Continuation Algorithms for Parameter-Dependent Compact Fixed Point Problems

C. T. Kelley
Joint work with
Yannis Kevrekidis, Matthew Lasater, Liang Qiao,
Andy Salinger, Dwight Woolard, Peiji Zhao

Department of Mathematics
Center for Research in Scientific Computation
North Carolina State University
Raleigh, North Carolina, USA

SANDIA

June 22, 2004

Supported by NSF, ARO.

Outline

- Fast introduction to compact fixed point problems
- Newton-GMRES and multilevel Newton-GMRES
- Path following: introduction
  Nonlinear solvers
  Pseudo-arclength continuation
- Three examples:
  - integral equation
    explicit integral operator
  - Wigner-Poisson Equation for RTDs
  - time-stepper for parabolic pde
    implicit integral operator
- Multilevel method.

Compact Fixed Point Problems

We're worried about problems like

$$F(u) = u - \mathscr{K}(u) = 0,$$

where

- F is Lipschitz continuously Frechét differentiable on a Banach space X.
- The "compact" part means that $\mathscr{K}'$ is a compact linear map on X.
- We want to exploit the compactness to design fast solvers.

How to exploit compactness

- Discretization

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers
  - Krylov solvers need no preconditioning (in theory).

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers
  - Krylov solvers need no preconditioning (in theory).
  - Multilevel methods are easy to design.

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers
  - Krylov solvers need no preconditioning (in theory).
  - Multilevel methods are easy to design.
  - No smoothers are needed.

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers
  - Krylov solvers need no preconditioning (in theory).
  - Multilevel methods are easy to design.
  - No smoothers are needed.
- Fast evaluation ($O(N \log(N))$) is common.

How to exploit compactness

- Discretization
  - Almost every reasonable scheme works, but
  - some approximations to $\mathscr{K}'$ converge in norm.
- Solvers
  - Krylov solvers need no preconditioning (in theory).
  - Multilevel methods are easy to design.
  - No smoothers are needed.
- Fast evaluation ($O(N \log(N))$) is common.
- Newton-Krylov, Newton-MG nonlinear solvers work with no surprises (most of the time).

World's Easiest Example

$$(I-K)u(x) = u(x) - \int_0^1 k(x,y)u(y)\,dy = f(x),$$

$f \in C[0,1]$, $k \in C([0,1]\times[0,1])$
Discretization: $V_h =$ piecewise linears/piecewise constants

$$u^h(x) - K_h u^h(x) = u^h(x) - \int_0^1 k_h(x,y)u^h(y)\,dy = P_h f(x)$$

where,

$$k_h(x,y) = \sum_{i,j=1}^{N_h} k(x_i,x_j)\phi_i(x)\phi_j(y)$$

P_h is a projection onto V_h, and we seek $u^h \in V_h$.

Properties of Discretization

- K_h operates on the function space

Properties of Discretization

- K_h operates on the function space
- $K_h \longrightarrow K$ in the operator norm

Properties of Discretization

- K_h operates on the function space
- $K_h \to K$ in the operator norm
- Lots of flexibility in P_h
  Strong convergence to I is all you need.

Properties of Discretization

- K_h operates on the function space
- $K_h \to K$ in the operator norm
- Lots of flexibility in P_h
  Strong convergence to I is all you need.
- If $I - K$ is nonsingular, then

$$u^h = (I - K_h)^{-1} P_h f \to (I - K)^{-1} f$$

Solve finite dimensional system for nodal values.

Properties of Discretization

- K_h operates on the function space
- $K_h \rightarrow K$ in the operator norm
- Lots of flexibility in P_h
  Strong convergence to I is all you need.
- If $I - K$ is nonsingular, then

$$u^h = (I - K_h)^{-1} P_h f \rightarrow (I - K)^{-1} f$$

  Solve finite dimensional system for nodal values.
- Other choices of K_h are possible
  Standard quadrature rule + fine-to-coarse by averaging

Performance of GMRES

Avoid the $O(N_h^3)$ cost of a direct solver, and compute

$$u^h = (I - K_h)^{-1} P_h f = \sum_{i=1}^{N_h} u_i^h \phi_i \in V_h.$$

with GMRES.

- Continuous problem: superlinear convergence
- Discrete problem: mesh independent performance
- Cost: One $K_h v$ evaluation/linear iteration
  Think $N_h \log N_h$ work if done slickly.

Nested iteration (aka grid sequencing) is a good idea.

