Stable and Scalable Spectral Element Methods

Paul Fischer
Mathematics and Computer Science Division
Argonne National Laboratory

Joint work with:
Fausto Cattaneo
James Lottes
Misun Min
Chaman Verma
Frank Loth UIC
Aleks Obabko U Chicago

and numerous others…
Overview

- High-order motivation: minimal dispersion/dissipation
- Efficiency – matrix-free factored forms
  - solvers: MG-preconditioned CG or GMRES
- Stability – high-order filters
  - dealiasing (i.e., “proper” integration)
- Scalability – long time integration
  - bounded iteration counts
  - scalable coarse-grid solvers (sparse-basis projection or AMG)
  - design for $P > 10^6$ ($P > 10^5$ already here...)
- Examples – vascular flows
  - MHD
  - Rod bundle flows
Navier-Stokes Time Advancement

\[
\frac{\partial u}{\partial t} + u \cdot \nabla u = -\nabla p + \nu \nabla^2 u \\
\nabla \cdot u = 0
\]

- **Nonlinear term**: *explicit*
  - \( k \) th-order backward difference formula / extrapolation
  - characteristics  \((\text{Pironneau '82, MPR '90})\)

- **Stokes problem – pressure/viscous decoupling**, \( P_N - P_{N-2} \) \((\text{Maday & Patera 89})\)
  - 3 Helmholtz solves for velocity – Jacobi-preconditioned CG
  - (consistent) Poisson equation for pressure  \((\text{computationally dominant})\)
Spatial Discretization: *Spectral Element Method*  

(Patera 84, Maday & Patera 89)

- Variational method, similar to FEM, using *GL* quadrature.

- Domain partitioned into *E* high-order quadrilateral (or hexahedral) elements  
  (decomposition may be nonconforming - *localized refinement*)

- Trial and test functions represented as *N*th-order tensor-product polynomials within each element. (*N* ~ 4 -- 15, typ.)

- *EN*³ gridpoints in 3D, *EN*² gridpoints in 2D.

- Converges *exponentially fast* with *N* for smooth solutions.

3D nonconforming mesh for arteriovenous graft simulations:  
*E* = 6168 elements, *N* = 7
Spectral Element Discretization

\[ u_t + c \cdot \nabla u = \nu \nabla^2 u \]

Find \( u \in X_0^N \subset H_0^1 \) such that

\[ (v, u_t)_N + (v, c \cdot \nabla u)_M = \nu (\nabla v, \nabla u)_N \quad \forall v \in X_0^N, \]

- \( (f, g)_M := \sum_{j=0}^{M} \rho_j^M f(\xi_j^M) g(\xi_j^M), \quad (1-D, \Omega = [-1, 1]) \)
- \( \xi_j^M, \rho_j^M \) — \( M \)th-order Gauss-Legendre points, weights.

2D basis function, \( N=10 \)
Accuracy
+
Costs
Spectral Element Convergence: Exponential with $N$

\[ \frac{||v - v_N||_{H^1}}{||v||_{H^1}} \]

$N$

\[ 10^{-14} \quad 10^{-13} \quad 10^{-12} \quad 10^{-11} \quad 10^{-10} \quad 10^{-9} \quad 10^{-8} \quad 10^{-7} \quad 10^{-6} \quad 10^{-5} \quad 10^{-4} \quad 10^{-3} \quad 10^{-2} \quad 10^{-1} \quad 10^{0} \]

\[ 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad 14 \]

Exact Navier-Stokes solution due to Kovazsnay(1948):

\[ v_x = 1 - e^{\lambda x} \cos 2\pi y \]

\[ v_y = \frac{\lambda}{2\pi} e^{\lambda x} \sin 2\pi y \]

\[ \lambda := \frac{Re}{2} - \sqrt{\frac{Re^2}{4} + 4\pi^2} \]

Mathematics and Computer Science Division, Argonne National Laboratory
Excellent transport properties, even for *non-smooth* solutions

Convection of non-smooth data on a 32x32 grid  ($K_1 \times K_1$ spectral elements of order $N$).

