Robust and Flexible Newton-Krylov based Solution Methods for Nonlinear Coupled Multiphysics Systems

John N. Shadid

Computational Science R&D Group
Sandia National Laboratories
Albuquerque, NM

Collaborators:

Paul T. Lin
Roger P. Pawlowski
Ray S. Tuminaro

Sandia National Laboratories

Luis Chacon

Oak Ridge National Laboratory

GIG Workshop on Mathematical and Computational Issues in the Solid Earth Geosciences. Santa Fe, Sept 15 -17, 2008

Outline

- **Motivation: Multiple-time-scale Multi-physics Nonlinear Systems**
- **Outline of Example Systems of Coupled Nonlinear PDEs**
- **Why Newton-Krylov Methods?**
    - **Multiple-time-scale Systems**
    - **Characterization of Complex Solution Spaces**
    - **Optimization**
- **Solution Algorithm Performance**
    - **Parallel and Algorithmic Scaling of DD preconditioners**
    - **Multi-level Preconditioners**
        - **N-level Aggressive Coarsening Graph-based Block DD/AMG**
- **Conclusions**

Sandia National Laboratories

Motivation: Achieving Predictive Simulations of Complex Highly Nonlinear Multi-physics Systems (PDEs)

Specific Driving/Focusing Application Areas: MHD and Transport/Reaction Systems

What are multi-physics systems? (A multiple-time-scale perspective)

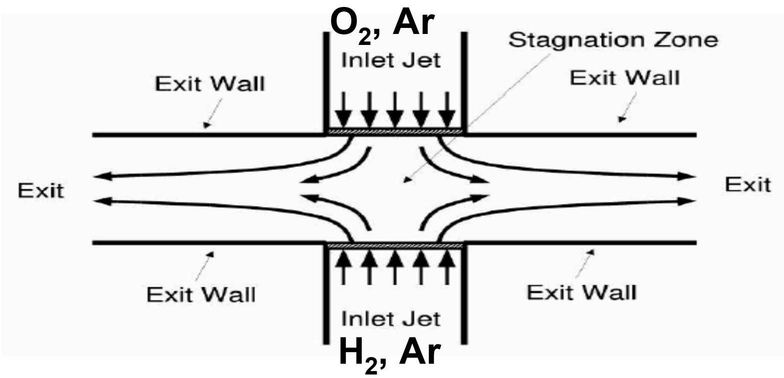
These systems are characterized by a myriad of complex, interacting, nonlinear multiple time- and length-scale physical mechanisms.

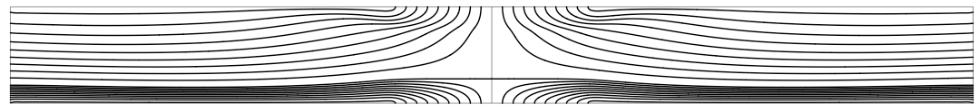
These mechanisms can balance to produce:

- **steady-state behavior,**
- **nearly balance to evolve a solution on a dynamical time scale that is long relative to the component time scales,**
- **or can be dominated by one, or a few processes, that drive a short dynamical time scale consistent with these dominating modes.**

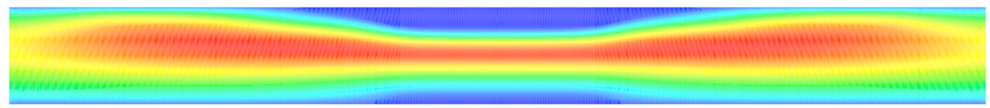
e.g. Fusion Reactors (Tokamak -ITER; Pulsed - NIF & Z-pinch); Fission Reactors (GNEP); Astrophysics; Combustion; Chemical Processing; Fuel Cells; etc.

Multiple-time-scale systems: Bifurcation Analysis of a Steady Reacting H_2, O_2, Ar, Opposed Flow Jet Reactor

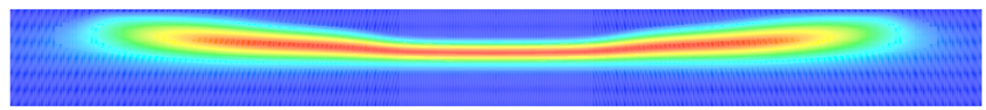




Streamlines

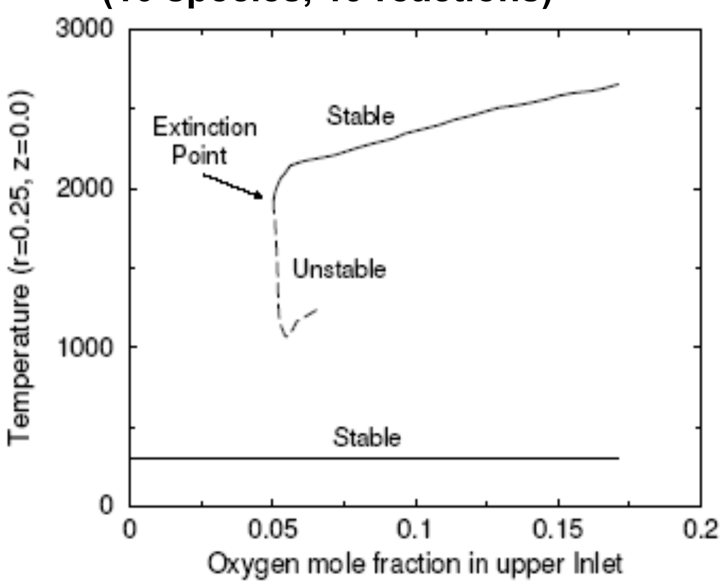


Temperature (Min. 300°K, Max 2727°K)



OH (Min. 0.0, Max 0.177)

70 steady state reacting flow solves (10 species, 19 reactions)



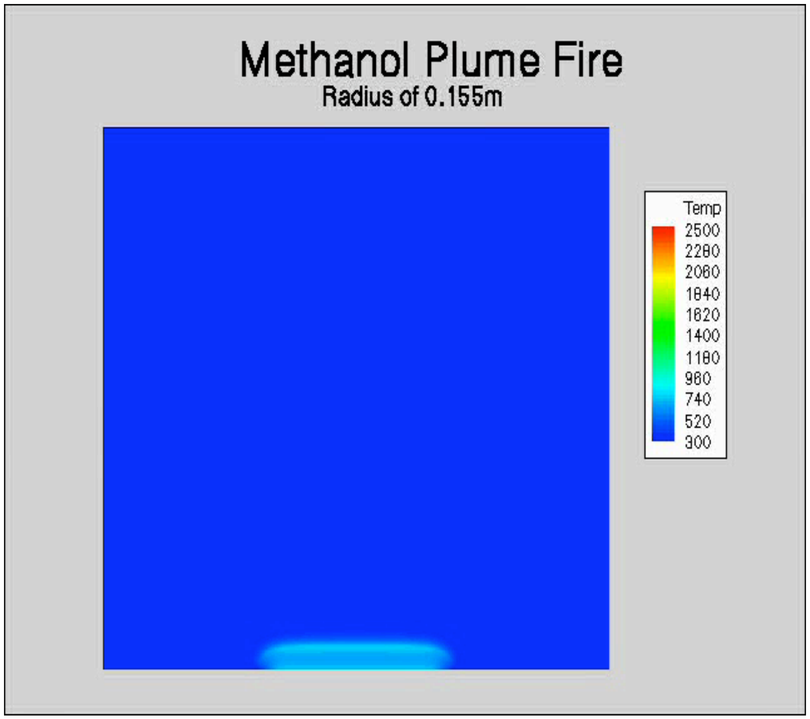
<u>**Approx. Physical Time scales (sec.):**</u>

- **Chemical kinetics: 10^{-12} to 10^{-4}**
- **Momentum diffusion: 10^{-6}**
- **Heat conduction: 10^{-6}**
- **Mass diffusion: 10^{-5} to 10^{-4}**
- **Convection: 10^{-5} to 10^{-4}**
- **Diffusion flame dynamics: ∞ (steady)**

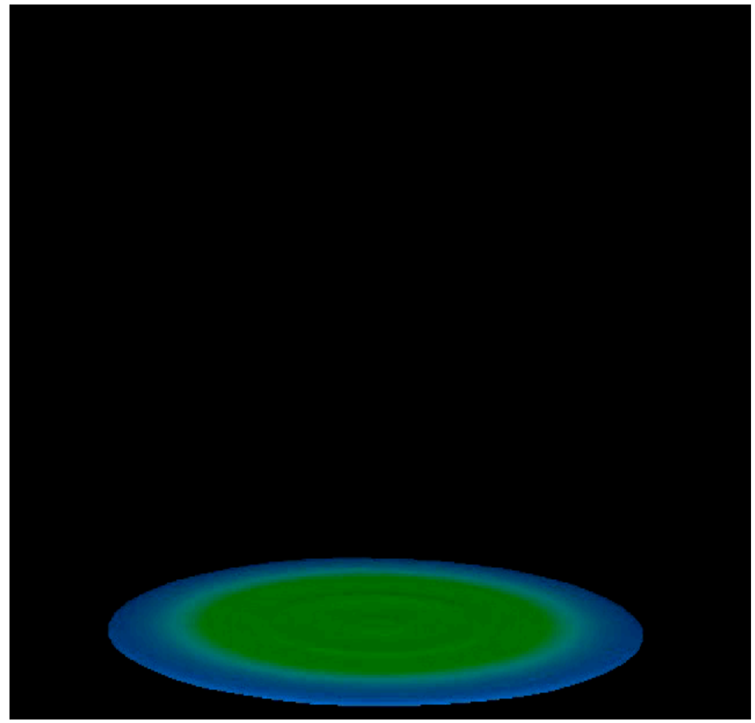
Sandia National Laboratories

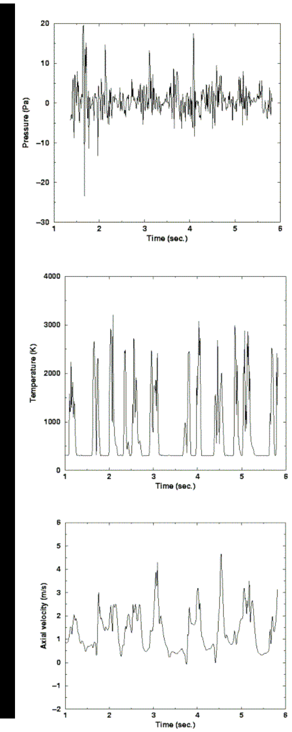
Multiple-time-scale systems: E.g. Methanol Pool Fire
LES-ksgs and Flamelet Combustion Model (w/ T. Smith – MPSalsa)

2D axisymmetric Simulation



Full 3D Simulation (note: non-axisymmetric mode)



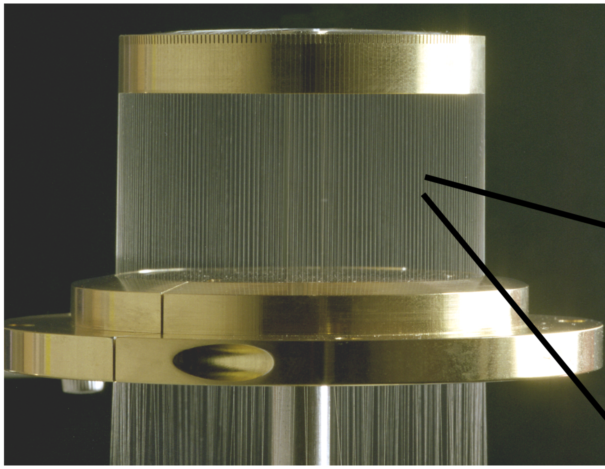


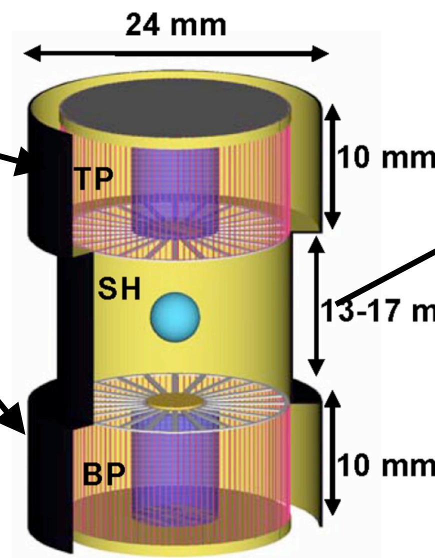
Approx. Physical Time scales (sec.):

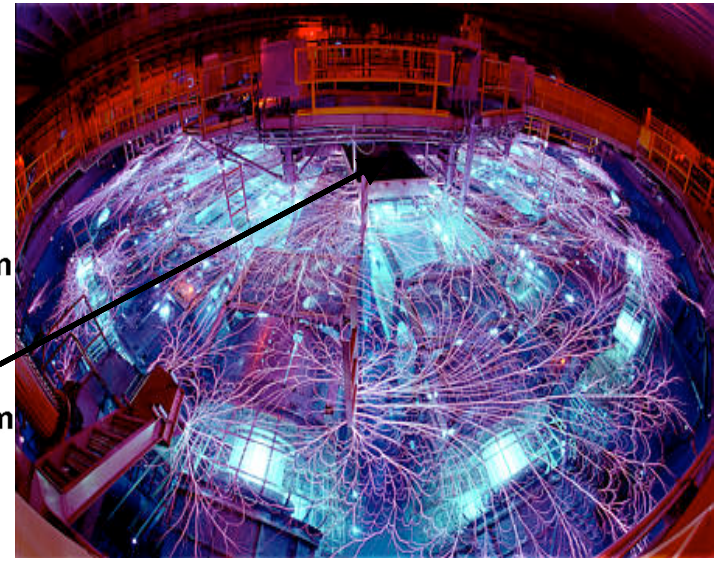
- **Chemical kinetics: 10^{-10} to 10^{-3}**
- **Momentum diffusion: 10^{-6}**
- **Heat conduction: 10^{-6}**
- **Convection: 10^{-3} to 10^{-1}**
- **Buoyancy (puffing freq. = 2.8Hz): 10^{-1} to 10^{0}**
- **Meandering mode: 10^{0}**

Sandia National Laboratories

Z-pinch Double Hohlraum Schematic







Z Machine (Approximate Ranges)

100ns current rise time for
20 MA Electrical Current

250 ns plasma shell collapse
and stagnation

10-30 ns X-ray power pulse
~280 TW power

A Recent Review: K. Matzen, et. al., POP 12, 055503 (2005)