Multilevel Method

Since $K_h \to K$ in the operator norm,

- $(I - K_H)$ ($h << H$) might be a good preconditioner for GMRES

Multilevel Method

Since $K_h \to K$ in the operator norm,

- $(I - K_H)$ ($h << H$) might be a good preconditioner for GMRES
- Richardson iteration is a better idea thanks to LOW STORAGE.

$$u \leftarrow u - (I - K_H)^{-1}((I - K_h)u - P_h f)$$

Multilevel Method

Since $K_h \to K$ in the operator norm,

- $(I - K_H)$ ($h << H$) might be a good preconditioner for GMRES
- Richardson iteration is a better idea thanks to LOW STORAGE.

$$u \leftarrow u - (I - K_H)^{-1}((I - K_h)u - P_h f)$$

- H suff small implies

Multilevel Method

Since $K_h \to K$ in the operator norm,

- $(I - K_H)$ ($h << H$) might be a good preconditioner for GMRES
- Richardson iteration is a better idea thanks to LOW STORAGE.

$$u \leftarrow u - (I - K_H)^{-1}((I - K_h)u - P_h f)$$

- H suff small implies
  - Krylovs independent of H.

Multilevel Method

Since $K_h \to K$ in the operator norm,

- $(I - K_H)$ ($h << H$) might be a good preconditioner for GMRES
- Richardson iteration is a better idea thanks to LOW STORAGE.

$$u \leftarrow u - (I - K_H)^{-1}((I - K_h)u - P_h f)$$

- H suff small implies
  - Krylovs independent of H.
  - One iteration/level suffices.

Nonlinear Problems

Generalization to the nonlinear case is easy,

$$u \leftarrow u - (I - \mathscr{K}_H'(u^H))^{-1} F_h(u)$$

if you're careful about the fine-to-coarse transfer.
If coarse mesh suff fine,

- Krylovs/Newton independent of H
- one Newton/level suffices.

Nested Iteration: Bottom up

$h = H$, $i = 0$
Solve $F_H(u^H) = 0$ to high accuracy.
$u \leftarrow u^H$
for $i = 1, \ldots m$ **do**
  $h \leftarrow h/2$
  $u \leftarrow u - (I - \mathscr{K}_H'(u^H))^{-1} F_h(u)$
end for

Nested Iteration: Bottom up

$h = H$, $i = 0$
Solve $F_H(u^H) = 0$ to high accuracy.
$u \leftarrow u^H$
for $i = 1, \dots m$ **do**
  $h \leftarrow h/2$
  $u \leftarrow u - (I - \mathscr{K}_H'(u^H))^{-1} F_h(u)$
end for

- All the linear solver work is on the coarse mesh.

Nested Iteration: Bottom up

$h = H$, $i = 0$
Solve $F_H(u^H) = 0$ to high accuracy.
$u \leftarrow u^H$
for $i = 1, \ldots m$ **do**
  $h \leftarrow h/2$
  $u \leftarrow u - (I - \mathscr{K}_H'(u^H))^{-1} F_h(u)$
end for

- All the linear solver work is on the coarse mesh.
- Only two grids H and h active at any time.

Nested Iteration: Bottom up

$h = H$, $i = 0$
Solve $F_H(u^H) = 0$ to high accuracy.
$u \leftarrow u^H$
for $i = 1, \ldots m$ **do**
  $h \leftarrow h/2$
  $u \leftarrow u - (I - \mathscr{K}_H'(u^H))^{-1} F_h(u)$
end for

- All the linear solver work is on the coarse mesh.
- Only two grids H and h active at any time.
- Cost of solve to truncation error:
  < 3 fine mesh evals, depending on cost of $\mathscr{K}_h$

Path Following

$F : X \times [a,b]$, F smooth, X a Banach space.
Objective: Solve $F(u,\lambda) = 0$ for $\lambda \in [a,b]$
Obvious approach:

Set $\lambda = a$, solve $F(u,\lambda) = 0$ with Newton-(MG, GMRES, ...) to obtain $u_0 = u(\lambda)$.
while $\lambda < b$ **do**
  Set $\lambda = \lambda + d\lambda$.
  Solve $F(u,\lambda) = 0$ with u_0 as the initial iterate.
  $u_0 \leftarrow u(\lambda)$
end while

What's the problem?

- Multiple solutions, hysteresis
- No solutions

What's the problem?

- Multiple solutions, hysteresis
- No solutions

A fix: Pseudo-arclength continuation.
Set $x = (u, \lambda)$ and solve $G(x, s) = 0$, where, for example

$$G(x,s) = \begin{pmatrix} F \\ N \end{pmatrix} = \begin{pmatrix} F(u(s), \lambda(s)) \\ \dot{u}^T (u - u_0) + \dot{\lambda}^T (\lambda - \lambda_0) - (s - s_0) \end{pmatrix}.$$

What's the problem?

