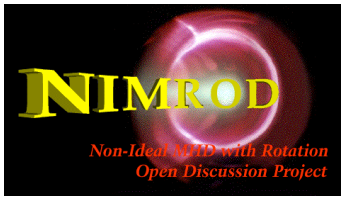
Preconditioning Tests

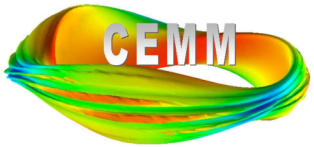
Carl Sovinec

NIMROD Team Meeting

Boulder, Colorado

March 29, 2008





Outline

- Test problem
- Review toroidal coupling with Hall
- Electron inertia
- Polynomial preconditioning

Implementation

- Terms appearing on the left sides of the equations are linear in the field being advanced (the Δ-factor in each).
- NIMROD uses finite Fourier series in the toroidal coordinate for spectral convergence.
  - With toroidally symmetric geometry and equilibria, linear analyses solve a separate linear system for each toroidal Fourier component, like separate 2D computations.
  - Matrices for nonlinear 3D computations have matrix elements that couple different Fourier components. They are smaller than the matrix elements for the poloidal-plane coupling by at least one factor of the perturbation amplitude.
  - Our use of FFTs to compute Fourier-component couplings scales well in production calculations, but algebraic representation of the complete '3D' matrix is not practical.
- Krylov-space solvers need matrix-vector product operations ($A \cdot v$) but not matrix elements of A. Matrix-elements of M are used to 'precondition.'

The NIMROD solver strategy has been to generate only the large matrix elements for poloidal-plane coupling.

- The poloidal-plane matrices are sparse with the finite-element representation.
- They are inverted with the parallel SuperLU_DIST library [Li, ACM Trans. Math. Software **29**, 110 (2003)].
- The set of inverses (one for each Fourier component) provides an approximate inverse for the 3D system.
- The Kryov-space iterations handle the smaller Fourier-component coupling.
- This works well for semi-implicit MHD with predictor/corrector advection where matrices are HPD [JCP **195**, 355 (2004)] and for semi-implicit MHD with implicit advection, where the non-Hermitian part is small.

A cylindrical 1/1 mode has been used to examine the solver's convergence properties in 3D Hall advances.

- The pressure profile is flat, and $q(r)=0.98+0.51(r/a)^2$ has the 1/1 resonant surface at $r=0.2a$.
- Other parameters are based on [Aydemir, PFB **4**, 3469 (1992)].

$$S = \tau_r/\tau_{Hp} = 10^6 \quad \tau_r \equiv \mu_0 a^2/\eta \quad \tau_{Hp} \equiv a\sqrt{\mu_0\rho}/B_p$$

$$\beta = 5\times10^{-3} \quad \delta = d_i/2 = 0.11$$

$$\mu_0\nu_{iso}/\eta = \mathrm{Pm} = 0.1 \quad \mu_0 D/\eta = 0.01 \quad T_i \cong 0$$

- Converged linear computations produce $\gamma_{MHD}\tau_{Hp}=0.0183$ and $\gamma_{2fl}\tau_{Hp}=0.0207$.
- Very small nonlinear computations (6×6 mesh, $0\le n\le 2$) started with 1% perturbations use 40-60 GMRES iterations for the Hall advance.

The number of Krylov iterations does not remain O(1) for the implicit 3D Hall advance where the non-HPD part is large.

- A computation with a 14×12 mesh of biquartic elements for the *r-z* plane and 3 Fourier components takes 40+ GMRES its with SLU preconditioning.
- Test case: 6×6 bicubic and a large initial perturbation of 1%.
- GMRES its for the HMHD B-advance on restarts after four steps:
  - With 3D n, $\mathbf{B}^{j+1/2}$, $\mathbf{J}^{j+1/2}$, and $\mathbf{V}^{j+1/2}$ coefficients: 58
  - With ϕ– averaged $\mathbf{B}^{j+1/2}$: 30
  - With ϕ– averaged $\mathbf{B}^{j+1/2}$ & $\mathbf{J}^{j+1/2}$: 29
  - With ϕ– averaged $\mathbf{B}^{j+1/2}$ & $\mathbf{J}^{j+1/2}$ and n: 7
  - With ϕ– averaged $\mathbf{B}^{j+1/2}$, $\mathbf{J}^{j+1/2}$ & $\mathbf{V}^{j+1/2}$ and n: 5
- GMRES orthogonalization accuracy was checked.

$$\frac{\Delta\mathbf{B}}{\Delta t} - \frac{1}{2}\nabla\times\left(\mathbf{V}^{j+1}\times\Delta\mathbf{B}\right) + \frac{1}{2}\nabla\times\frac{1}{\bar{n}e}\left(\mathbf{J}^{j+1/2}\times\Delta\mathbf{B} + \Delta\mathbf{J}\times\mathbf{B}^{j+1/2}\right) + \frac{1}{2}\nabla\times\eta\Delta\mathbf{J}$$

Left side of B-advance.

