

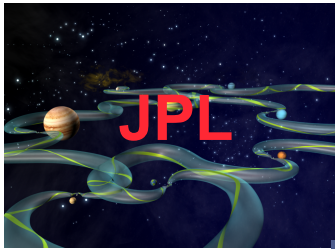


*Science Discovery with MTool on
Massive High Dimensional Data*

*AISRP Investigator Workshop 10/15/2009
Martin.Lo@jpl.nasa.gov*

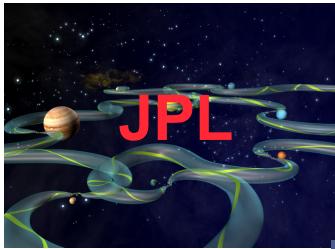
10/15/2009

JPL



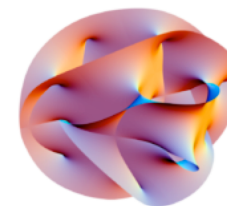
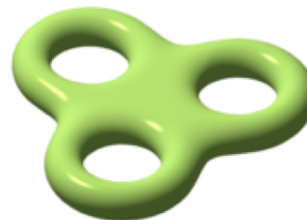
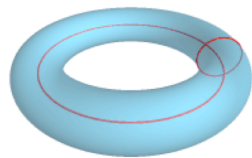
Outline

- **Geometric Approach to High Dimensional Data**
- **MTool Objectives**
- **Dynamical System Methods**
- **Topological Methods**

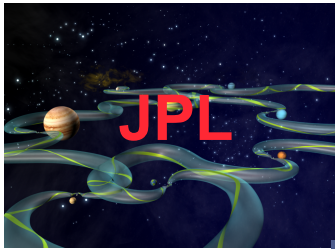


Geometric Approach to Hi-Dim Data

- **Segmentation Problem**
 - Extract hi-dim coherent structures from data
- **Reconstruction Problem**
 - Observation data = projection of original objects
- **Shape Problem**
 - Identify n-dim shapes which you can't visualize
- **Model Reduction Problem**
 - Find faithful projection to lower dimensions

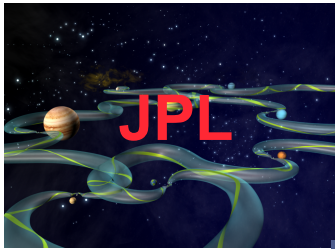


*Calabi-Yau
Manifold 6D*



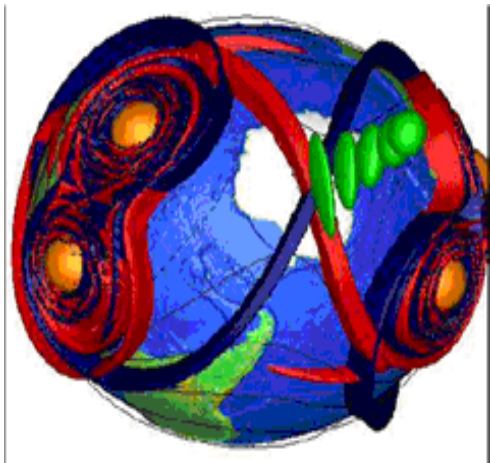
MTool Objectives

- ***Objective 1. Perform R&D on Hi-Dim SMD Data:***
 - Dynamical Systems Theory
 - Discrete Differential Geometry
 - Computational Topology
- ***Objective 2. Provide Working Prototype Tool: MTool 1.0***
 - Solve a series of significant problems with MTool 1.0
 - Use real SMD science data: astrophysics, Earth science
- ***Objective 3. Transfer Technology by:***
 - Distribute MTool end of year 2.