(cf. Gottlieb & Orszag 77)
Relative Phase Error for $h$ vs. $p$ Refinement: $u_t + u_x = 0$

- $X$-axis $= k / k_{max}$, $k_{max} := n / 2$ (Nyquist)
- Fraction of resolvable modes increased only through $p$-refinement
- Diagonal mass matrix (low $N$ significantly improved w/ full mass matrix)
- Polynomial approaches saturate at $k / k_{max} = 2 / \pi$
  \[ \Rightarrow N = 8-16 \sim \text{point of marginal return} \]
Costs

- Cost dominated by iterative solver costs, *proportional to*
  - iteration count
  - matrix-vector product + preconditioner cost

- Locally-structured tensor-product forms:
  - minimal indirect addressing
  - fast matrix-free operator evaluation
  - fast local operator inversion via fast diagonalization method (FDM) (Approximate, when element deformed.)
Matrix-Matrix Based Derivative Evaluation

- Local tensor-product form (2D),

\[ u(r, s) = \sum_{i=0}^{N} \sum_{j=0}^{N} u_{ij} h_i(r) h_j(s), \quad h_i(\xi_p) = \delta_{ip}, \; h_i \in \mathbb{P}_N \]

allows derivatives to be evaluated as matrix-matrix products:

\[ \frac{\partial u}{\partial r}_{|_{\xi_i, \xi_j}} = \sum_{p=0}^{N} u_{pj} \left. \frac{d h_p}{d r} \right|_{\xi_i} = \sum_{p} \tilde{D}_{ip} u_{pj} =: D_r u \]

\[ \text{mxm} \]
For a deformed spectral element, $\Omega^k$,

$$A^k u^k = \begin{pmatrix} D_r \\ D_s \\ D_t \end{pmatrix}^T \begin{pmatrix} G_{rr} & G_{rs} & G_{rt} \\ G_{sr} & G_{ss} & G_{st} \\ G_{tr} & G_{ts} & G_{tt} \end{pmatrix} \begin{pmatrix} D_r \\ D_s \\ D_t \end{pmatrix} u^k$$

$$D_r = (I \otimes I \otimes \hat{D}) \quad G_{rs} = J \circ B \circ \left( \frac{\partial r}{\partial x} \frac{\partial s}{\partial x} + \frac{\partial r}{\partial y} \frac{\partial s}{\partial y} + \frac{\partial r}{\partial z} \frac{\partial s}{\partial z} \right)$$

- Operation count is only $O(N^4)$ not $O(N^6)$ [Orszag ‘80]
- Memory access is 7 x number of points ($G_{rr}, G_{rs},$ etc., are diagonal)
- Work is dominated by matrix-matrix products involving $D_r, D_s,$ etc.
Summary: Computational Efficiency

- Error decays exponentially with $N$, typical $N \sim 5-15$

- For $n=EN^3$ gridpoints, require
  - $O(n)$ memory accesses
  - $O(nN)$ work in the form of matrix-matrix products

- Standard p-type implementation gives
  - $O(nN^3)$ memory accesses
  - $O(nN^3)$ work in the form of matrix-vector products

- Extensions to high-order tets:
  - Karniadakis & Sherwin (tensor-product quadrature)
  - Hesthaven & Warburton (geometry/canonical factorization: $D_r^T G^e D_r$)
  - Schoeberl et al. (orthogonal bases for linear operators)
Stability
Stabilizing High-Order Methods

In the absence of eddy viscosity, some type of stabilization is generally required at high Reynolds numbers.

Some options:

- high-order upwinding (e.g., DG, WENO)
- bubble functions
- spectrally vanishing viscosity
- filtering
- dealiasing
Filter-Based Stabilization

At end of each time step:
- Interpolate $u$ onto GLL points for $P_{N-1}$
- Interpolate back to GLL points for $P_N$

$$F_1(u) = I_{N-1} u$$

Results are smoother with linear combination:

$$F_\alpha(u) = (1-\alpha) u + \alpha I_{N-1} u$$

(\(\alpha \sim 0.05 - 0.2\))

Post-processing — no change to existing solvers

Preserves interelement continuity and spectral accuracy

Equivalent to multiplying by \((1-\alpha)\) the $N$th coefficient in the expansion

- $u(x) = \sum u_k \phi_k(x)$  \(\Rightarrow\)  $u^*(x) = \sum \sigma_k u_k \phi_k(x)$, $\sigma_k = 1$, $\sigma_N = (1-\alpha)$
- $\phi_k(x) := L_k(x) - L_{k-2}(x)$

(Gottlieb et al., Don et al., Vandeven, Boyd, ...)