The importance of having a symmetric system for preconditioning has been emphasized by Chacón [PoP **15**, 56103 (2008)]

- The HPD part of our system is increased by adding the $\frac{1}{\varepsilon_0 \omega_e^2}\frac{\partial}{\partial t}\mathbf{J}$ part of electron inertia.

$$\frac{\Delta\mathbf{B}}{\Delta t} - \frac{1}{2}\nabla\times\left(\mathbf{V}^{j+1}\times\Delta\mathbf{B}\right) + \frac{1}{2}\nabla\times\frac{1}{\bar{n}e}\left(\mathbf{J}^{j+1/2}\times\Delta\mathbf{B} + \Delta\mathbf{J}\times\mathbf{B}^{j+1/2}\right)$$

$$+\nabla\times\left(\frac{\eta}{2} + \frac{d_e^2}{\Delta t}\right)\nabla\times\Delta\mathbf{B} - \kappa_{divb}\nabla\nabla\cdot\Delta\mathbf{B}$$

for the lhs of the **B**-advance, also showing the divergence cleaning term. The electron skin depth is $d_e = c/\omega_e$.

- With $d_e = 5\times10^{-3}a$, the number of iterations decreases by 2-4 per time-step.
- With $d_e = 1\times10^{-2}a$, the number of iterations decreases by ~ 20%.
- With $d_e = 2\times10^{-2}a$, the number of iterations decreases by ~ 40%.

The changes with d_e implicate spatial resolution.

With a larger 12×12 mesh of bicubics, the base d_e=0 computation takes about 80 iterations to advance **B**.

- With $d_e=5\times10^{-3}a$, the number of iterations decreases by ~ 20%.
- With $d_e=1\times10^{-2}a$, the number of iterations decreases by ~ 40%.
- With $d_e=2\times10^{-2}a$, the number of iterations decreases by ~ 45%.
- Thus, the improvement occurs for smaller d_e with better spatial resolution.
- Of course, this changes the physical parameters of the problem.

If we just increase electron inertia in the preconditioner, leaving the matrix-vector product with d_e=0 (12×12 mesh):

- With $d_{e\text{-}pre}=5\times10^{-3}a$, the number of iterations decreases by ~ 1-2.
- With $d_{e\text{-}pre}=1\times10^{-2}a$, the number of iterations increases by ~ 3.
- With $d_{e\text{-}pre}=2\times10^{-2}a$, the number of iterations increase by > 40%.

From this result, see that just increasing the symmetric part of the preconditioner alone is not sufficient. Either:

- limiting the phase speed of the *R*-mode branch (or possible smoothing) helps, or
- the GMRES algorithm does not tolerate inaccuracy in $M \cong A$ for our matrices.

Another class of methods use polynomial approximation for preconditioning. [Again, see Saad's book.]

With ω as a scaling parameter and D a diagonal block of the original matrix A (for example, D is our set of poloidal-coupling submatrices),

$$\omega A = D - (D - \omega A) = D\left(I - \left(I - \omega D^{-1} A\right)\right)$$

$$(\omega A)^{-1} = \left(I - \left(I - \omega D^{-1} A\right)\right)^{-1} D^{-1}$$

If D itself is an approximate inverse, the norm of $N \equiv I - \omega D^{-1} A$ is small relative to the norm of I, so Taylor-expand (to a Neumann polynomial):

$$(\omega A)^{-1} \cong M^{-1} = \left(1 + N + N^2 + \ldots + N^{s-1}\right) D^{-1}$$

Relation to basic iteration with relaxation parameter ω:

Solve $Ax = r_0$ iterating from $x_0 = 0$:

$$\delta x_1 = \omega D^{-1} r_0$$

$$r_i = r_0 - A x_i = r_0 - A \sum_{j=1}^{i} \delta x_j = r_{i-1} - A \delta x_i \quad , \quad i \geq 1$$

$$\delta x_{i+1} = \omega D^{-1} r_i = \omega D^{-1}\left(r_{i-1} - A \omega D^{-1} r_{i-1}\right) = N \delta x_i$$

$$x_s = \left(I + N + N^2 + \ldots + N^{s-1}\right) \delta x_1 = \omega M^{-1} r_0$$

Applying this polynomial method for the 12×12 bicubic computation with $d_e=5\times10^{-3}$ reduces iterations in some cases.

s	ω	ave. its.	time
1	-	64	42
2	0.5	42	49
2	0.6	41	47
2	0.7	40.5	47
2	0.8	43.5	47
2	0.9	43.5	47
2	1	49.5	49
3	0.7	60	70
3	1	no conv.	-

- The matrix-vector product brings toroidal coupling into the preconditioner.
- Here, there is no savings in CPU time, because increasing s increases the number of matrix-vector products, like taking more GMRES iterations.
- If the number of GMRES iterations is very high, then polynomial approximation may help by reducing the amount work required for orthogonalization.

