TOPOLOGICAL TOOLS FOR IDENTIFYING ORDER IN COMPLICATED GEOMETRIES

AMIT PATEL

Since the dawn of history, we have aspired to understand the shapes around us. The earliest ideas include notions of length, area, and volume. The ancient Greeks understood basic shapes such as lines, circles, and polygons. Of course, mathematics has come a long way since then, but this line of work continues. Today, advances in information technology has revolutionized the way we record and simulate natural phenomena. We see complex shapes of which we have little geometric understanding. Using ideas from geometry, topology, and computer science, I build mathematical tools to identity order in complicated geometries. In building these tools, I keep in mind that data, whether recorded by probes or generated by simulation, is noisy.

One approach to studying the shape of a set $P \subset \mathbb{R}^n$ is to construct a function $f : \mathbb{R}^n \to \mathbb{R}$ that is somehow dependent on P and then studying its fibers $f^{-1}(r)$ or its sublevel sets $f^{-1}\left((-\infty, r]\right)$. For example, let f be the Euclidean distance to P. Reeb graphs [Ree46] and persistence diagrams [EH08] are two important topological tools for studying functions; see Figures 1 and 2. The persistence diagram is stable to perturbations of the function. Recently, we showed that the Reeb graph is also stable in the same (categorical) sense [dSMP].

Figure 1. The Reeb graph (right) of the function $f : X \to \mathbb{R}$ (left) identifies path connected components of fibers.

Figure 2. The d-th persistence diagram (right) of the function $f : X \to \mathbb{R}$ (left) records the evolution of the d-th homology group of its sublevel sets. Here $d = 0$.

Maps $g : \mathbb{R}^n \to \mathbb{R}^k$, for $k > 1$, arise naturally. For example, suppose the shape P deforms in $\mathbb{R}^n$ with time t. Then, a natural map to consider is the map $h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^2$ defined as $h(x,t) = (f_t(x), t)$, where $f_t : \mathbb{R}^n \to \mathbb{R}$ is the Euclidean distance to the deformed shape P_t. I am using constructible (co)sheaf theory to generalize and study Reeb graphs and persistence for for Whitney stratified maps to manifolds. Currently, Robert MacPherson and I are finishing a large project on persistent sheaves – our generalization of persistence. I am interested in using persistent sheaves to study the complicated geometries of nodal sets and gaussian random fields.

Date: November 2014.

School of Mathematics; Institute for Advanced Study; Princeton, NJ; amit@ias.edu.

I describe my work on Reeb spaces in Section 1, the persistent homology group in Section 2, and the persistent sheaf in Section 3. I conclude with future directions in Section 4.

1. Reeb Spaces

Reeb graphs are useful in many applications. For example, they are used in shape comparison [HSKK01, EHB13], data skeletonization [GSBW11, CS13], and surface denoising [WHDS04]. The Reeb space, defined as follows, is a generalization of the Reeb graph to continuous maps $f : X \to M$. Declare two points $x, y \in X$ equivalent, $x \sim y$, if $f(x) = f(y) = p$ are connected by a path in $f^{-1}(p)$. The **Reeb space** of f is the quotient space $X_f = X/\sim$ along with the map $\hat{f} : X_f \to M$ induced by f. In Figure 1, we see an example of a Reeb graph, $M = \mathbb{R}$.

I am interested in maps generated by finite data and representable on a computer. These maps are stratifiable. Stratified maps generalize many of the tameness properties of Morse functions. Roughly speaking, a stratified space $(X, \mathcal{S})$ is a space X and a decomposition $\mathcal{S}$ of X into manifolds called strata. A map $f : X \to M$ is **stratifiable** if there is a stratification $(M, \mathcal{S})$ such that f over each stratum in $\mathcal{S}$ is a fiber bundle.