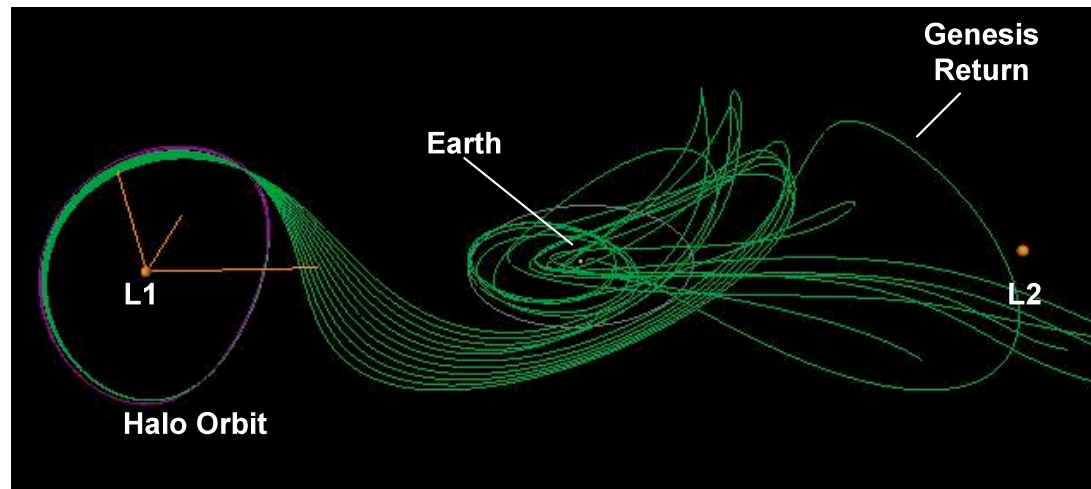


Segmentation Problem

- **Dynamical Systems Theory**
 - Compute Invariant Manifolds of Periodic Orbits
 - Extraction of Lagrangian Coherent Structures
 - Use velocity data from atmosphere, ocean



*Ozone Hole 2002
(Lekien 2005)*

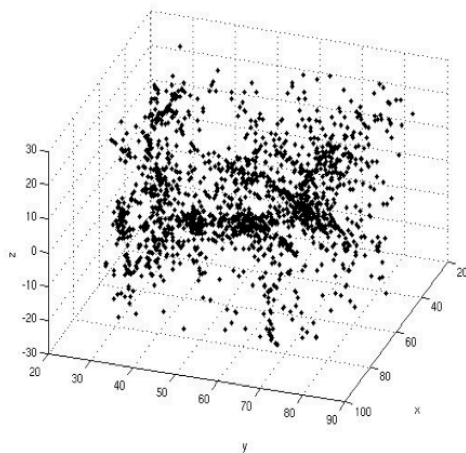


*Genesis Discovery Trajectory
(Lo et al. 1998)*

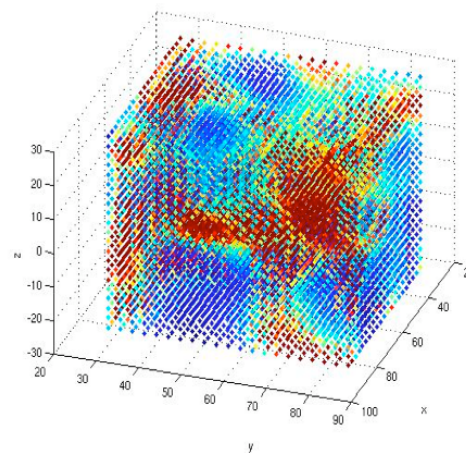


Shape Problem

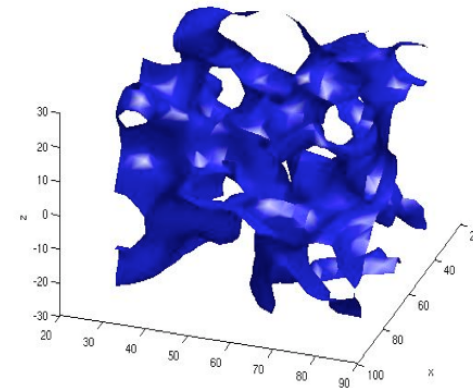
- **3D Large Scale Structure of Universe**
 - 2Mass Data: T. Jarrett, IPAC Caltech
- **Use Algebraic Topology to compute shape**
 - Identify & Count Filaments, Voids



Galaxies



Density Map



Density Contour

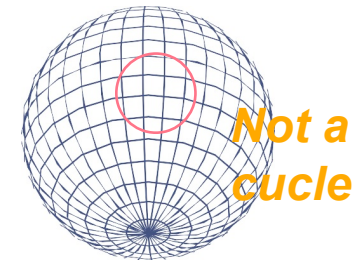


Counting Cycles & Holes with Homology

- Homology counts cycles & holes of a space M .