- Multiple solutions, hysteresis
- No solutions

A fix: Pseudo-arclength continuation.
Set $x = (u, \lambda)$ and solve $G(x, s) = 0$, where, for example

$$G(x,s) = \begin{pmatrix} F \\ N \end{pmatrix} = \begin{pmatrix} F(u(s), \lambda(s)) \\ \dot{u}^T (u - u_0) + \dot{\lambda}^T (\lambda - \lambda_0) - (s - s_0) \end{pmatrix}.$$

s is an artificial "arclength" parameter.
u_0 and λ_0 are from the previous step.
$\dot{u} \approx du/ds$ and $\dot{\lambda} \approx d\lambda/ds$,
(say by differences using s_0 and s_{-1}).

Watch out for scaling!

Simple Folds

We follow solution paths $\{x(s)\}$.
Assume that F is smooth and

- G_x is nonsingular (not always true) So implicit function theorem holds in s.

We are assuming that there is no true bifurcation and that the singularity in λ is a simple fold.

Arclength Continuation Algorithm

Set $\lambda = a$, $s = 0$ solve $F(u, \lambda) = 0$ with Newton-(MG, GMRES, ...) to obtain u_0.

Estimate ds, $\dot{u}$, $\dot{\lambda}$.

while $s < s_{max}$ **do**

  $s \leftarrow s + ds$.

  Solve $G(x, s) = 0$ with u_0 as the initial iterate.

  $x_0 \leftarrow x$

  Update ds, $\dot{u}$, $\dot{\lambda}$.

end while

Simple example: Chandreskhar H-Equation

$$H(\mu) = \left(1 - \frac{c}{2}\int_0^1 \frac{\mu H(\nu)\,d\nu}{\mu+\nu}\right)^{-1}$$

Simple example: Chandreskhar H-Equation

$$H(\mu) = \left(1 - \frac{c}{2}\int_0^1 \frac{\mu H(\nu)\,d\nu}{\mu + \nu}\right)^{-1}$$

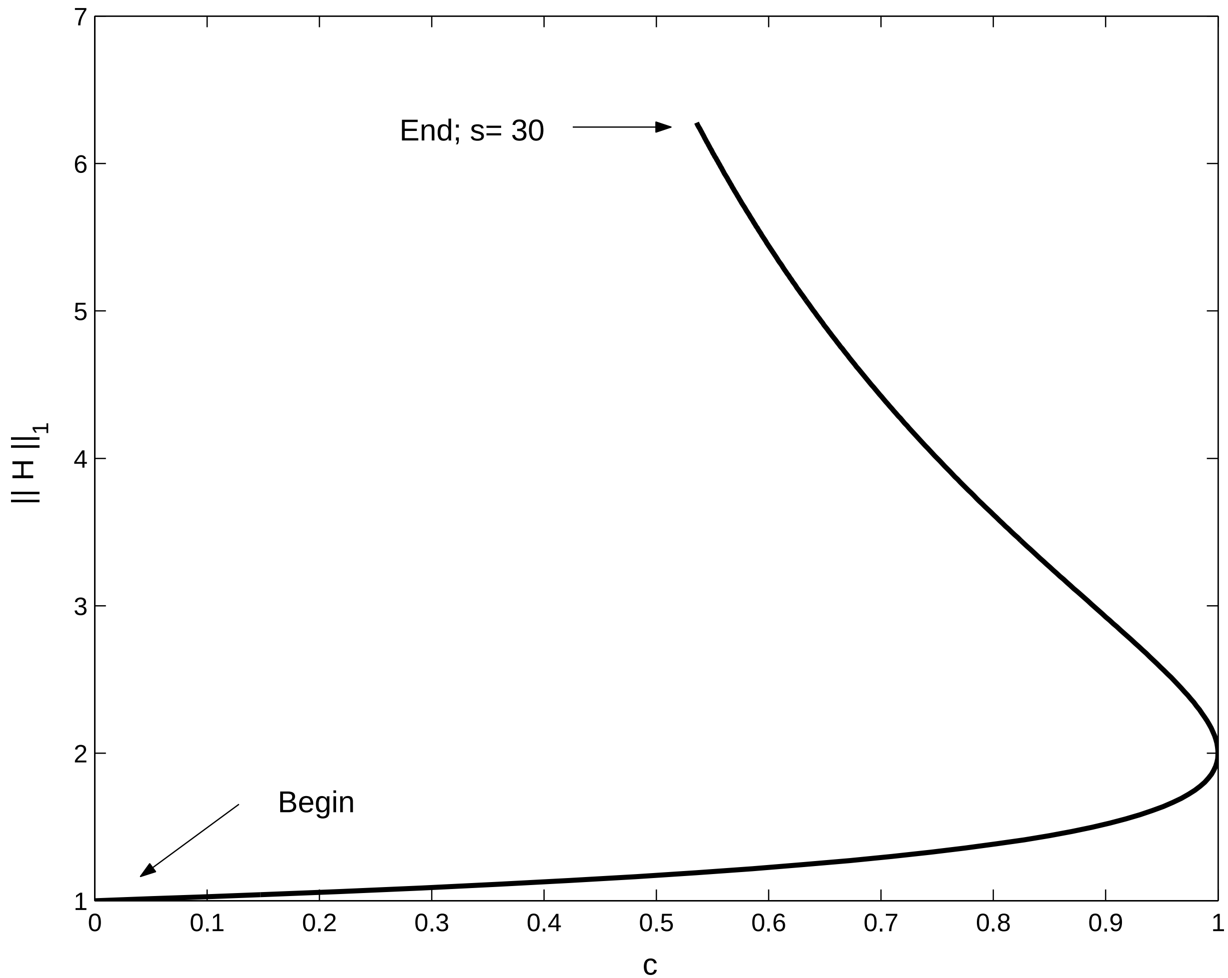
- Compact fixed point problem.
- Problem becomes harder as $H(1) \to \infty$
- Two solutions for $c \neq 0, 1$

