Discontinuous Galerkin and Finite Difference Methods for the Acoustic Equations with Smooth Coefficients

Mario Bencomo

TRIP '14 Annual Review Meeting

May 1, 2015

RICE

Mario Javier Bencomo

- Born in Mexico ... raised in El Paso TX
- B.S. in Physics and Applied Math (UTEP 2010)
- M.A. in Computational and Applied Math (Rice 2015)
- Ph.D in Computational and Applied Math (Rice ?)

Outline

- M.A. work: DG vs FD
- Ph.D. work: joint source and model inversion

Motivation of Study

Why DG?

- viable numerical method for forward modeling (discontinuous media)
- outperforms FD methods when using mesh aligning techniques for complex discontinuous media *(Wang 2010)*

Why smooth media?

- smooth trends in bulk modulus and density are observed in real data
- relevant for seismic imaging, i.e., the inverse problem

Comparison between FD and DG in smooth media has not been done before ... as far as we are aware.

Disclaimer

Limited comparison

- DG code is serial and in Matlab
- FD code is serial and in IWAVE (implemented in C)

What kind of comparison?

- counting FLOPs for a prescribed accuracy
- benefits to this type of comparison (hardware independent, and limits to FLOP rates)

Acoustic Equations (pressure-velocity form):

$$\rho(\boldsymbol{x})\frac{\partial \boldsymbol{v}}{\partial t}(\boldsymbol{x},t) + \nabla p(\boldsymbol{x},t) = 0 \tag{1a}$$

$$\beta(\boldsymbol{x})\frac{\partial p}{\partial t}(\boldsymbol{x},t) + \nabla \cdot \boldsymbol{v}(\boldsymbol{x},t) = f(\boldsymbol{x},t) \tag{1b}$$

for $\boldsymbol{x} = [x,y]^T \in \Omega$ and $t \in [0,T]$,

- $p =$ pressure
- $\boldsymbol{v} = [v_x, v_y]^T =$ velocity fields
- $\rho =$ density
- $\beta =$ compressibility $= 1/\kappa$
- $\boldsymbol{f}(\boldsymbol{x},t) =$ source term

Considering homogeneous boundary and initial conditions.

FD Methods: Staggered Grid FD

2-2k staggered FD method applied to 2D acoustic wave equation in first order form:

$$(v_x)^{n+1}_{i+\frac{1}{2}\,j} = (v_x)^{n}_{i+\frac{1}{2}\,j} + \Delta t \frac{1}{(\rho)_{i+\frac{1}{2}\,j}} \left\{ -D^{h,(k)}_x (p)^{n+\frac{1}{2}}_{i+\frac{1}{2}\,j} \right\}$$

$$(v_y)^{n+1}_{i\,j+\frac{1}{2}} = (v_y)^{n}_{i\,j+\frac{1}{2}} + \Delta t \frac{1}{(\rho)_{i\,j+\frac{1}{2}}} \left\{ -D^{h,(k)}_y (p)^{n+\frac{1}{2}}_{i\,j+\frac{1}{2}} \right\}$$

$$(p)^{n+\frac{1}{2}}_{ij} = (p)^{n-\frac{1}{2}}_{ij} + \Delta t \frac{1}{(\beta)_{ij}} \left\{ -D^{h,(k)}_x (v_x)^n_{ij} - D^{h,(k)}_y (v_y)^n_{ij} + (f)^n_{ij} \right\},$$

where $p^{n+\frac{1}{2}}_{ij} = p(ih, jh, (n+\frac{1}{2})\Delta t)$, and

$$D^{h,(k)}_x f(x_0) := \frac{1}{h} \sum_{n=1}^{k} a^{(k)}_n \left\{ f\left(x_0 + \left(n-\tfrac{1}{2}\right)h\right) - f\left(x_0 - \left(n-\tfrac{1}{2}\right)h\right) \right\}.$$