Mathematics and Computer Science Division, Argonne National Laboratory
Numerical Stability Test: Shear Layer Roll-Up
(Bell et al. JCP 89, Brown & Minion, JCP 95, F. & Mullen, CRAS 2001)

Figure 1: Vorticity for different $(K, N)$ pairings: (a–d) $\rho = 30$, $Re = 10^5$, contours from -70 to 70 by 140/15; (e–f) $\rho = 100$, $Re = 40,000$, contours from -36 to 36 by 72/13. (cf. Fig. 3c in [4]).
Error in Predicted Growth Rate for Orr-Sommerfeld Problem at Re=7500

(Malik & Zang 84)

Spatial and Temporal Convergence

(FM, 2001)

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta t = 0.003125$</th>
<th>$\Delta t = 0.003125$</th>
<th>$\Delta t = 0.003125$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.0$</td>
<td>$\alpha = 0.2$</td>
<td>$\Delta t$</td>
</tr>
<tr>
<td>7</td>
<td>0.23641</td>
<td>0.27450</td>
<td>0.20000</td>
</tr>
<tr>
<td>9</td>
<td>0.00173</td>
<td>0.11929</td>
<td>0.01000</td>
</tr>
<tr>
<td>11</td>
<td>0.00455</td>
<td>0.01114</td>
<td>0.00500</td>
</tr>
<tr>
<td>13</td>
<td>0.00004</td>
<td>0.00074</td>
<td>0.00250</td>
</tr>
<tr>
<td>15</td>
<td>0.00010</td>
<td>0.00017</td>
<td>0.00125</td>
</tr>
</tbody>
</table>

2nd Order

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta t$</th>
<th>$\alpha = 0.0$</th>
<th>$\alpha = 0.2$</th>
<th>$\alpha = 0.0$</th>
<th>$\alpha = 0.2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>0.12621</td>
<td>0.12621</td>
<td></td>
<td>171.370</td>
<td>0.02066</td>
</tr>
<tr>
<td>9</td>
<td>0.03465</td>
<td>0.03465</td>
<td></td>
<td>0.00267</td>
<td>0.00268</td>
</tr>
<tr>
<td>11</td>
<td>0.00910</td>
<td>0.00911</td>
<td></td>
<td>161.134</td>
<td>0.00040</td>
</tr>
<tr>
<td>13</td>
<td>0.00238</td>
<td>0.00238</td>
<td></td>
<td>1.04463</td>
<td>0.00012</td>
</tr>
<tr>
<td>15</td>
<td>0.00065</td>
<td>0.00066</td>
<td></td>
<td>0.00008</td>
<td>0.00008</td>
</tr>
</tbody>
</table>

3rd Order

Base velocity profile and perturbation streamlines

Mathematics and Computer Science Division, Argonne National Laboratory
Filtering permits $Re_{99} > 700$ for transitional boundary layer calculations.

Figure 1: Principal vortex structures identified by $\lambda_2 = -1$ isosurfaces at $Re_9 = 700$: standing horseshoe vortex (a), interlaced tails (b), hairpin head (c), and bridge (d). Colors indicate pressure. ($K=1021$, $N=15$).

$Re = 700$

$Re = 1000$

$Re = 3500$
Double shear layer example:

High-strain regions are troublesome…
Why Does Filtering Work?
(Or, Why Do the Unfiltered Equations Fail?)

