A Newton-Krylov solver for fully implicit 3D extended MHD

L. Chacón

Los Alamos National Laboratory P.O. Box 1663, Los Alamos, NM 87545

In collaboration with B. Philip (LANL), J. Shadid, R. Pawlowski, J. Banks (SNL)

2007 OASCR PI MEETING MAY 22-24, 2007 LIVERMORE, CA, USA



Outline

- Motivation: XMHD and the tyranny of scales
- Parabolization of XMHD: key for SCALABILITY
- Resistive MHD
- Hall MHD
- Migration to unstructured FE: status report (with SNL)
- Spatial adaptivity: Implicit + AMR (with B. Philip, LANL LDRD)



"The tyranny of scales" (SBES report, 2006)



(a) Time scales in fusion plasmas (FSP report)

(b) Length scales in a typical fusion plasma (Tang, *Phys. Plasmas*, 9 (5), 2002)

"The tyranny of scales will not be simply defeated by building bigger and faster computers" (SBES report, p. 30)



Algorithmic challenges in XMHD

- XMHD has mixed character, with strongly hyperbolic and parabolic components.
- Numerically, XMHD is a nonlinear algebraic system of very stiff equations:
 - Elliptic stiffness (diffusion): $\kappa(J) \sim \frac{\Delta t D}{\Delta x^2} \gg 1$
 - Hyperbolic stiffness (linear and dispersive waves): $\kappa(J) \sim \Delta t \, \omega_{fast} \sim \frac{\Delta t}{\Delta t_{CFL}} \gg 1$
- Brute-force algorithms will not be able to cover the span between disparate time/length scales, regardless of computer power (SBES report).
- Key algorithmic requirement: SCALABILITY [$CPU \sim O(N/n_p)$]!
 - Minimize number of degrees of freedom N: spatial adaptivity.
 - Follow slowest time scales (application dependent): implicit time stepping.
- Scalable implicit methods require MULTILEVEL approaches:

$$CPU \sim \mathcal{O}\left(\frac{N \, \log(N)}{n p^{\beta}}\right) \ , \ \beta \lesssim 1$$



XMHD and multilevel approaches

- A fundamental component of iterative ML methods is the SMOOTHER.
- XMHD is strongly hyperbolic \Rightarrow smoothing is a serious challenge (diagonally submissive for $\Delta t > \Delta t_{CFL}$).
 - Previous attempts to use multilevel methods (two-level NKS, MG-NKS) on XMHD have failed to demonstrate a scalable XMHD solver.

Our solution: parabolize XMHD! (multilevel-friendly)



Parabolization and Schur complement: an example

PARABOLIZATION EXAMPLE:

$$\partial_t u = \partial_x v , \ \partial_t v = \partial_x u.$$

$$u^{n+1} = u^n + \Delta t \partial_x v^{n+1}, v^{n+1} = v^n + \Delta t \partial_x u^{n+1}.$$

$$(I - \Delta t^2 \partial_{xx})u^{n+1} = u^n + \Delta t \partial_x v^n$$

• PARABOLIZATION via SCHUR COMPLEMENT:

$$\begin{bmatrix} D_1 & U \\ L & D_2 \end{bmatrix} = \begin{bmatrix} I & UD_2^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} D_1 - UD_2^{-1}L & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} I & 0 \\ D_2^{-1}L & I \end{bmatrix}.$$

Stiff off-diagonal blocks L, U now sit in diagonal via Schur complement $D_1 - UD_2^{-1}L$. The system has been "PARABOLIZED."

$$D_1 - UD_2^{-1}L = (I - \Delta t^2 \partial_{xx})$$



Our approach to a successful fully implicit algorithm for XMHD

- Even if a smoother exists, MG is remarkably temperamental.
- Combination of Krylov methods and MG is optimal:
 - MG provides scalability (as a preconditioner)
 - Krylov provides robustness

We seek to develop a successful algorithm for XMHD based on Newton-Krylov-MG

• Proof the concept in resistive MHD, and then move to XMHD.



