Visualization and Feature Extraction: A global View

Gordon Erlebacher
Department of Mathematics

Scientific Visualization

- Extract from large datasets more meaningful components (called *data extracts*)
  - Isosurface, streamlines, streaklines, vector field topology, vortex tubes, cracks, fault lines, etc.
  - Ionic fields, force topology, iconic representation of molecular structure, etc.
- Render this data with comprehension in mind, as opposed to visual realism

Topics of current interest in visualization community

- Data compression
- Feature-preserving data compression
- Temporal compression
- Automatic feature extraction
- Surface decimation (coarsening of data)
- Parameterization of complex surfaces

My Activities

- Scientific Visualization
- Visualization of vector fields
- Feature extraction
- Real time visualization
- Web-based visualization

- Other Activities
  - Face recognition
  - Applications of geometric algebra

Problem

- Proliferation of very large data sets
- Size of datasets is growing exponentially
  - – Massively parallel computers
    - FSU: 512 processors, 2.25 Teraflops, 30%/peak
    - Earth Simulator (Japan), 60 Teraflops 70%/peak
  - – Largest simulations: 2000x2000x2000
  - – Data generated by satellites: > 1 Terabyte/day

Problem

- High resolution simulations and new data acquisition technologies (In SAR) have created the a veritable data deluge
  (a data tsunami).
- There is an urgent need to scan through the heaps of data and extract out relevant features (plumes, faults, networks , rivers)
- Feature extraction goes hand in hand in this data-mining and visualization ventures.

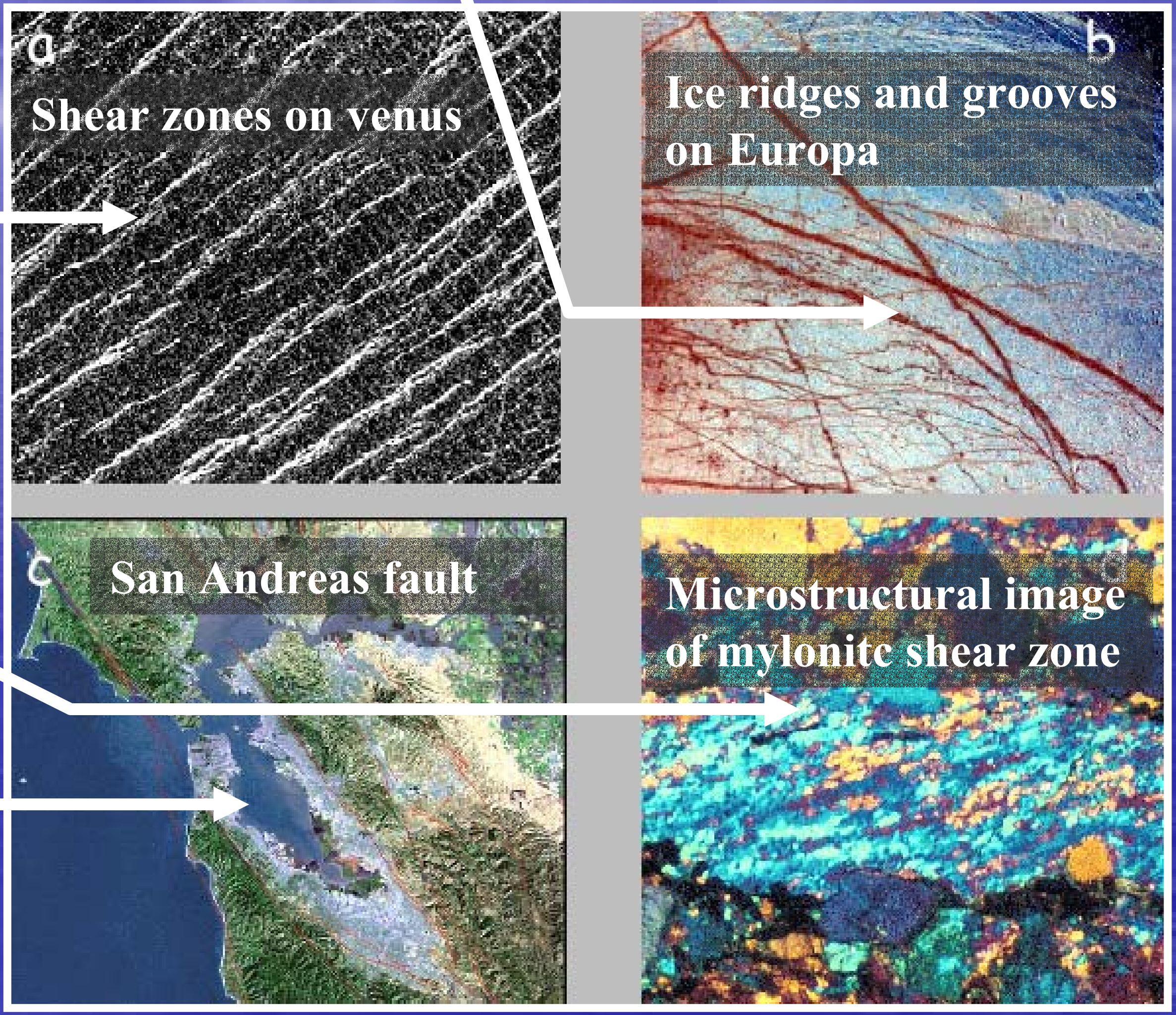
  Tools : wavelets, clustering etc.