Edelsbrunner, Harer, and I gave the first algorithm to compute the Reeb space of a piecewise linear map $f : |K| \to \mathbb{R}^n$ – a special case of a stratifiable map – from a finite combinatorial manifold K [EHP08]. The algorithm produces the coarsest stratification of the Reeb space $\hat{f}$. Since then, I have come to understand the Reeb space of a stratifiable map as a set valued constructible cosheaf. This categorification of the Reeb space allows me to port ideas of stability and simplification from the theory of persistence to Reeb spaces.

To give you a flavor, a set valued cosheaf over M is a functor $\mathsf{G} : \mathsf{O}(M) \to \mathsf{Set}$ that satisfies a continuity condition called the cosheaf axiom. The functor G assigns to each open set $U \subset M$ a set $\mathsf{G}(U)$ and to each pair of open sets $V \subset U$ a set map $\mathsf{G}(V \subset U) : \mathsf{G}(V) \to \mathsf{G}(U)$. Furthermore, for each triple $W \subset V \subset U$, $\mathsf{G}(W \subset U) = \mathsf{G}(V \subset U) \circ \mathsf{G}(W \subset V)$. The cosheaf G is constructible if there is a stratification $(M, \mathcal{S})$ such that G is locally constant over each stratum. If $g : X \to M$ is stratifiable, then the Reeb space $\hat{g}$ is equivalent to a constructible cosheaf G that assigns to each open set $U \subset M$ the set of path connected components in $g^{-1}(U)$. Once in the cosheaf setting, the *interleaving distance* [CSEH07, CCSG+09, BS12], which has roots in persistent homology, becomes the natural metric between Reeb spaces. Metrics are important as data is noisy and the computed Reeb space is not the "true" Reeb space but one that is close.

We illustrate these ideas for the Reeb graph.

Theorem 1 ([dSMP])**:** Let $f, g : X \to \mathbb{R}$ be two stratifiable functions, and suppose $\sup_{x \in X} |f(x) - g(x)| = \epsilon$. Then, the interleaving distance between their Reeb graphs is at most ϵ.

Along the way, we discovered a topological smoothing operator that works in perfect harmony with the interleaving distance. The **ϵ-simplification** of the Reeb graph $\hat{g}$ is the topologically simplest Reeb graph $\hat{h}$ within a distance of ϵ from $\hat{g}$; see Figure 3. Smoothing is useful when visualizing the Reeb graph of a very complicated or noisy function.

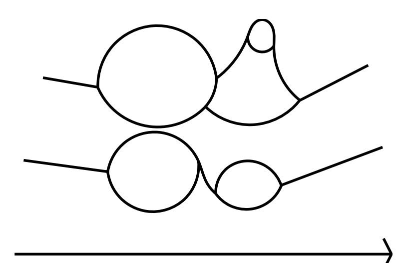

Figure 3. An ϵ-simplification (bottom) of the Reeb graph (top) of Figure 1.

Theorem 2 ([dSMP])**:** The ϵ-simplification of the Reeb graph of $g : X \to \mathbb{R}$ is the Reeb graph of the function $h : X \times [-\epsilon, \epsilon] \to \mathbb{R}$ defined as $h(x, t) = g(x) + t$.

2. Generalizing the Persistent Homology Group

Morse theory studies the critical points of a special but generic class of smooth functions $f : X \to \mathbb{R}$ called Morse functions. As the parameter r increases, the sublevel sets $f^{-1}((-\infty, r])$

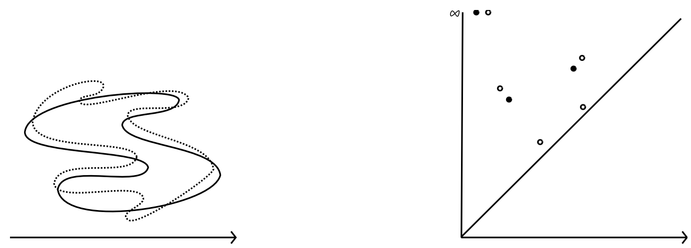

FIGURE 4. A small perturbation $g : \mathbb{S}^1 \to \mathbb{R}$ (dotted) of $f : \mathbb{S}^1 \to \mathbb{R}$ (solid) and their 0-dimensional persistence diagrams. The persistence diagram of g (circles) has more points, but they are close to the diagonal.

