Discontinous Galerkin Method for 1D Advection Equation

February 27, 2020

1 Weak Formulation

The problem that we want to solve is:

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} - bu = 0 \qquad x \in [-1, 1]$$
$$u(0, t) = g(t)$$
$$u(x, 0) = f(x)$$

We want to find u_h that is an approximation of u, and $u_h = \bigoplus_{k=1}^{K} u_h^k$, where k indexes elements. If we define the spaces V and V_f , V_f^k:

$$V = \{v(x) : \|v\|^2 < \infty, \langle v, v' \rangle < \infty\}$$
$$V_f = \{v(x) : v \in V, v \text{ is a linear combination of basis functions}\}$$
$$V_f^k = \{v(x) : v \in V_f, x \in D^k\}$$

Where D^k is the domain of element k. We can now consider the error between $u_h^k \in V_f^k$ and u^k. We get:

$$u_h^k - u^k = \left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k\right) - \left(\frac{\partial u^k}{\partial t} + a\frac{\partial u^k}{\partial x} - bu^k\right) =$$
$$\left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k\right) - 0 = \frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k$$

Similar to the Continious Galerkin (CG) Method we multiply the error by a test function and integrate, except now in Discontinious Galerkin (DG) the integral is over each element rather than the entire physical domain so that we get:

$$\int_{D_k} \left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k \right) v dx = 0 \qquad \forall v \in V_f^k, k \in K \tag{1}$$

Where K is the set of all elements on our domain. Also note that due to the definition of V this equations makes sense (the integral is finite), and that while u is both spatially dependent and time dependent, v is only a function of space.

What is the best way to think about Eq. 1? Eq. 1 is saying that the error of what will be our solution is orthogonal to every function in V_f^k. From linear algebra we know that if $\left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k \right) \neq \vec{0}$, and it is orthogonal to every vector in V_f^k then $\left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k \right) \notin V_f$.

The analog in $\mathbb{R}^n$ goes like this:
Let e_i be the standard basis in $\mathbb{R}^n$, and $\vec{u} \in \mathbb{R}^n$ such that $\vec{u} \cdot \vec{v} = 0, \forall v \in \mathbb{R}^n$. Then we can write:

$$\vec{u} \cdot \vec{v} = (a_1 e_1) \cdot (b_1 e_1) + (a_2 e_2) \cdot (b_2 e_2) + ... + (a_n e_n) \cdot (b_n e_n) = 0 \qquad \forall b_i \in \mathbb{R}$$

The only way we can zero out each term is if $a_i = 0$ otherwise there will be a nonzero b_i leading to a nonzero term. $a_i = 0$ then implies that $\vec{u} = 0$.

In other words Eq. 1 is "saying" the error on a single element is not in our finite dimensional space V_f^k, or that the error in V_f^k must be 0.

2 System of Linear equations

Further confining ourselves to the space:

$$V_0^k = \{v : v \in V_f^k, v \text{ is linear}\}$$

the problem becomes find $u_h \in V_0^k$ such that:

$$\int_{D_k} \left(\frac{\partial u_h^k}{\partial t} + a\frac{\partial u_h^k}{\partial x} - bu_h^k \right) v dx = 0 \qquad \forall v \in V_0, k \in K \tag{2}$$

We can write (2) using a basis for v. Let ϕ_1 and ϕ_2 be lagragian basis functions on any element k, where:

$$\phi_1(x) = -\frac{x_r^k - x}{h}$$
$$\phi_2(x) = \frac{x - x_l^k}{h}$$

Where $h = x_r^k - x_l^k$, and x_r^k is the x value of the right side of the element, and x_l^k is the left side. Then (2) is equivalent to:

$$\int_{D_k} \left(\frac{\partial u_h^k}{\partial t} + a \frac{\partial u_h^k}{\partial x} - b u_h^k \right) \phi_i dx = 0 \qquad i = 1, 2, k \in K \tag{3}$$

To see that (2) if and only if (3), first both equations for each element in (3) are just cases of (2) where if we write v as a linear combintation of basis functions $v = v_1\phi_1 + v_2\phi_2$, we have $v_1 = 0$ for one case and $v_2 = 0$ for the other. On the other hand if (3) is true then multiplying through by a constant, and adding the two equations of (3) we recover (2).

