Paul Fischer Mathematics and Computer Science Division Argonne National Laboratory

Joint work with: Fausto Cattaneo James Lottes Misun Min Henry Tufo Chaman Verma

Frank LothUICAleks ObabkoU 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

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

□ Nonlinear term: *explicit*

 \square k th-order backward difference formula / extrapolation

□ characteristics (Pironneau '82, MPR '90)

Stokes problem – pressure/viscous decoupling, P_N – P_{N-2} (Maday & Patera 89)
 3 Helmholtz solves for velocity – Jacobi-preconditioned CG

(consistent) Poisson equation for pressure (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. ($N \sim 4 15$, typ.)
- \Box *EN*³ gridpoints in 3D, *EN*² gridpoints in 2D.
- □ Converges *exponentially fast* with *N* for smooth solutions.



Mathematics and Computer Science Division, Argonne National Laboratory

Spectral Element Discretization

$$u_t + \mathbf{c} \cdot \nabla u = \nu \nabla^2 u$$

Find
$$u \in X_0^N \subset H_0^1$$
 such that
 $(v, u_t)_N + (v, \mathbf{c} \cdot \nabla u)_M = \nu (\nabla v, \nabla u)_N \quad \forall v \in X_0^N,$
• $(f, g)_M := \sum_{j=0}^M \rho_j^M f(\xi_j^M) g(\xi_j^M), \quad (1-D, \Omega = [-1, 1])$

• ξ_j^M , ρ_j^M —*M*th-order Gauss-Legendre points, weights.



Accuracy + Costs

Spectral Element Convergence: Exponential with N



Excellent transport properties, even for *non-smooth* solutions



 $K_1 = 8, N = 4$

 $K_1 = 4, N = 8$

Convection of non-smooth data on a 32x32 grid (K₁ x K₁ spectral elements of order N).

(cf. Gottlieb & Orszag 77)

Relative Phase Error for *h* vs. *p* Refinement: $u_t + u_x = 0$



 $\square X-axis = k / k_{max}, k_{max} := n / 2 \quad (Nyquist)$

- 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 ~ 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.*)

□ 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:

$$\frac{\partial u}{\partial r}\Big|_{\xi_i,\xi_j} = \sum_{p=0}^N u_{pj} \frac{dh_p}{dr}\Big|_{\xi_i} = \sum_p \widehat{D}_{ip} u_{pj} =: D_r \underline{u}$$

 \Box For a deformed spectral element, Ω^k ,

$$\begin{aligned} A^{k}\underline{u}^{k} &= \begin{pmatrix} D_{r} \\ D_{s} \\ D_{t} \end{pmatrix}^{T} \begin{pmatrix} G_{rr} & G_{rs} & G_{rt} \\ G_{rs} & G_{ss} & G_{st} \\ G_{rt} & G_{st} & G_{tt} \end{pmatrix} \begin{pmatrix} D_{r} \\ D_{s} \\ D_{t} \end{pmatrix} \underline{u}^{k} \end{aligned}$$
$$\begin{aligned} D_{r} &= (I \otimes I \otimes \hat{D}) \qquad G_{r} &= J \otimes B \otimes (\frac{\partial r}{\partial s} + \frac{\partial r}{\partial s} + \frac{\partial r}{\partial s}) \end{aligned}$$

$$D_r = \langle I \otimes I \otimes D \rangle$$
 $G_{rs} = J \cup D \cup \langle \partial x \partial x \top \partial y \partial y \top \partial z \partial z \rangle$

Operation count is only O (N⁴) not O (N⁶) [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.

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

For n=EN³ gridpoints, require
 O(n) memory accesses

 \bigcirc *O*(*nN*) work in the form of matrix-matrix products

Standard p-type implementation gives

- \Box $O(nN^3)$ memory accesses
- \Box $O(nN^3)$ 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

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

At end of each time step:

- □ Interpolate *u* onto GLL points for P_{N-1}
- \Box Interpolate back to GLL points for P_N

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



□ Results are smoother with linear combination: (F. & Mullen 01) $F_{\alpha}(u) = (1-\alpha) u + \alpha I_{N-1} u$ ($\alpha \sim 0.05 - 0.2$)

- Post-processing no change to existing solvers
- Preserves interelement continuity and spectral accuracy

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

 $u(x) = \sum u_k \phi_k(x) \rightarrow u^*(x) = \sum \sigma_k u_k \phi_k(x), \quad \sigma_k = 1, \quad \sigma_N = (1 - \alpha)$ $\phi_k(x) := L_k(x) - L_{k-2}(x)$ (Boyd 98)