Computational Stability Constraints:

Hyperbolic Operators: $\Delta t < \Delta x/2c$
- **Alfven waves**
- **Magneto-sonic waves**
- **Material transport**
- **Radiation transport**

Parabolic Operators: $\Delta t < \Delta x^2/D$
- **Magnetic Diffusion**
- **Heat Conduction**

Hall Physics: Whistler waves
-> $\Delta t < \Delta x^2/(V_A d_i)$

Sandia National Laboratories

Globalized Inexact Newton Method (incomplete citations)

Globalized Newton Methods

- **backtracking (line-search)**
- **trust region (dogleg)**
- **Dennis-Schnabel 1983**

Inexact Newton Methods

- **local theory**
  Dembo-Eisenstat-Steihaug 1982
- **global theory**
  Eisenstat-Walker 1994, Brown-Saad 1994
- **Linkage to linear solver criteria**
  Dembo-Eisenstat-Steihaug 1982, Eisenstat-Walker 1996

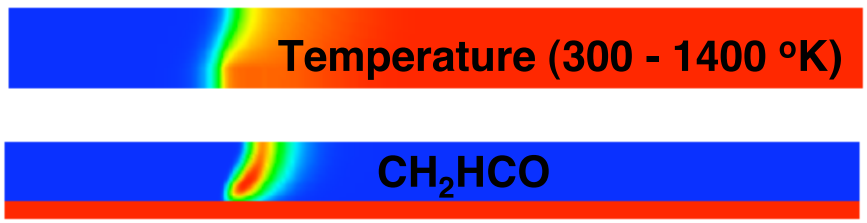
Globalized Newton-Krylov Methods

- **Use Krylov solvers to determine inexact Newton steps**
- **Backtracking and trust region-like globalizations- Robustness**
  Brown-Saad 1990, Shadid-Tuminaro-Walker 1997, Pernice-Walker 1998
- **Review: Jacobian free Newton methods:** Keyes-Knoll, 2003
- **Guide: Algorithms and implementation:** T. Kelley 2003
- **Review: Globalization techniques for Newton-Krylov:**
  Pawlowski-Shadid-Simonis-Walker, 2007
- **General algorithms and software**
  - **NKSOL (later KINSOL),** Brown-Saad 1990
  - **NITSOL,** Pernice-Walker 1998
  - **PETSc,** Balay-Gropp-Curfman McInnes-Smith 2001
  - **NOX** (Trilinos Solver Framework), Pawlowski-Kolda-Hooper 2002

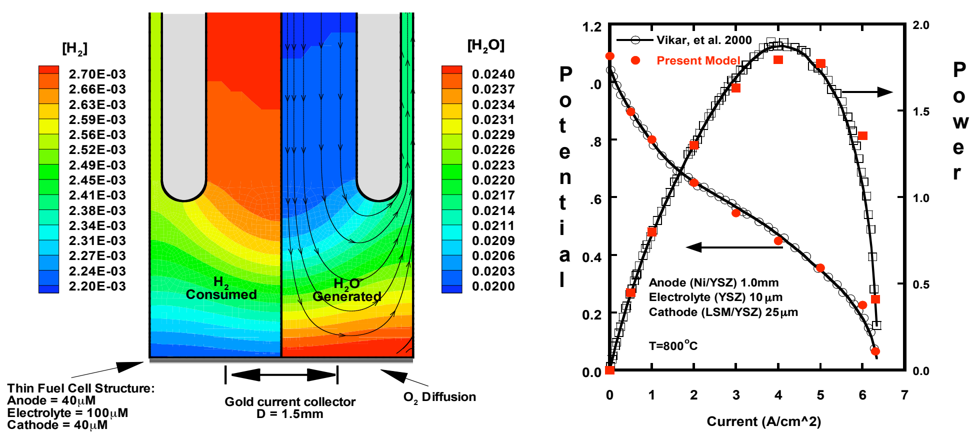
Sandia National Laboratories

A very broad range of scientific and engineering applications require the high-resolution computational analysis of strongly coupled nonlinear multiple-time-scale multiphysics systems.

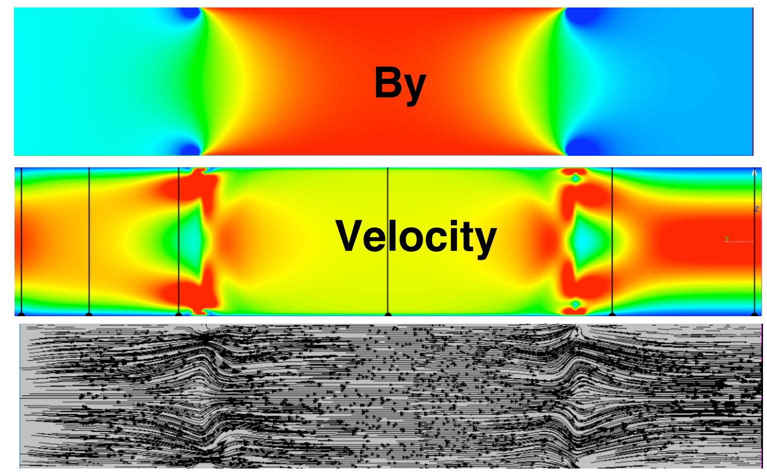
E. g. Transport / Reaction Systems, MHD



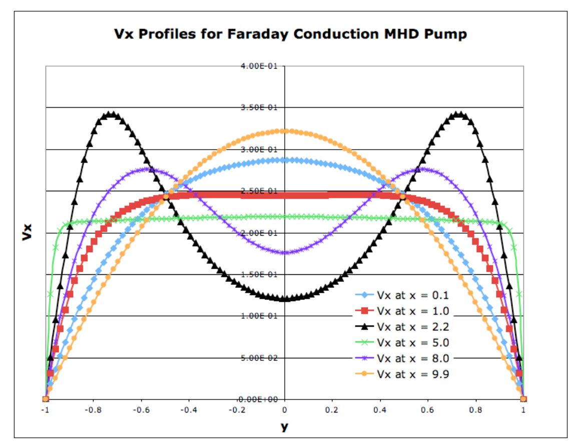
Partial Catalytic Oxidation of Ethane on Pt
22 gas-phase species, 77 reactions
17 surface-phase species, 35 reactions



Simulation of Experimental Solid Oxide H_2 Fuel Cell



MHD Pump



Sandia National Laboratories

Transport / Reaction and Resistive MHD Models

Navier Stokes

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla \cdot [\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{T}] - \rho \mathbf{g} = 0; \quad \mathbf{T} = -\left(P + \frac{2}{3}\mu(\nabla \bullet \mathbf{u})\right)\mathbf{I} + \mu[\nabla \mathbf{u} + \nabla \mathbf{u}^T]$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot [\rho \mathbf{v}] = 0$$

$$\frac{\partial(\rho e)}{\partial t} + \nabla \cdot [\rho \mathbf{v} + \mathbf{q}] = 0$$

Discretization - Extensions of Stabilized FE (Hughes et. al)
Q1/Q1 V-P elements, SUPG like terms and
Discontinuity Capturing type operators

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{v}} \\ \dot{\mathbf{P}} \end{bmatrix} + \begin{bmatrix} \mathbf{A} & -\mathbf{B}^T \\ \mathbf{BR} & \mathbf{K} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{P} \end{bmatrix}$$

General Case a Strongly Coupled, Multiple Time- and Length-Scale, Nonlinear, Nonsymmetric System with Parabolic and Hyperbolic Character

Transport / Reaction and Resistive MHD Models

Navier Stokes + Transport / Reaction Physics

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla \cdot [\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{T}] - \rho \mathbf{g} = 0; \quad \mathbf{T} = -\left(P + \frac{2}{3}\mu(\nabla \bullet \mathbf{u})\right)\mathbf{I} + \mu[\nabla \mathbf{u} + \nabla \mathbf{u}^T]$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot [\rho \mathbf{v}] = 0$$

$$\frac{\partial(\rho e)}{\partial t} + \nabla \cdot [\rho \mathbf{v} + \mathbf{q}] + \sum_{k=1}^{N} \mathbf{j}_k \cdot \hat{C}_{p,k} \nabla T - \sum_{k=1}^{N} h_k W_k \omega_k = 0$$

Species Transport / Reaction Equation

$$\frac{\partial(\rho Y_k)}{\partial t} + \nabla \bullet (\mathbf{u} Y_k + \mathbf{j}_k) - W_k \dot{\omega}_k; \qquad k = 1, 2, \ldots, N-1 \; ; \; \sum_{k=1}^{N} Y_k = 1$$

General Case a Strongly Coupled, Multiple Time- and Length-Scale, Nonlinear, Nonsymmetric System with Parabolic and Hyperbolic Character

Sandia National Laboratories

Transport / Reaction and Resistive MHD Models

Navier Stokes + Electro-magnetics

$$\frac{\partial(\rho \mathbf{v})}{\partial t} + \nabla \cdot [\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{T}] - \mathbf{J} \times \mathbf{B} - \rho \mathbf{g} = 0 \;\; ; \quad \mathbf{T} = -\left(P + \frac{2}{3}\mu(\nabla \bullet \mathbf{u})\right)\mathbf{I} + \mu[\nabla \mathbf{u} + \nabla \mathbf{u}^T]$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot [\rho \mathbf{v}] = 0$$

$$\frac{\partial(\rho e)}{\partial t} + \nabla \cdot [\rho \mathbf{v} + \mathbf{q}] - \eta \|\mathbf{J}\|^2 = 0$$

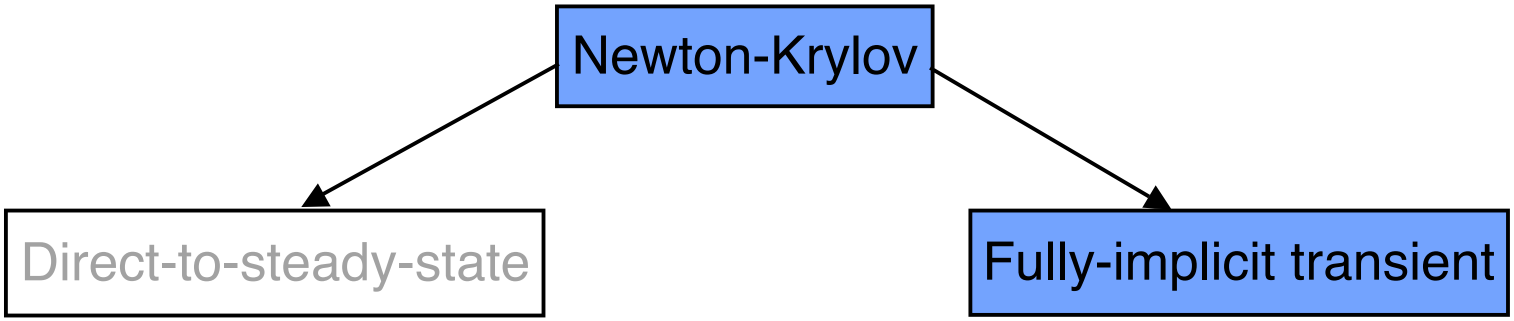
Reduced form of Maxwell's Equations

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times [\mathbf{v} \times \mathbf{B}] + \nabla \times (\eta \mathbf{J}) = 0 \; ; \qquad \mathbf{J} = \frac{1}{\mu_0} \nabla \times \mathbf{B}$$

General Case a Strongly Coupled, Multiple Time- and Length-Scale, Nonlinear, Nonsymmetric System with Parabolic and Hyperbolic Character

Sandia National Laboratories

Why Newton-Krylov Methods?



$$\mathbf{F}(\dot{\mathbf{x}}, \mathbf{x}, \lambda_1, \lambda_2, \lambda_3, ..) = \mathbf{0}$$

e.g.

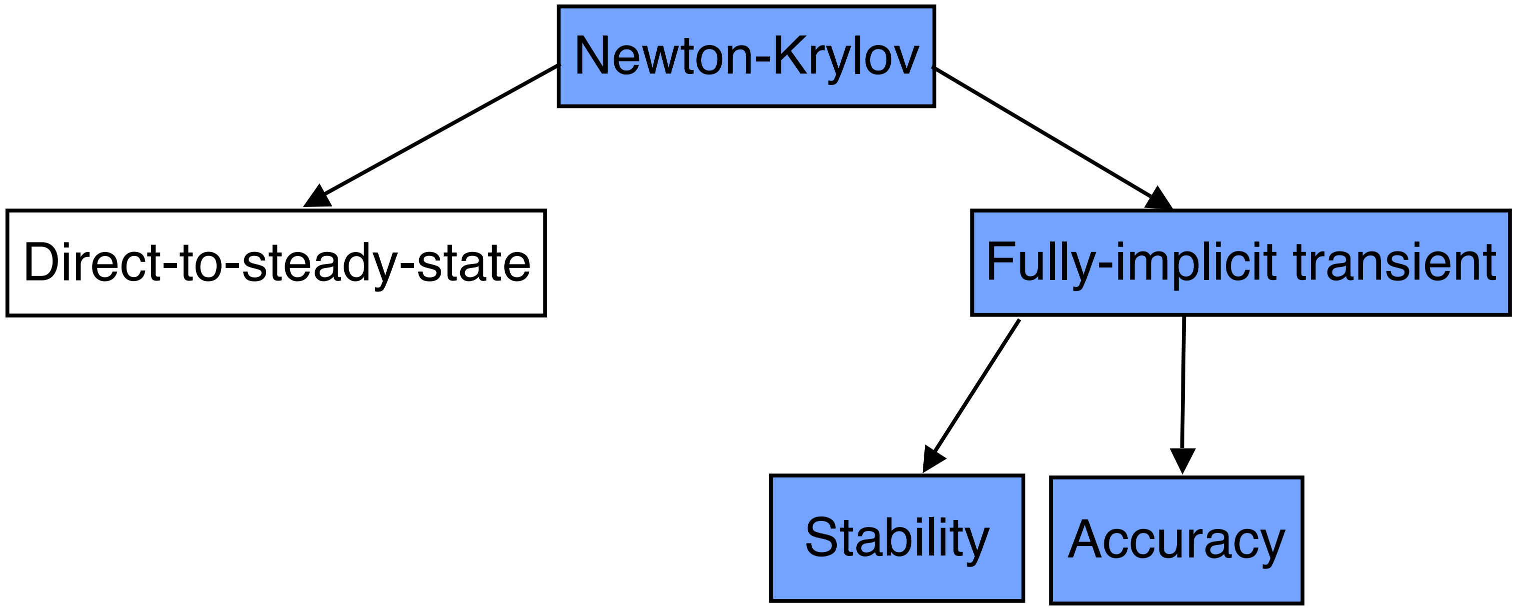
$$\left.\frac{\partial c}{\partial t}\right|^{n+1} + \nabla \bullet \left([\rho c \mathbf{u}]^{n+1} \right) - \nabla \bullet \left[D^{n+1} \nabla c^{n+1} \right] + S_c^{n+1} = 0$$

<u>Stability and Accuracy Properties</u>

- Stable (stiff systems)
- High order methods
- Variable order techniques
- Local and global error control possible
- Can be stable and accurate run at the dynamical time-scale of interest in multiple-time-scale systems

Sandia National Laboratories

Why Newton-Krylov Methods?



