

Department of Mathematical Sciences Colloquium

Ross E. STAFFELDT

Department of Mathematical Sciences
New Mexico State University

TOPOLOGY FROM DATA SETS

Given a large sample of high-dimensional data, the problem of finding meaningful structures of relatively low dimension hidden in the high-dimensional observations arises immediately. Certain classical techniques, such as principal component analysis and multi-dimensional scaling, find linear structures approximating the data, according to certain criteria. But, by their nature, these methods overlook non-linear phenomena that may be expected or important. On the other hand, by its nature, topology provides a set of tools for describing non-linear phenomena. The application of algebraic topology to data sets is a fairly recent development, so there are lots of interesting questions to explore.

We will describe how one uses combinatorial structures called simplicial complexes to provide topological approximations to a sample space. There will be a demonstration of software that enables one to manipulate simplicial complexes in Matlab, the World-Wide-Web and Matlab permitting. Finally, we introduce persistent homology, a powerful new idea that unifies topological phenomena in data at different scales of resolution. Time permitting, we will see how persistent homology has been used to find meaningful non-trivial topology in a couple of intrinsically interesting data sets.

**Friday, April 3, 2009 at 3pm in Bell Hall 143
The University of Texas at El Paso**

Refreshments will be served in front of the colloquium room,
15 minutes before the start of the colloquium.

For further information, please contact Dr. Andrzej Pownuk, Bell Hall 201.
Phone: (915) 747-6759, e-mail: ampownuk@utep.edu.

