Discretization, Solvers, and Statistics in Computational Geodynamics

http://59A2.org/files/20130423-EarthCube.pdf

Jed Brown

Mathematics and Computer Science Division, Argonne National Laboratory

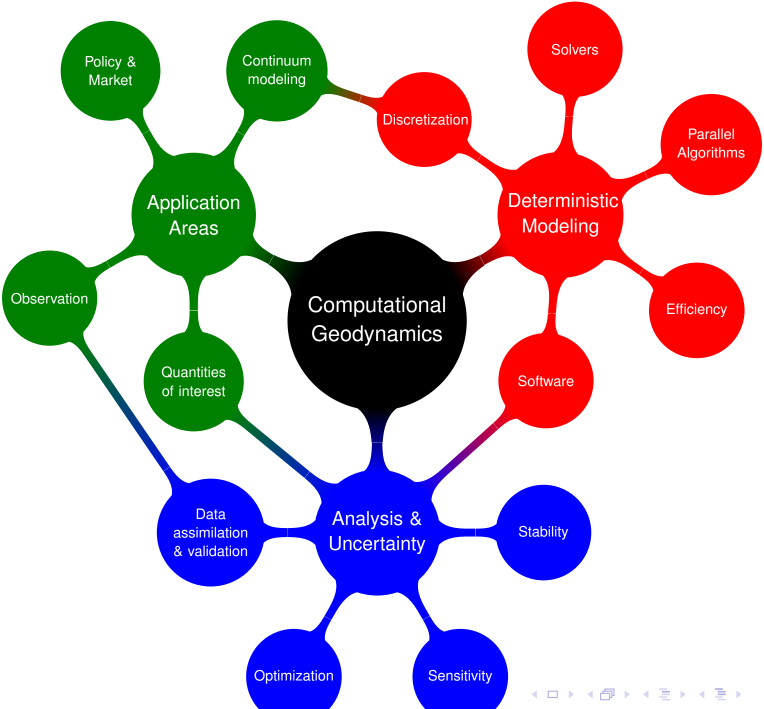
EarthCube, Boulder, CO, 2013-04-23

Challenges

- Discretization
  - high accuracy
  - heterogeneity and homogenization
  - tracers for material properties
- Solvers
  - stiff transient systems
  - elliptic problems
  - globalization for nonlinear problems
- Statistics
  - Seismic tomography
  - Data assimilation and validation
  - Experimental design
- Reusability and reproducibility
  - Libraries[1]
  - Common formats
  - Shared simulation software

---

[1] Disclaimer: I am a developer of PETSc.



Computational Geodynamics

Application Areas

Policy & Market

Continuum modeling

Observation

Quantities of interest

Deterministic Modeling

Discretization

Solvers

Parallel Algorithms

Efficiency

Software

Analysis & Uncertainty

Data assimilation & validation

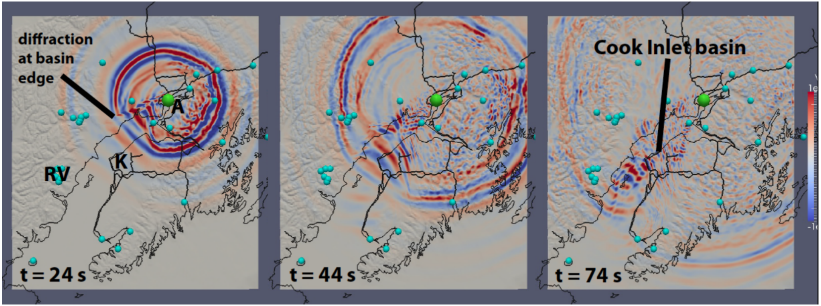
Stability

Optimization

Sensitivity

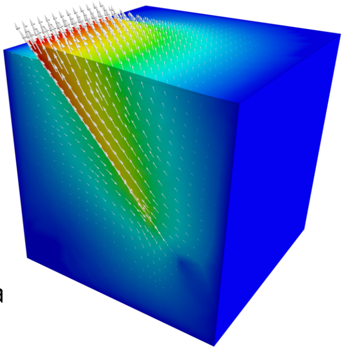
SPECFEM3D: Seismic wave propagation and tomography

- Spectral element methods: accurate, local, smooth solutions
- Linear materials
- Adjoint-based tomography
- `http://geodynamics.org/cig/software/specfem3d`