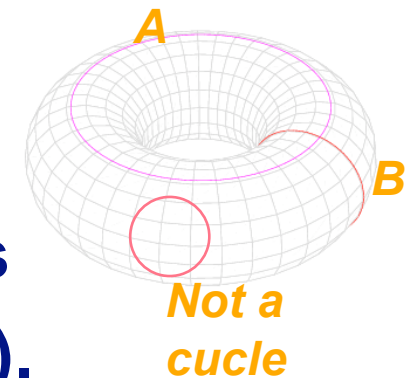
- Cycle C is $\sim k$ -dimensional surface in M which cannot be deformed into a point in M :

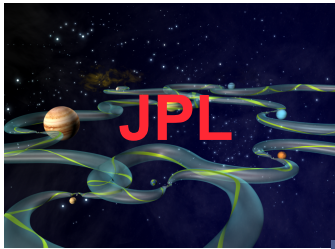
- Circles on the sphere are not cycles
- Circles A, B on the torus are cycles



- K -dim cycles form a group $H_k(M)$

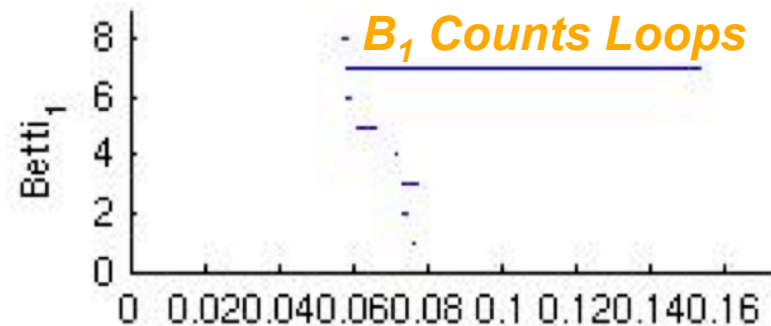
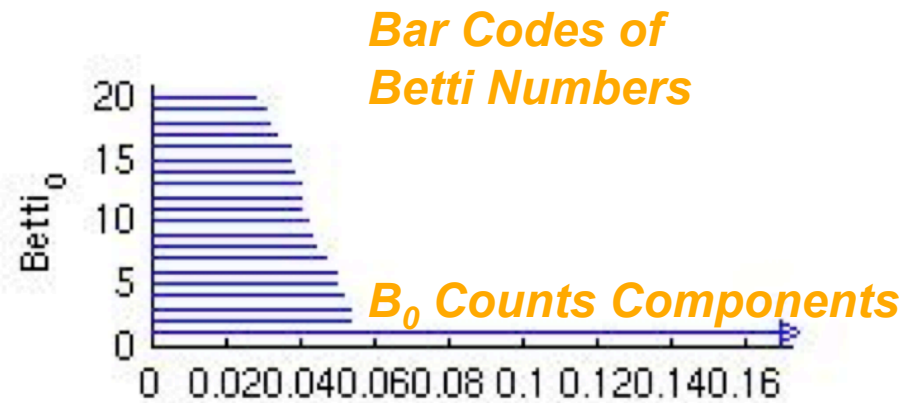
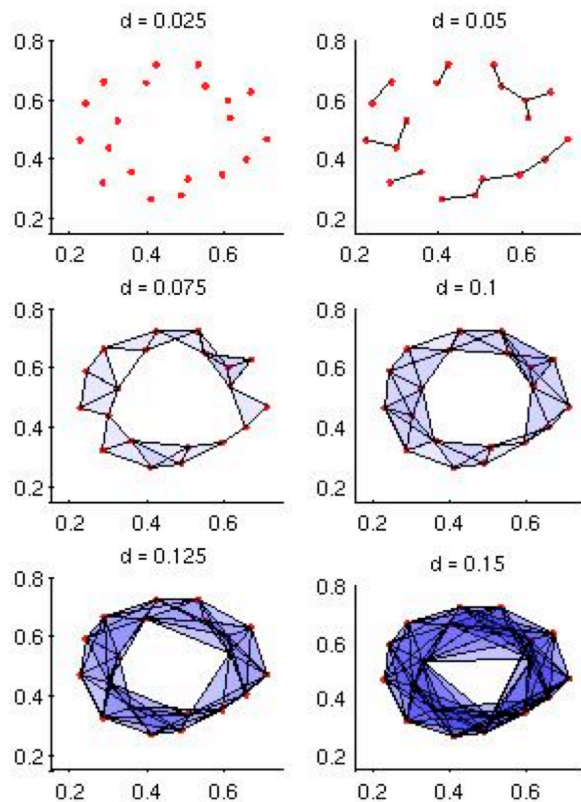
- $\text{Dim}(H_k(M)) = B_k = k^{\text{th}}$ Betti Number
- Counts number of independent cycles
- Generalizes Euler Characteristics $\chi(M)$.