- use second-generation wavelets to extract coherent features from 3-D mantle convection generated via numerical simulation

Data Deluge

Fault Extraction via Beamlets

Image from Regenauer & Yuen 2002



Feature extraction via wavelets

Thermal Convection

Physical Problem

- T = deviation from reference state
- p is pressure in excess of hydrostatic

$$\frac{DT}{Dt} = \frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T = \nabla \cdot \left(\kappa \left(\mathbf{x} \right) \nabla T \right)$$

$$\nabla^2 \mathbf{u} - \nabla p - Ra\, T\, e_z = 0$$

Velocity time scale set by viscosity

$$Ra = \frac{\alpha g \Delta T d^3}{\upsilon \kappa_{ref}}$$

Rayleigh number
(hermal Reynolds number)

Challenges

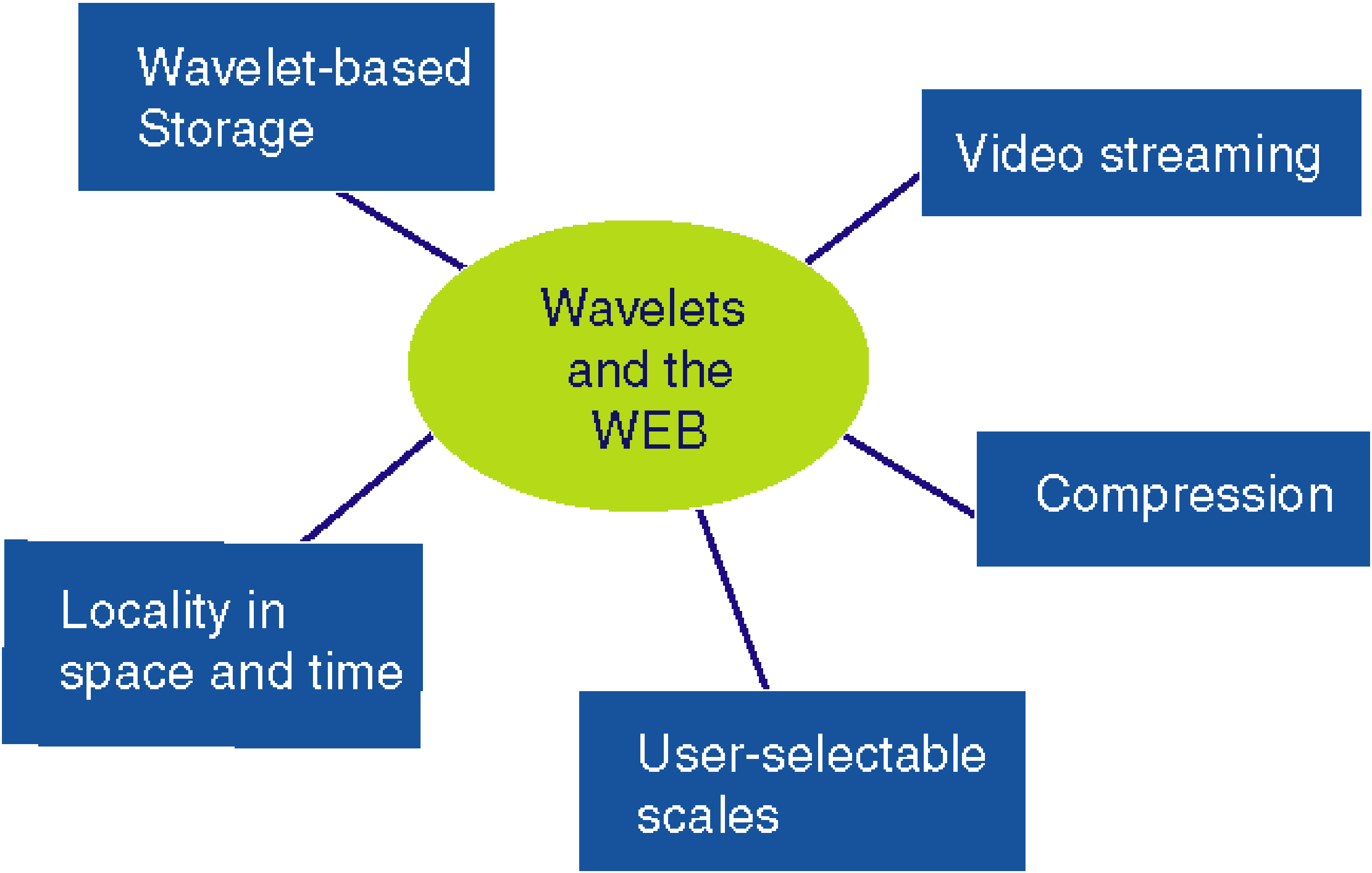
- Large datasets
- Time-dependent flow
- Strong nonlinearities
- Phenomena of interest occupies small fraction of domain
- Laminar to turbulent transition (of the thermal field) as the *Ra* increases