Simple example: Chandreskhar H-Equation

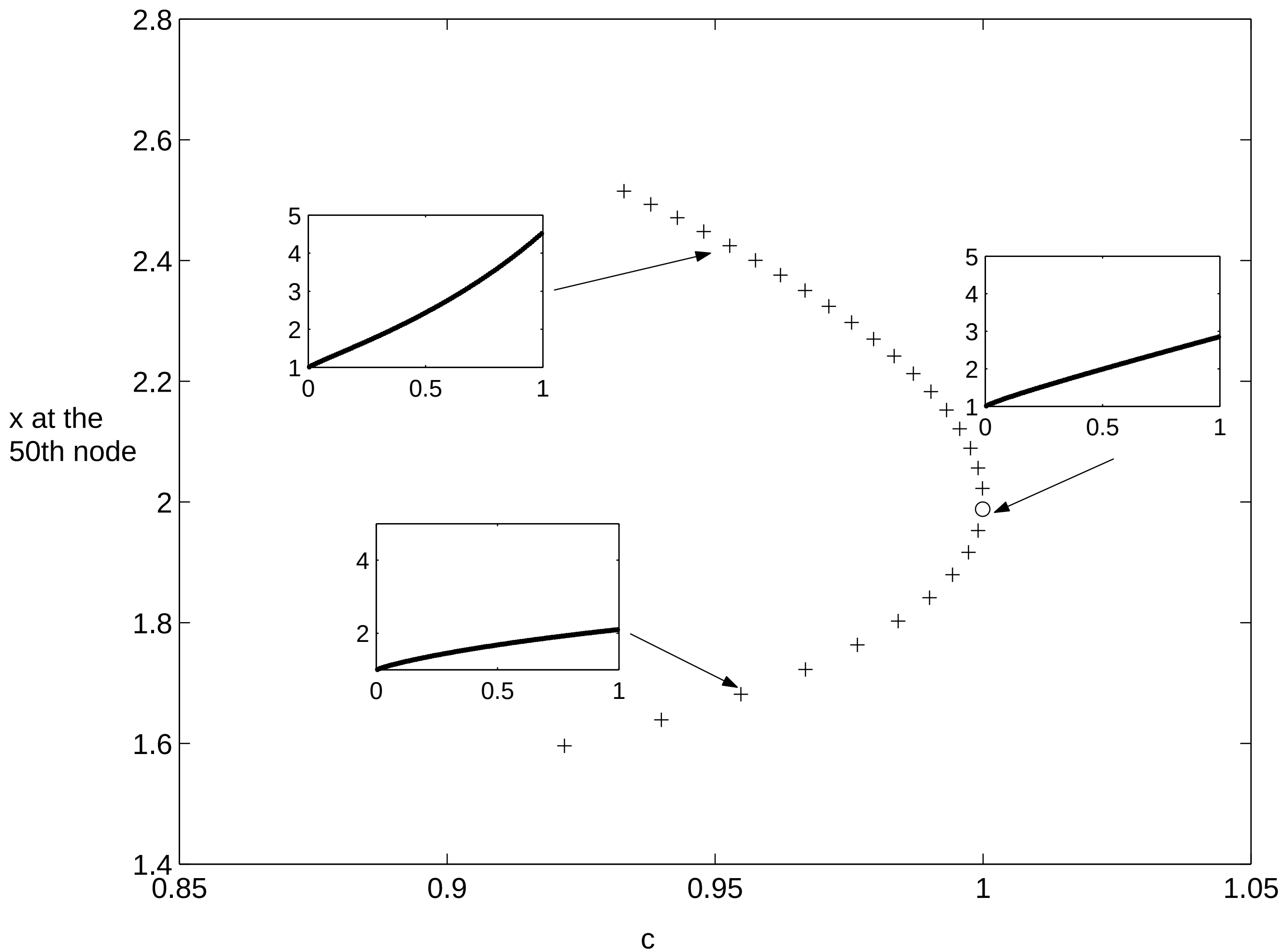
$$H(\mu) = \left(1 - \frac{c}{2}\int_0^1 \frac{\mu H(\nu)\,d\nu}{\mu + \nu}\right)^{-1}$$