(Dan Reynolds previous talk)

Multiple-time-scale systems: Numerical Experiments
Chemical Dynamics (Brusselator)

$$\frac{\partial T}{\partial t} = D_1 \frac{\partial^2 T}{\partial x^2} + \alpha - (\beta + 1)T + T^2 C$$

$$\frac{\partial C}{\partial t} = D_2 \frac{\partial^2 C}{\partial x^2} + \beta T - T^2 C$$

$D_1 = D_2 = 1/40$

$\alpha = 0.6$

$\beta = 2.0$

$\Delta x = 1/100$

$T_{min} \approx 10.0$

Fully-implicit Method: Trapezoidal Rule

2nd order (FI 2nd):

$$M_k(\dot{\chi}^{n+1}) + D_k^{n+1}(\chi^{n+1}) + S_k^{n+1}(\chi^{n+1}) + F_k = 0$$

$$\dot{\chi}^{n+1} = 2\left(\frac{\chi^{n+1} - \chi^n}{\Delta t}\right) - \dot{\chi}^n$$

Strang Splitting (SS):

to advance solution over $\left[t^n, t^n + \Delta t\right]$

$$M_k(\dot{\chi}^*) + D_k^*(\chi^*) + F_k = 0 \quad \text{on } [0, \Delta t/2]$$

$$M_k(\dot{\chi}^{**}) + S_k^{**}(\chi^{**}) = 0 \quad \text{on } \left[0, \Delta t\right]$$

$$M_k(\dot{\chi}^{***}) + D_k^{**}(\chi^{***}) + F_k = 0 \quad \text{on } [0, \Delta t/2]$$

$$\chi^{n+1} = \chi^{***}(\Delta t) \longrightarrow \chi^{n+1} = \tilde{D}_{\Delta t/2} \tilde{S}_{\Delta t} \tilde{D}_{\Delta t/2} \chi^n$$

(w/David Ropp, C. Ober)

G. Strang, SIAM J. Numer. Anal. 5,3, 1968

Sandia National Laboratories

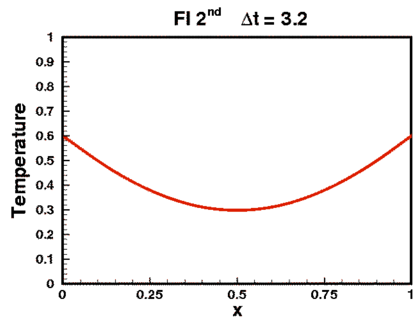
Diffusion/Reaction System

Operator Split Component solvers:

- **Diffusion: 2nd order Crank-Nicholson Galerkin FE (A-stable)**
  2nd order SDIRK Galerkin FE (A & L -stable)

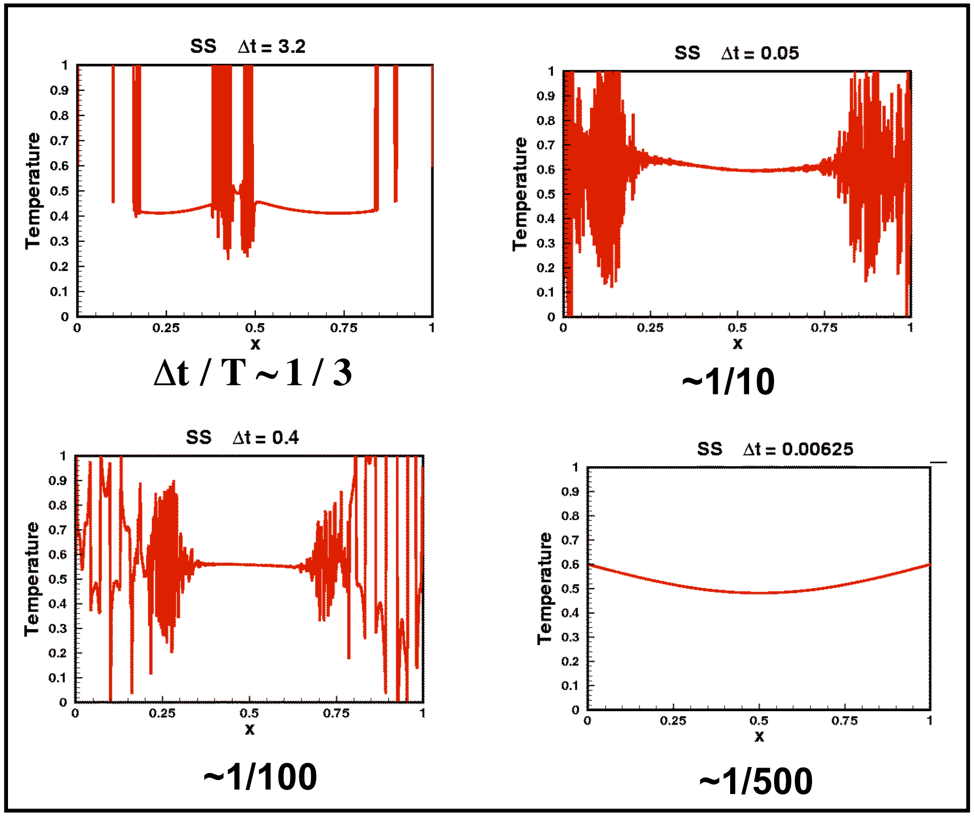
- **Reaction: CVODE Variable order - High accuracy tolerances**

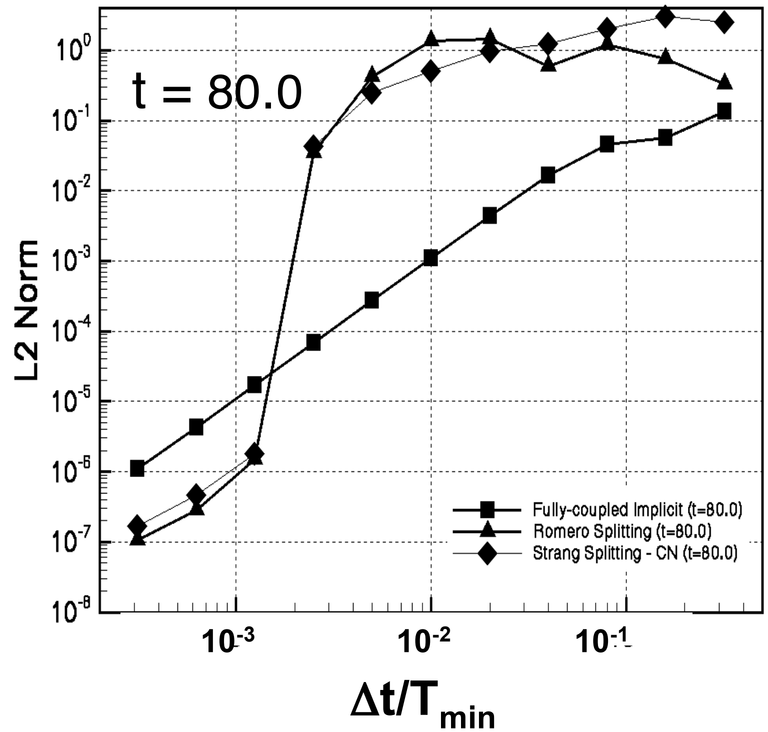
Brusselator: Comparison of Spatial and Temporal Profiles for Strang Split and Fully Implicit Solvers



FI

$\Delta t / T \sim 1/3$





Multiple time scales:
Knoll, Chacon, Margolin, Mousseau, JCP 2003
Ropp, S., JCP 2004, 2005
Ober, S. JCP 2004

Sandia National Laboratories

Brusselator: L-stability of diffusion solve is critical for stability (SDIRK)

- Parameter γ determines limit of amplification factor "*R* " as $\lambda \Delta t \to -\infty$

Case 1: A-stable, 2nd order

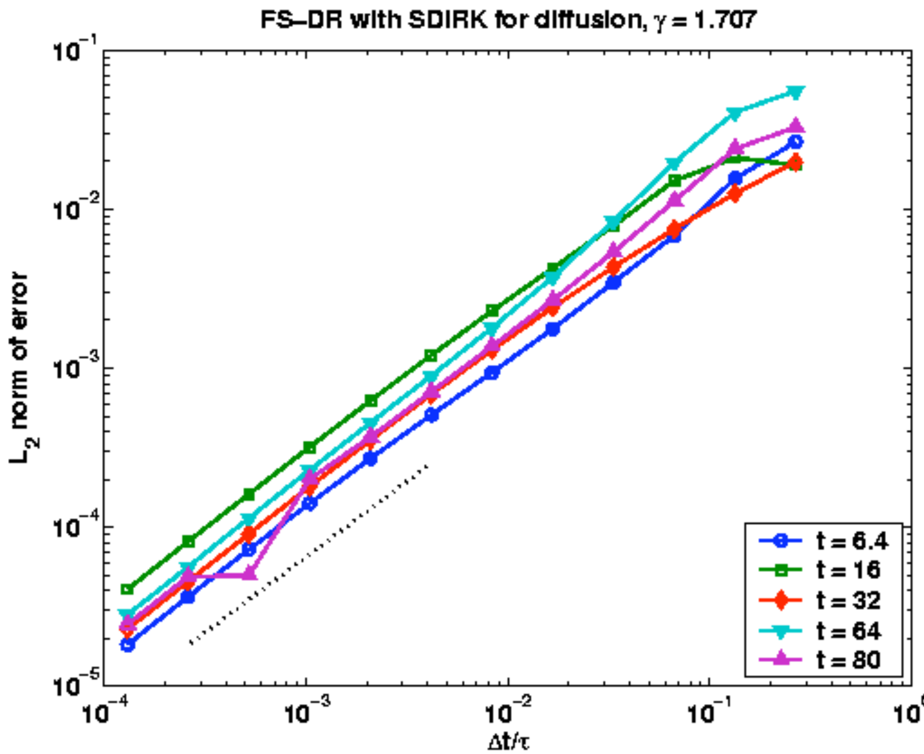
$$\gamma = 0.5, \lim_{z \to -\infty} R(z) = -1$$

Case 2: A-stable, 3rd order

$$\gamma = 0.789, \lim_{z \to -\infty} R(z) = -0.455$$

Case 3: A- and L-stable, 2nd order

$$\gamma = 1.707, \lim_{z \to -\infty} R(z) = 0$$



First order splitting with A- and L-stable diffusion solves demonstrate effect of damping of high wavenumber instability

Ropp, S., JCP 2004, 2005
Ober, S., JCP 2004

Sandia National Laboratories

Convection/Diffusion/Reaction System

Operator Split Component solvers:

- **Advection: 2nd order implicit FE-FCT Kuzmin et. al. (2000)**
- **Diffusion: 2nd order Crank-Nicholson Galerkin FE (A-stable)**
  2nd order SDIRK Galerkin FE (A & L -stable)
- **Reaction: CVODE Variable order - High accuracy tolerances**

A-stability of Operator Split Integration of Convection/Diffusion/Reaction System: Initial Results

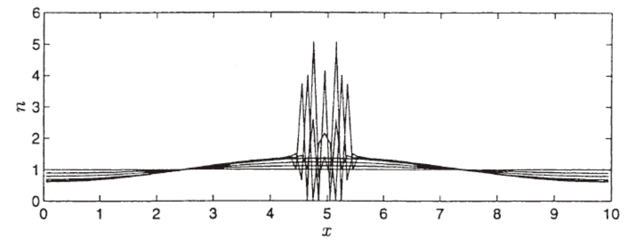
$$\frac{\partial n}{\partial t} + \nabla \bullet \left(\left[\alpha \nabla c \right] n \right) - D_n \nabla^2 n = 0$$

$$\frac{\partial c}{\partial t} - D_c \nabla^2 c + nc = 0$$

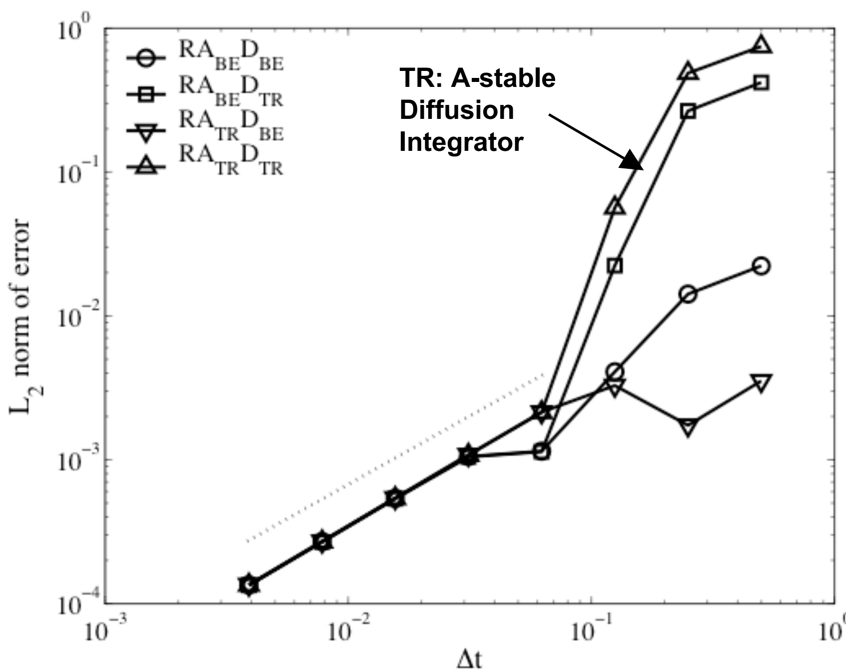
n - cell density;
C - chemo-attractant concentration;