Second Generation Wavelets

Advantages

- Operate on non-uniform meshes
- Curvilinear grids handled by tensor product
- Operate on manifolds generated via subdivision methods
- There is a 1-1 identification between 2nd generation wavelets and points in physical space
- Useful for numerical methods of the “collocation” type

Multiscale Physics



Wavelets and the WEB

Wavelet-based Storage

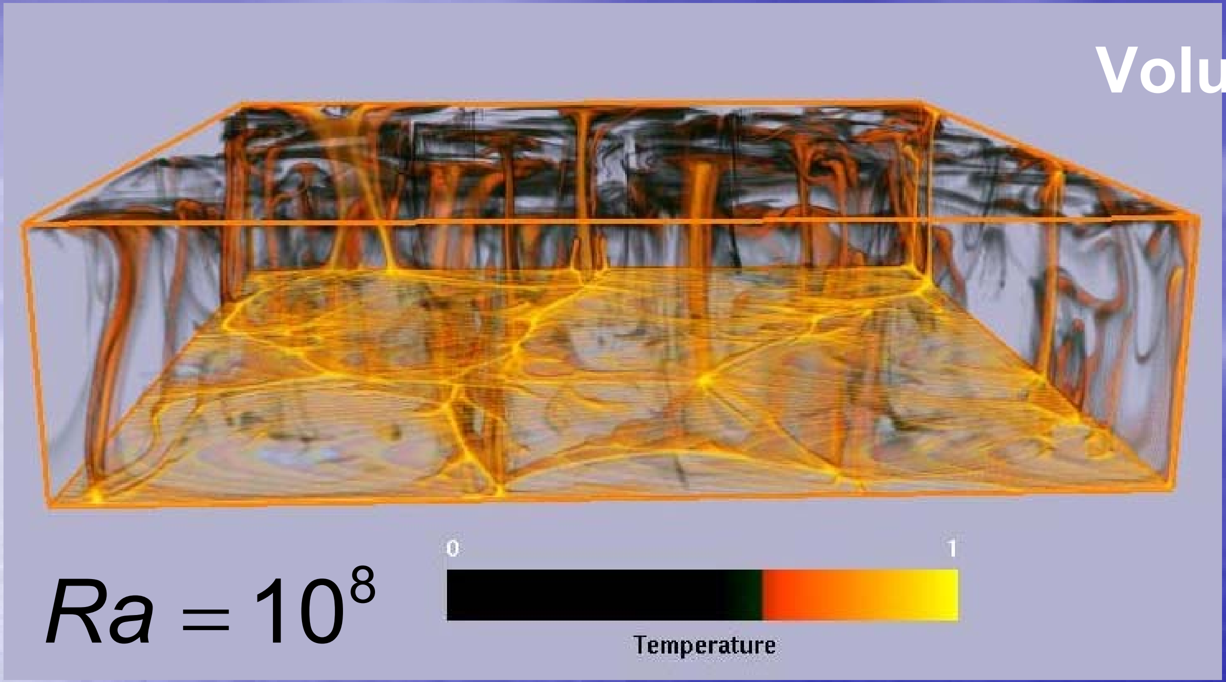
Video streaming

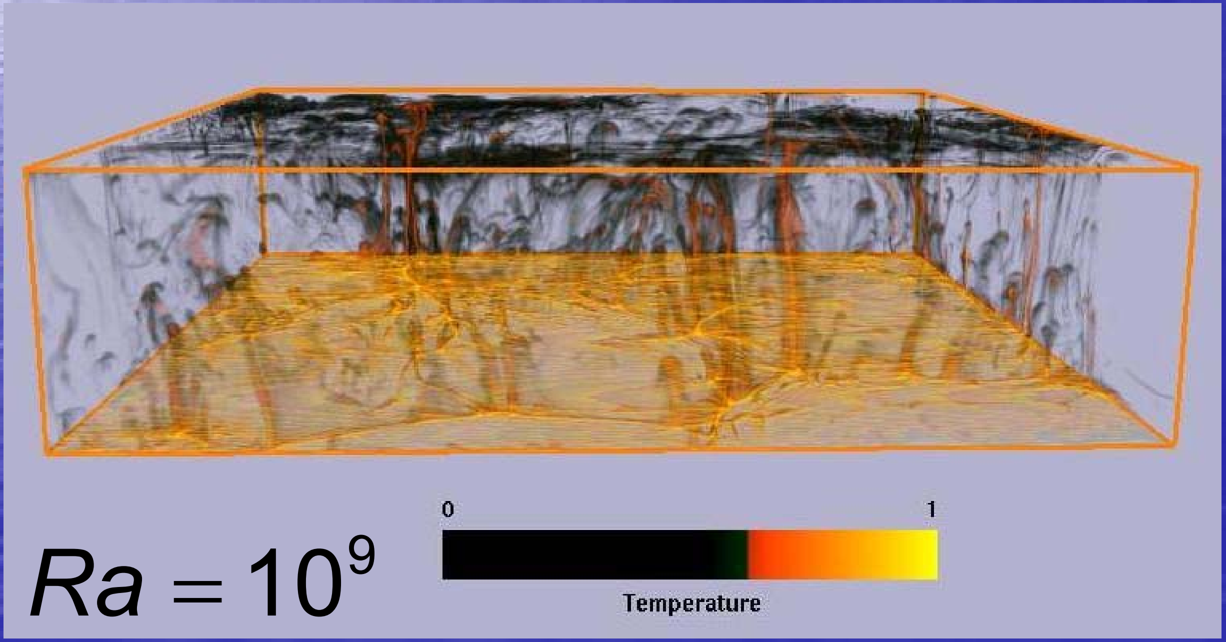
Compression

User-selectable scales

Locality in space and time

Volume rendering





Complex structures

Plumes get thinner at higher Ra

How do you quantify these structures

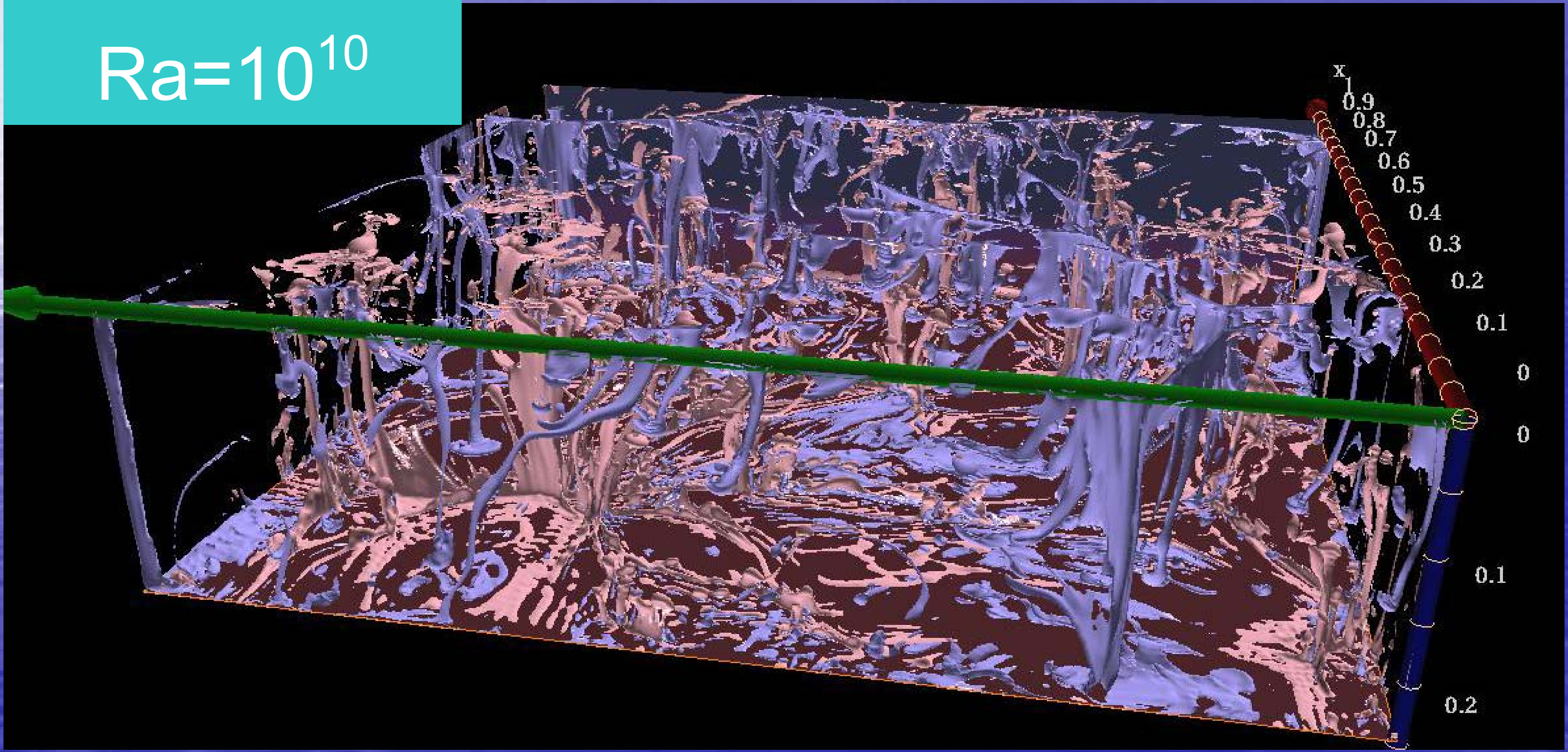
How to compute color maps automatically?