FD Methods: Staggered Grid FD

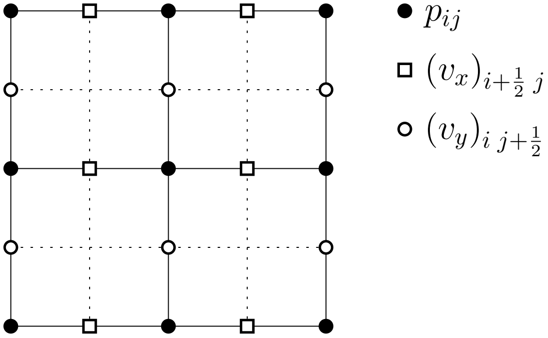


Figure 1: Staggered grid points for 2D acoustics.

DG Methods: Semi-Discrete Scheme

Define:

- $\mathscr{T}_h =$ triangulation/mesh
- $\mathscr{W}_h =$ approximation space (piecewise polynomial)
- $\{\ell_i^{(\tau)}\}_{i=1}^{N^*} =$ local basis functions on triangle $\tau \in \mathscr{T}_h$, where $N^* := \frac{1}{2}(N+1)(N+2)$ for polynomial order N (Lagrange polynomials)

From PDE to **strong formulation**: find $p, v_x, v_y \in \mathscr{W}_h$ such that

$$\int_\tau \rho \frac{\partial v_x}{\partial t} w \, d\boldsymbol{x} + \int_\tau \frac{\partial p}{\partial x} w \, d\boldsymbol{x} + \int_{\partial\tau} \hat{n}_x (p^* - p) w \, d\sigma = 0$$

$$\vdots$$

for all $w \in \mathscr{W}_h$ and all $\tau \in \mathscr{T}_h$.

Numerical flux p^*: provides numerical stability and transmits information between elements (upwind flux)

DG Methods: Semi-Discrete Scheme

After introducing basis functions, solve for coefficients $\mathbf{v}_x^{(\tau)}, \mathbf{v}_y^{(\tau)}, \mathbf{p}^{(\tau)}$
$\Longrightarrow$ **Semi-discrete scheme**:

$$M[\rho]\frac{d}{dt}\mathbf{v}_x^{(\tau)}(t) + S^x\mathbf{p}^{(\tau)}(t) + \sum_{e\in\partial\tau}\hat{n}_x M^{(e)}\left((\mathbf{p}^{(e)})^* - \mathbf{p}^{(e)}\right)(t) = 0,$$

$$\vdots$$

for each $\tau \in \mathscr{T}_h$.

DG operators:

weighted mass matrix	$M[\omega]_{ij}$	$:=$	$\int_\tau \omega \ell_i^{(\tau)}\ell_j^{(\tau)}\,d\boldsymbol{x}$,	in $\mathbb{R}^{N^*\times N^*}$
edge mass matrix	$M^{(e)}{}_{ij}$	$:=$	$\int_e \ell_i^{(\tau)}\ell_j^{(e)}\,d\sigma$,	in $\mathbb{R}^{N^*\times(N+1)}$
α-stiffness matrix	$S^\alpha{}_{ij}$	$:=$	$\int_\tau \ell_i^{(\tau)}\frac{\partial \ell_j^{(\tau)}}{\partial\alpha}\,d\boldsymbol{x}$,	in $\mathbb{R}^{N^*\times N^*}$

for $\omega \in \{\rho,\beta\}$ and $\alpha \in \{x,y\}$.