Next, since $u_h \in V_0$ we can also write it as a combination of the basis functions on a single element so that:

$$u_h^k = \sum_{j=1}^{2} \bar{u}_j^k(x_j^k, t)\phi_j(x)$$

Where $\bar{u}$ is a coordinate. It is important here to realize what we will be solving for in the long run. The basis functions are only functions of space like the test functions were. The coordinates of the basis functions are functions of time though, and at each timestep we want to solve for a new set of coordinates. Secondly the notation $\bar{u}_j^k(x_j^k, t)$ is not great, x_j^k does not mean that $\bar{u}_j^k$ is a function of space (their locations do not very in time), but only that this coordinate has an associated point on x, because the basis functions are interpolating polynomials. Plugging in this representation of u_h^k to (3) we get:

$$\sum_{j=1}^{2} \frac{\partial \bar{u}_j^k}{\partial t} \int_{D_k} \phi_j \phi_i dx + a \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j' \phi_i dx - b \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j \phi_i dx = 0 \qquad i = 1, 2, k \in K$$

At this point we could make a systems of equations for each element, but there would be no connection between any of the elements, and no way for boundary data to enter the system. Doing integration by parts on the second term solves this problem, giving:

$$\sum_{j=1}^{2} \frac{\partial \bar{u}_j^k}{\partial t} \int_{D_k} \phi_j \phi_i dx + a f^* \phi_i \Big|_{x_l^k}^{x_r^k} - a \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j \phi_i' dx - b \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j \phi_i dx = 0 \qquad i = 1, 2, k \in K$$

Where instead of taking the boundaries of the element, we replace them with fluxes f^*. This is one of the really centeral pieces to DG. The DG user then defines what their flux is with their problem in mind (with convergance in mind too). We then integrate by parts again, to get SAT or penalty parameter terms for the boundaries of each element, so that we have:

$$\sum_{j=1}^{2} \frac{\partial \bar{u}_j^k}{\partial t} \int_{D_k} \phi_j \phi_i dx + a f^* \phi_i \Big|_{x_l^k}^{x_r^k} - a u_h^k \phi_i \Big|_{x_l^k}^{x_r^k} + a \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j' \phi_i dx - b \sum_{j=1}^{2} \bar{u}_j^k \int_{D_k} \phi_j \phi_i dx = 0 \qquad i = 1, 2, k \in K$$

Writing this in matrix form we get:

$$M^k \frac{\partial \vec{u}_h^k}{\partial t} + a S^k \vec{u}_h^k - b M^k \vec{u}_h^k = (a u_h^k(x_r^k) - a f_r^*) \vec{\phi}(x_r^k) - (a u_h^k(x_l^k) - a f_l^*) \vec{\phi}(x_l^k) \qquad \forall k \in K \tag{4}$$

$$M^k = \langle \phi_j, \phi_i \rangle \qquad S^k = \langle \phi_j', \phi_i \rangle \tag{5}$$

The only thing left to figure out to have a linear system of equations is what the fluxes can and should be. For each adjacent element we have that $x_r^{k-1} = x_l^k$ and $x_r^k = x_l^{k+1}$ so that on each element boundary u_h is multiply defined. So, the solution vector has length $2|K|$. It therefore seems reasonable to make each flux/penalty parameter be equal to some weighted average of values defined on element $k-1$ and element k for the left boundary, and k and $k+1$ for the right boundary, so that:

$$f_l^* = \alpha u_h^{k-1}(x_r^{k-1}) + (1-\alpha) u_h^k(x_l^k)$$
$$f_r^* = \alpha u_h^k(x_r^k) + (1-\alpha) u_h^{k+1}(x_l^k)$$

Where $0 \leq \alpha \leq 1$. To include the physical boundary condtions, on the left side of the first element we can make $f_l^* = g(t)$, and at the right boundary we have no $k+1$ element so $f_r^* = u_h^k(x_r^k)$.

3 Stability for Fluxes

To find the stability conditions for the flux we can use the energy method. Similar to SBP-SAT we want S^k to mimic summation by parts and M^k to mimic the energy. Before showing this though the energy over the whole domain without any estimates is:

$$\frac{\partial}{\partial t} \|u\|^2 = \frac{\partial}{\partial t} \int_{-1}^{1} u^2 dx = \int_{-1}^{1} 2u \frac{\partial u}{\partial t} dx = \int_{-1}^{1} 2u \left(bu - a \frac{\partial u}{\partial x} \right) dx = 2b \int_{-1}^{1} u^2 dx - 2a \int_{-1}^{1} u \frac{\partial u}{\partial x} dx$$
$$= 2b \int_{-1}^{1} u^2 dx - a \int_{-1}^{1} u \frac{\partial u}{\partial x} dx + a \int_{-1}^{1} u \frac{\partial u}{\partial x} dx - a u^2 \Big|_{-1}^{1} = 2b \|u\|^2 + a(u^2(-1) - u^2(1))$$