Mantle Convection

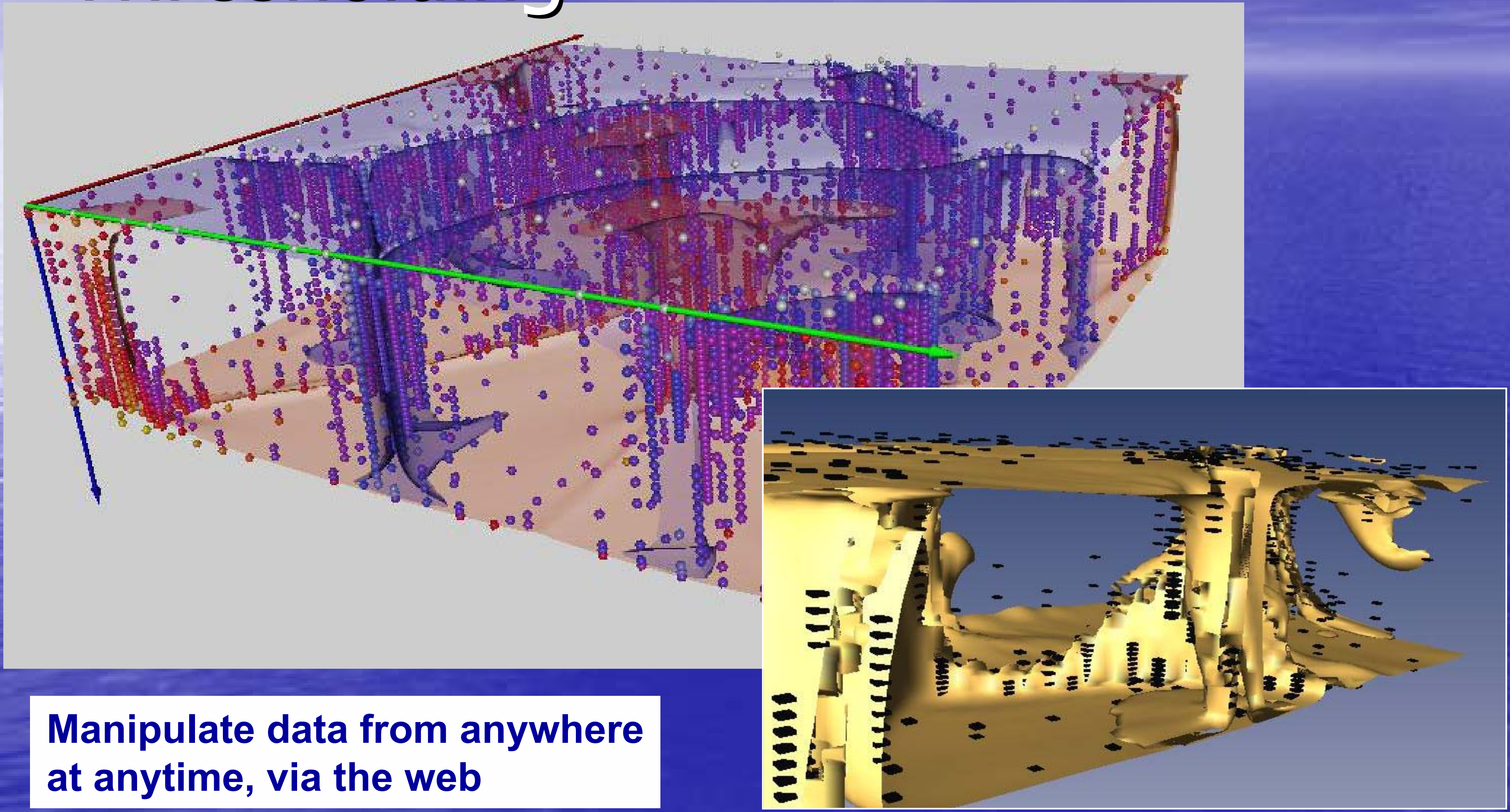
(simulations by F. Dubuffet and D. Yuen, U. Minnesota)