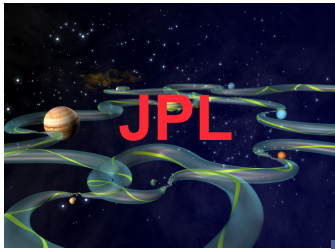


Persistent Homology & Bar Codes

- Cycles which persists are “real” features

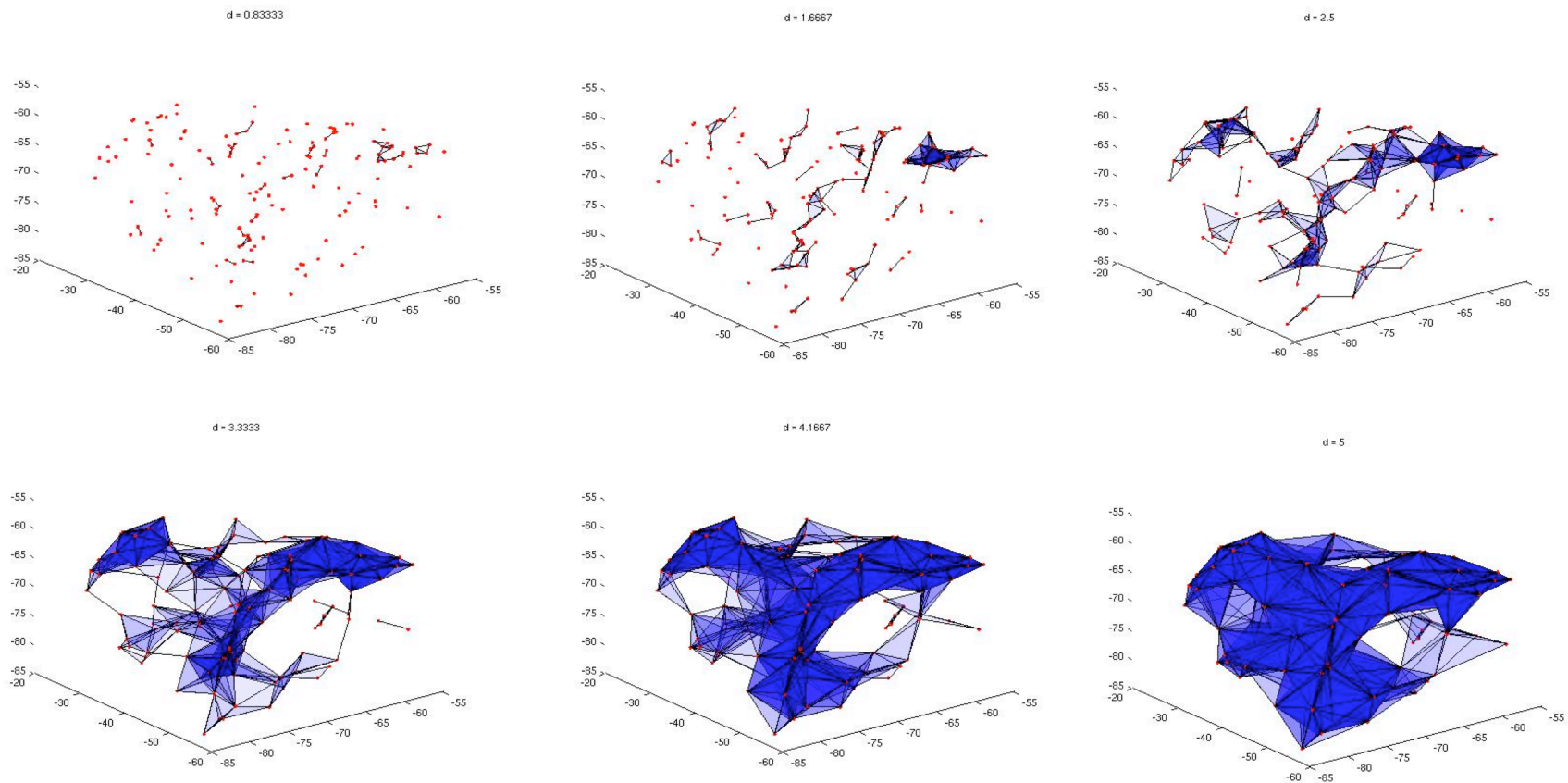


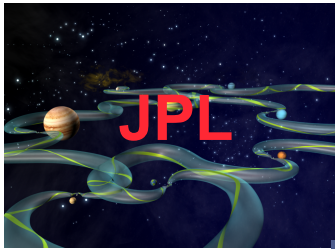