[c/o Carl Tape, UAF]

PyLith: Short-term Lithosphere



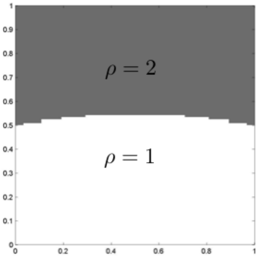
- Unstructured finite element methods
- Faults meshed-in (CUBIT, LaGriT)
- Cohesive cells and Lagrange multipliers
- Nonlinear materials and non-smooth beha
- Extensible material models and boundary
- Long time scales requires implicit solvers: fieldsplit and multigrid
- Libraries: PETSc (mesh and solvers), spatialdata (proj), numpy, FIAT (elements), HDF5
- `http://geodynamics.org/cig/software/pylith`

Stokes problems are ubiquitous in long-term geodynamics

$$\nabla \cdot (-\eta D u + p 1) = \rho g$$
$$\nabla \cdot u = c$$

- $Du = \frac{1}{2}\left[\nabla u + (\nabla u)^T\right]$, rheology $\eta(Du, \ldots)$
- Mantle, lithosphere, magma
- Coupled to other processes
  - Thermodynamics
  - Multi-material transport, chemistry
  - Plasticity/brittle failure: difficult non-smooth
  - Elasticity: typical Maxwell time of 1000 years
- Discontinuous coefficients: 10^{10} jumps
- Material properties defined using markers
- Discretization is difficult
  - Trade-offs between accuracy, robustness, and efficiency
  - What can go wrong? Next sequence from Dave May (ETHZ)

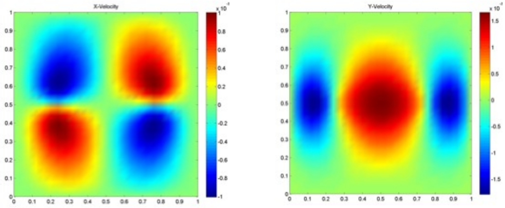
Q1Q1stab :: Incompressibility



Isoviscous

25x25
Q2Pm1

Reference solution



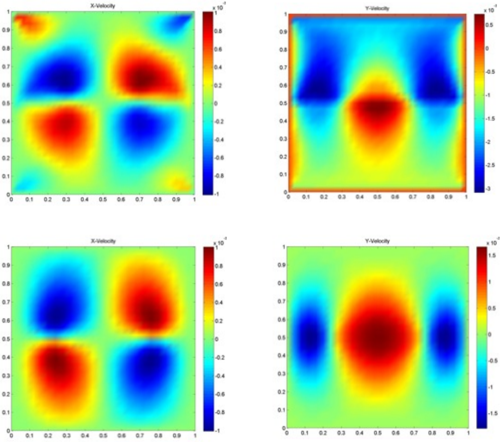
Q1Q1stab :: Incompressibility

Vx

Vy

30x30
Q1Q1stab

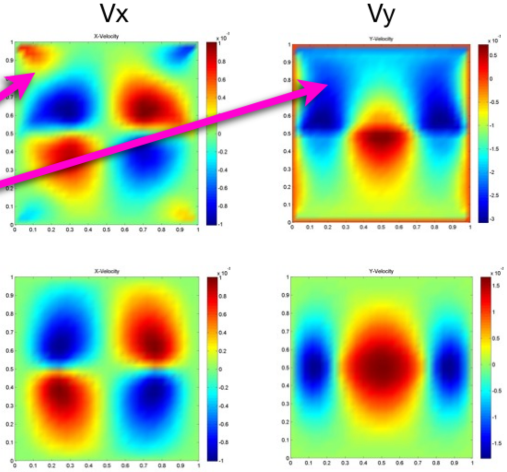
25x25
Q2Pm1



Yury Mishin

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Q1Q1stab :: Incompressibility



30x30
Q1Q1stab

1. Corner anomalies

2. Signs of artificial compaction due to mesh dependent incompressibility

Result of forward evolution is incorrect.