Jacobian-Free Newton-Krylov Methods

- Objective: solve nonlinear system $\vec{G}(\vec{x}^{n+1}) = \vec{0}$ efficiently (scalably).
- Converge nonlinear couplings using Newton-Raphson method:
- $\left. rac{\partial ec G}{\partial ec x}
 ight|_k \delta ec x_k = -ec G(ec x_k) \; \; .$

• Jacobian-free implementation:

$$\left(\frac{\partial \vec{G}}{\partial \vec{x}}\right)_k \vec{y} = J_k \vec{y} = \lim_{\epsilon \to 0} \frac{\vec{G}(\vec{x}_k + \epsilon \vec{y}) - \vec{G}(\vec{x}_k)}{\epsilon}$$

- Krylov method of choice: GMRES (nonsymmetric systems).
- Right preconditioning: solve equivalent Jacobian system for $\delta y = P_k \delta \vec{x}$:

$$J_k P_k^{-1} \underbrace{\underline{P_k \delta \vec{x}}}_{\delta \vec{y}} = -\vec{G}_k$$

APPROXIMATIONS IN PRECONDITIONER DO NOT AFFECT ACCURACY OF CONVERGED SOLUTION; THEY ONLY AFFECT EFFICIENCY!



Implicit resistive MHD solver



Resistive MHD model equations

$$\begin{split} \frac{\partial \rho}{\partial t} &+ \nabla \cdot (\rho \vec{v}) = 0, \\ \frac{\partial \vec{B}}{\partial t} &+ \nabla \times \vec{E} = 0, \\ \frac{\partial (\rho \vec{v})}{\partial t} + \nabla \cdot \left[\rho \vec{v} \vec{v} - \vec{B} \vec{B} &- \rho \nu \nabla \vec{v} + \overleftarrow{I} \left(p + \frac{B^2}{2} \right) \right] = 0, \\ \frac{\partial T}{\partial t} + \vec{v} \cdot \nabla T &+ (\gamma - 1) T \nabla \cdot \vec{v} = 0, \end{split}$$

- Plasma is assumed polytropic $p \propto n^{\gamma}.$
- Resistive Ohm's law:

$$\vec{E} = -\vec{v} \times \vec{B} + \eta \nabla \times \vec{B}$$



Resistive MHD Jacobian block structure

• The linearized resistive MHD model has the following couplings:

$$\begin{split} \delta \rho &= L_{
ho}(\delta
ho, \delta ec v) \ \delta T &= L_{T}(\delta T, \delta ec v) \ \delta ec B &= L_{B}(\delta ec B, \delta ec v) \ \delta ec v &= L_{v}(\delta ec v, \delta ec B, \delta
ho, \delta T) \end{split}$$

• Therefore, the Jacobian of the resistive MHD model has the following coupling structure:

$$J\delta \vec{x} = \begin{bmatrix} D_{\rho} & 0 & 0 & U_{v\rho} \\ 0 & D_{T} & 0 & U_{vT} \\ 0 & 0 & D_{B} & U_{vB} \\ L_{\rho v} & L_{Tv} & L_{Bv} & D_{v} \end{bmatrix} \begin{pmatrix} \delta \rho \\ \delta T \\ \delta \vec{B} \\ \delta \vec{v} \end{pmatrix}$$

• Diagonal blocks contain advection-diffusion contributions, and are "easy" to invert using MG techniques. Off diagonal blocks L and U contain all hyperbolic couplings.



PARABOLIZATION: Schur complement formulation

• We consider the block structure:

$$J\delta\vec{x} = \begin{bmatrix} M & U \\ L & D_v \end{bmatrix} \begin{pmatrix} \delta\vec{y} \\ \delta\vec{v} \end{pmatrix} ; \ \delta\vec{y} = \begin{pmatrix} \delta\rho \\ \deltaT \\ \delta\vec{B} \end{pmatrix} ; \ M = \begin{pmatrix} D\rho & 0 & 0 \\ 0 & D_T & 0 \\ 0 & 0 & D_B \end{pmatrix}$$

• *M* is "easy" to invert (advection-diffusion, MG-friendly).