- Compact fixed point problem.
- Problem becomes harder as $H(1) \to \infty$
- Two solutions for $c \neq 0, 1$
  - Two continuous solutions for $0 < c < 1$.
  - Complex conjugate pairs for $c > 1$.
  - One continuous, one unbounded for $c < 0$.

$\|H\|_1$ vs c



H and the path



Wigner-Poisson Equation for $f(t,x,k)$

$$\frac{\partial f}{\partial t} = -\frac{hk}{2\pi m^*}\frac{\partial f}{\partial x} - V(f) + \left.\frac{\partial f}{\partial t}\right|_{coll},$$

Wigner-Poisson Equation for $f(t,x,k)$

$$\frac{\partial f}{\partial t} = -\frac{hk}{2\pi m^*}\frac{\partial f}{\partial x} - V(f) + \left.\frac{\partial f}{\partial t}\right|_{coll},$$

$$V(f)(x,k) = \frac{1}{h}\int dk' f(x,k')\int dy[U(x+y)-U(x-y)]sin[2y(k-k')].$$

$$U(z) = u(z) + \Delta_c(z), \frac{d^2}{dx^2}u(x) = \frac{q^2}{\varepsilon}\left[N_d(x) - \int_{-\infty}^{\infty}\frac{dk}{2\pi}f(x,k)\right].$$

Wigner-Poisson Equation for $f(t,x,k)$

$$\frac{\partial f}{\partial t} = -\frac{hk}{2\pi m^*}\frac{\partial f}{\partial x} - V(f) + \left.\frac{\partial f}{\partial t}\right|_{coll},$$

$$V(f)(x,k) = \frac{1}{h}\int dk' f(x,k') \int dy[U(x+y) - U(x-y)] sin[2y(k-k')].$$

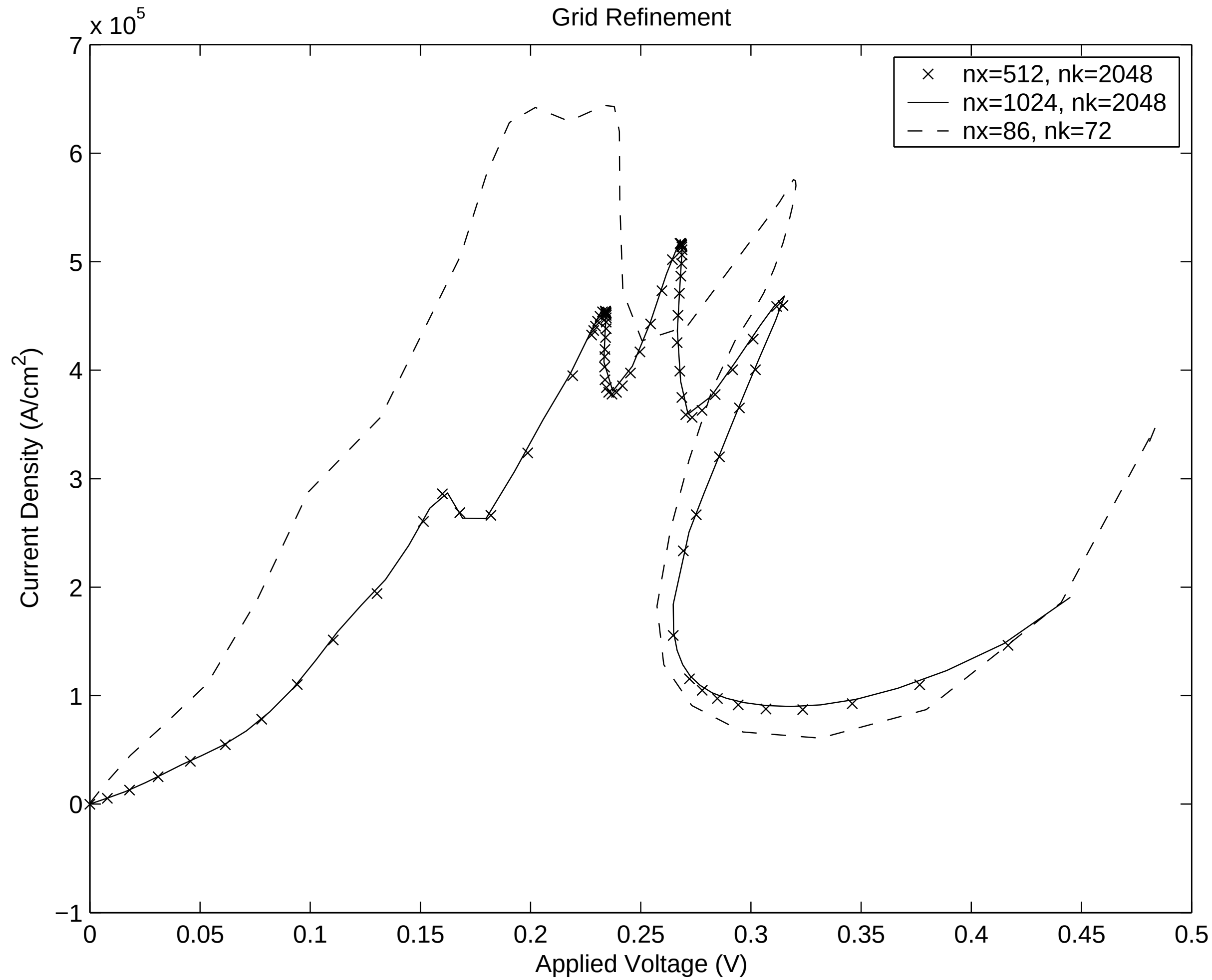
$$U(z) = u(z) + \Delta_c(z), \frac{d^2}{dx^2}u(x) = \frac{q^2}{\varepsilon}\left[N_d(x) - \int_{-\infty}^{\infty}\frac{dk}{2\pi}f(x,k)\right].$$

