is either zero or two. A historical update: all seven bridges were destroyed during World War II, but only five were later rebuilt. As a consequence, now it is actually possible to devise an Eulerian walk throughout the five bridges of Kaliningrad.

More than 250 years after Euler, the idea of translating a practical question into a topological problem whose solution reduces to computing a mathematical gadget, has prompted myriad applications in science and engineering. In fact this is exactly what researchers in the rapidly growing field of applied topology do routinely. The scientific questions we deal with today are more involved and, consequently, the appropriate topological constructions are more intricate. The books under review cover exactly these ideas: modern applications of algebraic topology, and the mathematics behind these applications.

2. Applied algebraic topology

Many questions in modern science and engineering can be approached using insights from algebraic and geometric topology. For instance in neuroscience [10],

2010 *Mathematics Subject Classification.* Primary 55N99, 68W05; Secondary 55U99.
This work was partially supported by the NSF (DMS-1622301) and DARPA (HR0011-16-2-003).

given measurements of co-activation strength between neurons, it is possible to quantify the degree of geometric organization in neural activity [21]. This is a fundamental step in understanding the functionality of brain circuitry. When the neurons are place-cells (e.g., in rodent hippocampus), one can infer physical/layout properties of the environment from spike data alone [11,12]. In evolutionary biology, gene expression data can be leveraged to test the validity of Darwin's phylogenetic tree as the right descriptor for, say, viral evolution. *Spoiler*: it isn't [6]. In computer vision, topological methods can help elucidate the structure and statistics of the space of visual stimuli that humans are exposed to. It turns out that there is a Klein bottle parameterizing one of the most fundamental portions of this space [3] and that this geometric model can be used for image compression [23] and classification tasks [25]. In bio-robotics, swarms of insects (Madagascar hissing cockroaches) equipped with wireless neuro-stimulation backpacks, can be used to infer topological maps of physical environments for search and rescue [14]. In mechanical robotics, tools from bundle theory and configuration spaces lead to impossibility theorems, as well as complexity measures, for the existence of solutions to continuous motion planning problems [19]. In theoretical computer science, computational complexity [29] and concurrency theory [18] have also benefited from topological ideas, and the list of modern applications of topology goes on and on.

The types of applications described above have prompted the development of new mathematics, and making these tools accessible to researchers from various applied areas calls for novel types of algebraic topology books. Robert Ghrist's book (*Elementary applied topology*) takes an important step in fulfilling this need. It takes the topics covered in the typical first courses on algebraic, geometric, and differential topology (e.g., manifolds, homotopy, (co)homology, Morse theory, sheaves, etc.) and presents them through the lens of applications. It is by no means a detailed development of the underlying mathematics, but rather a rich compendium of examples of ways in which these ideas have made their way into applications. Again, think robotics, signal analysis, target enumeration, consensus problems, computational chemistry, etc. Let me illustrate this interplay with a specific example.

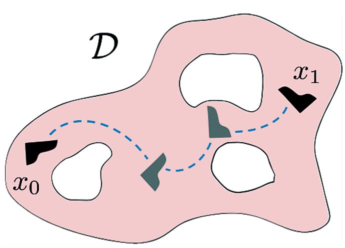

FIGURE 1. The piano movers' problem

Let $\mathcal{D} \subset \mathbb{R}^2$ be a bounded planar domain, and consider the problem of continuously moving a piano between two configurations in $\mathcal{D}$; this is called the *piano movers' problem* in motion planning; see Figure 1 and [27]. By a configuration we mean an orientation $\theta \in S^1 = \{z \in \mathbb{C} : |z| = 1\}$ for the piano, together with its position in $\mathcal{D}$. If $X \subset \mathcal{D} \times S^1$ denotes the space of all possible piano configurations or, more generally, the configuration space for an arbitrary mehcanical system (e.g., a robotic arm), then a motion plan between two configurations $x_0, x_1 \in X$ is a continuous path $\gamma : [0,1] \longrightarrow X$ so that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. The geometry of X might be very complicated, which makes finding motion plans a difficult algorithmic problem in practice. If PX denotes the space of all continuous paths in X (endowed with the compact-open topology), and