Numerical Experiments

- 2-2 and 2-4 FD staggered grid schemes; implemented in C, IWAVE *(Symes et al., 2009)*

- RK-DG with $N = 2, 4$; implemented in Matlab *(Hesthaven & Warburton, 2007)*
  - considered upwind flux
  - considered quadrature-free and quadrature-based implementations
  - considered mesh refinement for lower velocity zones
  - triangular meshes

- Numerical results were compared to a highly discretized 2-4 FD solution ($h = 0.5m, dt = 0.0442ms$)

- Comparing FLOP count for achieving prescribed accuracy ($RMS < 5\%, \max < 6\%$)

Numerical Experiments: Defining Error

Relative error:

$$E_h(\boldsymbol{x}_r) = \frac{\|p_h(\boldsymbol{x}_r, \cdot) - p(\boldsymbol{x}_r, \cdot)\|}{\|p(\boldsymbol{x}_r, \cdot)\|},$$

with p is a high fidelity solution (2-4 FD with $h_x = h_y = 0.5m$), where

$$\|p(\boldsymbol{x}_r, \cdot)\| = \left(\sum_i |p(\boldsymbol{x}_r, t_i)|^2 \right)^{\frac{1}{2}}$$

Accuracy conditions:

$$RMS\ E_h(\boldsymbol{x}_r) < 5\%$$

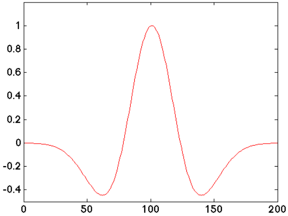
$$\max\ E_h(\boldsymbol{x}_r) < 6\%$$

For all simulations:

- source term $f(\boldsymbol{x}, t) = \chi(\boldsymbol{x})\Psi(t)$, where

$$\begin{aligned} \Psi(t) &= \Psi(t; t_c, f_{peak}) &= \text{Ricker wavelet} \\ \chi(\boldsymbol{x}) &= \chi(\boldsymbol{x}; \boldsymbol{x}_c, d_x) &= \text{cosine bump function} \end{aligned}$$

  with $f_{peak} = 10\ Hz$ and $d_x = [50\ m, 50\ m]$
- density is assumed to be constant, $\rho = 2.3\ g/cm^3$



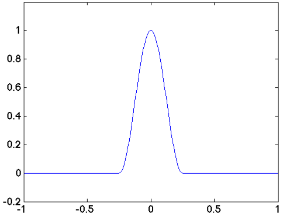


Figure 2: ($\Leftarrow$) sample Ψ; ($\Rightarrow$) sample χ.

Numerical Experiments: Negative-Lens Velocity Model

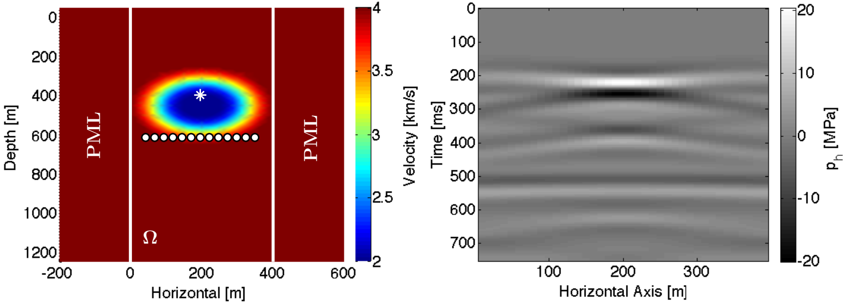


Figure 3: (⇐) Velocity model; (⇒) traces of **p**

Results: Negative-Lens Velocity Model

- discretization parameters (dt, h) tuned to satisfy accuracy conditions $(RMS < 5\%, max < 6\%)$
- GPW $= c_{min}/(f_{peak} h)$ [FD] or $N \times c_{min}/(f_{peak} h)$ [DG]

	$dt[ms]$	$h[m]$	GPW	GFLOPs
FD 2-2	0.838	6	33.33	0.6296
FD 2-4	1.565	15	13.33	0.0820
no mesh ref.				
DG N=2, Q-free	1.003	40	10	19.72
DG N=2, with Q	0.963	60	6.66	7.72
DG N=4, Q-free	0.655	50	16	99.92
DG N=4, with Q	1.199	80	10	19.99
mesh ref.				
DG N=2, Q-free	0.983	80:40	10	7.44
DG N=2, with Q	0.852	100:50	8	3.61
DG N=4, Q-free	0.655	100:50	16	32.19
DG N=4, with Q	1.205	150:75	10.66	8.29

Table 1: Results for negative-lens test case.