$$\left.\frac{\partial f}{\partial t}\right|_{coll} = \frac{1}{\tau}\left[\frac{f_0(x,k)}{\int dk\, f_0(x,k)}\int dk\, f(x,k) - f(x,k)\right].$$

Path following for Wigner Poisson Eq

- Use LOCA (Salinger-Phipps)
  NOX, AztecOO, Anasazi, Epetra
- Precondition with inverse of spatial differential operator
- Uniformly bounded, not quite compact
- Folds, hysteresis, Hopf bifurcation

Latest LOCA results



Time-stepper

Model Problem: Investigate steady-state solutions of the Chafee-Infante equation

$$u_t - \nu u_{xx} + u^3 - u = 0, x \in [0, \pi], u(0,t) = u(\pi,t) = 0,$$

as functions of ν.

Time-stepper

Model Problem: Investigate steady-state solutions of the Chafee-Infante equation

$$u_t - \nu u_{xx} + u^3 - u = 0, x \in [0, \pi], u(0,t) = u(\pi,t) = 0,$$

as functions of ν.
Method: Let $K(T,u,\nu)$ be the solution of the PDE at time T with initial data u. Solve

$$F(u,\nu) = u - K(T,u,\nu).$$

If $\nu > 0$, K is a smoother.
T becomes an algorithmic parameter.
More complex examples of this idea are in Schroff-Keller(93), Gear-Kevrekidis(03) ...

General Timesteppers

- $u \to K(T, u, \lambda)$ is “almost” finite rank.

General Timesteppers

- $u \longrightarrow K(T, u, \lambda)$ is “almost” finite rank.
  - Finitely many important modes (inertial manifold)

General Timesteppers

- $u \to K(T, u, \lambda)$ is "almost" finite rank.
  - Finitely many important modes (inertial manifold)
  - Size of T affects number of "slow modes"

General Timesteppers

- $u \rightarrow K(T, u, \lambda)$ is "almost" finite rank.
  - Finitely many important modes (inertial manifold)
  - Size of T affects number of "slow modes"
- Map may come from

General Timesteppers

- $u \rightarrow K(T, u, \lambda)$ is “almost” finite rank.
  - Finitely many important modes (inertial manifold)
  - Size of T affects number of “slow modes”
- Map may come from
  - Black-box codes