$$\begin{array}{rccc} \pi: & PX & \longrightarrow & X \times X \\ & \gamma & \mapsto & \big(\gamma(0), \gamma(1)\big), \end{array}$$

then a motion planner is a function $\mathsf{m} : X \times X \longrightarrow PX$ so that $\pi \circ \mathsf{m}$ is the identity of $X \times X$. A simple-minded question is whether or not one can choose a *continuous* motion planner. That is, one where replacing a configuration pair $(x_0, x_1) \in X \times X$ with nearby configurations $(\widetilde{x}_0, \widetilde{x}_1)$, always leads to similar motion plans $\mathsf{m}(x_0, x_1), \mathsf{m}(\widetilde{x}_0, \widetilde{x}_1) : [0,1] \longrightarrow X$. It turns out that this is surprisingly difficult to do (see [19, Theorem 1]):

Theorem 2.1. *A continuous motion planner* $\mathsf{m} : X \times X \longrightarrow PX$ *exists if and only if the configuration space* X *is contractible.*

As an application of this theorem, it follows that the piano movers' problem does not admit a continuous motion planner. The proof of Theorem 2.1 is actually very short, and it is presented together with other ideas in Chapter 8 of Ghrist's book in the context of homotopy theory, fibrations, and fiber bundles. Other applications in that chapter include DNA topology, social choice, and computational chemistry. Again, the book does not delve too deeply into all the mathematical details; there are excellent references which already do this, and Ghrist provides a list of relevant sources, historical comments, and further reading in the notes at the end of each chapter. My impression after reading this book is that, in addition to serving as an overview for applied topology broadly, it is a great source of inspiration for future directions given the variety of topics it covers. A beginner will certainly find this useful, especially if paired with more detailed and structured references.

3. Persistent homology: a bird's-eye view

The *shape of a topological space*, intuitively speaking, refers to those properties which remain unchanged under deformations that do not tear holes or glue portions of the space together. For instance, the number of connected components is one of these properties. A central question in topology is how to determine if two spaces are topologically equivalent, and algebraic topology offers computational tools for systematically approaching this question. The main idea is to associate algebraic objects—such as numbers, vector spaces, and other structures—to a space. This is done in such a way that if two spaces are topologically equivalent, then their associated algebraic objects are isomorphic. Such objects are called invariants.

The homology of a space is one of the prime examples of a rich, yet algorithmically computable, invariant [15]. Given a topological space B, an integer $n \geq 0$ and a field $\mathbb{F}$ (e.g., the rational numbers $\mathbb{Q}$ or the integers modulo a prime), the nth homology of B with coefficients in $\mathbb{F}$, denoted $H_n(B;\mathbb{F})$, is a vector space over $\mathbb{F}$ and its dimension $\beta_n(B;\mathbb{F})$, the nth Betti number of B over $\mathbb{F}$, provides a count for the number of essentially distinct n-dimensional holes in B. Specifically, and when $\mathbb{F} = \mathbb{Q}$, $\beta_0(B)$ corresponds to the number of connected components of B, $\beta_1(B)$ measures the number of essentially distinct closed loops that do not bound a filled-in region, $\beta_2(B)$ counts the number of closed 2-dimensional portions of B which bound a 3-dimensional void, and so on. For example, if S^2 denotes the 2-sphere and $T = S^1 \times S^1$ the torus, then we have the following Betti numbers. Both S^2 and T are connected (i.e., $\beta_0 = 1$); every closed loop on the surface of S^2 bounds a filled-in region ($\beta_1(S^2) = 0$), but there are two essentially distinct loops on T

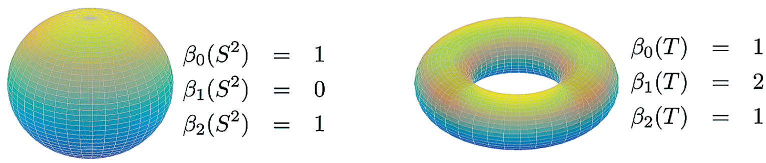

FIGURE 2

which do not: one vertical and one horizontal ($\beta_1(T) = 2$). Both S^2 and T are themselves 2-dimensional closed regions bounding a 3-dimensional void, and hence $\beta_2(S^2) = \beta_2(T) = 1$; see Figure 2.