Joint work with R. Hodos 2009
PLEX Tool: Carlsson 2005.



Shape/Reconstruction Problem

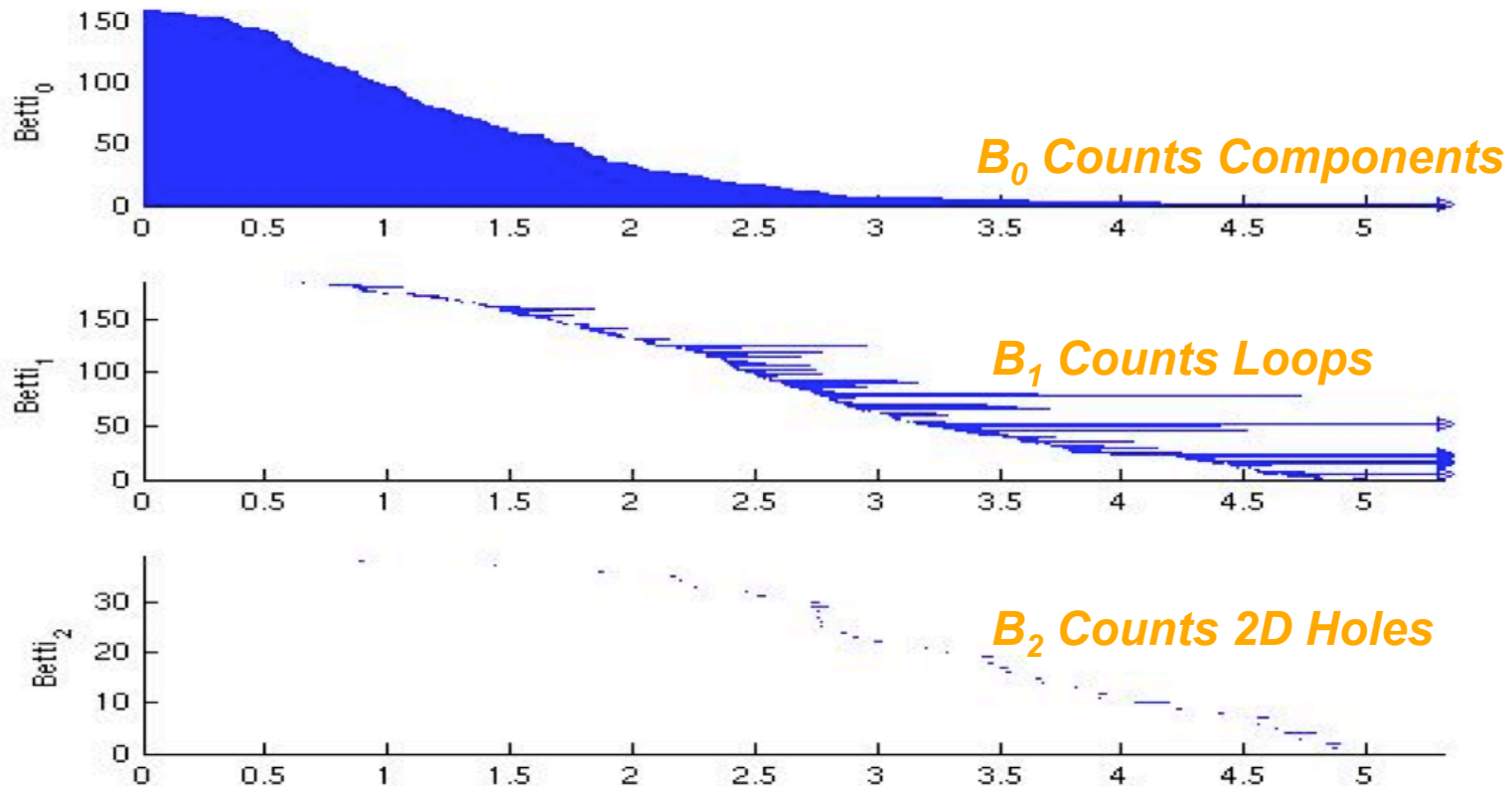
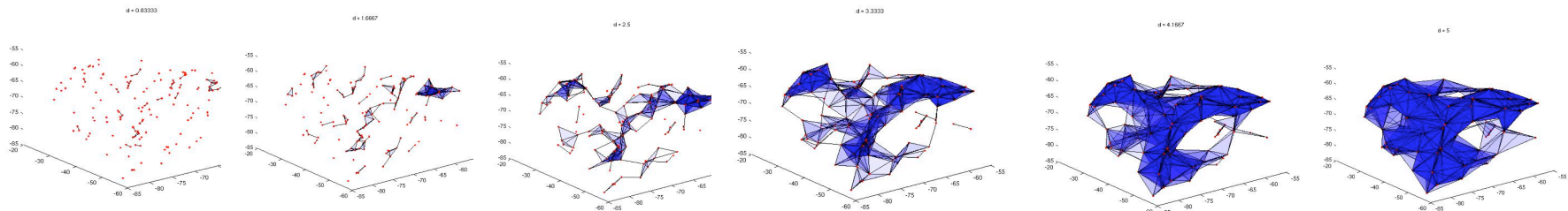
- Extract Shape from Point Cloud 2MASS Data