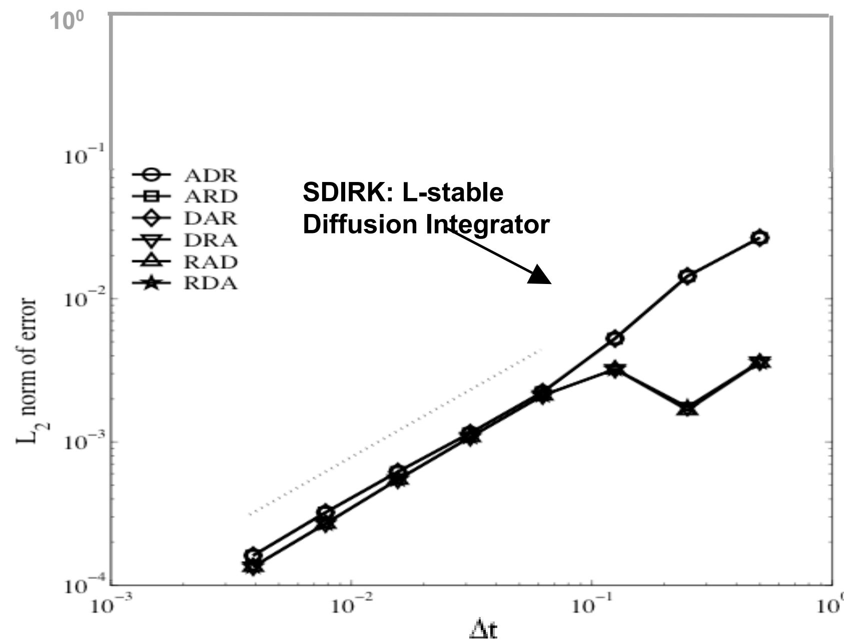
$D_n = D_c = 0.1$ and $\alpha = 2$

$n(t=0) = 1, \quad c(t=0) = 1 - \dfrac{\cos(\pi x/5)}{4}.$



Rebecca Tyson · L.G. Stern · Randall J. LeVeque
Fractional step methods applied to a chemotaxis model
J. Math. Biol. 41, 455–475 (2000)

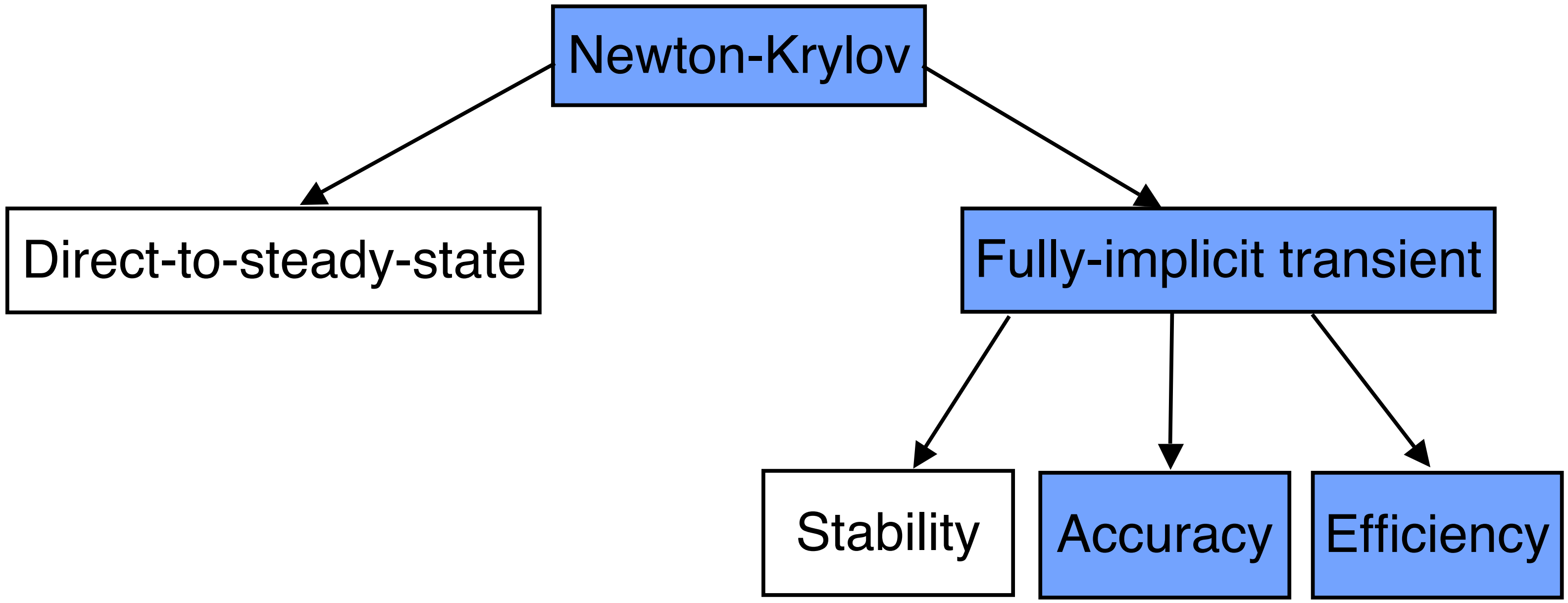




Ropp, S., Submitted to JCP

Sandia National Laboratories

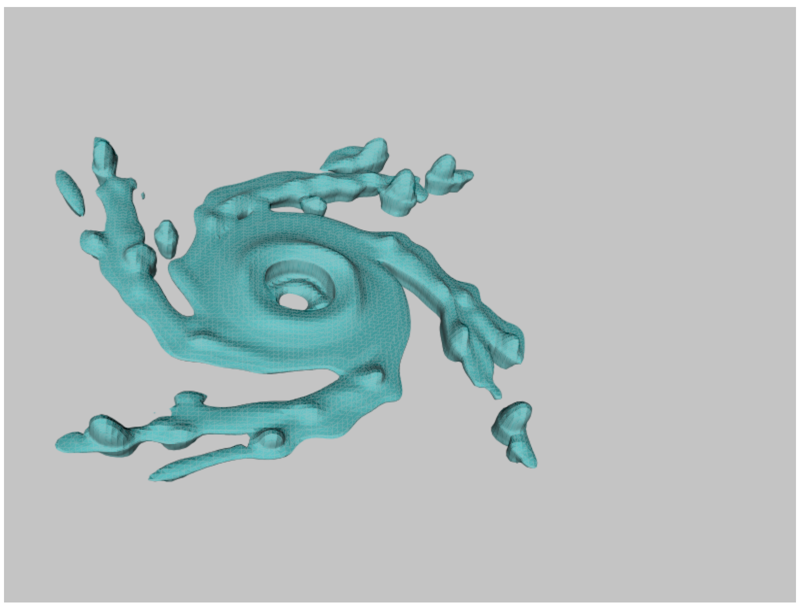
Why Newton-Krylov Methods?



Multiple-time-scale Systems: Newton-Krylov Methods for Hurricane Simulations

(Riesner, Mousseau, Wyszogrodzki, Knoll, MWF 2004)

- **3D compressible N-S & phase change**
- **Error/CPU time Comparison of**
  - **Semi-implicit (SI)**
  - **JFNK with SI as preconditioner**
- **Study transient hurricane intensification to ramped increase in sea surface temperature**



(Courtesy of D. Knoll - LANL)

Hurricane Equation Set

$$\frac{\partial u\rho}{\partial t} + \frac{\partial uu\rho}{\partial x} + \frac{\partial vu\rho}{\partial y} + \frac{\partial wu\rho}{\partial z} = -\frac{\partial p'}{\partial x}$$
$$+ f\rho(v - v_e) - \tilde{f}w + \frac{\partial \kappa\rho\tau^{11}}{\partial x} + \frac{\partial \kappa\rho\tau^{12}}{\partial y} + \frac{\partial \kappa\rho\tau^{13}}{\partial z}, \quad (1)$$

$$\frac{\partial v\rho}{\partial t} + \frac{\partial uv\rho}{\partial x} + \frac{\partial vv\rho}{\partial y} + \frac{\partial wv\rho}{\partial z} = -\frac{\partial p'}{\partial y}$$
$$- f\rho(u - u_e) + \frac{\partial \kappa\rho\tau^{21}}{\partial x} + \frac{\partial \kappa\rho\tau^{22}}{\partial y} + \frac{\partial \kappa\rho\tau^{23}}{\partial z}, \quad (2)$$

$$\frac{\partial w\rho}{\partial t} + \frac{\partial uw\rho}{\partial x} + \frac{\partial vw\rho}{\partial y} + \frac{\partial ww\rho}{\partial z} = -\frac{\partial p'}{\partial z}$$
$$+ \tilde{f}\rho(u - u_e) - (\rho + q_c)g + \frac{\partial \kappa\rho\tau^{31}}{\partial x} + \frac{\partial \kappa\rho\tau^{32}}{\partial y} + \frac{\partial \kappa\rho\tau^{33}}{\partial z}, \quad (3)$$

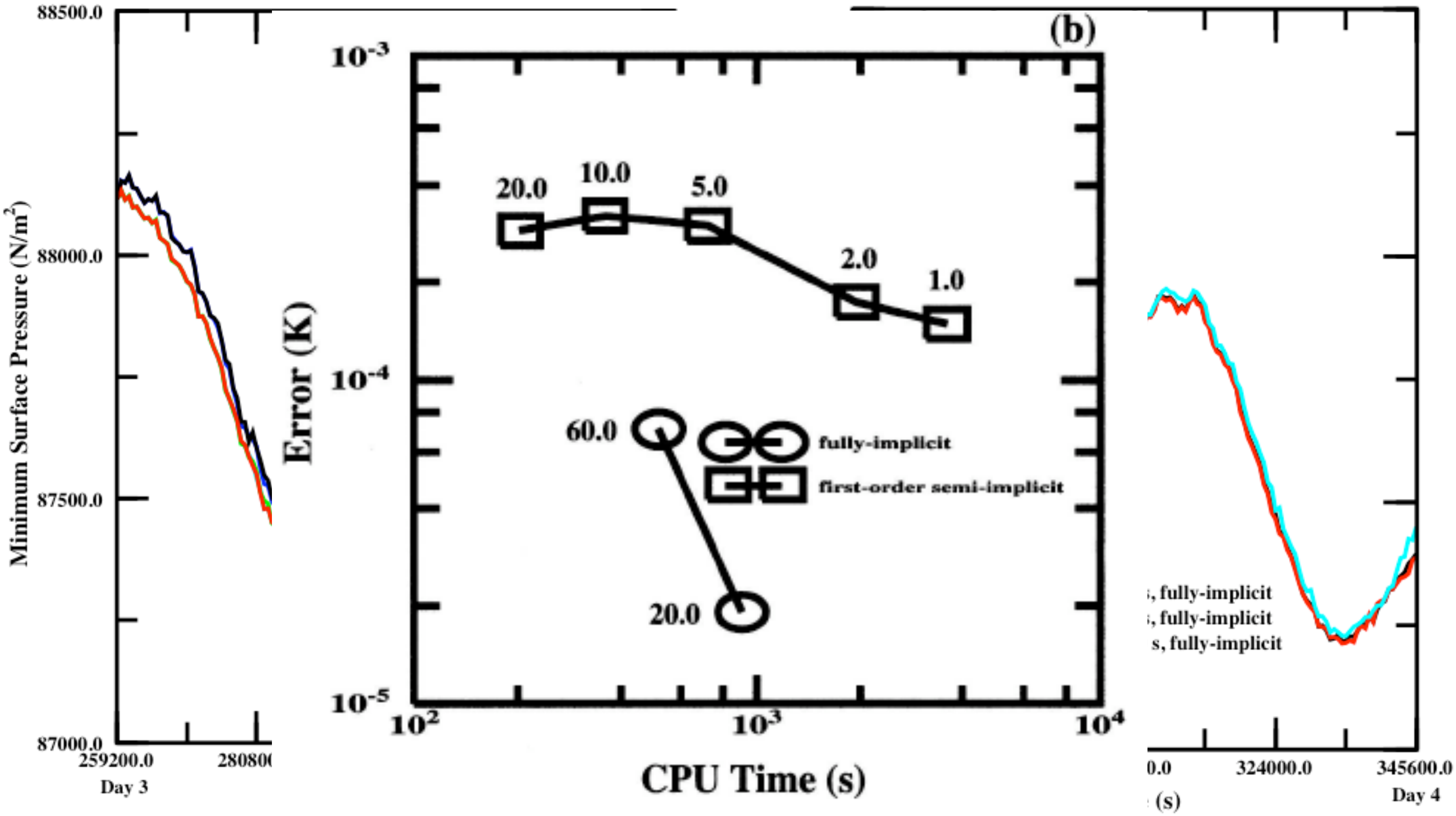
$$\frac{\partial \theta\rho}{\partial t} + \nabla\cdot(\mathbf{V}\theta\rho) = \frac{\theta\rho L}{TC_p} f_{cloud} + f_{surface-energy} + \nabla\cdot(\mathbf{F}_\theta) \quad (4)$$

$$\frac{\partial q_v\rho}{\partial t} + \nabla\cdot(\mathbf{V}q_v\rho) = -f_{cloud} + f_{surface-gas} + \nabla\cdot(\mathbf{F}_{\mathbf{q_v}}) \quad (5)$$

$$\frac{\partial q_c\rho}{\partial t} + \nabla\cdot(\mathbf{V}q_c\rho) = f_{cloud} - f_{fall} + \nabla\cdot(\mathbf{F}_{\mathbf{q_c}}) \quad (6)$$

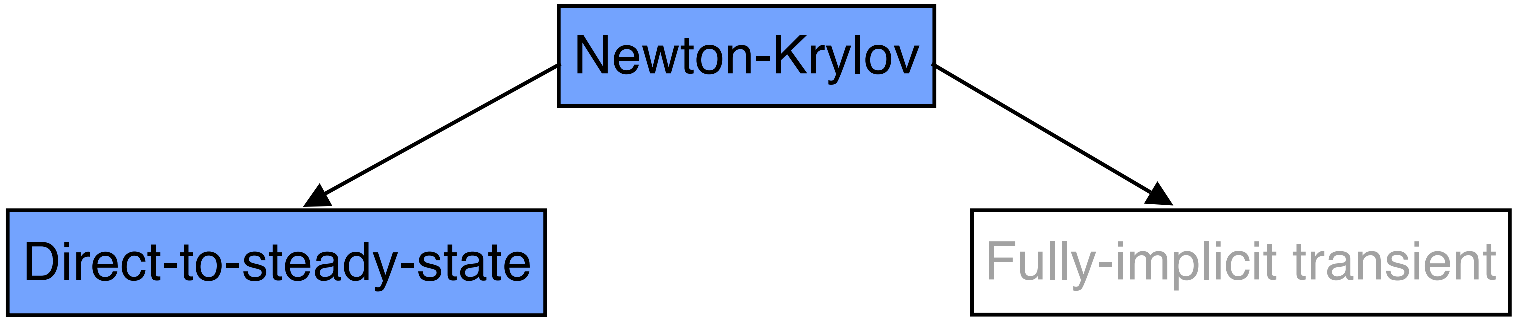
$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\mathbf{V}\rho) = -f_{cloud} + f_{surface-gas} \quad (7)$$

Multiple-time-scale Systems: Newton-Krylov Methods for Hurricane Simulations (Riesner, Mousseau, Wyszogrodzki, Knoll, MWR 2004)



SI - needs to run at stiff wave CFL; JFNK - dynamical time scale

Why Newton-Krylov Methods?



