# MATH 583B: TOPOLOGICAL DATA ANALYSIS

| | |
|---:|:---|
| Place: | Mary Gates Hall (MGH) 085 |
| Time: | MW 9-10:20a.m. |
| Instructor: | Kyle Ormsby (ormsbyk@uw.edu) |
| Drop-in Hours: | MW 10:30–11:30a.m. in PDL C-524 |
| Website: | kyleormsby.github.io/583/ |

**Course description.** What does data look like? What features does it have at various scales? How can we statistically infer answers to these questions? Topological data analysis (TDA) aims to answer these question *topologically*: without choosing parameters, and with an aim towards qualitative descriptions that advance structural understanding.

This course will focus on persistent homology, the preeminent tool used by topological data analysts. Mathematically, we may view this pipeline as beginning with a finite metric space and concluding with a statistically interpretable qualitative summary of that space's characteristics. In applications, the finite metric space might be a point cloud in high-dimensional Euclidean space, and, under the *manifold hypothesis* — that the data lies on some lower-dimensional manifold —, its persistent homology should capture relevant homological features of the underlying manifold.

Time permitting, we will also explore the related topics of (1) multiparameter persistence and (2) hierarchical clustering (via the moduli space of ultrametric trees).

While our explorations will be primarily theoretical, we will also introduce the `Ripser` persistent homology package for Python and related tools. These will allow us to explore examples along with applications of persistent homology to both real and synthetic data.

**Prerequisites.** The course will assume basic familiarity with point-set topology, simplicial complexes, homology, and abstract algebra. Students who have taken a first course in algebraic topology that included singular or simplicial homology will be well-prepared. That said, students who have not formally studied these topics can probably still get a lot from the course. Prior exposure to the Python programming language and Jupyter notebooks would be helpful but not required.

**Course meetings.** This course meets Mondays and Wednesdays 9–10:20am. For the initial portion of the course, I will give lectures on fundametnal topics in persistent homology and clustering theory. The final portion of the course will be devoted to student presentations of advanced topics and/or projects.

**Texts.** We do not have an assigned textbook for this course, but I will be drawing from several references to prepare lectures. These inclue [CVJ22, DW22, Les, Oud15, RB19, Vir]. I plan to produce notes for each lecture meeting.

**Homework.** I will assign a number of optional homework problems during the lecture portion of the course. Students do not need to turn these in, but they are welcome to discuss solutions during drop-in hours.

**Presentations.** During our final course meetings, enrolled students are expected to give a 35-minute presentation on an advanced topic from topological data analysis. I will work with students to develop these topics, find relevant resources, and prepare presentations. You may report on a topic or paper not covered in lecture, or perform your own topological or geometric analysis of real or synthetic data. Here are some potential topics (but students are welcome to develop their own proposal):

→ topological methods for signal analysis,
→ persistent homology of point clouds generated by discrete-time dynamical systems,

$\rightarrow$ applications of persistent homology to symplectic topology,
$\rightarrow$ dimensionality reduction with Eilenberg-MacLane coordinates (following [PST23]),
$\rightarrow$ decorated merge trees (following [CHM$^+$22]),
$\rightarrow$ density-scaled Vietoris-Rips complexes (following [Hic21]).

Students will need to turn in slides, a short paper, or a Github repository to accompany their presentation; this resource will be shared with the rest of the class for input and feedback.

**Evaluation.** Grades will be based on engagement with the course and final presentations.

**Drop-in hours.** I will hold weekly drop-in hours in my office, Padelford (PDL) C-524, Monday and Wednesday 10:30–11:30a.m.. During these times, I am available to discuss course material and help with homework problems. I am also available via email (ormsbyk@uw.edu) and by appointment — if in doubt, please reach out!

**Schedule.** Below you will find a rough schedule for the course. It is highly likely that this will evolve based on student background and interest.

| Week | Day | Topic |
|---|---|---|
| 1 | M III.25 | welcome & warmup |
| | W III.27 | topology, simplicial complexes, and homology review |
| 2 | M IV.1 | persistence modules, barcodes, and persistence diagrams |
| | W IV.3 | persistent homology |
| 3 | M IV.8 | distance and stability for persistence diagrams |
| | W IV.10 | point clouds via Čech and Vietoris–Rips complexes |
| 4 | M IV.15 | Morse theory, Reeb graphs, and zigzag persistence |
| | W IV.17 | topological analysis of weighted graphs |
| 5 | M IV.22 | multiparameter persistence I |
| | W IV.24 | multiparameter persistence II |
| 6 | M IV.29 | NO CLASS |
| | W V.1 | hierarchical clustering and ultrametric spaces |
| | | *presentation proposal due* |
| 7 | M V.6 | CAT(0) cubical complexes, the geodesic treepath problem, and Fréchet means |
| | W V.8 | case study: Topological network analysis of beta-amyloid and tau in Alzheimer's disease using PET imaging data |
| 8 | M V.13 | [flex / additional case studies] |
| | W V.15 | [flex / additional case studies] |
| 9 | M V.20 | student presentations |
| | W V.22 | student presentations |
| 10 | M V.27 | NO CLASS — Memorial Day |
| | W V.29 | student presentations |

## References

[CHM⁺22]  Justin Curry, Haibin Hang, Washington Mio, Tom Needham, and Osman Berat Okutan. Decorated merge trees for persistent topology. *J. Appl. Comput. Topol.*, 6(3):371–428, 2022.

[CVJ22]  Gunnar Carlsson and Mikael Vejdemo-Johansson. *Topological data analysis with applications*. Cambridge University Press, Cambridge, 2022.

[DW22]  Tamal Krishna Dey and Yusu Wang. *Computational topology for data analysis*. Cambridge University Press, Cambridge, 2022.

[Hic21]  Abigail Hickok. A family of density-scaled filtered complexes. arXiv:2112.03334, 2021.

[Les]  Michael Lesnick. Notes on multiparameter persistence. `https://www.albany.edu/~ML644186/840_2022/Math840_Notes_22.pdf`. Accessed on 19 March 2024.

[Oud15]  Steve Y. Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[PST23]  Jose A. Perea, Luis Scoccola, and Christopher J. Tralie. Dreimac: Dimensionality reduction with eilenberg-maclane coordinates. *Journal of Open Source Software*, 8(91):5791, 2023.

[RB19]  Raul Rabadan and Andrew J. Blumberg. *Topological Data Analysis for Genomics and Evolution: Topology in Biology*. Cambridge University Press, Cambridge, 2019.

[Vir]  Žiga Virk. Introduction to persistent homology. `https://zalozba.fri.uni-lj.si/virk2022.pdf`. Accessed on 19 March 2024.