Numerical Experiments: Mixed Model

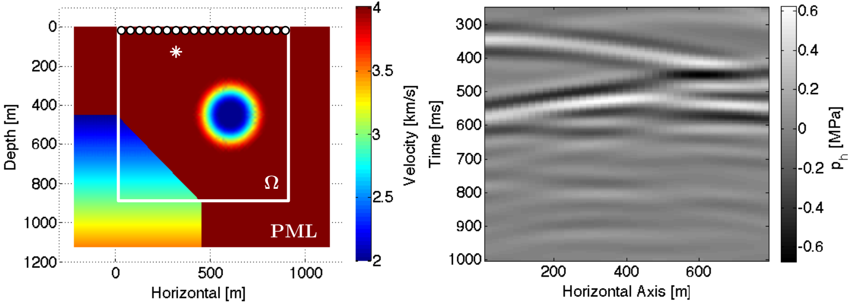


Figure 4: (⇐) velocity model; (⇒) traces of **p**

Results: Mixed Model

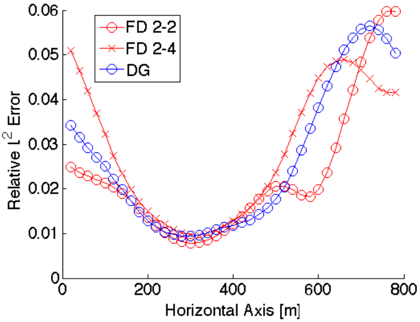


Figure 5: Relative errors for mixed velocity model.

	$dt[ms]$	$h[m]$	GPW	GFLOPs
FD 2-2	0.742	6	33.33	1.4308
FD 2-4	1.130	8	25	0.7793
DG	1.038	112.5:56.25	14.22	25.68

Table 2: Results for mixed test case.

Results

	hom.	linear	lens	mixed
GFLOP (DG/FD)	119	76	44	33

Table 3: Approximate GFLOP ratios between best of DG over FD, for each test case.

- smaller FLOP counts for quadrature vs quadrature-free DG
- overall FD methods yield smaller FLOP counts than DG, at the least by a factor of 33 for the mixed model test case

Conclusions

Overview:

- Goal of thesis is to compare DG and FD in the context of 2D acoustics, with smooth coefficients.
- Incorporated methodology for dealing with variable media (quadrature vs quadrature-free DG and mesh refinement).
- Limited comparison due to implementations of numerical methods (DG in Matlab and FD in C).
- Comparison is done by looking at FLOP counts.

Conclusions

On FLOP count ...

- $20\% \sim 30\%$ peak machine performance[1] can be achieved for FD methods, via vectorization and cache optimization *(Zhou 2014)*
  $\Longrightarrow$ GFLOP count is a crude metric for computation time

$$T_{FD} = \frac{GFLOPs}{0.2 * GLFOPs/sec}$$

$$\Longrightarrow T_{DG}/T_{FD} = \frac{33 * GFLOPs}{\varepsilon * GFLOPs/sec} \Big/ \frac{GFLOPs}{0.2 * GFLOPs/sec}$$

$$= 33\frac{0.2}{\varepsilon} \geq 6.6$$

[1] Sandy Bridge Xeon E5-2660 processor

Further Remarks

On accuracy condition ....

What if you want higher accuracy?

- recall, FD schemes were $O(\Delta t^2)$ while RK-DG was $O(\Delta t^4)$
  $\Longrightarrow$ FD will not scale as well as RK-DG
- increase the time discretization (Lax-Wendroff schemes)
  $\Longrightarrow$ expect increase in FLOP count for new FD methods
  $\Longrightarrow$ How will FD compare to RK-DG?