Consider the model problem:

\[
\frac{\partial u}{\partial t} = -c \cdot \nabla u
\]

Weighted residual formulation:

\[
B \frac{du}{dt} = -Cu
\]

\[
B_{ij} = \int_{\Omega} \phi_i \phi_j \, dV = \text{symm. pos. def.}
\]

\[
C_{ij} = \int_{\Omega} \phi_i c \cdot \nabla \phi_j \, dV
\]

\[
= -\int_{\Omega} \phi_j c \cdot \nabla \phi_i \, dV - \int_{\Omega} \phi_j \phi_j \nabla \cdot c \, dV
\]

\[
= \text{skew symmetric, if } \nabla \cdot c \equiv 0.
\]

\[
B^{-1}C \quad \text{Imaginary eigenvalues}
\]

Discrete problem should never blow up.
Why Does Filtering Work?
(Or, Why Do the Unfiltered Equations Fail?)

Weighted residual formulation vs. spectral element method:

\[ C_{ij} = (\phi_i, c \cdot \nabla \phi_j) = -C_{ji} \]

\[ \bar{C}_{ij} = (\phi_i, c \cdot \nabla \phi_j)_N \neq -\bar{C}_{ji} \]

This suggests the use of over-integration (dealiasing) to ensure that skew-symmetry is retained

\[ C_{ij} = (J\phi_i, (Jc) \cdot J\nabla \phi_j)_M \]

\[ J_{pq} := h^N_q(\xi^M_p) \quad \text{interpolation matrix (1D, single element)} \]

(Orszag '72, Kirby & Karniadakis '03, Kirby & Sherwin '06)
Aliased / Dealiased Eigenvalues: \( u_t + \mathbf{c} \cdot \nabla u = 0 \)

- Velocity fields model first-order terms in expansion of straining and rotating flows.
- For straining case, \( \frac{d}{dt}|u|^2 \sim |\tilde{u}_N|^2 - |\tilde{u}_N|^2 \)
- Rotational case is skew-symmetric.
- Filtering attacks the leading-order unstable mode.

\[ \mathbf{c} = (-x,y) \quad \mathbf{c} = (-y,x) \]
Stabilization Summary

- Filtering acts like well-tuned hyperviscosity
  - Attacks only the fine scale modes (that, numerically speaking, shouldn’t have energy anyway…)
  - Can precisely identify which modes in the SE expansion to suppress (unlike differential filters)
  - Does not compromise spectral convergence

- Dealiasing of convection operator recommended for high Reynolds number applications to avoid spurious eigenvalues
  - Can run double shear-layer roll-up problem forever with
    - $\nu = 0$,
    - no filtering
Dealiased Shear Layer Roll-Up Problem, $128^2$

- $\nu = 0$, no filter
- $\nu = 10^{-5}$, no filter
- $\nu = 0$, filter = (.1,.025)
Linear Solvers
Linear Solvers for Incompressible Navier-Stokes

- Navier-Stokes time advancement:

\[
\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = -\nabla P + \frac{1}{Re} \nabla^2 \vec{u}
\]

\[-\nabla . \vec{u} = 0\]

- Nonlinear term: explicit
  - \(k\) th-order backward difference formula / extrapolation
  - characteristics (Pironneau ’82, MPR ‘90)

- Stokes problem: pressure/viscous decoupling:
  - 3 Helmholtz solves for velocity ("easy" w/ Jacobi-precond. CG)
  - (consistent) Poisson equation for pressure (computationally dominant)
$P_N - P_{N-2}$ Spectral Element Method for Navier-Stokes (MP 89)

WRT: Find $u \in X^N, p \in Y^N$ such that:

$$
\frac{1}{Re} (\nabla u, \nabla v)_{GL} + \frac{1}{\Delta t} (u, v)_{GL} - (p, \nabla \cdot v)_G = (f, v)_{GL} \quad \forall \, v \in X^N \subset H^1
$$

$$
- (q, \nabla \cdot u)_G = 0 \quad \forall \, q \in Y^N \subset L^2
$$

Velocity, $u$ in $P_N$, continuous
Pressure, $p$ in $P_{N-2}$, discontinuous

Gauss-Lobatto Legendre points (velocity)