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)	
	$\Delta t = 0.003125$		N = 17	2nd Order		3rd Order	
N	lpha=0.0	lpha=0.2	Δt	lpha=0.0	lpha=0.2	lpha=0.0	lpha=0.2
7	0.23641	0.27450	0.20000	0.12621	0.12621	171.370	0.02066
9	0.00173	0.11929	0.10000	0.03465	0.03465	0.00267	0.00268
11	0.00455	0.01114	0.05000	0.00910	0.00911	161.134	0.00040
13	0.00004	0.00074	0.02500	0.00238	0.00238	1.04463	0.00012
15	0.00010	0.00017	0.01250	0.00065	0.00066	0.00008	0.00008



Base velocity profile and perturbation streamlines

Filtering permits Re_{899} > 700 for transitional boundary layer calculations



Mathematics and Computer Science Division, Argonne National Laboratory

Why Does Filtering Work ? (Or, Why Do the Unfiltered Equations Fail?)

Double shear layer example:



Why Does Filtering Work ? (Or, Why Do the Unfiltered Equations Fail?)

 \sim

Consider the model problem:
$$\frac{\partial u}{\partial t} = -\mathbf{c} \cdot \nabla u$$

Weighted residual formulation: $B\frac{d\underline{u}}{dt} = -C\underline{u}$

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

$$C_{ij} = \int_{\Omega} \phi_i \mathbf{c} \cdot \nabla \phi_j \, dV$$
$$= -\int_{\Omega} \phi_j \mathbf{c} \cdot \nabla \phi_i \, dV - \int_{\Omega} \phi_j \phi_j \nabla \cdot \mathbf{c} \, dV$$

= skew symmetric, if $\nabla \cdot \mathbf{c} \equiv 0$.

 $\longrightarrow B^{-1}C$ = skew symmetric

Discrete problem should never blow up.

Why Does Filtering Work ? (Or, Why Do the Unfiltered Equations Fail?)

Weighted residual formulation vs. spectral element method:

$$C_{ij} = (\phi_i, \mathbf{c} \cdot \nabla \phi_j) = -C_{ji}$$
$$\tilde{C}_{ij} = (\phi_i, \mathbf{c} \cdot \nabla \phi_j)_N \neq -\tilde{C}_{ji}$$

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

$$C_{ij} = (J\phi_i, (J\mathbf{c}) \cdot J\nabla\phi_j)_M$$

 $J_{pq} := h_q^N(\xi_p^M)$ 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, $\frac{d}{dt}|u|^2 \sim |\hat{u}_{N}|^2 |\hat{u}_N|^2$
- Rotational case is skew-symmetric.
- □ Filtering attacks the leading-order unstable mode.



N=19, M=19

N=19, M=20

Mathematics and Computer Science Division, Argonne National Laboratory

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

v = 0, no filter

$v = 10^{-5}$, no filter

v = 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
 - \Box 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
$$\mathbf{u} \in X^N$$
, $p \in Y^N$ such that:

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

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

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



Gauss-Lobatto Legendre points (velocity)



Semi-implicit: explicit treatment of nonlinear term.
 Leads to Stokes saddle problem, which is algebraically split

MPR 90, Blair-Perot 93, Couzy 95

$$\begin{bmatrix} \mathbf{H} & -\mathbf{D}^T \\ -\mathbf{D} & 0 \end{bmatrix} \begin{pmatrix} \underline{\mathbf{u}}^n \\ \underline{p}^n - \underline{p}^{n-1} \end{pmatrix} = \begin{pmatrix} \mathbf{B}\underline{\mathbf{f}} + \mathbf{D}^T \underline{p}^{n-1} \\ \underline{f}_p \end{pmatrix}$$

$$\begin{bmatrix} \mathbf{H} & -\frac{\Delta t}{\beta_0} \mathbf{H} \mathbf{B}^{-1} \mathbf{D}^T \end{bmatrix} \begin{pmatrix} \mathbf{\underline{u}}^n \\ E \end{bmatrix} \begin{pmatrix} \mathbf{\underline{p}}^n - \mathbf{\underline{p}}^{n-1} \end{pmatrix} - \begin{pmatrix} \mathbf{B} \mathbf{\underline{f}} + \mathbf{D}^T \mathbf{\underline{p}}^{n-1} \\ \underline{g} \end{bmatrix} + \begin{pmatrix} \mathbf{\underline{r}} \\ \underline{0} \end{pmatrix} + \begin{pmatrix} \mathbf{\underline{r}} \\ \underline{0} \end{pmatrix} + ,$$

$$E := \frac{\Delta t}{\beta_0} \mathbf{D} \mathbf{B}^{-1} \mathbf{D}^T , \qquad \mathbf{\underline{r}} = O(\Delta t^2)$$