The *shape of a data set* can be studied in much the same way as that of a topological space. Most scientific data (e.g., collections of images, text documents, or molecular compounds) can be endowed with the structure of a finite metric space. Instead of connected components, the appropriate notion here would be clustering [4]. Homology can also be adapted to this setting; the relevant construction is called *persistent homology*, and what it estimates from a data set, for each integer $n \geq 0$, is the presence of n-dimensional holes in the underlying continuous space around which the data accumulates. Persistent homology, denoted PH_n in dimension $n \geq 0$, is often visualized as a collection of horizontal bars (a barcode), denoted $\mathsf{bc}(PH_n)$. More explicitly, $\mathsf{bc}(PH_n)$ is a collection of intervals $[\epsilon, \epsilon + \delta) \subset [-\infty, \infty]$, which may appear with repetitions. Each interval $[\epsilon, \epsilon + \delta) \in \mathsf{bc}(PH_n)$ represents an n-dimensional hole in the data, and its length $\delta > 0$ indicates the likelihood that said hole corresponds to a true homological feature of the underlying continuous space from which the data was sampled. In Figure 3 we show a data set sampled from the torus (left), and the barcodes for its persistent homology in dimensions $n = 0, 1, 2$ (right). The number of long intervals in $\mathsf{bc}(PH_n)$ is akin to the Betti numbers β_n for the homology of a topological space: a single long interval in the barcode for 0-dimensional persistence is indicative that the underlying data space is connected; the two long intervals in $\mathsf{bc}(PH_1)$ indicate that there are two 1-dimensional holes; and the single long interval in $\mathsf{bc}(PH_2)$ corresponds to a significant void in the data. If one were to start from these barcodes, then a strong hypothesis would be that the data has been sampled from a space with the topology of a torus. Of course one would have to confirm this with further investigations, but it is in this fashion that persistent homology, and its barcodes, can be used to measure and infer the shape of data.

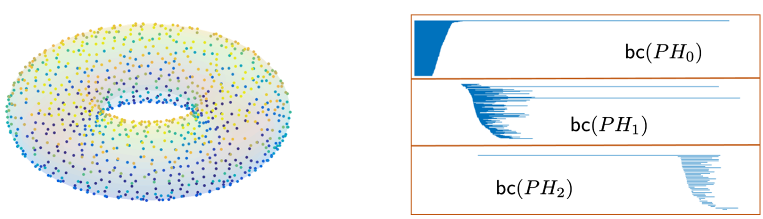

FIGURE 3

Persistent homology is currently one of the most widely used tools in applied topology and topological data analysis (TDA). This is due in equal parts to its usefulness, a good code base for the computation of barcodes [24], and a strong theoretical foundation. As for usefulness, many of the applications mentioned in section 2 of this review rely on barcodes. In fact, the Klein bottle model of fundamental visual features mentioned earlier [3], was discovered through topological inference with barcodes. Persistent homology has also been especially effective in modeling, classification, and prediction tasks where the shape of data is relevant. For instance, in molecular biology—where the geometry of molecules is tightly related to function—deep learning on barcodes yields state-of-the-art results for computer-aided drug discovery [31]. It is worth noting that some of the ideas behind persistent homology are also starting to make an impact in pure mathematics; see for instance [26, 30] for recent applications to symplectic geometry.

As for theoretical foundations, first recall that the nth Betti number $\beta_n(B;\mathbb{F})$ is defined as the dimension of the vector space $H_n(B;\mathbb{F})$. Since two vector spaces (over the same field) are isomorphic if and only if they have the same dimension, it follows that $\beta_n(B;\mathbb{F})$ uniquely determines the isomorphism type of $H_n(B;\mathbb{F})$. It turns out that persistent homology is also an algebraic object, albeit one a bit more complicated than a vector space (in its simplest form it is a graded module over the ring of polynomials $\mathbb{F}[t]$ in one variable), and that in favorable situations the barcodes exist and uniquely determine its isomorphism type.