Gauss Legendre points (pressure)
Navier-Stokes Solution Strategy

- Leads to Stokes saddle problem, which is algebraically split

\[
\begin{bmatrix}
H & -D^T \\
-D & 0
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}^n \\
\mathbf{p}^n - \mathbf{p}^{n-1}
\end{bmatrix}
= 
\begin{bmatrix}
\mathbf{Bf} + D^T \mathbf{p}^{n-1} \\
\mathbf{f}_p
\end{bmatrix}
\]

\[
\begin{bmatrix}
H & -\frac{\Delta t}{\beta_0} \mathbf{H} \mathbf{B}^{-1} D^T \\
0 & E
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}^n \\
\mathbf{p}^n - \mathbf{p}^{n-1}
\end{bmatrix}
= 
\begin{bmatrix}
\mathbf{Bf} + D^T \mathbf{p}^{n-1} \\
\mathbf{g}
\end{bmatrix}
+ \begin{bmatrix}
\mathbf{r} \\
\mathbf{0}
\end{bmatrix}
\]

\[E := \frac{\Delta t}{\beta_0} \mathbf{DB}^{-1} D^T, \quad \mathbf{r} = O(\Delta t^2)\]

- \(E\) - consistent Poisson operator for pressure, SPD
  - Stiffest substep in Navier-Stokes time advancement
  - Most compute-intensive phase
  - Spectrally equivalent to SEM Laplacian, \(\mathbf{A}\)
Pressure Solution Strategy: $E p^n = g^n$

1. **Projection**: compute best approximation from previous time steps
   - Compute $p^*$ in span\{ $p^{n-1}$, $p^{n-2}$, …, $p^{n-l}$ \} through straightforward projection.
   - Typically a 2-fold savings in Navier-Stokes solution time.
   - Cost: 1 (or 2) matvecs in $E$ per timestep

2. **Preconditioned CG or GMRES to solve**
   
   \[ E \Delta p = g^n - E p^* \]
Two-Level Overlapping Additive Schwarz Preconditioner

\[ z = M r = \sum_{e=1}^{E} R_e^T A_e^{-1} R_e r + R_0^T A_0^{-1} R_0 r \]

**Local Overlapping Solves:** FEM-based Poisson problems with homogeneous Dirichlet boundary conditions, \( A_e \).

**Coarse Grid Solve:** Poisson problem using linear finite elements on entire spectral element mesh, \( A_0 \) (GLOBAL).
Solvers for Overlapping Schwarz / Multigrid

Local Solves: fast diagonalization method (Rice et al. ‘64, Couzy ‘95)

- $A^{-1} = (S \otimes S) (I \otimes \Lambda_x + \Lambda_y \otimes I)^{-1} (S \otimes S)^T$
- Complexity $< A P$
- For deformed case, approximate with nearest rectangular brick

Coarse Grid Solver: cast solution as projection onto $A_0$-conjugate basis (PF ‘96, Tufo & F ‘01)

- $x_0 = X_l X_l^T b_0$
- Matrix-vector products inherently parallel
- Here, choose basis $X_l = (x_1, x_2, \ldots, x_l)$ to be sparse.
- Use Gram-Schmidt to fill remainder of $X_l$ as $l \rightarrow n$
- Properly ordered, $X_n X_n^T = A_0^{-1}$ is a quasi-sparse factorization of $A_0^{-1}$
- Sublinear in $P$, minimal number of messages.

Mathematics and Computer Science Division, Argonne National Laboratory
Two-Level Schwarz Heuristics

- Local solves eliminate fine-scale error.

- Remaining error, due to Green’s functions from incorrect BCs on the local solves, is at scale $O(H)$, which is corrected by the coarse-grid solve.

- Additive preconditioning works in CG / GMRES contexts because eigenvalues of (preconditioned) fine and coarse modes are pushed towards unity.
Importance of weighting by $W$: Poisson eqn. example

- **Error after a single Schwarz smoothing step**
  - $M_{\text{Schwarz}}$
  - $\sigma M_{\text{Schwarz}}$
  - $W M_{\text{Schwarz}}$

- **Error after coarse-grid correction**
  - $\sigma (2.0)$
  - $W (0.8)$
  - $M (0.15)$

- **Weighting the additive-Schwarz step is essential to ensuring a smooth error**
  (Szyld has recent results)
$E$-Based Schwarz vs. SEMG for High-Aspect Ratio Elements