change homotopy type every time the level of a critical point is passed. Morse theory is useful in understanding the topology of spaces. Persistence, or, dare I say it, modern Morse theory, is the study of the birth and death of topology in the sublevel set filtration of f. However, for maps to manifolds, the notion of birth and death, which is one-dimensional, does not make sense. *Much of my work has focused on generalizing the persistent homology group – the most basic idea in the theory of persistence – to maps.*

The **persistence** of a function $f : X \to \mathbb{R}$ is a collection of measurements, one for each interval $(p, s) \subset \mathbb{R}$ called a persistent homology group. For the interval (p, s), we have an inclusion of sublevel sets $f^{-1}((-\infty, p]) \subset f^{-1}((\infty, s])$ inducing a homomorphism

$$f_*(p, s) : \mathsf{H}_* \left(f^{-1}((-\infty, p])\right) \to \mathsf{H}_* \left(f^{-1}((-\infty, s])\right)$$

of homology groups. The **persistent homology group** associated to the interval (p, s) is the image $P^f_*(p, s) = \mathrm{im}\ f_*(p, s)$. Suppose we use field coefficients k for homology. Then, the **persistence diagram** visualizes the ranks of all the persistent homology groups.

Let $g : X \to \mathbb{R}$ be a second function, and let $\epsilon = \inf_{x \in X} |f(x) - g(x)|$. For two intervals $(q, r) \subset (p, s)$ such that $q - p > \epsilon$ and $s - r > \epsilon$, the following diagram, where all the homomorphisms are induced by inclusion of spaces, commutes:

$$\begin{array}{ccc} \mathsf{H}_* \left(f^{-1}(-\infty, p]\right) & \longrightarrow & \mathsf{H}_* \left(g^{-1}(-\infty, q]\right) \\ {\scriptstyle f_*(p,s)} \downarrow & & \downarrow {\scriptstyle g_*(q,r)} \\ \mathsf{H}_* \left(f^{-1}(-\infty, s]\right) & \underset{i}{\longleftarrow} & \mathsf{H}_* \left(g^{-1}(-\infty, r]\right). \end{array} \tag{1}$$

By commutativity, $P^f_*(p, s) \subset i(P^g_*(q, r))$. That is, the persistent homology group $P^f_*(p, s)$ lives over (q, r), for all functions within ϵ of f. In this way, the persistent homology group is a robust measurement. The stability of the persistence diagram follows [CSEH07]; see Figure 4.

Now consider a stratifiable map $h : X \to M$ to an oriented m-manifold. For each open set $U \subset M$, we want a measurement that is a generalization of the persistent homology group. First of all, sublevel sets no longer make sense. Instead, I focus on the fibers of h. However, nothing is lost as the sublevel sets of a function can be expressed as the fibers of a slightly modified function. The orientation on M generalizes the notion of birth followed by death.

An orientation on M is the choice of a generator $\mu \in \mathsf{H}^m(M)$ in the top dimensional cohomology group, which is assumed to be isomorphic to k. If M is not compact, then we want to use cohomology with compact support. The cap product, which may be interpreted as a geometric intersection [Gor76], defines a bilinear product:

$$\mathsf{H}_{*+m} \left(X, X - h^{-1}(U)\right) \times \mathsf{H}^m \left(X, X - h^{-1}(U)\right) \xrightarrow{\ \frown\ } \mathsf{H}_* \left(h^{-1}(U)\right).$$