Grid: 500^3

$Ra=10^{10}$

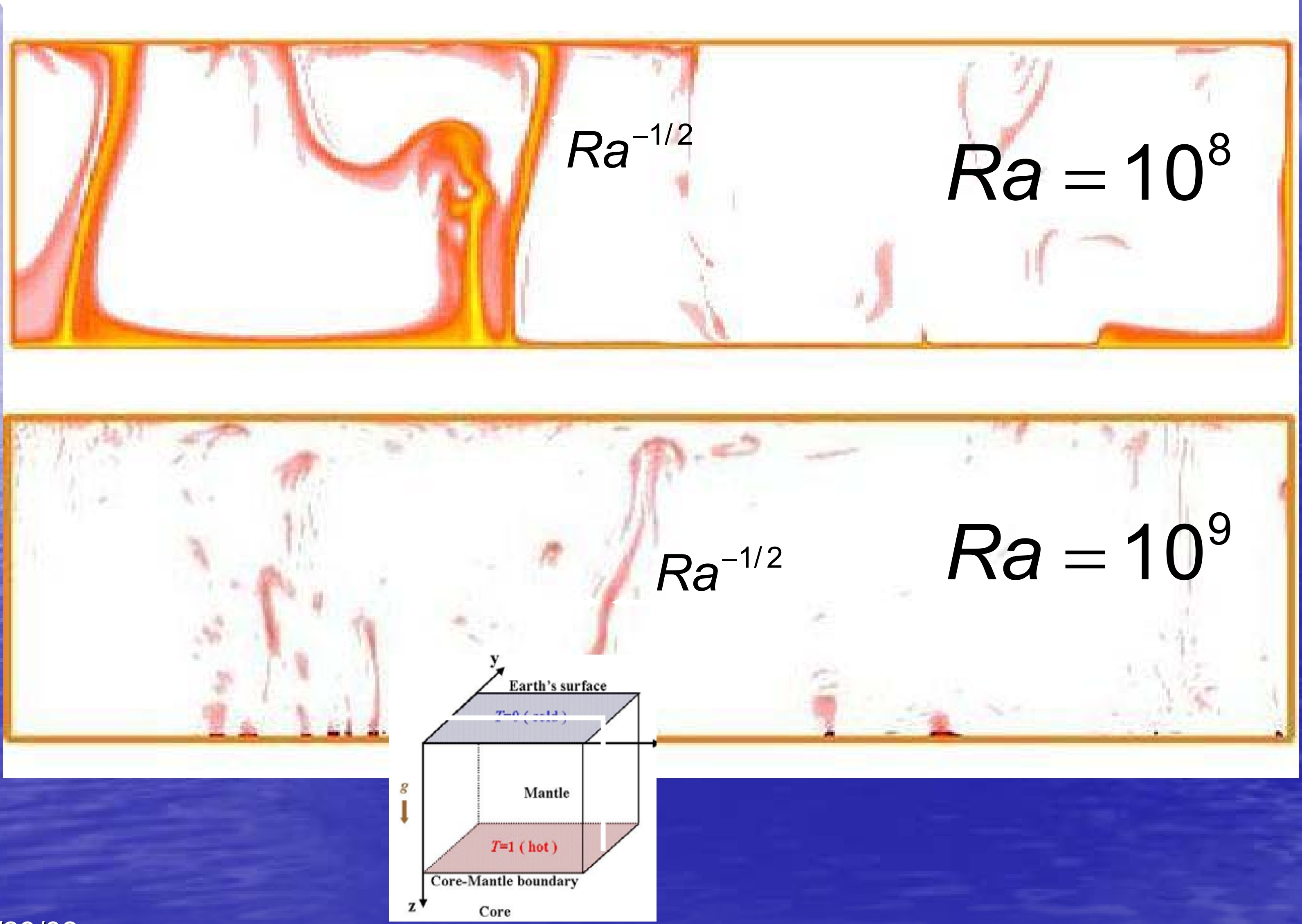


Data Reduction via Wavelet Thresholding

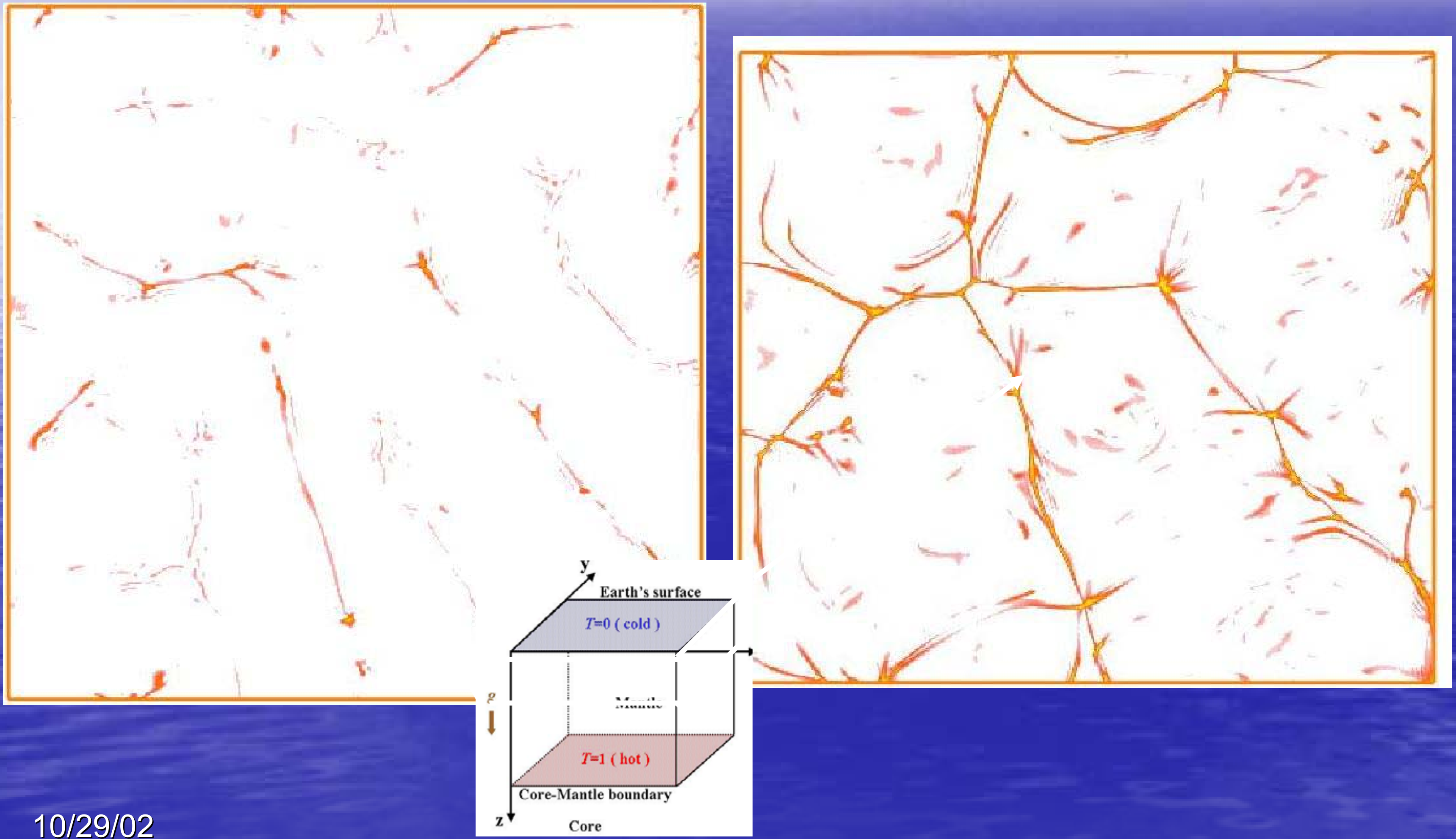


Manipulate data from anywhere at anytime, via the web

X-Z Cuts


$Ra^{-1/2}$
$Ra = 10^8$
$Ra^{-1/2}$
$Ra = 10^9$
y
Earth's surface
T=0 (cold)
Mantle
g
T=1 (hot)
Core-Mantle boundary
z
Core

$Ra = 10^8$, Grid: 401^3



Phase Diagrams