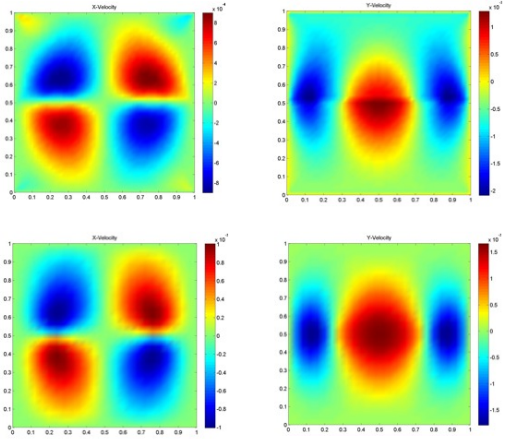
25x25
Q2Pm1

Q1Q1stab :: Incompressibility

Vx　Vy

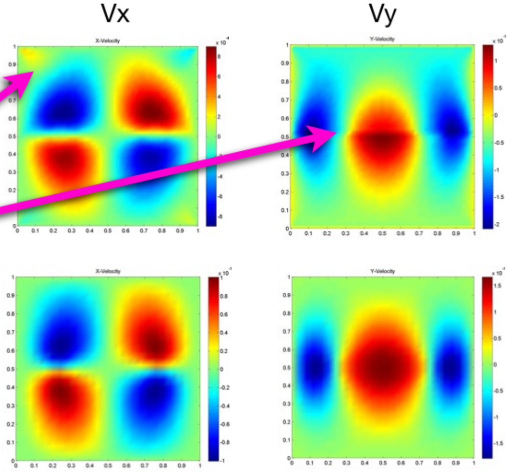
50x50
Q1Q1stab

25x25
Q2Pm1



Yury Mishin

Q1Q1stab :: Incompressibility



50x50
Q1Q1stab

1. Corner anomaly reduced

2. Artifacts at interface still present

Result of forward evolution is incorrect.