One of the main questions in the theory of persistent homology deals with the existence and computability of barcode-like invariants [2, 5, 16, 32]. Recently, it has been observed that this problem is related to the representation theory of quivers [13], and the book under review by Steve Oudot, *Persistence theory: from quiver representations to data analysis*, utilizes this point of view as a unifying framework in which to present the theory of persistent homology. Another big theme in the theory is stability [7, 9]; that is, the question of what happens to the barcodes as the input data is changed slightly. Other facets of the theory deal with connections to measure theory [8], statistics [20], random topology [1], machine learning [28], etc.

Oudot's book does a very nice job of putting all these ideas in context, and it provides a clear treatment of the main themes in the field. Specifically, the computational challenges and expected topological reconstruction guarantees, the existence and computability of useful isomorphism invariants, as well as their stability with respect to several geometric noise models. The book is a pleasant and mostly self-contained read that, in my opinion, would serve as an excellent textbook for both students and researchers interested in the field. It streamlines several results scattered throughout the literature (especially for topological inference and stability), and includes various computational examples which serve to illustrate what the best practices are in applying persistent homology to data analysis; e.g., choosing the right version of persistence, data preprocessing, dealing with noise, computational (memory and speed) requirements, etc. The topological data analysis practitioner will certainly appreciate this.

References

[1] O. Bobrowski and M. Kahle, *Topology of random geometric complexes: a survey*, J. Appl. and Comput. Topology, (2018), 1–34.

[2] Gunnar Carlsson and Vin de Silva, *Zigzag persistence*, Found. Comput. Math. **10** (2010), no. 4, 367–405, DOI 10.1007/s10208-010-9066-0. MR2657946

[3] Gunnar Carlsson, Tigran Ishkhanov, Vin de Silva, and Afra Zomorodian, *On the local behavior of spaces of natural images*, Int. J. Comput. Vis. **76** (2008), no. 1, 1–12, DOI 10.1007/s11263-007-0056-x. MR3715451

[4] Gunnar Carlsson and Facundo Mémoli, *Characterization, stability and convergence of hierarchical clustering methods*, J. Mach. Learn. Res. **11** (2010), 1425–1470. MR2645457

[5] Gunnar Carlsson and Afra Zomorodian, *The theory of multidimensional persistence*, Discrete Comput. Geom. **42** (2009), no. 1, 71–93, DOI 10.1007/s00454-009-9176-0. MR2506738

[6] Joseph Minhow Chan, Gunnar Carlsson, and Raul Rabadan, *Topology of viral evolution*, Proc. Natl. Acad. Sci. USA **110** (2013), no. 46, 18566–18571, DOI 10.1073/pnas.1313480110. MR3153945

[7] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot, *Proximity of persistence modules and their diagrams*, in Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry, Association for Computing Machinery, New York, 2009, pp. 237–246.

[8] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot, *The structure and stability of persistence modules*, SpringerBriefs in Mathematics, Springer, [Cham], 2016. MR3524869

[9] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, Discrete Comput. Geom. **37** (2007), no. 1, 103–120, DOI 10.1007/s00454-006-1276-5. MR2279866

[10] Carina Curto, *What can topology tell us about the neural code?*, Bull. Amer. Math. Soc. (N.S.) **54** (2017), no. 1, 63–78, DOI 10.1090/bull/1554. MR3584098

[11] Carina Curto and Vladimir Itskov, *Cell groups reveal structure of stimulus space*, PLoS Comput. Biol. **4** (2008), no. 10, e1000205, 13, DOI 10.1371/journal.pcbi.1000205. MR2457124

[12] Y. Dabaghian, F. Mémoli, L, Frank, and G. Carlsson, *A topological paradigm for hippocampal spatial map formation using persistent homology*, PLoS Computational Biology **8** (2012), e1002581.