Convergence properties

- Strongly coupled multi-physics often requires a strongly coupled nonlinear solver
- Quadratic convergence near solutions (backtracking, adaptive convergence criteria)
- Often only require a few iterations to converge, if close to solution, independent of problem size

$$\mathbf{F}(\mathbf{x}, \lambda_1, \lambda_2, \lambda_3, ..) = \mathbf{0}$$

Inexact Newton-Krylov

Solve $\mathbf{J}\mathbf{p}_k = -\mathbf{F}(\mathbf{x}_k)$; until $\dfrac{\|\mathbf{J}\mathbf{p}_k + \mathbf{F}(\mathbf{x}_k)\|}{\|\mathbf{F}(\mathbf{x}_k)\|} \le \eta_k$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Theta \mathbf{p}_k$$

Jacobian Free N-K Variant

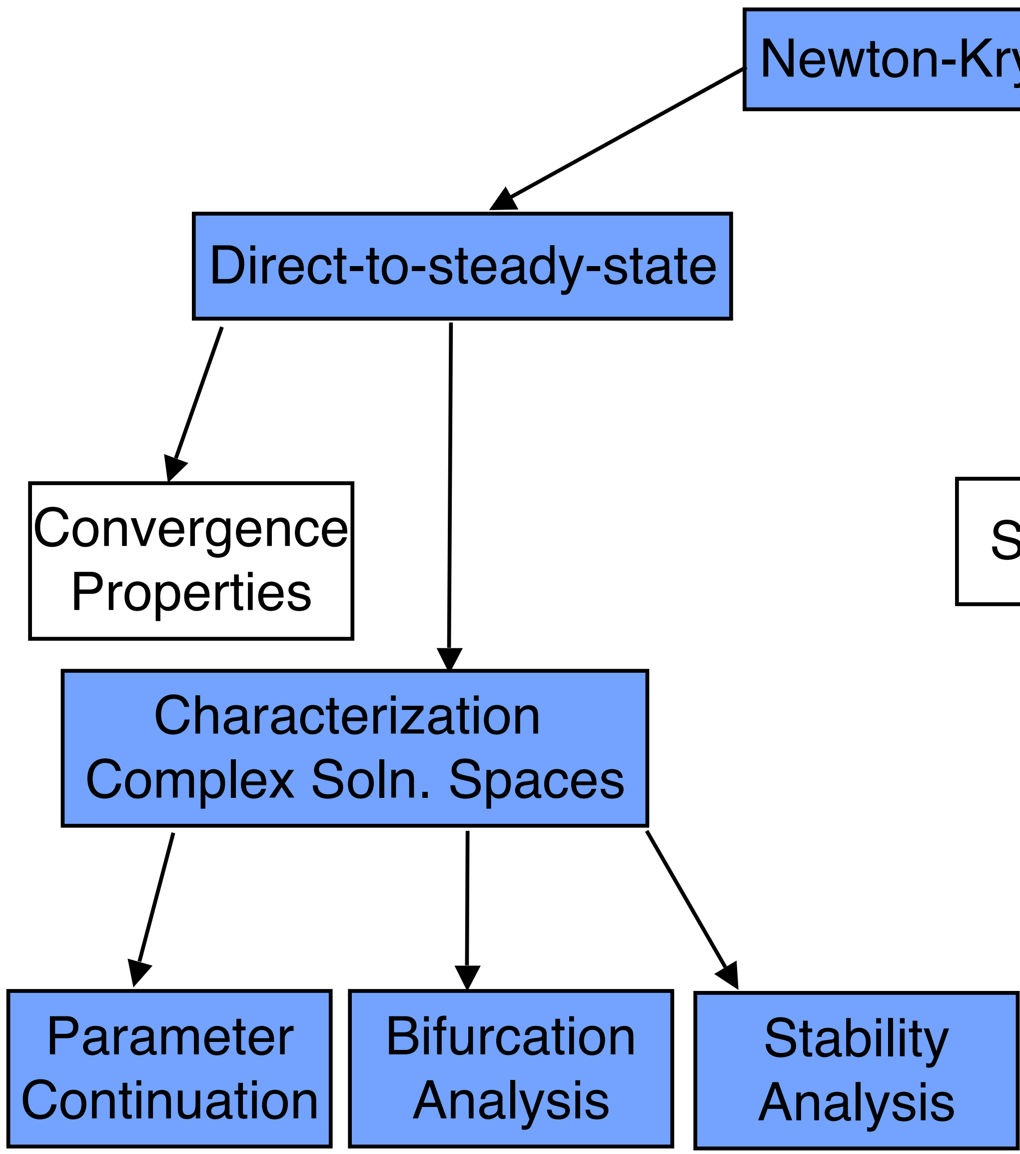
$$\mathbf{M}\mathbf{p}_k = \mathbf{v}$$

$$\mathbf{J}\mathbf{p}_k = \frac{\mathbf{F}(\mathbf{x} + \delta \mathbf{p}_k) - \mathbf{F}(\mathbf{x})}{\delta} \text{ ; or by AD}$$

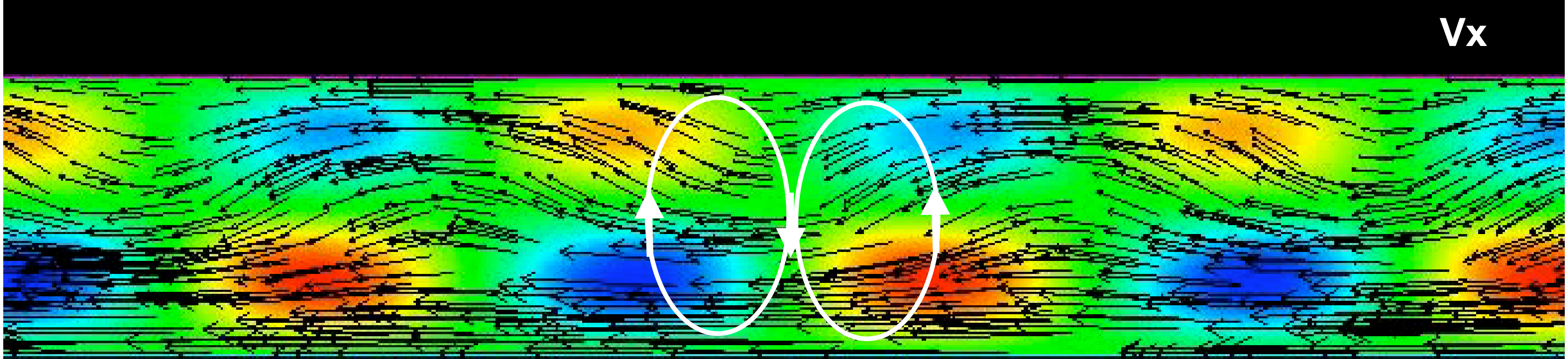
See e.g. Knoll & Keyes, JCP 2004

Sandia National Laboratories

Why Newton-Krylov Methods?



Sandia National Laboratories

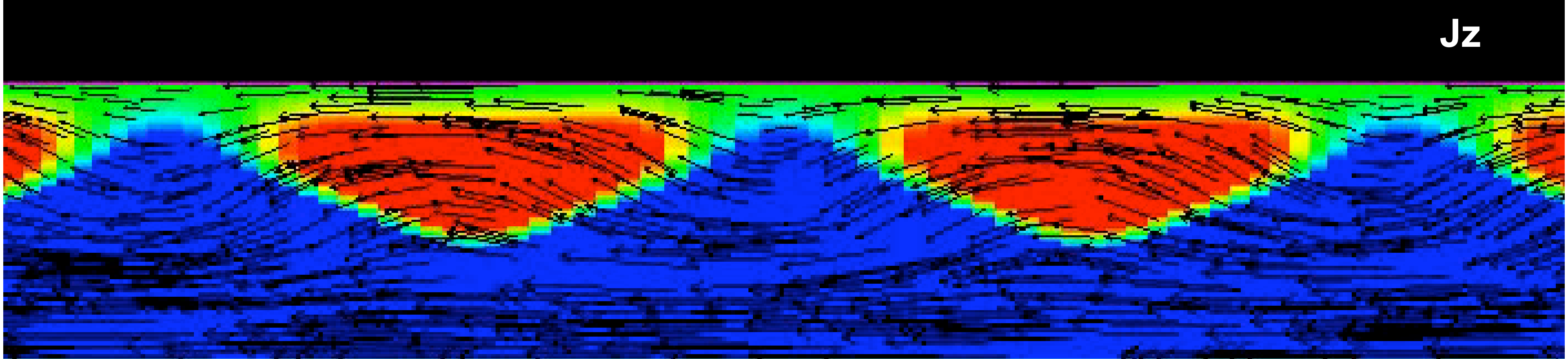


Hydro-Magnetic Rayleigh-Bernard Stability

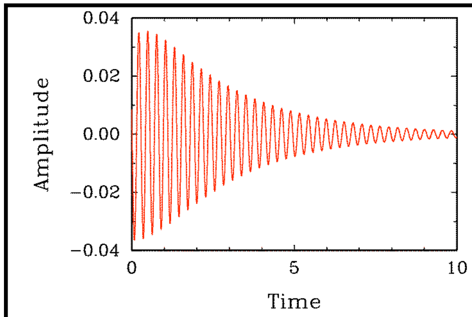
$$Ra = \frac{g\beta}{\nu\alpha}\Delta T d^3 \qquad and \qquad Q = \frac{B_0^2 d^2}{\mu_0 \rho \nu \eta}$$

Stable Fields/Flow at Ra = 4000, Q = 81

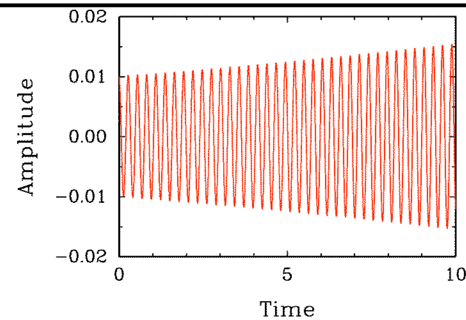
Unstable Flow at Ra = 4000, Q = 144



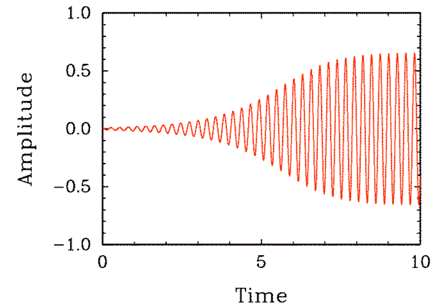
Characterizing Complex Nonlinear Solution Spaces with a Transient Code is Difficult



p=5



p=10 Hopf Bifurcation?

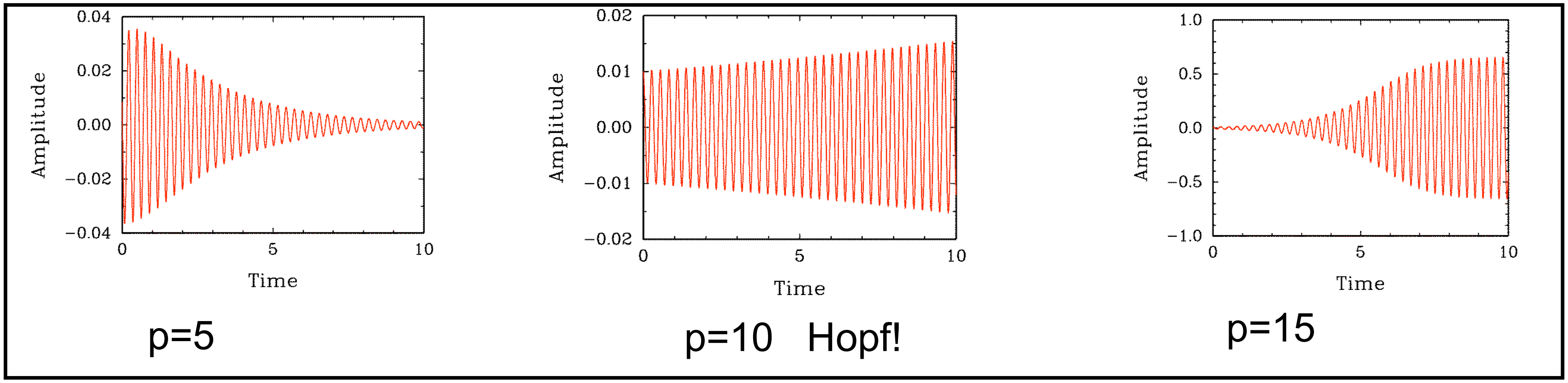


p=15

Various discrete time integration methods:

- **can produce “spurious” stable and unstable steady solutions and limit cycles**
- **can stabilize unstable solutions of the ODE/PDE**
- **can produce very different dynamics and bifurcation behavior than ODE/PDE**

Characterizing Complex Nonlinear Solution Spaces with a Transient Code is Difficult



Various discrete time integration methods: (can also be said of discrete spatial approx)