Schur complement analysis of 2x2 block J yields:

$$\begin{bmatrix} M & U \\ L & D_v \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -LM^{-1} & I \end{bmatrix} \begin{bmatrix} M^{-1} & 0 \\ 0 & P_{Schur}^{-1} \end{bmatrix} \begin{bmatrix} I & -M^{-1}U \\ 0 & I \end{bmatrix},$$
$$P_{Schur} = D_v - LM^{-1}U.$$

- EXACT Jacobian inverse only requires M^{-1} and P_{Schur}^{-1} .
- Schur complement formulation is fundamentally unchanged in Hall MHD!



Physics-based preconditioner (I)

 The Schur complement analysis translates into the following 3-step EXACT inversion algorithm:

Predictor : $\delta \vec{y}^* = -M^{-1}G_y$ Velocity update : $\delta \vec{v} = P_{Schur}^{-1}[-G_v - L\delta \vec{y}^*], P_{Schur} = D_v - LM^{-1}U$ Corrector : $\delta \vec{y} = \delta \vec{y}^* - M^{-1}U\delta \vec{v}$

• MG treatment of P_{Schur} is impractical due to M^{-1} .

Need suitable simplifications (SEMI-IMPLICIT)!

- We consider the small-flow-limit case: $M^{-1} \approx \Delta t$
- This approximation is equivalent to splitting flow in original equations.



Physics-based preconditioner (II)

• Small flow approximation: $M^{-1} \approx \Delta t$ in steps 2 & 3 of Schur algorithm:

$$\begin{split} \delta \vec{y}^* &= -M^{-1} G_y \\ \delta \vec{v} &\approx P_{SI}^{-1} \left[-G_v - L \delta \vec{y}^* \right] ; \ P_{SI} = D_v - \Delta t L U \\ \delta \vec{y} &\approx \delta \vec{y}^* - \Delta t U \delta \vec{v} \end{split}$$

where:

$$P_{SI} = \rho^{n} \left[\overleftarrow{I} / \Delta t + \theta (\vec{v}_{0} \cdot \nabla \overleftarrow{I} + \overleftarrow{I} \cdot \nabla \vec{v}_{0} - \nu^{n} \nabla^{2} \overleftarrow{I}) \right] + \Delta t \theta^{2} W(\vec{B}_{0}, p_{0})$$
$$W(\vec{B}_{0}, p_{0}) = \vec{B}_{0} \times \nabla \times \nabla \times \left[\overleftarrow{I} \times \vec{B}_{0} \right] - \vec{j}_{0} \times \nabla \times \left[\overleftarrow{I} \times \vec{B}_{0} \right] - \nabla \left[\overleftarrow{I} \cdot \nabla p_{0} + \gamma p_{0} \nabla \cdot \overleftarrow{I} \right]$$

- *P*_{SI} is block diagonally dominant by construction!
- We employ multigrid methods (MG) to approximately invert P_{SI} and M: 1 V(4,4) cycle



Efficiency: Δt scaling (2D tearing mode)

32×32

Δt	Newton/ Δt	$GMRES/\Delta t$	CPU (s)	CPU_{exp}/CPU	$\Delta t/\Delta t_{CFL}$
2	5.9	20.9	115	3.1	354
3	5.9	25.6	139	3.8	531
4	6.0	30.5	163	4.3	708
6	6.0	34.7	184	5.8	1062

128×128

Δt	Newton/ Δt	$GMRES/\Delta t$	CPU (s)	CPU_{exp}/CPU	$\Delta t/\Delta t_{CFL}$
0.5	4.9	8.4	764	8.0	380
0.75	5.7	10.2	908	10.0	570
1.0	5.0	11.5	1000	12.7	760
1.5	5.6	14.7	1246	14.6	1140



Efficiency: grid scaling

$\Delta t \approx 1100 \Delta t_{CFL}$, 10 time steps

Grid	Δt	Newton/ Δt	GMRES/ Δt	CPU	\widehat{CPU}
32x32	6	6.0	34.7	184	5.3
64x64	3	5.8	22.9	468	20.4
128x128	1.5	5.6	14.8	1246	84.2

Why does GMRES/ Δt decrease with resolution?