To mimic this with M^k, and S^k we have:

$$(\vec{u}_h^k)^T M^k \vec{u}_h^k = \begin{bmatrix}\bar{u}_1^k & \bar{u}_2^k\end{bmatrix} \begin{bmatrix}\int_{D^k} \phi_1\phi_1 & \int_{D^k} \phi_1\phi_2 \\ \int_{D^k} \phi_2\phi_1 & \int_{D^k} \phi_2\phi_2\end{bmatrix} \begin{bmatrix}\bar{u}_1^k \\ \bar{u}_2^k\end{bmatrix}$$
$$= \begin{bmatrix}\int_{D_k} \sum_{j=1}^2 \bar{u}_1^k\phi_j\phi_1 & \int_{D_k} \sum_{j=1}^2 \bar{u}_2^k\phi_j\phi_2\end{bmatrix} \begin{bmatrix}\bar{u}_1^k \\ \bar{u}_2^k\end{bmatrix} = \left\|u_h^k\right\|^2$$

$$(\vec{u}_h^k)^T S^k \vec{u}_h^k = \begin{bmatrix}\bar{u}_1^k & \bar{u}_2^k\end{bmatrix} \begin{bmatrix}\int_{D^k} \phi_1\phi_1' & \int_{D^k} \phi_1\phi_2' \\ \int_{D^k} \phi_2\phi_1' & \int_{D^k} \phi_2\phi_2'\end{bmatrix} \begin{bmatrix}\bar{u}_1^k \\ \bar{u}_2^k\end{bmatrix}$$
$$\begin{bmatrix}\int_{D_k} \sum_{j=1}^2 \bar{u}_1^k\phi_j\phi_1' & \int_{D_k} \sum_{j=1}^2 \bar{u}_2^k\phi_j\phi_2'\end{bmatrix} \begin{bmatrix}\bar{u}_1^k \\ \bar{u}_2^k\end{bmatrix} = \int_{D_k} u_h^k \frac{\partial u_h^k}{\partial x} dx = \frac{1}{2}(u_h^k)^2\Big|_{x_l^k}^{x_r^k}$$

If we multiple (4) by $(\vec{u}_h^k)^T$, and move all of the terms to the right then we can get an energy estimate for the discrete system:

$$\frac{\partial}{\partial t}\left\|u_h^k\right\|^2 = b\left\|u_h^k\right\|^2 - a\frac{1}{2}(u_h^k)^2\Big|_{x_l^k}^{x_r^k} + \vec{u}_h^k\vec{\phi}(x_r^k)(au_h^k(x_r^k) - af_r^*) - \vec{u}_h^k\vec{\phi}(x_l^k)(au_h^k(x_l^k) - af_l^*)$$
$$= 2b\left\|u_h^k\right\|^2 - a(u_h^k)^2\Big|_{x_l^k}^{x_r^k} + 2u_h^k(x_r^k)(au_h^k(x_r^k) - af_r^*) - 2u_h^k(x_l^k)(au_h^k(x_l^k) - af_l^*)$$
$$= 2b\left\|u_h^k\right\|^2 + (a(u_h^k)^2(x_r^k) - 2u_h^k(x_r^k)af_r^*) - (a(u_h^k)^2(x_l^k) - 2u_h^k(x_l^k)af_l^*)$$

For our scheme to be stable we need $\frac{\partial}{\partial t}\left\|u_h^k\right\|^2 \leq 0$ over the whole domain. We can ignore the expotential growth term (since unless $b \leq 0$ we know this will grow our solution). It is confusing to try to figure out global stability though in terms of local solutions. If we plug in our fluxes here, we will have terms involving u^{k-1} and u^{k+1}. A way around this is to shift our persective from each element to each boundary between elements. If we define $u^- = u_r^{k-1}$ as the value on the boundary of the left element, and $u^+ = u_l^k$ as the value on the boundary of the right element. Then we need on every boundary:

$$a(u^-)^2 - 2(u^-)af^{-*} - a(u^+)^2 + 2(u^+)af^{+*} \leq 0$$

Plugging in the fluxes we get:

$$a(u^-)^2 - 2u^-a(\alpha u^- + (1-\alpha)u^+) - a(u^+)^2 + 2u^+a(\alpha u^- + (1-\alpha)u^+) \leq 0$$
$$(a - 2a\alpha)(u^-)^2 - 4\alpha u^+u^- + 2u^+u^- + (2a(1-\alpha) - a)(u^+)^2 \leq 0$$