- Phases are well known in thermodynamics, e.g., pressure versus specific volume generates three phases (for water)
  - Water
  - Ice
  - Vapor
- This diagram has a triple point.

- More generally, diagrams in more complex systems might look like (in 2D): …

Phase Diagrams

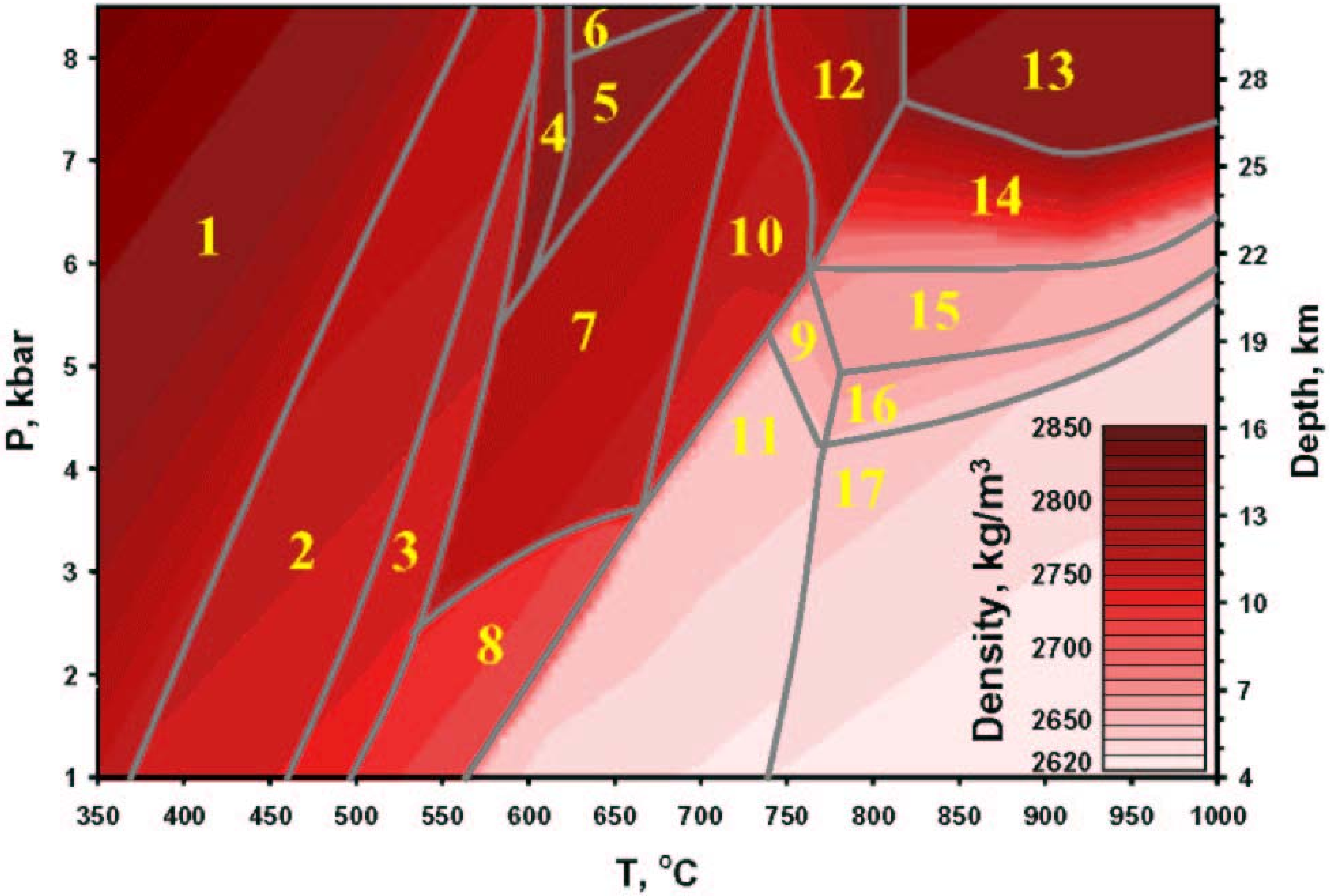
- Thermodynamic variables p, T, a, b, c, …
- Define a thermodynamic function