Effect of spatial truncation error





Implicit extended MHD solver



Extended MHD model equations

$$\begin{split} \frac{\partial \rho}{\partial t} &+ \nabla \cdot (\rho \vec{v}) = 0, \\ \frac{\partial \vec{B}}{\partial t} &+ \nabla \times \vec{E} = 0, \\ \frac{\partial (\rho \vec{v})}{\partial t} + \nabla \cdot \left[\rho \vec{v} \vec{v} - \vec{B} \vec{B} &- \rho \nu \nabla \vec{v} + \overleftarrow{I} \left(p + \frac{B^2}{2} \right) \right] = 0, \\ \frac{\partial T_e}{\partial t} + \vec{v} \cdot \nabla T_e &+ (\gamma - 1) T_e \nabla \cdot \vec{v} = 0, \end{split}$$

- Plasma is assumed polytropic $p \propto n^{\gamma}.$
- We assume cold ion limit: $T_i \ll T_e \Rightarrow \boxed{p \approx p_e}$
- Generalized Ohm's law:

$$ec{E} = -ec{v} imes ec{B} + \eta
abla imes ec{B} - rac{d_i}{
ho} (ec{j} imes ec{B} -
abla p_e)$$



Extended MHD Jacobian block structure

• The linearized extended MHD model has the following couplings:

$$\begin{split} \delta \rho &= L_{\rho}(\delta \rho, \delta \vec{v}) \\ \delta T &= L_{T}(\delta T, \delta \vec{v}) \\ \delta \vec{B} &= L_{B}(\delta \vec{B}, \delta \vec{v}, \delta \rho, \delta T) \\ \delta \vec{v} &= L_{v}(\delta \vec{v}, \delta \vec{B}, \delta \rho, \delta T) \end{split}$$

• Jacobian coupling structure:

$$J\delta \vec{x} = \begin{bmatrix} D_{\rho} & 0 & 0 & U_{v\rho} \\ 0 & D_{T} & 0 & U_{vT} \\ L_{\rho B} & L_{TB} & D_{B} & U_{vB} \\ L_{\rho v} & L_{Tv} & L_{Bv} & D_{v} \end{bmatrix} \begin{pmatrix} \delta \rho \\ \delta T \\ \delta \vec{B} \\ \delta \vec{v} \end{pmatrix}$$

• We have added off-diagonal couplings.



Extended MHD Jacobian block structure (cont.)

• The coupling structure can be substantially simplified if we note $(p \approx p_e)$:

$$\frac{1}{\rho}(\vec{j} \times \vec{B} - \nabla p_e) \approx \frac{D\vec{v}}{Dt}$$

and therefore:

$$\vec{E} \approx -\vec{v} \times \vec{B} + \frac{\eta(T)}{\mu_0} \nabla \times \vec{B} - d_i \frac{D\vec{v}}{Dt}$$

• This transforms jacobian coupling structure to:

$$J\delta \vec{x} \approx \begin{bmatrix} D_{\rho} & 0 & 0 & U_{v\rho} \\ 0 & D_{T} & 0 & U_{vT} \\ 0 & 0 & D_{B} & U_{vB}^{R} + U_{vB}^{H} \\ L_{\rho v} & L_{Tv} & L_{Bv} & D_{v} \end{bmatrix} \begin{pmatrix} \delta \rho \\ \delta T \\ \delta \vec{B} \\ \delta \vec{v} \end{pmatrix}$$

We can therefore reuse ALL resistive MHD PC framework!