Which is only stable for $\frac{1}{2} \leq \alpha \leq 1$

4 Local and Global Matrices

The local scheme (rewriting equation 4 by dividing my M^{-1}) looks like this:

$$\frac{\partial \vec{u}_h^k}{\partial t} = (-aM^{-1}S^k)\vec{u}_h^k + b\vec{u}_h^k + M^{-1}((au_h^k(x_r^k) - f_r^*)\vec{\phi}(x_r^k) - (au_h^k(x_l^k) - f_l^*)\vec{\phi}(x_l^k)) \qquad \forall k \in K \tag{6}$$

Assuming we can integrate all the combintations of basis function that we need on a reference element then scale them to each element, we don't need anymore matrices. Since:

$$\vec{\phi}(x_l^k) = \begin{bmatrix} \phi_1(x_l^k) \\ \phi_2(x_l^k) \end{bmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and the reverse for $\vec{\phi}(x_r^k)$, along with the expressions for the flux (6) becomes:

$$\frac{\partial \vec{u}_h^k}{\partial t} = (-aM^{-1}S^k + bI)\vec{u}_h^k + M^{-1}F\vec{u}_h^{k\prime} \qquad \forall k \in K$$

$$F = \begin{bmatrix} a\alpha & -a\alpha & 0 & 0 \\ 0 & 0 & a(1-\alpha) & -a(1-\alpha) \end{bmatrix}$$

$$\vec{u}_h^{k\prime} = \begin{bmatrix} u_r^{k-1} \\ u_l^k \\ u_r^k \\ u_l^{k+1} \end{bmatrix}$$

Constructing the global matrix, the RHS becomes a diagonal single banded matrix, with a vector for the boundary condition:

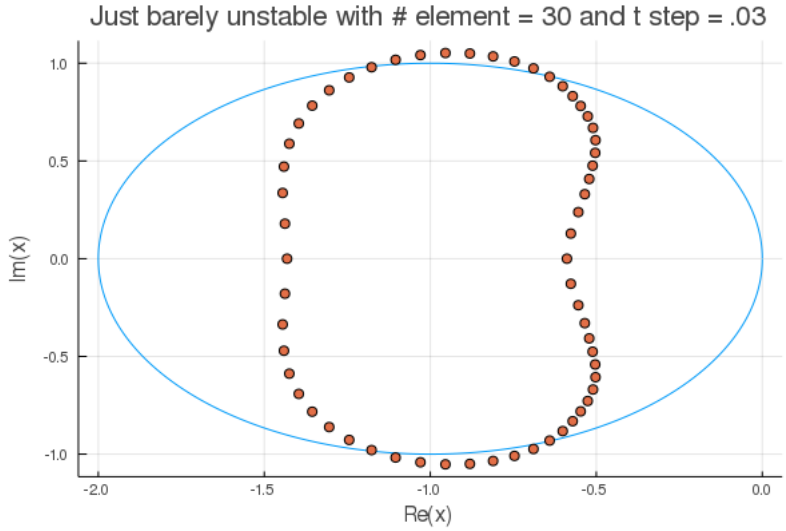
$$\frac{\partial \vec{u}_h^k}{\partial t} = G\vec{u}_h + b \qquad b = \begin{bmatrix} aM_{11}g(t) \\ aM_{21}g(t) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