  - Phi(p,T,a,b,c,…) = integer

  Cost to evaluate phi(…) is very expensive.

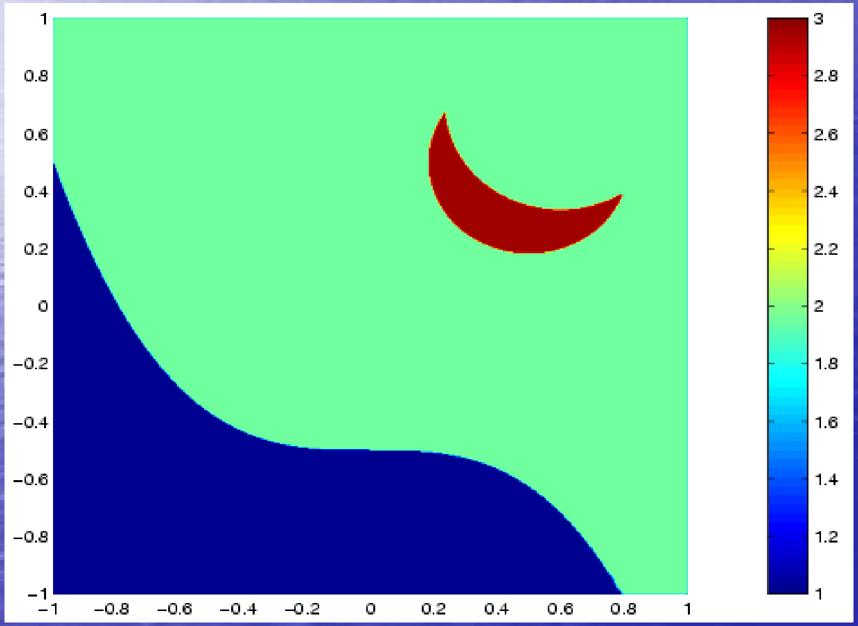
  Therefore: must minimize the number of evaluations.

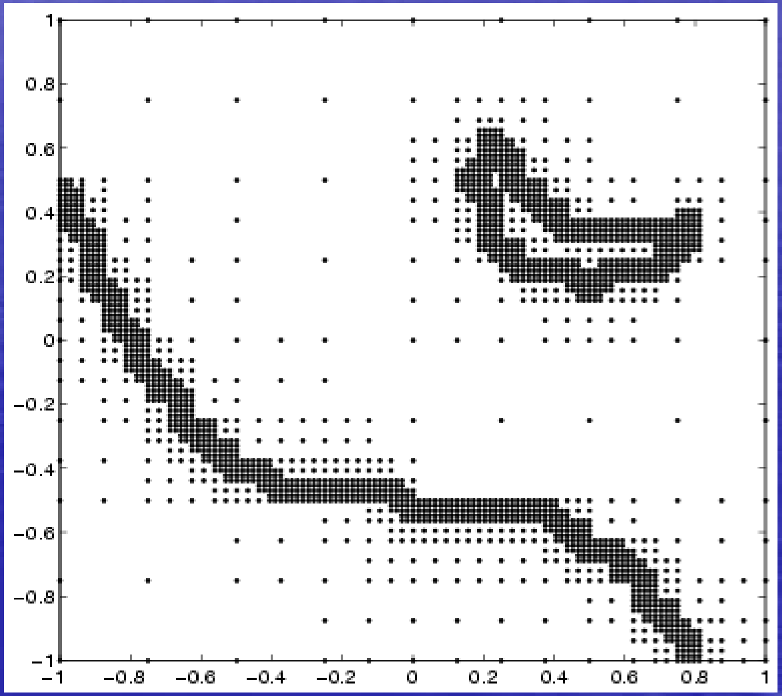
Phase diagram



Phase Diagram Extraction

Vasilyev, Yuen





Based on secondary wavelets

Challenge

- How to visualize phase diagrams in 3D, 4D and 5D space?
- How to understand the structure of the phases?

Future of Visualization

- Visualization is/has become multidisciplinary
- Successful visualization system must address
  - I/O
  - Maintainability
  - Flexibility (via plugins for example)
  - Accessibility (low cost and easy to use/install)
  - Robust
  - Standardization
- The above features are not consistent with each other

Visualization Ubiquity

- Collaboration through visualization
- Office walls become visualization displays (E-Ink: thin, pliable medium capable of electronic encoding)
- Exchange of visual data becomes as ubiquitous as exchange of text documents in 2001