Extended MHD preconditioner

- Use same Schur complement approach.
- *M* block contains ion scales only! Approximation $M^{-1} \approx \Delta t$ is very good in extended MHD (ion scales do NOT contribute to numerical stiffness).
- Additional block U_{vB}^{H} results, after the Schur complement treatment, in systems of the form:

$$\partial_t \delta \vec{v} - d_i \vec{B_0} \times (\nabla \times \nabla \times \delta \vec{v}) = rhs$$

- This system supports dispersive waves $\omega \sim k^2!$
- We have shown analytically that damped JB is a smoother for these systems!

We can use classical MG!



Preliminary efficiency results (2D tearing mode)

 $d_i = 0.05$

1 time step, $\Delta t = 1.0$, V(3,3) cycles, mg_tol=1e-2

Grid	Newton/ Δt	$GMRES/\Delta t$	CPU (s)	CPU_{exp}/CPU	$\Delta t/\Delta t_{exp}$
32x32	5	22	25	0.44	110
64x64	5	12	66	1.4	238
128x128	5	8	164	6.2	640
256x256	4	7	674	30	3012

Again, GMRES/ Δt decreases with resolution!



Effect of spatial truncation error



Residual history vs. GMRES it# with fixed time step Dt=1

Parallel performance with PETSc Toolkit (unpreconditioned, 3D, weak scaling with 32^3 nodes per processor)

Migration to unstructured FE

(In collaboration with J. Shadid, R. Pawlowski, J. Banks, SNL)

Currently: Initial Single Fluid Resistive MHD Unstructured FE Formulation

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \bullet \mathbf{F} + \mathbf{S} = \mathbf{0}$$

$$E = e + \frac{1}{2} \|\mathbf{v}\|^{2}$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho \mathbf{v} \\ \Sigma_{tot} \\ \mathbf{B} \end{bmatrix} \mathbf{F} = \begin{bmatrix} \rho \mathbf{v} \\ \rho \mathbf{v} \otimes \mathbf{v} - \frac{1}{\mu_{0}} \mathbf{B} \otimes \mathbf{B} - \mathbf{T} + \frac{1}{2\mu_{0}} \|\mathbf{B}\|^{2} \mathbf{I} \\ \rho E \mathbf{v} - \mathbf{T} \cdot \mathbf{v} + \mathbf{E} \times \mathbf{B} + \mathbf{q} \\ \mathbf{v} \otimes \mathbf{B} - \mathbf{B} \otimes \mathbf{v} - \frac{\eta}{\mu_{0}} (\nabla \mathbf{B} - \nabla \mathbf{B}^{T}) \end{bmatrix} \mathbf{S} = \begin{bmatrix} 0 \\ \mathbf{0} \\ Q^{rad} + Q \\ \mathbf{0} \end{bmatrix}$$

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} + \eta \mathbf{J} + \frac{1}{\underline{e}n} (\mathbf{J} \times \mathbf{B} - \nabla \mathbf{P}_{\mathbf{e}}) + \frac{m_{e}}{\underline{e}^{2} n_{e}} \frac{d \mathbf{J}}{dt}.$$

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} + \eta \mathbf{J} + \frac{1}{\underline{e}n} (\mathbf{J} \times \mathbf{B} - \nabla \mathbf{P}_{\mathbf{e}}) + \frac{m_{e}}{\underline{e}^{2} n_{e}} \frac{d \mathbf{J}}{dt}.$$

Project Goals:

• Develop stable, accurate, physics compatible, scalable and efficient fully-implicit computational formulations for xMHD and PTR (e.g. SNL Cray XT3 12.5K nodes, 25K cores)

Develop and evaluate scalable physics-based preconditioners, based on multi-level methods

• Produce comprehensive accuracy, convergence, stability and scalability studies employing challenging prototype problems.