- Base mesh of $E=93$ elements
  - Quad refine to generate $E=372$ and $E=1488$ elements,
  - $N=4,\ldots,16$
  - SEMG reduces $E$ and $N$ dependence
  - 2.5 X reduction in Navier-Stokes CPU time for $N=16$

2D Navier-Stokes Model Problem

Overlapping Schwarz

Weighted Schwarz/MG

Mathematics and Computer Science Division, Argonne National Laboratory
Iteration Histories for 3D Unsteady Navier-Stokes \((n \sim 10^6)\)

- **Std.**  — 2-level additive Schwarz \(R_e^T A_e R_e\)
- **Mod.** — 2-level additive Schwarz, based on \(WR_e^T E_e R_e\)
- **Add.** — 3-level additive scheme
- **Hyb.** — 3-level multiplicative scheme
SEM Examples
Comparison of spectral element and measured velocity distributions in an arteriovenous graft, $Re_G=1200$

Coherent structures in arteriovenous graft @ $Re_G = 1200$

(Computations by S.W. Lee, UIC. Experiments by D. Smith, UIC)
Experimental data is low-pass (< 250 Hz) filtered to remove spurious fluctuations inherent in LDA measurements of regions of high shear flow.
Influence of Reynolds Number and Flow Division on $u_{rms}$
Study of turbulent magneto-rotational instabilities (w/ F. Cattaneo & A. Obabko, UC)

$E=97000, \ N=9 \ ( n = 71 \ M )$

$P=32768$

$\sim .8 \ sec/\text{step}$

$\sim 8 \ \text{iterations/step for U & B}$

Similar behavior for $n=112 \ M$
Numerical Magneto-Rotational Instabilities

w/ Fausto Cattaneo (ANL/UC) and Aleks Obabko (UC)

- SEM discretization of incompressible MHD (112 M gridpoints)
- Hydrodynamically stable Taylor-Couette flow

- Distributions of excess angular velocity at inner, mid, and outer radii
- Computations using 16K processors on BGW

- Simulation Predicts:
  - MRI
  - Sustained dynamo
MRI Angular Velocity Perturbation \( (v' = v - \langle v \rangle) \)

- **Axisymmetric**
- **3D**

Inner Wall

Outer Wall
Rod Bundle Flow at $Re=30,000$ w/ C. Tzanos ‘05

Low-speed streaks and log-law velocity profiles

$N = 9$  $N = 11$  $N = 15$

Mathematics and Computer Science Division, Argonne National Laboratory
Wire Wrapped Rod Bundles

- Uniformity of temperature controls peak power output
- A better understanding of flow distribution (interior, edge, corner) can lead to improved subchannel models.
- Wire wrap geometry is relatively complex
Single Rod in a Periodic Array, $Re=20,000$

$E=26000$, $N=7$, 8.7 M gridpoints
5 hours on $P=2048$ of 700 MHz BG/L for flow-through time.

Figure 3. Mean velocity distributions at $z = 0$ for (left) $Re = 14052$, $H/D=20.1$, $N=7$ and (right) $Re = 28104$, $H/D=13.4$, $N=9$. Only 1/4 of the vectors are shown.
7 Pin Mesh:

\[ E = 132,000, \quad N = 7 \]
\[ n_v \sim 44 \, M \]
\[ n_p \sim 28 \, M \]
\[ n_{\text{iter}} \sim 30 \, / \, \text{step} \]
7 Pin Configuration

Time-averaged axial (top) and transverse (bottom) velocity distributions.

Snapshot of axial velocity
High-order SEM formulation

- Stable formulation – dealiasing / filtering
  - Investigating relationship to SGS modeling
    - (e.g., RT model, Schlatter ’04, comparisons with D-Smagorinsky)

Scalable solvers

- Low iteration counts for typical “spectral-type” domains
- Iteration counts higher for very complex geometries
  - (e.g., multi-rod bundles) – work in progress
- We will need to switch to AMG for coarse-grid solve soon
  \[ E > 100,000; \ P > 10,000 \]

Future

- Significant need for scalable, conservative, design codes
  - Developing conservative DG variant