



Analyzing Data

AMS Podcast Series



Robert Ghrist,
University of Pennsylvania

Much of modern research—from genome sequencing to digital surveys of outer space—generates tremendous amounts of multi-dimensional data. Unfortunately, visualizing dimensions higher than three is not easy, which makes analyzing and understanding the data difficult. Topology, a branch of mathematics concerned with the properties of geometrical structures, helps make sense of large data sets by providing a way of classifying the shapes of these sets. It's especially useful for locating groups of similar points called “clusters”, which can, for example, distinguish between distinct types of a given disease, each requiring its own treatment.

Topology (specifically algebraic topology) is also important in the operation of wireless sensor networks, which are used in applications as diverse as monitoring automobile traffic and controlling irrigation. Combined with numerical integration, results from algebraic topology provide the complete picture based on strictly local data. The advantage is that such sensor networks, maintained without GPS or other distance measures, are generally much cheaper to operate. So, in the case of irrigation, mathematical discoveries made almost a century before the advent of today's technology save money while helping us use precious water wisely. In topological terms, just like the Möbius strip: What goes around, comes around.

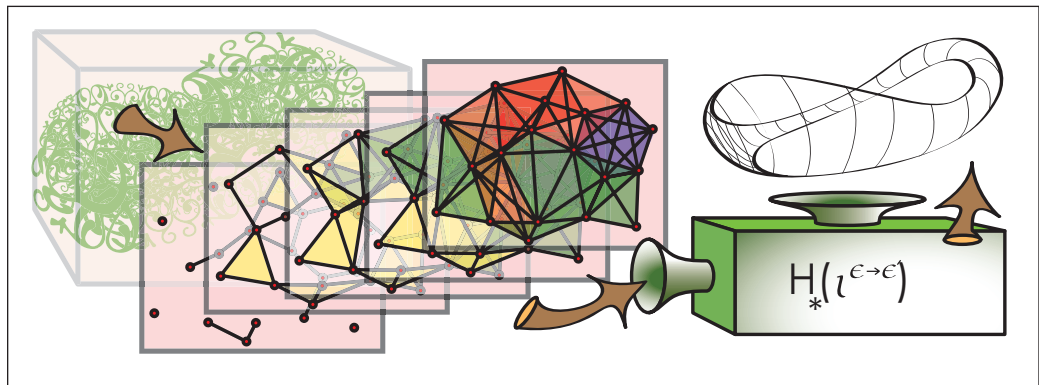


Image: Persistent homology of a simplicial approximation finds hidden structures in large data sets, courtesy of Robert Ghrist.

For More Information: “Topology and Data”, Gunnar Carlson, *Bulletin of the American Mathematical Society* (Vol. 46, No. 2), April 2009.



The *Mathematical Moments* program promotes appreciation and understanding of the role mathematics plays in science, nature, technology, and human culture.

www.ams.org/mathmoments