[13] Harm Derksen and Jerzy Weyman, *Quiver representations*, Notices Amer. Math. Soc. **52** (2005), no. 2, 200–206. MR2110070

[14] Alireza Dirafzoon, Alper Bozkurt, and Edgar Lobaton, *Geometric learning and topological inference with biobotic networks*, IEEE Trans. Signal Inform. Process. Netw. **3** (2017), no. 1, 200–215, DOI 10.1109/TSIPN.2016.2623093. MR3615702

[15] Jean-Guillaume Dumas, Frank Heckenbach, David Saunders, and Volkmar Welker, *Computing simplicial homology based on efficient Smith normal form algorithms*, Algebra, geometry, and software systems, Springer, Berlin, 2003, pp. 177–206. MR2011758

[16] Emerson G. Escolar and Yasuaki Hiraoka, *Persistence modules on commutative ladders of finite type*, Discrete Comput. Geom. **55** (2016), no. 1, 100–157, DOI 10.1007/s00454-015-9746-2. MR3439262

[17] L. Euler, *Solutio problematis ad geometriam situs pertinentis*, Commentarii academiae scientiarum Petropolitanae **8** (1741), 128–140.

[18] Lisbeth Fajstrup, Martin Raußen, and Eric Goubault, *Algebraic topology and concurrency*, Theoret. Comput. Sci. **357** (2006), no. 1-3, 241–278, DOI 10.1016/j.tcs.2006.03.022. MR2242768

[19] Michael Farber, *Topological complexity of motion planning*, Discrete Comput. Geom. **29** (2003), no. 2, 211–221, DOI 10.1007/s00454-002-0760-9. MR1957228

[20] Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, and Aarti Singh, *Confidence sets for persistence diagrams*, Ann. Statist. **42** (2014), no. 6, 2301–2339, DOI 10.1214/14-AOS1252. MR3269981

[21] Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov, *Clique topology reveals intrinsic geometric structure in neural correlations*, Proc. Natl. Acad. Sci. USA **112** (2015), no. 44, 13455–13460, DOI 10.1073/pnas.1506407112. MR3429279

[22] C. Hierholzer, *Ueber Kegelschnitte im Raume* (German), Math. Ann. **2** (1870), no. 4, 563–586, DOI 10.1007/BF01444042. MR1509680

[23] A. Maleki, M. Shahram, and G. Carlsson, *A near optimal coder for image geometry with adaptive partitioning*, in Proceedings of the 15th IEEE International Conference on Image Processing, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 2008, pp. 1061–1064.

[24] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, *A roadmap for the computation of persistent homology*, EPJ Data Science **6** (2017), no. 1, 17.

[25] Jose A. Perea and Gunnar Carlsson, *A Klein-bottle-based dictionary for texture representation*, Int. J. Comput. Vis. **107** (2014), no. 1, 75–97, DOI 10.1007/s11263-013-0676-2. MR3165575

[26] Leonid Polterovich and Egor Shelukhin, *Autonomous Hamiltonian flows, Hofer's geometry and persistence modules*, Selecta Math. (N.S.) **22** (2016), no. 1, 227–296, DOI 10.1007/s00029-015-0201-2. MR3437837

[27] Jacob T. Schwartz and Micha Sharir, *On the "piano movers'" problem. I. The case of a two-dimensional rigid polygonal body moving amidst polygonal barriers*, Comm. Pure Appl. Math. **36** (1983), no. 3, 345–398, DOI 10.1002/cpa.3160360305. MR697469

[28] P. Škraba, *Persistent homology and machine learning*, Informatica **42** (2018), 2.

[29] Steve Smale, *On the topology of algorithms. I*, J. Complexity **3** (1987), no. 2, 81–89, DOI 10.1016/0885-064X(87)90021-5. MR907191

[30] Michael Usher and Jun Zhang, *Persistent homology and Floer-Novikov theory*, Geom. Topol. **20** (2016), no. 6, 3333–3430, DOI 10.2140/gt.2016.20.3333. MR3590354

[31] Wei, G.-W. Persistent homology analysis of biomolecular data. *Journal of Computational Physics 305*, 276–299.

[32] Afra Zomorodian and Gunnar Carlsson, *Computing persistent homology*, Discrete Comput. Geom. **33** (2005), no. 2, 249–274, DOI 10.1007/s00454-004-1146-y. MR2121296

José A. Perea
Department of Computational Mathematics, Science & Engineering
Department of Mathematics, Michigan State University
East Lansing, Michigan
Email address: joperea@msu.edu