- **E** consistent Poisson operator for pressure, SPD
 - □ Stiffest substep in Navier-Stokes time advancement
 - Most compute-intensive phase
 - \square Spectrally equivalent to SEM Laplacian, A

- 1. Projection: compute best approximation from previous time steps
 - □ Compute \underline{p}^* in span{ \underline{p}^{n-1} , \underline{p}^{n-2} , ..., \underline{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. Schwarz or multigrid preconditioned CG or GMRES to solve $E \Delta \underline{p} = \underline{g}^n - E \underline{p}^*$

Initial guess for $E\underline{p}^n = \underline{g}^n$ via projection onto previous solutions

(F '93, '98)

 $\square \| \underline{p}^n - \underline{p}^* \|_E = O(\Delta t^l) + O(\varepsilon_{\text{tol}})$

Results with/without projection (1.6 million pressure nodes):



• 4 fold reduction in iteration count, 2-4 in typical applications

Similar results for pulsatile carotid artery simulations – 10⁸-fold reduction in initial residual

- 1. Projection: compute best approximation from previous time steps
 - **Compute** \underline{p}^* in span{ \underline{p}^{n-1} , \underline{p}^{n-2} , ..., \underline{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. Schwarz or multigrid preconditioned CG or GMRES to solve $E \Delta \underline{p} = \underline{g}^n - E \underline{p}^*$

Two-Level Overlapping Additive Schwarz Preconditioner

(Dryja & Widlund 87, Pahl 93, PF 97, FMT 00)



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*)

$$A_e^{-1} = (S \otimes S) (I \otimes \Lambda_x + \Lambda_y \otimes I)^{-1} (S \otimes S)^T$$

- □ Complexity < A <u>p</u>
- For deformed case, approximate with nearest rectangular brick



Coarse Grid Solver: cast solution as projection onto A₀-conjugate basis (PF '96, Tufo & F '01)

- $\Box \quad \underline{x}_0 = X_l X_l^T \underline{b}_0$
- Matrix-vector products inherently parallel
- □ Here, choose basis $X_l = (\underline{x}_1, \underline{x}_2, ..., \underline{x}_l)$ to be *sparse*.
- □ Use Gram-Schmidt to fill remainder of X_l as $l \rightarrow n$
- **D** 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.
- Good up to $P \sim 10^4$, $E/P \sim 1$. Must be revisited for $P > 10^5$.



Coarse Grid Solver Timings: 127² Poisson Problem on ASCI Red

(Tufo & F 01)



Coarse-Grid Solve Times

□ 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.

Natural sequence of nested grids — intra-element multigrid

No problem with restriction / prolongation (variational MG)

Difficulty is selection of *smoother*, *M*







E=9, N=8

E=9, N=4

E=9, N=2

High aspect-ratio cells resulting from tensor-product of GLL grids degrades performance of pointwise smoothers (Zang et al. 82, Ronquist & Patera '87, Heinrichs '88, Shen et al. 00,...)

Use overlapping Schwarz (w/ FDM solver...)
 eliminates the local-cell aspect ratio problem

Can also be cured by FEM on SEM grid + AMG (CU group...)



Importance of weighting by W: Poisson eqn. example

Error after a single Schwarz smoothing step



Error after coarse-grid correction



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

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



2D Navier-Stokes Model Problem



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



Scalability of these algorithms is well-understood and validated through numerous examples

■ Bottom line: need $n / P \sim 10^3 - 10^4$ ■ For P=10⁵ → $n \sim 10^9$

SEM Examples

Transition in Vascular Flows

w/ F. Loth, UIC

Comparison of spectral element and measured velocity distributions in an arteriovenous graft, Re_G =1200

Mean Axial Velocity

RMS for Re 1200, 70:30 flow division

Influence of Reynolds Number and Flow Division on u_{rms}

Low-speed streaks and log-law velocity profiles

7 Pin Configuration

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

Snapshot of axial velocity

SEMG Scalability: Incompressible MHD

- Study of turbulent magneto-rotational instabilities (w/ F. Cattaneo & A. Obabko, UC)
- \Box E=97000, N=9 (n = 71 M)
- **□** *P*=*32768*
- **□** ~ .8 sec/step
- $\square \sim 8$ iterations / step for U & B
- Similar behavior for n=112 M

Mathematics and Computer Science Division, Argonne National Laboratory

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 ($v' = v - \langle v \rangle$)

Summary / Future Effort

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 James Lottes talk

Future

❑ Significant need for scalable, conservative, design codes
→ Developing conservative DG variant