25x25
Q2Pm1

Yury Mishin

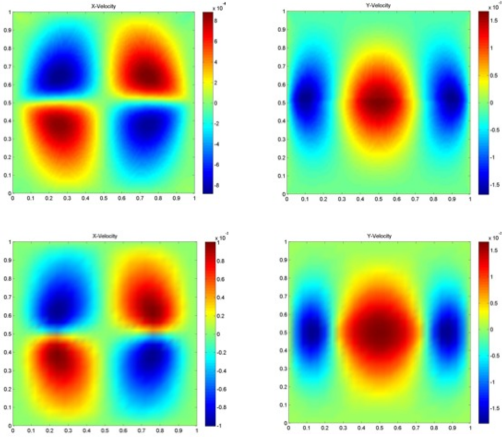
Q1Q1stab :: Incompressibility

Vx

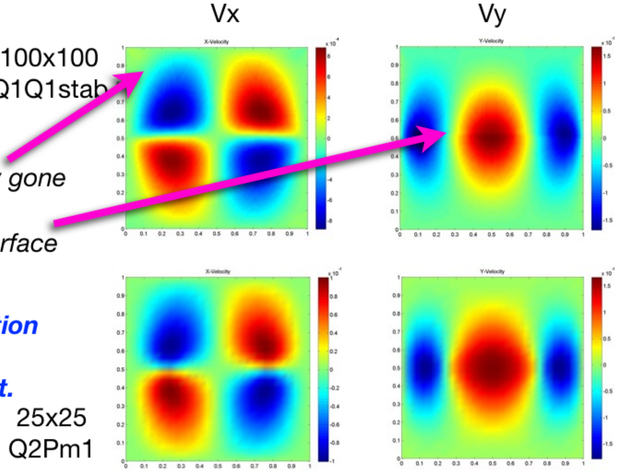
Vy

100x100
Q1Q1stab

25x25
Q2Pm1



Q1Q1stab :: Incompressibility



100x100
Q1Q1stab

1. Corner anomaly gone

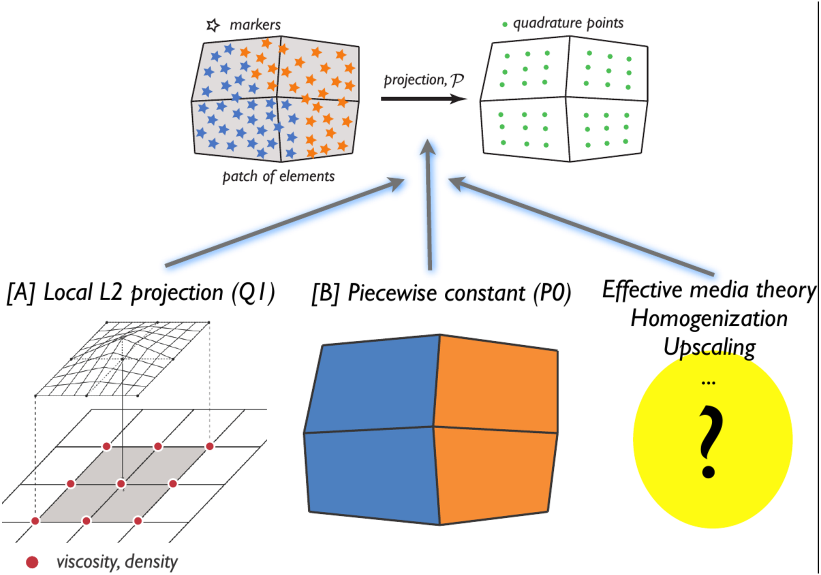
2. Artifacts at interface are milder

Forward evolution starts to be consistent.

25x25
Q2Pm1

storage(Q1Q1- 100x100) ≈ 2 x storage(Q2Pm1 - 25x25)

Material transport using markers



[A] Local L2 projection (Q1)

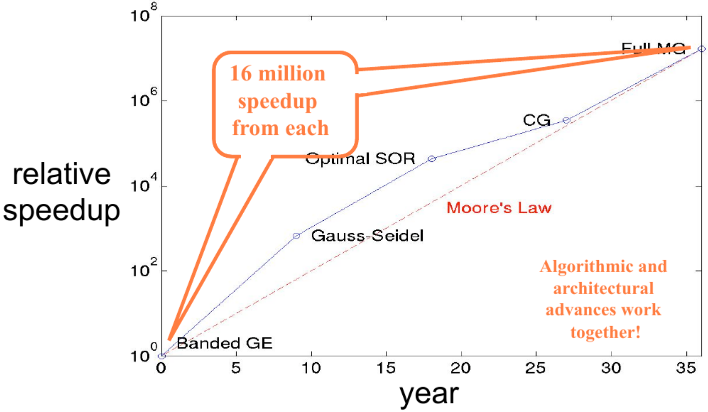
[B] Piecewise constant (P0)

Effective media theory
Homogenization
Upscaling

[c/o Dave May, ETHZ]

Algorithms keep pace with computing

- Consider an elliptic PDE on an $n \times n \times n$ grid
- Banded Gaussian Elimination: $\mathcal{O}(n^7)$
- Full Multigrid: $\mathcal{O}(n^3)$
- Optimal algorithms become more critical as we solve larger problems



16 million speedup from each

Algorithmic and architectural advances work together!

[c/o David Keyes, KAUST]

The Great Solver Schism: Monolithic or Split?

Monolithic

- Direct solvers
- Coupled Schwarz
- Coupled Neumann-Neumann (need unassembled matrices)
- Coupled multigrid

X Need to understand local spectral and compatibility properties of the coupled system

Split

- Physics-split Schwarz (based on relaxation)
- Physics-split Schur (based on factorization)
  - approximate commutators SIMPLE, PCD, LSC
  - segregated smoothers
  - Augmented Lagrangian
  - “parabolization” for stiff waves

X Need to understand global coupling strengths

- Preferred data structures depend on which method is used.
- Interplay with geometric multigrid.

Splitting for Multiphysics

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

- Relaxation: `-pc_fieldsplit_type [additive,multiplicative,symmetric_multiplicative]`

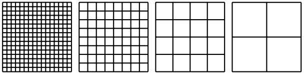
$$\begin{bmatrix} A & \\ & D \end{bmatrix}^{-1} \qquad \begin{bmatrix} A & \\ C & D \end{bmatrix}^{-1} \qquad \begin{bmatrix} A & \\ & 1 \end{bmatrix}^{-1} \left(1 - \begin{bmatrix} A & B \\ & 1 \end{bmatrix} \begin{bmatrix} A & \\ C & D \end{bmatrix}^{-1} \right)$$

  - Gauss-Seidel inspired, works when fields are loosely coupled
- Factorization: `-pc_fieldsplit_type schur`

$$\begin{bmatrix} A & B \\ & S \end{bmatrix}^{-1} \begin{bmatrix} 1 & \\ CA^{-1} & 1 \end{bmatrix}^{-1}, \qquad S = D - CA^{-1}B$$