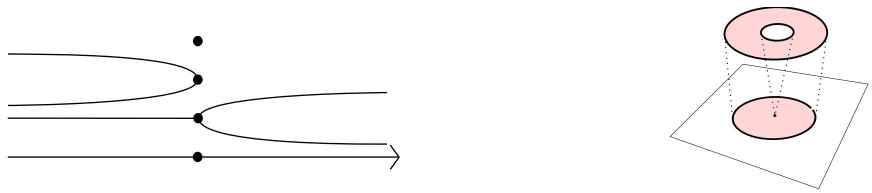

FIGURE 5. For the function $f : X \to \mathbb{R}$ (left) and an open interval $U \subset \mathbb{R}$ containing the 0-stratum (bold dot), the persistent homology group $P_0^f(U) \cong \mathsf{k}$. That is, of the three components above the 0-stratum, only one spreads out over U. For the map $h : X \to \mathbb{R}^2$ (right) and a small open disk V containing the 0-stratum (center dot) , $P_0^h(V) = 0$. That is, the component above the 0-stratum is not robust to perturbations of the map. Also, $P_1^h(V) = 0$ because the 1-cycle above the 0-stratum is not robust.

The map h induces a homomorphism $h^\star : \mathsf{H}^m(M, M-U) \to \mathsf{H}^m\left(X, X - h^{-1}(U)\right)$ on cohomology. Suppose, for simplicity, $U \subset M$ is path connected. Then, the map $z : \mathsf{H}^m(M, M-U) \to \mathsf{H}^m(M)$ on cohomology is an isomorphism. Define

$$\phi_*^h(U) : \mathsf{H}_{*+m}\left(X, X - h^{-1}(U)\right) \to \mathsf{H}_*\left(h^{-1}(U)\right)$$

as $\phi_*^h(U)(v) = v \frown \left(h^\star \circ z^{-1}\right)(\mu)$. The **cap image group** $P_*^h(U)$ is the image of $\phi_*^h(U)$ [MP]; see Figure 5. The cap image group generalizes the persistent homology group.

Proposition 1: Let $H : X \times [0,1] \to M$ be a homotopy taking $H_0 = h$ to a stratifable map $H_1 = \tilde{h}$. Suppose for an open set $V \subset U$, $X - h^{-1}(U) \subset X - H_t^{-1}(V)$, for each $t \in [0,1]$. Then, the following diagram, where the horizontal maps are induced by inclusion of spaces, commutes:

$$\begin{array}{ccc}
\mathsf{H}_{*+m}\left(X, X - h^{-1}(U)\right) & \longrightarrow & \mathsf{H}_{*+m}\left(X, X - \tilde{h}^{-1}(V)\right) \\
\downarrow \phi_*^h(U) & & \downarrow \phi_*^{\tilde{h}}(V) \\
\mathsf{H}_*\left(h^{-1}(U)\right) & \xleftarrow{\;j\;} & \mathsf{H}_*\left(\tilde{h}^{-1}(V)\right).
\end{array}$$

Compare this diagram to Diagram 1. By commutativity, $P_*^h(U) \subset j(P_*^{\tilde{h}}(V))$.

My interpretation of the persistent homology group as a cap product is the end result of many years of work. The **well group**, which I introduced with Edelsbrunner and Morozov [EMP11] and studied further in [EMP10, BEMP13, CPS12], has the desired property of robustness. Unfortunately, the well group is not computable outside a few special cases. Nonetheless, the well group has drawn attention from computational homotopy theorists [FK14] and data visualization scientists [WRS+13, SWCR14, WS14].

3. Persistent Sheaves

Let $f : X \to M$ be a stratifable map to an oriented manifold. The **persistence** of f is a collection of cap image groups, one for each open set of M. Is persistence a sheaf? Robert MacPherson and I started working on this question in the fall of 2011. It has been a long journey of ups and downs through Whitney stratification theory, category theory, two-category theory, and Quillen's Q-construction. The answer is the persistent sheaf.