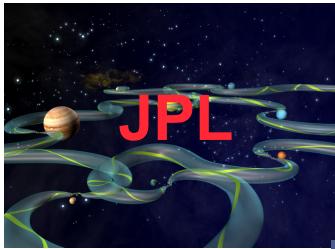




Bar Code On-Going Work ...

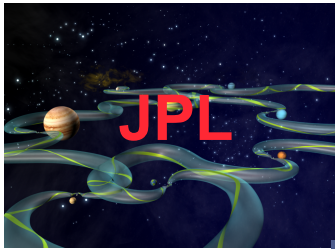
- Not so easy to identify shapes here.





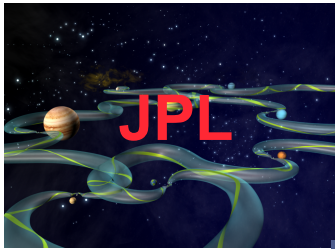
Next Steps

- **Combine with Principal Component Analysis & Bayesian Methods**
- **Explore Other Homology Codes**
- **Visualization**



Other Projects for FY10

- **Dynamical Systems Methods**
 - Galactic morphology via invariant manifolds
 - Atmospheric dynamics via Lagrangian Coherent Structures
- Tomographic methods for processing astrophysical images



References

- Carlsson, G., *Persistent homology and the analysis of high dimensional data*,
Symposium on the Geometry of Very Large Data Sets, Fields
Institute for Research in Mathematical Sciences, February 24, 2005.
- Hodos, R., *Applications of Computational Algebraic Topology to Data Analysis*, USRP Internship Final Report, Spring Session 2009.
- Lekien, F., Ross, S. D., and Newton, P. K. (in preparation, 2005d).
<http://www.lekien.com/~francois/research/>
- Lo, M. et al., *Genesis Mission Design*, AIAA Astrodynamics
Conference, August, 1998, Boston, MA. Paper No. 98-4468.
<http://www.gg.caltech.edu/~mwl>



Acknowledgements

- This work was conducted for the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.
- Many thanks to AISR Program for supporting this work.
- Many thanks to USRP Internship Program and the JPL Education Office for supporting this work.
- Thanks to T. Jarrett for providing the latest *2MASS* data.
- Persistent Homology Tool, *PLEX*, provided by Gunner Carlsson's group, Stanford University