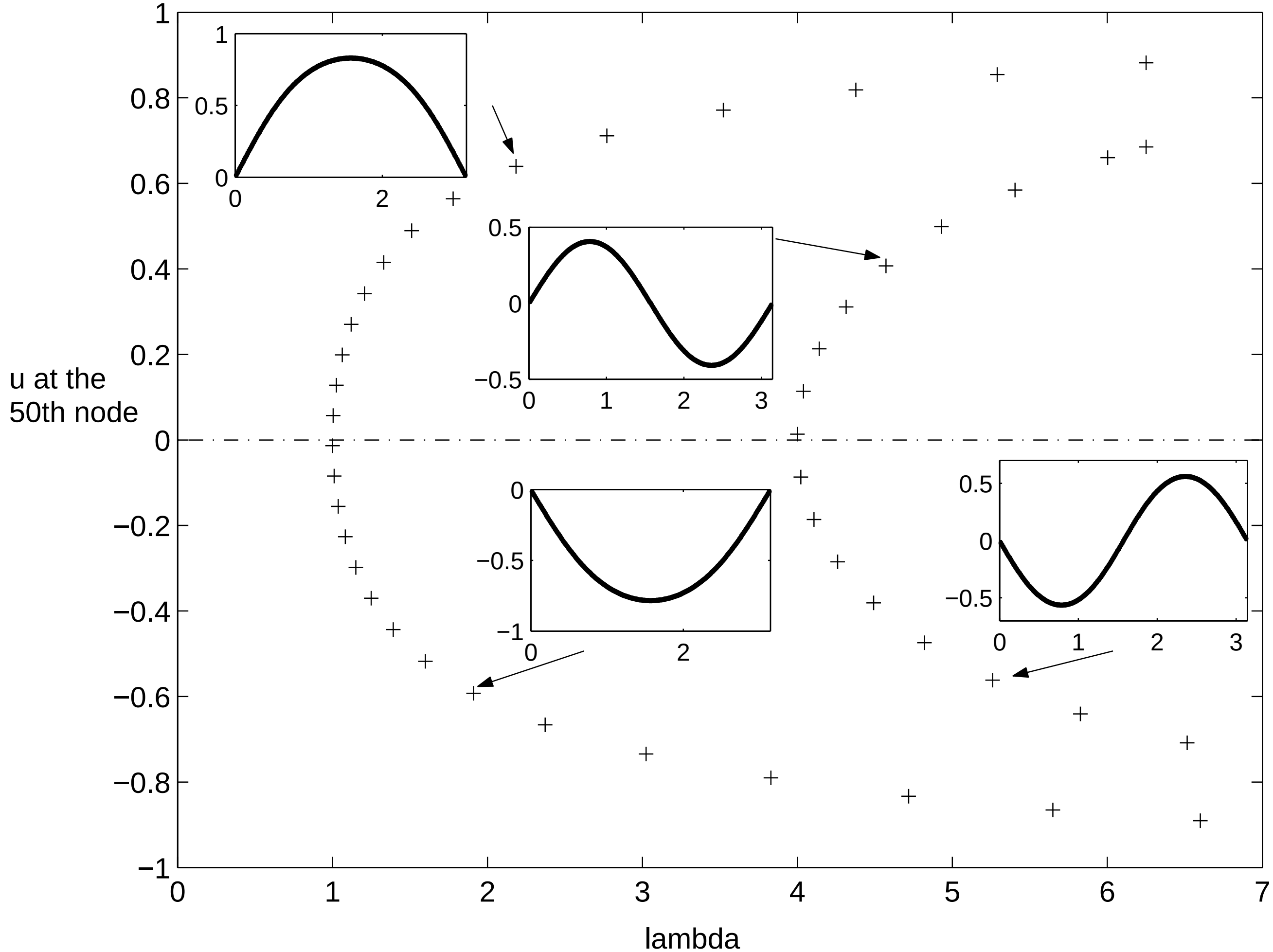
General Timesteppers

- $u \rightarrow K(T, u, \lambda)$ is “almost” finite rank.
  - Finitely many important modes (inertial manifold)
  - Size of T affects number of “slow modes”
- Map may come from
  - Black-box codes
  - Microscale simulations scales using non-DE methods

General Timesteppers

- $u \to K(T, u, \lambda)$ is “almost” finite rank.
  - Finitely many important modes (inertial manifold)
  - Size of T affects number of “slow modes”
- Map may come from
  - Black-box codes
  - Microscale simulations scales using non-DE methods
  - Large codes that are hard to modify and/or understand

u and the path



Branch Switching

These were not simple folds.

- Simple bifurcations (the forks) $\rightarrow$ sign change in determinant.
  How do you compute that determinant?
- Matrix-free detection $\rightarrow$
  - generalized eigenvalue problem $\rightarrow$
  - s^* and $w \neq 0$ such that $G_x(x(s^*))w = 0$
- At the bifurcation point s^*: choice of directions.
  $\dot{x}$ or the new direction $\pm w$.

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)
    - Mesh-independent performance for compact ranges of s,

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)
  - Mesh-independent performance for compact ranges of s,
  - Preconditioning easy or unnecessary(?).