Produce first-of-a-kind large-scale computational demonstrations on selected science / technology problems

- Science
 - Magnetic Reconnection Studies
 - Hydro-Magnetic Rayleigh-Taylor (e.g. Z-pinch [HEDP])
- Technology (e.g. advanced materials processing)
 - Plasma arc jet CVD, Plasma CVD/ Etching

(J. N. Shadid, R. P. Pawlowski, J. W. Banks - SNL)

Example Unstructured Mesh Solutions

Implicit NK-AMR

B. Philip, M. Pernice, and L. Chacón, Lecture Notes in Computational Science and Engineering, accepted (2006).

Current-Vorticity Formulation of Reduced Resistive MHD¹

$$(\partial_t + \mathbf{u} \cdot \nabla - \eta \Delta) J + \Delta E_0 = \mathbf{B} \cdot \nabla \omega + \{\Phi, \Psi\}$$
$$(\partial_t + \mathbf{u} \cdot \nabla - \nu \Delta) \omega + S_\omega = \mathbf{B} \cdot \nabla J$$
$$\Delta \Phi = \omega$$
$$\Delta \Psi = J$$

$$\mathbf{u} = \vec{z} \times \nabla \Phi , \ \mathbf{B} = \vec{z} \times \nabla \Psi$$
$$\{\Phi, \Psi\} = 2[\Phi_{xy}(\Psi_{xx} - \Psi_{yy}) - \Psi_{xy}(\Phi_{xx} - \Phi_{yy})]$$

Preconditioner is an extension of Chacón, Knoll and Finn, JCP, **178** (2002).

¹Strauss and Longcope. JCP, **147**, 1998

Implicit Structured Adaptive Mesh Refinement (SAMRAI-PETSc-*hypre*)

• *Structured* adaptive mesh refinement (SAMR) represents a locally refined mesh as a union of logically rectangular meshes.

- The mesh is organized as a hierachy of refinement levels.
- Each refinement level defines a region of uniform resolution.
- Each refinement level is the union of logically rectangular patches.

AMR-grids and multilevel methods are fundamentally compatible approaches!

Performance (tearing mode)

• Generalized 2D reduced MHD PC [Chacon et al., JCP (2002)] for SAMR (MG \Rightarrow FAC).

			NNI					NLI		
Levels	1	2	3	4	5	1	2	3	4	5
32×32	1.5	2.0	2.0	2.1	2.5	3.4	7.9	12.0	19.3	33.7
64×64	1.8	2.0	2.0	2.4	—	6.5	11.7	19.1	33.2	-
128×128	1.8	2.0	2.4	—	_	12.5	20.1	27.2	_	-
256×256	1.9	2.0	_	_	_	19.9	27.5	_	_	_
512×512	1.9	_	_	_	_	26.3	_	_	_	_

 $\Delta t = 1$ (fixed), $\eta_k = 0.1$, $\epsilon_{rel} = \epsilon_{abs} = 10^{-7}$, 2 SI iterations, V(3,3) cycles

- Fixed implicit time step (problem gets harder with refinement)
- Performance does not degrade with grid-refinement levels

Island Coalescence Results at t=8

Tilt Instability Results at t=7

Conclusions

• Developed a scalable, multilevel-based, fully implicit NK-MG solver for XMHD.

Key algorithmic breakthrough: PARABOLIZATION + MG.

- Equivalence between parabolization and the Schur decomposition:
 - Provides a rigorous foundation for the parabolization step.
 - Provides a path to generalize approach when more complete XMHD models are considered.
- Demonstrated algorithmic viability of implicit AMR by generalizing single-grid preconditioning approaches for MHD.
- Future work:
 - Massively parallel test of 3D resistive MHD algorithm (NERSC).
 - Bring Hall MHD to production stage (high-order dissipation required).
 - Implicit AMR on 3D resistive MHD (B. Philip).
 - Multilevel-based PC on unstructured FE (SNL).