A sheaf requires a morphism for each pair of open sets $V \subset U$ of M. The following diagram, where the horizontal homomorphisms are induced by inclusions of spaces, commutes:

$$\begin{array}{ccc} \mathsf{H}_{*+m}\left(X, X - f^{-1}(U)\right) & \xrightarrow{\quad i \quad} & \mathsf{H}_{*+m}\left(X, X - f^{-1}(V)\right) \\ \Big\downarrow \phi_*(U) & & \Big\downarrow \phi_*(V) \\ \mathsf{H}_*\left(f^{-1}(U)\right) & \longleftarrow & \mathsf{H}_*\left(f^{-1}(V)\right). \end{array} \tag{2}$$

We hoped for a linear map between the persistent homology groups $P_*(U) = \text{im } \phi_*(U)$ and $P_*(V) = \phi_*(V)$. Unfortunately, this diagram admits no such map. There is, however, a multimap we call a Quillen injection.

A **Quillen injection** $\alpha : A \hookrightarrow B$ between vector spaces is a subspace $Q_\alpha \subset A \times B$ such that the projection $p : Q_\alpha \to A$ is surjective and the projection $q : Q_\alpha \to B$ is injective. We call this a Quillen *injection* because $q\left(p^{-1}(a)\right) \cap q\left(p^{-1}(a')\right) = \emptyset$, for $a \neq a'$. Composition $\beta \circ \alpha : A \hookrightarrow C$ of two Quillen injections $\alpha : A \hookrightarrow B$ and $\beta : B \hookrightarrow C$ is given by the following subspace:

$$Q_{\beta\circ\alpha} = \{(a, c) \in A \times C \mid \exists b \in B \text{ where } (a, b) \in Q_\alpha \text{ and } (b, c) \in Q_\beta\}.$$

Diagram 2 gives rise to a Quillen injection $\alpha : P_*(U) \hookrightarrow P_*(V)$ given by the following subspace:

$$Q_\alpha = \left\{(a, b) \in P_*(U) \times P_*(V) \;\middle|\; b \in \left(\phi_*(V) \circ i \circ \phi_*^{-1}(U)(a)\right)\right\}.$$

For three open sets $W \subset V \subset U$, suppose $\alpha : P_*(U) \hookrightarrow P_*(V)$, $\beta : P_*(V) \hookrightarrow P_*(W)$, and $\gamma : P_*(U) \hookrightarrow P_*(W)$. Unfortunately, $\gamma \neq \beta \circ \alpha$, and therefore, this construction in not functorial. Fortunately, $Q_\gamma \subset Q_{\beta\circ\alpha}$.

Persistence does not form a sheaf, but it forms a higher sheaf called a lax sheaf. The **persistent sheaf** of f is a lax sheaf over M taking values in the Quillen two-category – a two-categorification of Quillen's Q-construction from higher K-theory [Qui73] – defined as follows. Its objects are vector spaces, its one-morphisms are Quillen injections $\alpha : A \hookrightarrow B$, and its two-morphisms $\alpha \Rightarrow \beta$ are containment relations $Q_\alpha \supset Q_\beta$. The robustness of the cap image group extends to a robustness of the persistent sheaf.

4. Future Work

I want to use the persistent sheaf to study complicated geometries. This will be experimental work. My inspiration is the work of MacPherson and Schweinhart, where they use persistence to study fractals [MS12]. There are few tools to study fractals. This is not surprising considering we only started seeing them in the 60s with the help of computers. I am interested in using 2-D persistence (maps to the plane) to study the geometry of eigenfunctions and Gaussian random fields.

Let M be a Riemannian manifold, and let ∇ be its Laplace-Beltrami operator. The **nodal set** is the complement of the pre-image $f^{-1}(0)$ of an eigenfunction, $\nabla^2 f = \lambda f$. Nodal sets become increasingly complex with increasing eigenvalues λ; see Figure 6. In [SW13], Sarnak and Wigman study the behavior of the number of connected components of $f^{-1}(0)$ with increasing eigenvalues λ. In light of persistence, counting connected components of a single fiber seems crude. I am interested in the shape of $f^{-1}(p)$ in M, for each $p \in \mathbb{R}$. This requires a 2-D theory of persistence; we have f and a thickening parameter $\epsilon > 0$ of the sets $f^{-1}(p)$ in M.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $T \subset \mathbb{R}^n$ a topological space. A random field is a measurable map $f : \Omega \to \mathbb{R}^T$, where $\mathbb{R}^T$ is the space of all functions on T. A random Gaussian field on T is a random field such that $f(\omega)$, for each $\omega \in \Omega$, is Gaussian distribution. The level sets of a realization of Gaussian random field have complex geometries and can even be fractal [Xia13]; see Figure 7. In this case, any measurement we invent will be a statistical measurement.