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)
  - Mesh-independent performance for compact ranges of s,
  - Preconditioning easy or unnecessary(?).
- Multilevel solvers

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)
  - Mesh-independent performance for compact ranges of s,
  - Preconditioning easy or unnecessary(?).
- Multilevel solvers
  - Easy to build. Compactness smooths for you.

How should compactness help?

- Newton-Krylov solvers: Ferng-K(00), K, Kevrekidis, Qiao (04)
  - Mesh-independent performance for compact ranges of s,
  - Preconditioning easy or unnecessary(?).
- Multilevel solvers
  - Easy to build. Compactness smooths for you.
  - Appropriate coarse grid data depend on s.

Timesteppers and Compactness

Let D have dimension d

$$F_u(u,v) = I - K + E$$

where

- $K = P_D K P_D$, where P_D is a projection onto D
- $\|E\|$ is small, and
- we solve $F_u(u,v)s = -F(u,v)$ with GMRES.

Dimension of D will depend on T.
T should be selected with thought.

Convergence of GMRES

Let r_m be the mth GMRES residual.
Set

$$p(z) = p_M(z)/p(0).$$

where p_M is the minimal polynomial of $I - K$.

Convergence of GMRES

Let r_m be the mth GMRES residual.
Set

$$p(z) = p_M(z)/p(0).$$

where p_M is the minimal polynomial of $I - K$. Since

$$\|p(F_u)\| = O(\|E\|) \text{ so } \|p(F_u)^m\| = O(\|E\|^m)$$

Convergence of GMRES

Let r_m be the mth GMRES residual.
Set

$$p(z) = p_M(z)/p(0).$$

where p_M is the minimal polynomial of $I-K$. Since

$$\|p(F_u)\| = O(\|E\|) \text{ so } \|p(F_u)^m\| = O(\|E\|^m)$$

we can apply standard GMRES theory to show

$$\|r_{m(d+1)}\| \leq \|p(F_u)^m r_0\| = O(\|E\|^m),$$

for all $m \geq 1$.

Inflated system

Same results for

$$G(x,s) = \begin{pmatrix} F \\ N \end{pmatrix} = \begin{pmatrix} F(u(s),\lambda(s)) \\ \dot{u}^T(u-u_0) + \dot{\lambda}^T(\lambda-\lambda_0) - (s-s_0) \end{pmatrix}.$$

with d replaced by $d+2$.
Meaning: cost of solve is independent of discretization,
unless d begins to increase with s.

Multilevel Approach

Pathfollowing on coarse mesh + nested iteration fails.

- $F(u, \lambda) = u - \mathscr{K}(u, \lambda)$
- $\lambda(s)$ is sensitive to the mesh.
- Track path on fine mesh.
- Use coarse mesh problem to approximate $\mathscr{K}_u$
  Apply GMRES to new problem.

Coarse mesh problem construction

For continuation in λ

- $x^h = x^h + dx$, Euler predictor on fine mesh.
- $u^H = I_h^H(u^h)$, $\lambda = \lambda^H = \lambda^h$.
- Build $K_H = I_H^h \mathscr{K}_u^H(u^H, \lambda) I_h^H$
- Norm convergent (K, 1995) if I_h^H is done right degenerate kernel approximation
- Approximate Newton step by solving $s - K_H s = -F_h(u^H, \lambda)$.
  Fine mesh residual and coarse mesh solve.

Continuation in s

Approximate G_x by

$$G^{H,h}_{u,\lambda}(u,\lambda) \equiv \begin{pmatrix} I - \partial \mathscr{K}_H(I^H_h u, \lambda)/\partial u & -\partial \mathscr{K}_H(I^H_h u, \lambda)/\partial \lambda \\ \\ (I^H_h \dot{u})^T & \dot{\lambda} \end{pmatrix}.$$

and apply GMRES.

Continuation in s

Approximate G_x by

$$G^{H,h}_{u,\lambda}(u,\lambda) \equiv \begin{pmatrix} I - \partial \mathscr{K}_H(I^H_h u, \lambda)/\partial u & -\partial \mathscr{K}_H(I^H_h u, \lambda)/\partial \lambda \\ \\ (I^H_h \dot{u})^T & \dot{\lambda} \end{pmatrix}.$$

and apply GMRES.

- Operator-function product is now on coarse mesh.
- Works for "black-box" functions. Flexible choice of $\mathscr{K}^H$.
- Theory follows from older work,
  if you coarsen only in $\mathscr{K}$, not in G.

Conclusions

- Exploitation of compactness in path following
  - Simple folds
    - 6 coarse mesh Krylovs/Newton for H-equation
    - Multilevel Chafee-Infante results in progress
    - GMRES working for Wigner-Poisson Eq
  - Branching and Hopf in the works
    Wigner-Poisson results for Hopf almost there
- Scaling F vs N important as path grows