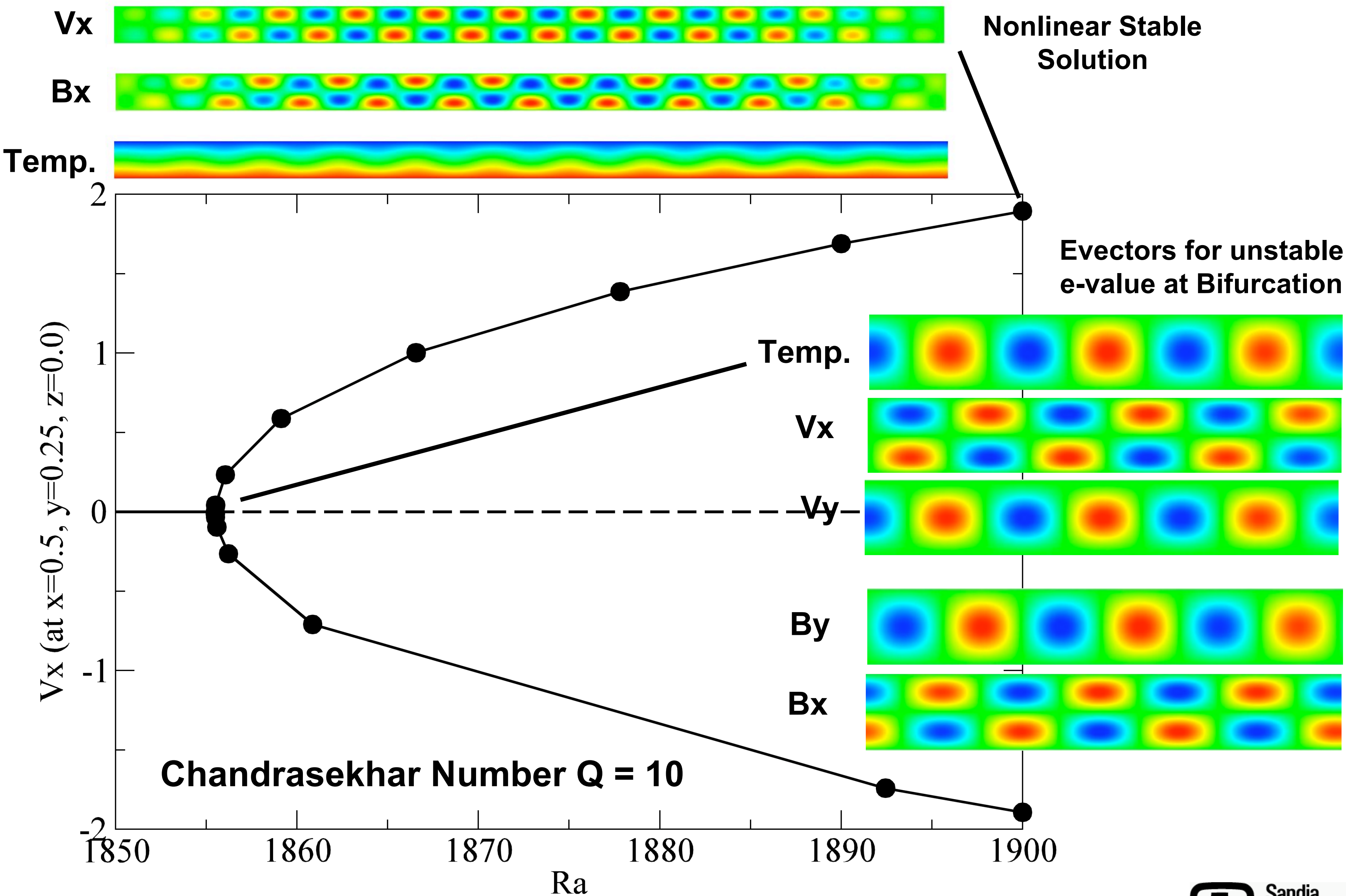
- **can produce "spurious" stable and unstable steady solutions and limit cycles**
- **can stabilize unstable solutions of the ODE/PDE**
- **can produce very different dynamics and bifurcation behavior than ODE/PDE**

In addition:

- **turn a BVP -> IBVP with unknown initial data (basin of attraction of solutions)**
- **require very long time integration near critical points**
- **require a detailed sampling of parameter space to characterize a solution space**
- **produce complex interactions between temporal and spatial discretizations**
- **cannot be used to efficiently "track" location of critical points with multiple parameters**

Sandia National Laboratories

Hydro-Magnetic Rayleigh-Bernard Stability: Direct Determination of Linear Stability and Nonlinear Equilibrium Solutions (Steady State Solves)



Sandia National Laboratories

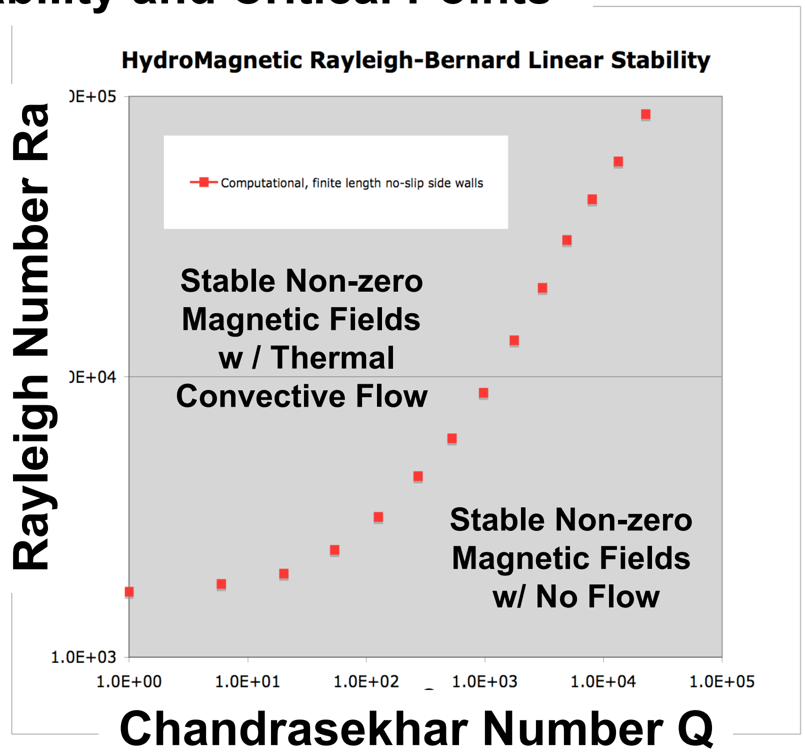
Hydro-Magnetic Rayleigh-Bernard: Determining Critical Stability and Critical Points

Linear Stability of Computational Solution by Normal Mode Analysis

$$\sigma_i \mathbf{B} \mathbf{q}_i = \mathbf{F}^{'} \mathbf{q}_i$$

$$(\mathbf{F}^{'} - \eta_c \mathbf{B})^{-1} (\mathbf{F}^{'} - \mu_c \mathbf{B}) \mathbf{w} = \nu \mathbf{w}$$

Approximately invert by ML preconditioned Krylov solve



Turning Point Tracking:

$$\mathbf{F}\left(\mathbf{x}, Ra^{*}, Q^{*}\right) = \mathbf{0}$$

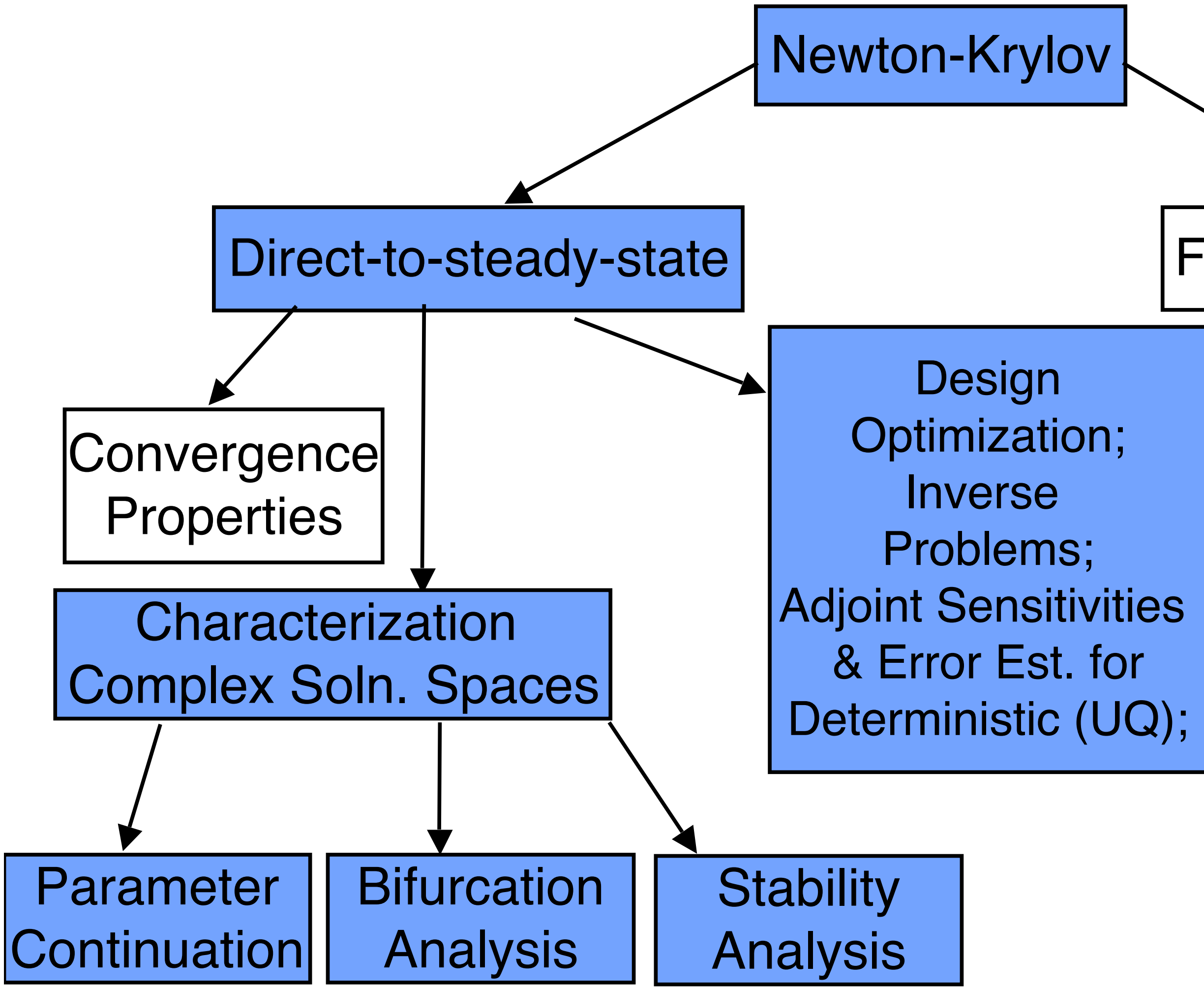
$$\mathbf{F}^{'} \mathbf{v} = \mathbf{0}$$

$$\mathbf{\Gamma}^{\mathbf{T}} \mathbf{v} - 1 = 0$$

Solve extended system with Newton's method

Sandia National Laboratories

Why Newton-Krylov Methods?



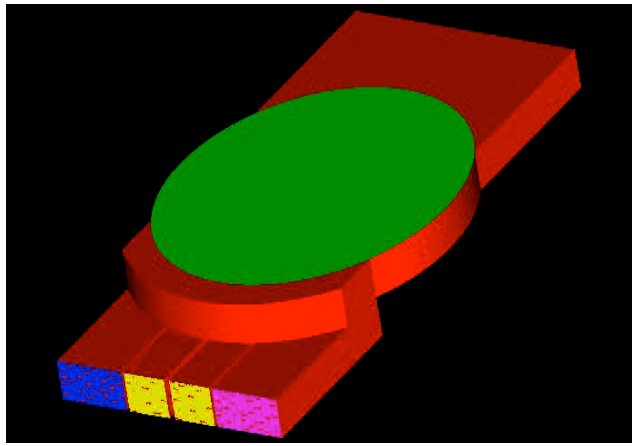
Optimization:
George Biros talk

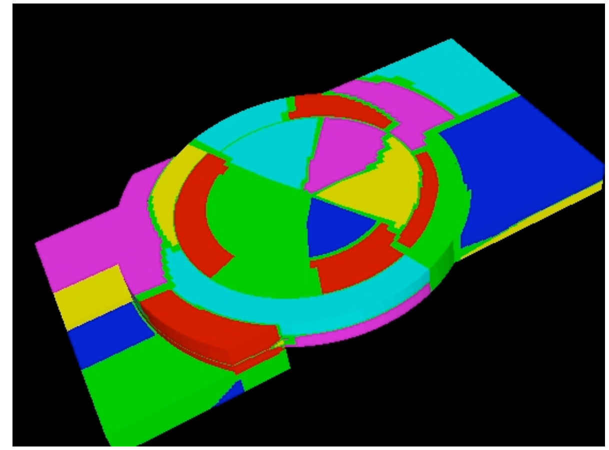
Error Estimation:
Don Estep Talk

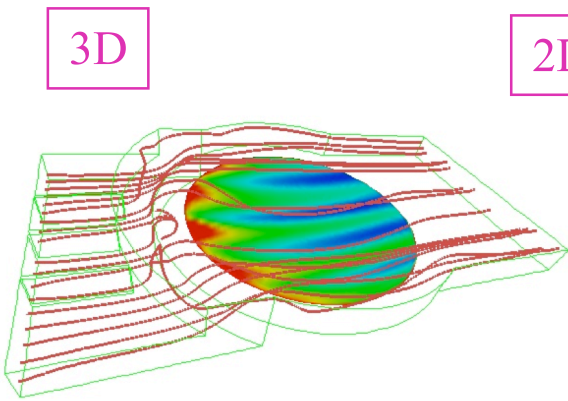
PDE Constrained Optimization of Poly-Silicon CVD Reactor
Unstructured FE Reacting Flow MPSalsa code

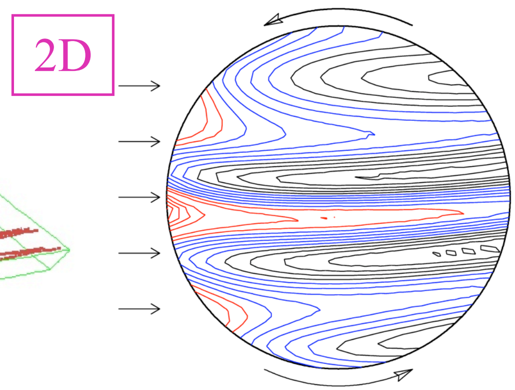
Poly-Silicon Epitaxy from Trichlorosilane in Hydrogen Carrier;