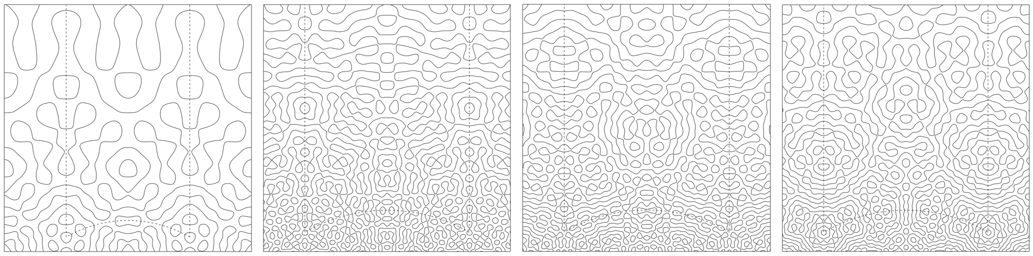

Figure 6. Nodal sets of eigenfunctions of increasing eigenvalues from left to right. Pictures taken from [HR92].

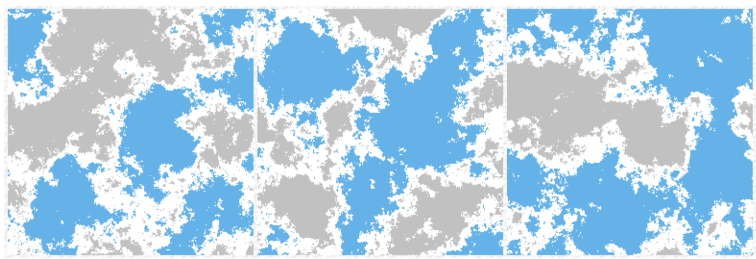

Figure 7. Three realizations of a random Guassian field. The random fields are clipped at three values indicated by the three colors: blue, gray, and white. Pictures taken from [Ked].

References

[BEMP13] Paul Bendich, Herbert Edelsbrunner, Dmitriy Morozov, and Amit Patel. Homology and robustness of level and interlevel sets. *Homology, Homotopy and Applications*, 15(1):51–72, 2013.

[BS12] Peter Bubenik and Jonathan A Scott. Categorification of persistent homology. *arXiv:1205.3669*, 2012.

[CCSG+09] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the 25th annual symposium on Computational geometry*, SCG '09, pages 237–246, New York, NY, USA, 2009. ACM.

[CPS12] Frédéric Chazal, Amit Patel, and Primoz Skraba. Computing well diagrams for vector fields on $\mathbb{R}^n$. *Applied Mathematics Letters*, 25:1725–1728, Nov 2012.

[CS13] Frédéric Chazal and Jian Sun. Gromov-Hausdorff approximation of metric spaces with linear structure. *arXiv:1305.1172*, 2013.

[CSEH07] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 37(1):103–120, 2007.

[dSMP] Vin de Silva, Elizabeth Munch, and Amit Patel. Categorified Reeb graphs. In preparation.

[EH08] H. Edelsbrunner and J. Harer. Persistence homology – a survey. *Contemporary Mathematics*, 453, 2008.

[EHB13] Francisco Escolano, Edwin R. Hancock, and Silvia Biasotti. Complexity fusion for indexing Reeb digraphs. In Richard Wilson, Edwin Hancock, Adrian Bors, and William Smith, editors, *Computer Analysis of Images and Patterns*, volume 8047 of *Lecture Notes in Computer Science*, pages 120–127. Springer Berlin Heidelberg, 2013.