Future Work (Ph.D.)

Idea: Joint source-model inversion, for anisotropic sources, via variable projection.

- source estimation and representation
  - an accurate estimation of source wavelet is crucial for the reconstruction of impedance profiles *(Delprat & Lailly 2005)*
  - anisotropy is real!
    $p - \tau$ data set from Gulf of Mexico *(Minkoff & Symes 1997)*
- variable projection (VP) method *(Golub & Pereyra 1973)*
  - reduces dimensionality of problem while perserving global minimizer
  - better conditioned problem in most instances *(Ruhe & Wedin 1980)*
  - outperforms alternating direction and simultaneous descent *(Rickett 2013)*

Future Work (Ph.D.)

Source representation: multipole-point-source approximation
(Santosa & Symes 2000)

$$f_j(\boldsymbol{\eta}, t) = \sum_{n=0}^{N} (-1)^n F^{(n)}_{j;k_1\cdots k_n}(t) \frac{\partial}{\partial \eta_{k_1}} \cdots \frac{\partial}{\partial \eta_{k_n}} \delta(\boldsymbol{\eta} - \boldsymbol{\eta}^*)$$

$$\begin{aligned} \Longrightarrow u_i(\boldsymbol{x}, t) &= \int dV(\boldsymbol{\eta})\, f_j(\boldsymbol{\eta}, t) * G_{ij}(\boldsymbol{x}, t; \boldsymbol{\eta}) \\ &= \sum_{n=0}^{N} F^{(n)}_{j;k_1\cdots k_n}(t) * G_{ij,k_1\cdots k_n}(\boldsymbol{x}, t; \boldsymbol{\eta}^*) \end{aligned}$$

where

$$G_{ij,k_1\cdots k_n}(\boldsymbol{x}, t; \boldsymbol{\eta}^*) := \frac{\partial}{\partial \eta_{k_1}} \cdots \frac{\partial}{\partial \eta_{k_n}} G_{ij}(\boldsymbol{x}, t; \boldsymbol{\eta}) \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^*}$$

and $\mathbf{F}^{(n)}$ is the n^{th} degree *force moment tensor*, related to the seismic moment tensor from earthquake source representation.

Future Work (Ph.D.)

source paramters $\mathbf{f}$ (i.e., $\mathbf{F}^{(n)}$), model parameters m

OLS Formulation: minimize $J_{OLS}[\mathbf{f}, m]$,

$$
\begin{aligned}
J_{OLS}[\mathbf{f}, m] &:= \frac{1}{2}\sum_r \sum_k \left| u_{i_r}(\boldsymbol{x}_r, \omega_k) - d(\boldsymbol{x}_r, \omega_k) \right|^2 \\
&= \frac{1}{2}\sum_r \sum_k \left| \sum_{n=0}^{N} F^{(n)}_{j;k_1\cdots k_n}(\omega_k) G_{i_r j, k_1 \cdots k_n}(\boldsymbol{x}_r, \omega_k; \boldsymbol{\eta}^*) - d(\boldsymbol{x}_r, \omega_k) \right|^2 \\
&= \frac{1}{2}\left\| \mathbf{G}[m]\mathbf{f} - \mathbf{d} \right\|^2
\end{aligned}
$$

VP Formulation: minimize $J_{VP}[m]$,

$$J_{VP}[m] := J_{OLS}[\mathbf{f}(m), m],$$

where

$$\mathbf{f}(m) := \operatorname*{argmin}_{\mathbf{f}} J_{OLS}[\mathbf{f}, m].$$

Future Work (Ph.D.)

Questions:

- How difficult is the joint inversion problem, via VP method, in comparison to the non-reduced problem and the ideal case where source is known?

  The key is in the Hessian? ...

- Can source parameters be determined? Uniquely? Stably? What data do I need?

Acknowledgements

- TRIP sponsors
- NSF support, Graduate Research Fellowship (grant no. 1450681)