Possible improvements: toroidal preconditioning requires another approximate matrix that can be inverted relatively easily.

- Poloidal and toroidal operations can be applied sequentially

  $AM^{-1}v \rightarrow AM_{pol}^{-1}M_{tor}^{-1}v$ or $AM_{tor}^{-1}M_{pol}^{-1}v$ or additively.

- FFTs can be used to transform the operand and resultant vectors.

- Finite differences in toroidal angle may provide a useful approximation for the spectral derivatives, leading to tridiagonal or block-tridiagonal matrices over toroidal angle.

- NIMROD's layer parallelism and domain swapping will work.

- A module of utility functions and some simple operations based on 1D-in-ϕ-like equations have been started.

- Limited poloidal coupling may be incorporated if needed.

Conclusions

- Slowly converging GMRES iterations are not the result of a bug or poor-quality orthogonalization.
    - A “no preconditioning” option and separate residual check help verification.
- A small test case shows how iterations grow as coupling among Fourier components arises.
- Electron inertia helps reduce iteration for the implicit Hall advance by smoothing or limiting wave-speeds.
- Polynomial methods do not provide much benefit.

Future Work

- Test 1D tridiagonal systems as a basis for toroidal preconditioning.
- Implement limited poloidal coupling in the toroidal operation if needed.

Solver basics: the generalized minimal residual (GMRES) method finds an orthonormal bases for a Krylov space, starting from the initial residual, then finds the linear combination of bases that minimizes the residual.

- The required Krylov space K_m is $\left(r_0, Ar_0, AAr_0, \ldots, A^{m-1}r_0\right)$

  solving $Ax = b, \quad r_0 \equiv b - Ax_0 \Rightarrow x = x_0 + \delta x, \quad A\delta x = r_0$

- The Arnoldi algorithm generates the next basis vector (MGrSchm) as soon as the next Krylov vector is computed. [from Saad's book]

- For A of rank n, the first m+1 basis vectors define an $n \times m$+1 matrix V_{m+1} and the m+1$\times m$ Hessenberg matrix that satisfy

  $AV_m = V_{m+1}H_m$ typically $m << n$

- The best possible solution, $\delta x = V_m y_m$, is found from a relatively small least-squares problem.

$$\min\|r_0 - A\delta x\|_2 \Rightarrow \min\|r_0 - AV_m y\|_2 = \min\|\beta v_1 - V_{m+1}H_m y\|_2 = \min\|V_{m+1}(\beta e_1 - H_m y)\|_2$$

where $\beta \equiv \|r_0\|$ and Givens rotations provide the remaining error after each step.

- With right preconditioning, we solve $AM^{-1}u = r_0, \ \delta x = M^{-1}u, \ M \cong A$

For reference, we show the preconditioned Arnoldi algorithms with modified Gram-Schmidt orthogonalization. [Yousef Saad, Iterative Methods for Sparse Linear Systems, 2nd ed. SIAM.]

Choose v_1, $\|v_1\|_2 = 1$
for $j = 1, \ldots m$ do
  $w_j = M^{-1} A v_j$
  for $i = 1, \ldots j$ do (MGS)
    $h_{ij} = \left(w_j, v_i\right)$
    $w_j = w_j - h_{ij} v_i$
  end
  $h_{j+1,j} = \|w_j\|_2$
  $v_{j+1} = w_j / h_{j+1,j}$
end

to construct the basis set that spans

$$\left\{v_1, M^{-1}Av_1, \left(M^{-1}A\right)^2 v_1, \ldots\right\}$$

left preconditioning
[with $v_1 \sim r_0 = M^{-1}(b - Ax_0)$]

Choose v_1, $\|v_1\|_2 = 1$
for $j = 1, \ldots m$ do
  $w_j = AM^{-1} v_j$
  for $i = 1, \ldots j$ do (MGS)
    $h_{ij} = \left(w_j, v_i\right)$
    $w_j = w_j - h_{ij} v_i$
  end
  $h_{j+1,j} = \|w_j\|_2$
  $v_{j+1} = w_j / h_{j+1,j}$
end

to construct the basis set that spans

$$\left\{v_1, AM^{-1}v_1, \left(AM^{-1}\right)^2 v_1, \ldots\right\}$$

right preconditioning
[with $v_1 \sim r_0 = b - Ax_0$]

Recent *_store routines allow the coefficient centerings to be largely transparent at the calculation level.

- These operations save spatially interpolated data at numerical quadrature locations.
- **V** is only interpolated at the integer time levels.
- The spatially interpolated temporal average of n is saved after its advance to the next half-integer level. Its fully advanced value is interpolated at the start of the next time-step.
- The interpolated average of the temperature prediction is saved for the **B**-advance. Its fully advanced value (after a possible correction) is saved at the start of the next step.
- The interpolated average of **B** is used for anisotropic thermal conduction and Ohmic heating in temperature correction steps.
- A recent change moves viscous heating out of vrhs for centering. Some stress computations have also been moved to reuse code.