[EHP08] Herbert Edelsbrunner, John Harer, and Amit Patel. Reeb spaces of piecewise linear mappings. In *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, SCG '08, pages 242–250, New York, NY, USA, 2008. ACM.

[EMP10] Herbert Edelsbrunner, Dmitriy Morozov, and Amit Patel. The stability of the apparent contour of an orientable 2-manifold. In V. Pascucci, X. Tricoche, H. Hagen, and J. Tierny, editors, *Topological Methods in Data Analysis and Visualization: Theory, Algorithms, and Applications*. Springer-Verlag, 2010.

[EMP11] Herbert Edelsbrunner, Dmitriy Morozov, and Amit Patel. Quantifying transversality by measuring the robustness of intersections. *Foundations of Computational Mathematics*, 11:345–361, June 2011.

[FK14] Peter Franek and Marek Krčál. Robust satisfiability of systems of equations. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 193–203. SIAM, 2014.

[Gor76] Robert Mark Goresky. *Geometric Cohomology and Homology of Stratified Objects*. PhD thesis, Brown University, June 1976.

[GSBW11] Xiaoyin Ge, Issam I. Safa, Mikhail Belkin, and Yusu Wang. Data skeletonization via Reeb graphs. In J. Shawe-Taylor, R.S. Zemel, P. Bartlett, F.C.N. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 837–845. 2011.

[HR92] Dennis A. Hejhal and Barry N. Rackner. On the topography of Maass waveforms for $\mathrm{PSL}(2,\mathbf{Z})$. *Experimental Mathematics*, 1(4):275–305, 1992.

[HSKK01] Masaki Hilaga, Yoshihisa Shinagawa, Taku Kohmura, and Tosiyasu L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '01, pages 203–212, New York, NY, USA, 2001. ACM.

[Ked] Benjamin Kedem. Gaussian and clipped gaussian random fields. `http://www.math.umd.edu/~bnk/CLIP/clip.gauss.htm`. Accessed: 10-27-2014.

[MP] Robert MacPherson and Amit Patel. Persistent sheaves. In preparation.

[MS12] Robert MacPherson and Benjamin Schweinhart. Measuring shape with topology. *Journal of Mathematical Physics*, 53(7), 2012.

[Qui73] Daniel Quillen. Higher algebraic k-theory: I. In H. Bass, editor, *Higher K-Theories*, volume 341 of *Lecture Notes in Mathematics*, pages 85–147. Springer Berlin Heidelberg, 1973.

[Ree46] Georges Reeb. Sur les points singuliers d'une forme de Pfaff complèment intégrable ou d'une fonction numérique. *Comptes Rendus de L'Académie ses Séances*, 222:847–849, 1946.

[SW13] Peter Sarnak and Igor Wigman. Topologies of nodal sets of random band limited functions. 2013.

[SWCR14] Primoz Skraba, Bei Wang, Guoning Chen, and Paul Rosen. 2d vector field simplification based on robustness. In *IEEE Pacific Visualization Symposium*, PacificVis, 2014.

[WHDS04] Zoë Wood, Hugues Hoppe, Mathieu Desbrun, and Peter Schröder. Removing excess topology from isosurfaces. *ACM Trans. Graph.*, 23(2):190–208, April 2004.

[WRS+13] Bei Wang, Paul Rosen, Primoz Skraba, Harsh Bhatia, and Valerio Pascucci. Visualizing robustness of critical points for 2d time-varying vector fields. In *Computer Graphics Forum (EuroVis)*, 2013.

[WS14] Bei Wang and Primoz Skraba. *Approximating Local Homology from Samples*, chapter 13, pages 174–192. 2014.

[Xia13] Yimin Xiao. Recent developments on fractal properties of gaussian random fields. In Julien Barral and Stphane Seuret, editors, *Further Developments in Fractals and Related Fields*, Trends in Mathematics, pages 255–288. Birkhäuser Boston, 2013.