  - robust (exact factorization), can often drop lower block
  - how to precondition S which is usually dense?
    - interpret as differential operators, use approximate commutators

Multigrid Preliminaries

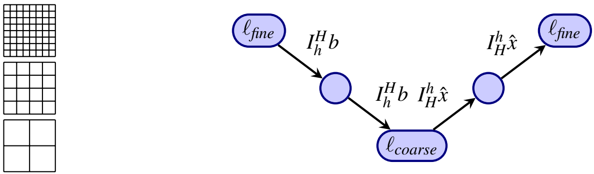


Multigrid is an $O(n)$ method for solving algebraic problems by defining a hierarchy of scale. A multigrid method is constructed from:

1. a series of discretizations
   - coarser approximations of the original problem
   - constructed algebraically or geometrically
2. intergrid transfer operators
   - residual restriction I_h^H (fine to coarse)
   - state restriction $\hat{I}_h^H$ (fine to coarse)
   - partial state interpolation I_H^h (coarse to fine, ‘prolongation’)
   - state reconstruction $\mathbb{I}_H^h$ (coarse to fine)
3. Smoothers (S)
   - correct the high frequency error components
   - Richardson, Jacobi, Gauss-Seidel, etc.
   - Gauss-Seidel-Newton or optimization methods

Linear Multigrid

- **Multigrid** methods use coarse correction for long-range influence



Algorithm $MG(A,b)$ for the solution of $Ax = b$:

$$x = S^m(x,b) \qquad \text{pre-smooth}$$

$$b^H = I_h^H(r - Ax) \qquad \text{restrict residual}$$

$$\hat{x}^H = MG(I_h^H A I_H^h, b^H) \qquad \text{recurse}$$

$$x = x + I_H^h \hat{x}^H \qquad \text{prolong correction}$$

$$x = x + S^n(x,b) \qquad \text{post-smooth}$$

Status quo for implicit solves in lithosphere dynamics

- global linearization using Newton or Picard
- assembly of a sparse matrix
- “block” factorization preconditioner, approximate Schur complement
- algebraic or geometric multigrid on positive-definite systems

Why is this bad?

- nonlinearities (e.g., plastic yield) are mostly local
  - feed back through nearly linear large scales
  - frequent visits to fine-scales even in nearly-linear regions
  - no way to locally update coarse grid operator
  - Newton linearization introduces anisotropy
- assembled sparse matrices are terrible for performance on modern hardware
  - memory bandwidth is very expensive compared to flops
  - fine-scale assembly costs a lot of memory
  - assembled matrices are good for algorithmic experimentation
- block preconditioners require more parallel communication

Reproducibility

- Geometry, Boundary, and Initial conditions
- Model configuration has poor reproducibility and automation
  - CAD software to create geometry
  - Interactive meshing (CUBIT)
  - Observational metadata
    - lack of uncertainties, correlation
    - diverse data sources, hard to quantify value
  - Interactive postprocessing
- Model execution *can* be reproducible
  - Exact versions in SCM (Git, Subversion)
  - Compilers, dependencies, configure- and run-time options
  - Postprocessing scripts

Data assimilation and experimental design

- Impact of geodynamics
  - Fundamental science questions
  - Hazards, safety, construction
  - Industry: minerals, petroleum
- Analysis tools more mature for faster processes
  - Short time scales and “single-physics” processes
  - Seismic tomography serves both science and industry
- More ad-hoc for longer term processes
  - More diverse data sources
  - Extremely indirect observations
  - Little meaning inferrable using single-physics models
  - Uncertainty propagation is under-developed
  - Non-smooth processes are troublesome for adjoints
- What measurements provide the most information?

Looking forward

- Is it good for everyone to write their own models?
  - Diversity is good for improving models
  - Creating a complete model from scratch is a lot of mundane work
  - Common interfaces allow users to compare multiple models
  - Libraries are a maintainable way to provide long-term reuse
  - Few models start out as libraries, some become libraries
  - Coupling necessary to understand long-term processes
- Scaling people
  - “Experts in everything” are valuable, but hard to find
  - The best algorithms remove comfortable abstractions like sparse matrices
  - Many open research topics: difficult to establish interfaces
- Postprocessing
  - Status quo is to write entire state to disk — not sustainable
  - Think like an engineer: ask precise questions — good for reproducibility