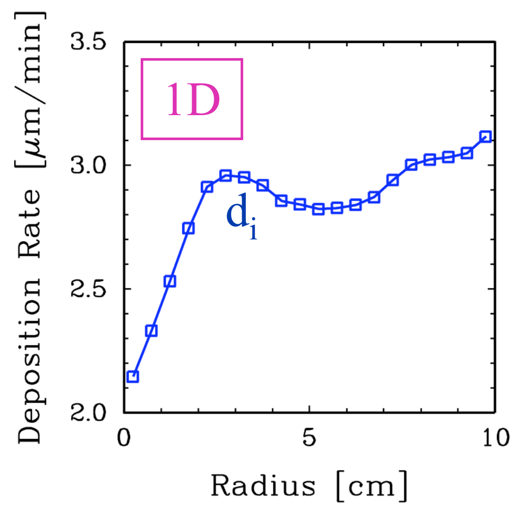
3D (u,v,w,P,T)
3 chemical species
1.2M unknowns









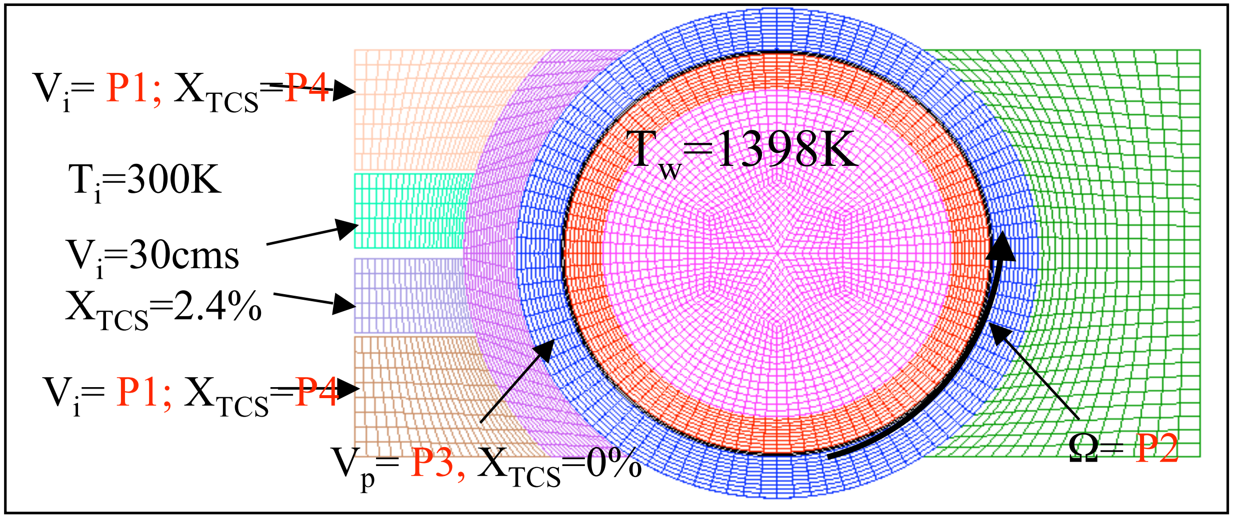


0D Objective Function:

$$f = \frac{1}{2}\sum_{\text{radii}} \left(d_i / d_{ave} - 1 \right)^2$$

Sandia National Laboratories

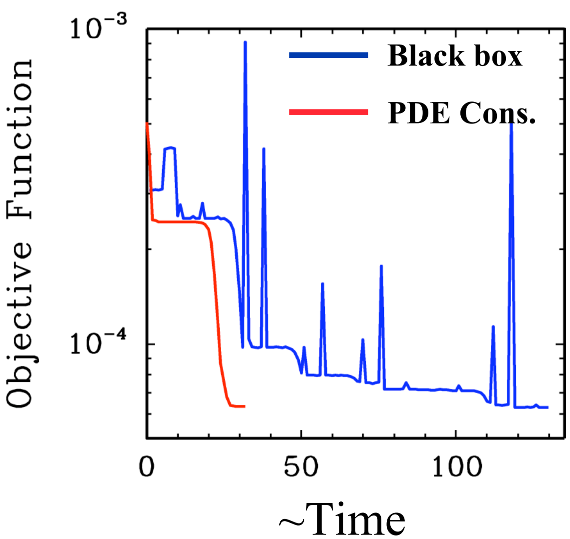
PDE Constrained Optimization of Poly-Silicon CVD Reactor

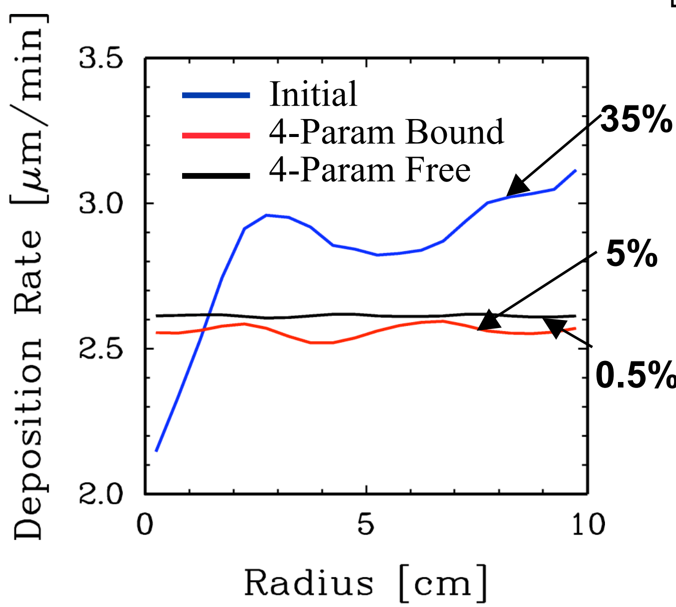


PDE Constrained Optimization:

Minimize: $f(\mathbf{x},\mathbf{p})$
such that: $\mathbf{F}(\mathbf{x},\mathbf{p})=\mathbf{0}$

Use Newton's Method solve KKT system



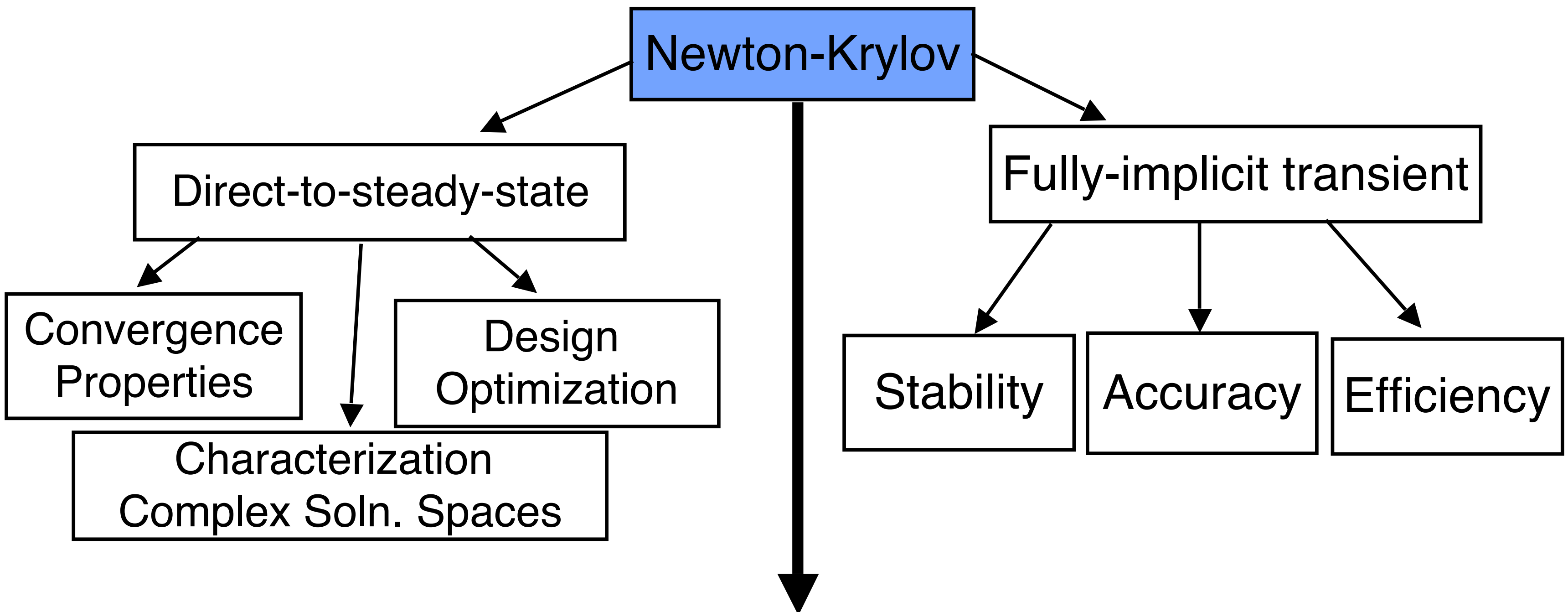


Unks	Procs	Time (hrs.)
1.2	48	6.2 (3GHz Cluster)
4.8M	128	~ 6 (Red Storm: XT3)
38M	1024	~ 7 (Red Storm: XT3)

W/Pawlowski, Salinger, van Bloemen Waanders, Bartlett, Lin - SNL

Sandia National Laboratories

Why Newton-Krylov Methods?



Very Large Problems -> Parallel Iterative Solution of Sub-problems

Krylov Methods - Robust, Scalable and Efficient Parallel Preconditioners

- Approximate Block Factorizations
- Physics-based Preconditioners
- Multi-level solvers for systems and scalar equations

Ray Tuminaro, Dan Reynolds: earlier talks

Sandia National Laboratories

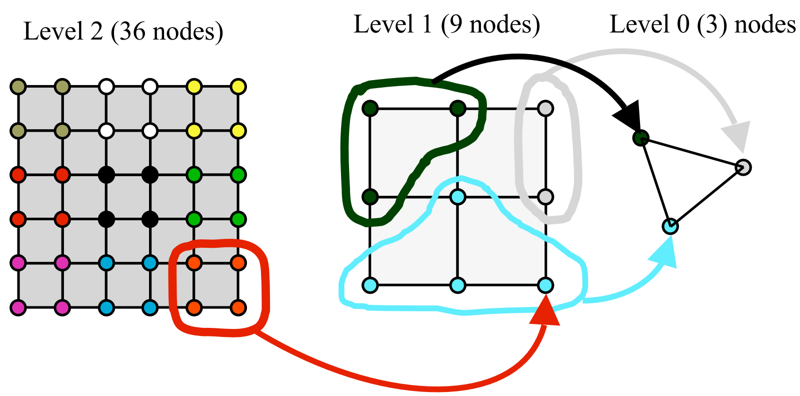
ML library: Multilevel Preconditioners

(R. Tuminaro, M. Sala, J. Hu, M. Gee (UT Munich)]

2-level and N-level Aggressive Coarsening Graph-based Block AMG

- Aggregation is used to produce a coarse operator
  - **Create graph where vertices are block nonzeros in matrix A_k**
  - **Edge between vertices i and j included if block $B_k(i,j)$ contains nonzeros**
  - **Decompose graph into aggregates (subgraphs) [Metis/ParMetis]**
- Construction of simple restriction/interpolation operators (e.g. piecewise constants on agg.)
- Construction of A_{k-1} as $A_{k-1} = R_{k-1} A_k I_{k-1}$

- Nonsmoothed aggregation
- Domain decomposition smoothers (sub-domain GS and ILU)
- Coarse grid solver can use fewer processors than for fine mesh solve (direct/approximate/iterative)

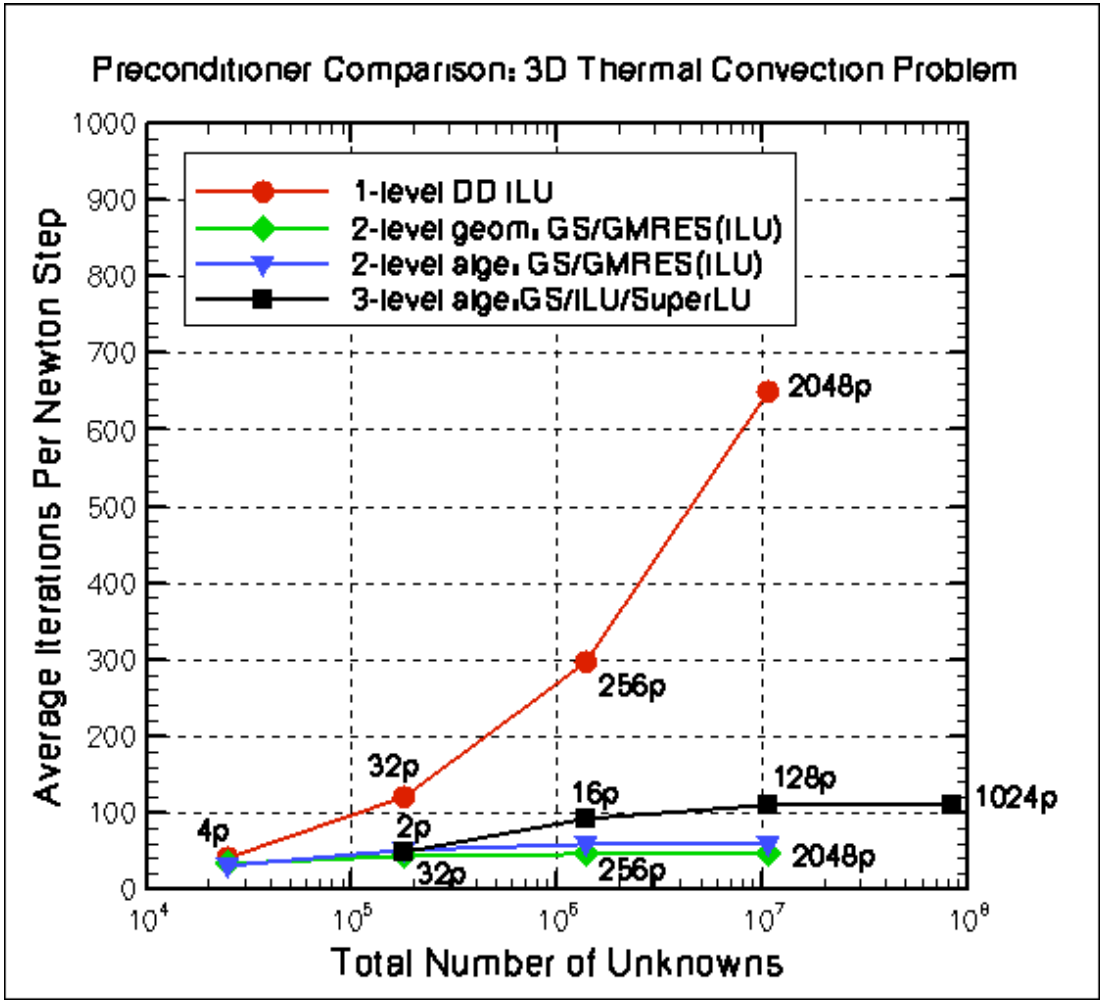


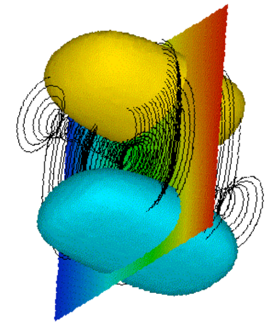
Visualization of effect of partition of matrix graph on mesh

Aggregation based Multigrid:
- **Vanek, Mandel, Brezina, 1996**
- **Vanek, Brezina, Mandel, 2001**

Aggregation used in DD:
- **Paglieri, Scheinine, Formaggia, Quateroni, 1997**
- **Jenkins, Kelley, Miller, Kees, 2000**
- **Toselli, Lasser, 2000**
- **Sala, Formaggia, 2001**

Multilevel Preconditioner Scaling Study: 3D Thermal Buoyancy Driven Convection





Sandia National Laboratories

Comparison of 1-level with 2-level geometric & algebraic 2D & 3D Thermal Convection Problem

proc	fine grid unknowns	1 - level Method Ilu DD		coarse unknowns		2-level: ilu-superlu			
						geometric		algebraic	
		avg its per Newt step	time (sec)	geometric	algebraic	avg its per Newt step	time (sec)	avg its per Newt step	time (sec)
1	4356	41	23	100	96	29	18	28	20
4	16,900	98	62	324	320	37	25	40	27
16	66,564	251	275	1156	1088	40	34	50	39
64	264,196	603	1,399	4356	4096	38	57	57	69
256	1,052,676	1,478	8,085	16900	16384	37	151	63	191

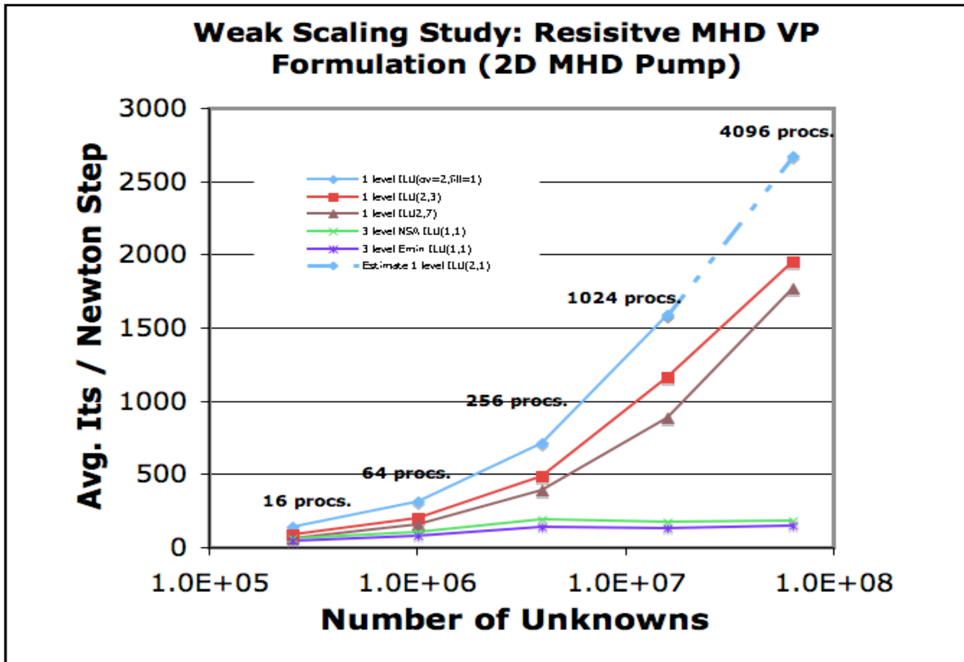
proc	fine grid unknowns	1 - level Method Ilu DD		coarse unknowns		2-level: gs2-superlu			
						geometric		algebraic	
		avg its per Newt step	time (sec)	geometric	algebraic	avg its per Newt step	time (sec)	avg its per Newt step	time (sec)
4	24.565	40[5]	123	135	120	36[5]	101	30[4]	71
32	179,685	112[5]	282	625	480	44[4]	107	50[4]	109
256	1.373.125	296[5]	863	3,645	2560	47[5]	179	58[4]	152
2048	10,733,445	650[5]	2,915	24,565		47[4]	546	59[4]	681

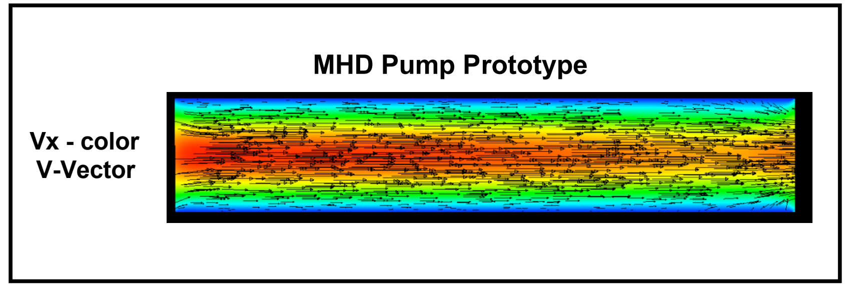
Analysis: Sala; Math. Modeling and Numer. Anal., 2004
Sala, Shadid, Tuminaro; accepted in SIMAX
Numerical Exp:
Lin, Sala, Shadid, Tuminaro; accepted in IJNME

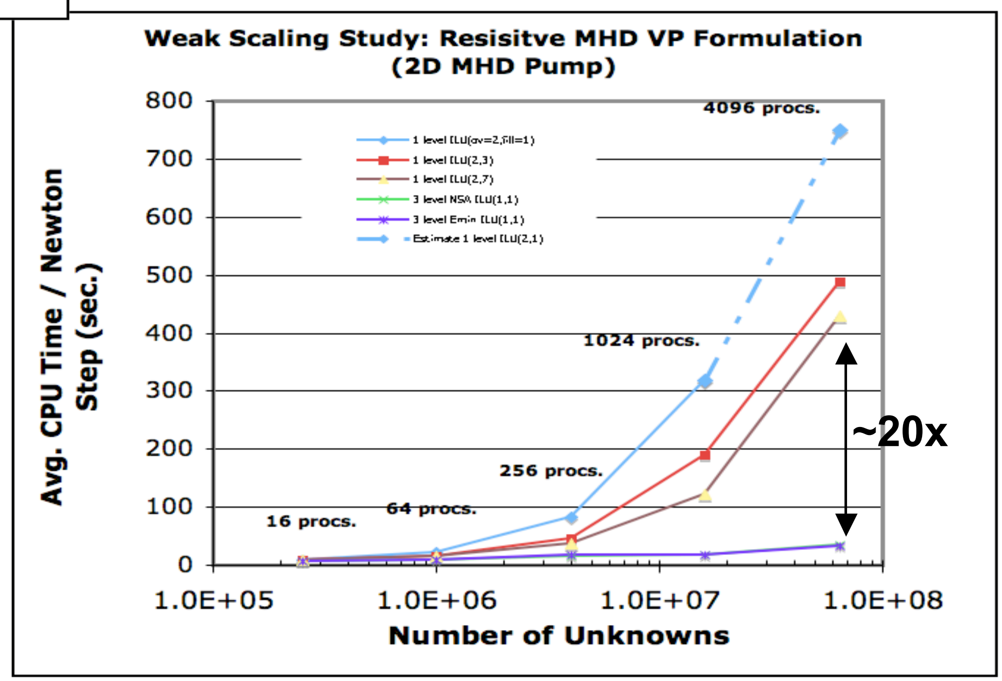
- Coarse mesh: SuperLU direct solver
- Run on Sandia ASCI Red machine

Sandia National Laboratories

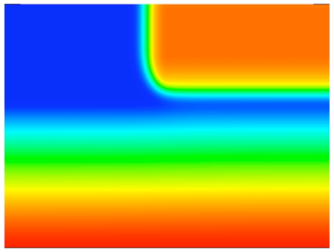
Red Storm Block AMG Scaling Results - Cray XT3: (2D MHD Pump)







Scaling Study: Steady-State NPN BJT 1- and 3-level Preconditioners



- Steady-state 2D drift diffusion bias 0.3V; initial guess NLP solution
- Smoothers/solvers: ILU, ILU, KLU
- 85 nodes per aggregate; nonsmoothed aggregation
- Run on Sandia Red Storm machine (Cray XT3)

Electric potential

$$\lambda^2 \nabla \cdot (\epsilon_r \mathbf{E}) = p - n + C \qquad \mathbf{E} = -\nabla \psi$$

Current conservation

$$\nabla \cdot \mathbf{J}_n = \frac{\partial n}{\partial t} + R$$

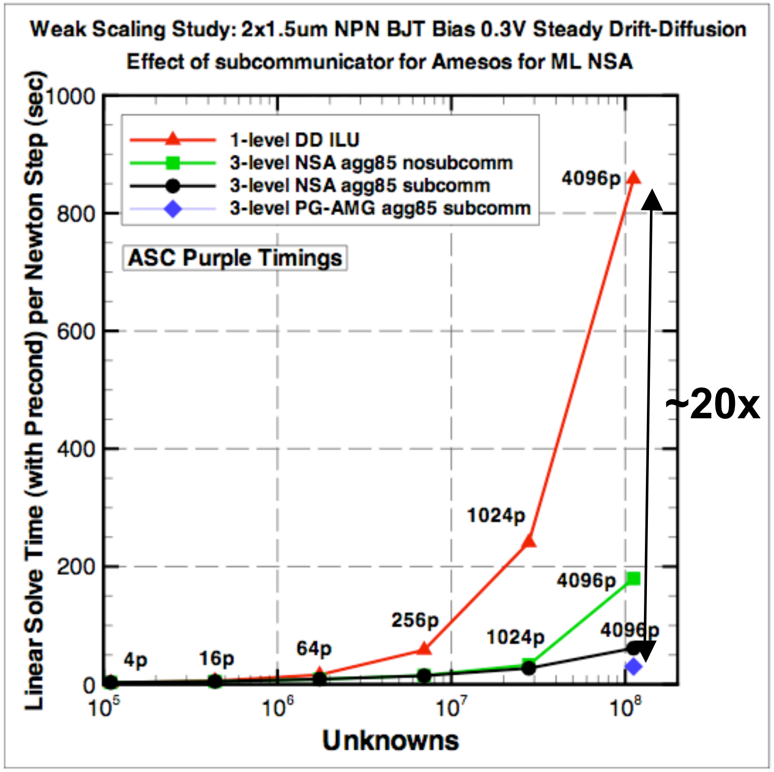
$$-\nabla \cdot \mathbf{J}_p = \frac{\partial p}{\partial t} + R$$

"constitutive" relation

$$\mathbf{J}_n = \mu_n n \mathbf{E} + D_n \nabla n$$

$$\mathbf{J}_p = \mu_p p \mathbf{E} - D_p \nabla p$$

$||\Delta C|| = 10^{19}$



Largest Calculation

- 8192 processing nodes (single core per node)
- ~333M FE nodes
- 1.01B unknowns
- Solve time per Newton ~570 seconds (ML 4 lPG evel)

Stabilized FE method (Charon - Hennigan, Hoekstra, Lin, S)

Sandia National Laboratories

Trilinos: Full Vertical Solver Coverage
(Part of DOE: TOPS SciDAC Effort)

Optimization Unconstrained: Constrained:	Find $u \in \Re^n$ that minimizes $g(u)$ Find $x \in \Re^m$ and $u \in \Re^n$ that minimizes $g(x,u)$ s.t. $f(x,u) = 0$	**MOOCHO**
Bifurcation Analysis	Given nonlinear operator $F(x,u) \in \Re^{n+m} \rightarrow \Re^n$ For $F(x,u) = 0$ find space $u \in U \ni \frac{\partial F}{\partial x}$ singular	**LOCA**
Transient Problems DAEs/ODEs:	Solve $f(\dot{x}(t), x(t), t) = 0$ $t \in [0,T], x(0) = x_0, \dot{x}(0) = x_0'$ for $x(t) \in \Re^n, t \in [0,T]$	**Rhythmos**
Nonlinear Problems	Given nonlinear operator $F(x,u) \in \Re^{n+m} \rightarrow \Re^n$ Solve $F(x) = 0 \quad x \in \Re^n$	**NOX**
Linear Problems Linear Equations: Eigen Problems:	Given Linear Ops (Matrices) $A, B \in \Re^{m \times n}$ Solve $Ax = b$ for $x \in \Re^n$ Solve $A\nu = \lambda B \nu$ for (all) $\nu \in \Re^n, \quad \lambda \in \Re$	**AztecOO** **Belos** **Ifpack, ML, etc...** **Anasazi**
Distributed Linear Algebra Matrix/Graph Equations: Vector Problems:	Compute $y = Ax; A = A(G); A \in \Re^{m \times n}, G \in \Im^{m \times n}$ Compute $y = \alpha x + \beta w; \alpha = \langle x, y \rangle; x, y \in \Re^n$	**Epetra** **Tpetra**

Conclusions

• Newton-Krylov methods can provide a very effective, robust and flexible solution technology for analysis and characterization of complex nonlinear solution spaces. For steady state, time dependent and optimization type solutions. (e.g. Transport/reaction, resistive MHD)

• High parallel efficiencies for fully-implicit fully coupled Newton-Krylov iterative solvers for a wide range of problems are possible.

• Parallel multilevel aggressive coarsening block AMG preconditioners for systems have shown promising results for algorithmic scalability and CPU time performance of transport solutions.

(Issues: Strong convection, reaction and FE aspect ratios for multilevel methods. -> Physics-based for efficient transient solution)

• Cray XT3 very capable parallel computing platform. Very good scaling results.

Sandia National Laboratories