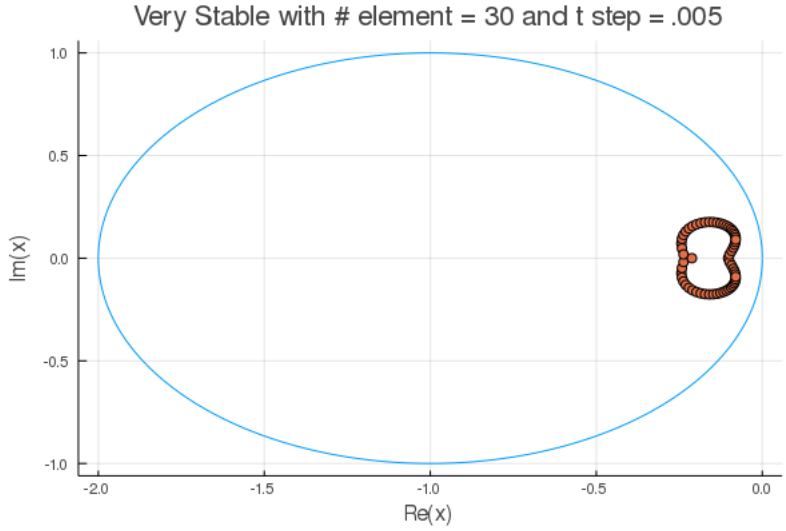
5 Time Stability

Using Euler to discretize in time, we can derive the region of absolute stability by inserting the test equation $\frac{\partial \vec{u}}{\partial t} = \lambda I \vec{u}$ into the foward Euler method:

$$\vec{u_n} = \vec{u}_{n-1} + k\frac{\partial \vec{u}}{\partial t} = \vec{u}_{n-1} + k\lambda I \vec{u}_{n-1} = (\vec{1} + k\lambda I)\vec{u}_{n-1} = (\vec{1} + k\lambda I)^n u_0$$

For stability we need $(\vec{1} + k\lambda I) < 1$ or in our case $|\vec{1} + k$ eigenvalues(G)$| < 1$ where the magnitude is elementwise.
Here are a few plots of the Eigenvalues:





6 Convergence Tests

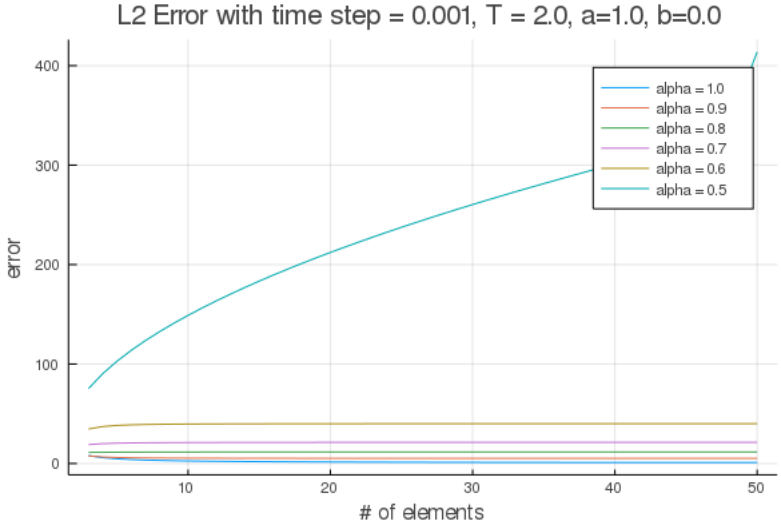
To test for convergence I found the analytic solution to the problem with:

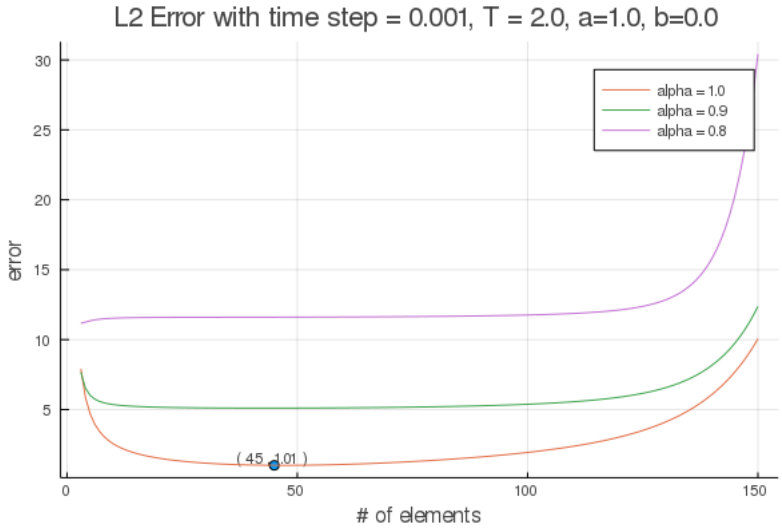
$$u(0,t) = 0$$
$$u(x,0) = \cos\left(\frac{\pi x}{2}\right)$$
$$u(x,t) = \cos\left(\frac{\pi(x-at)}{2}\right)e^{bt}$$

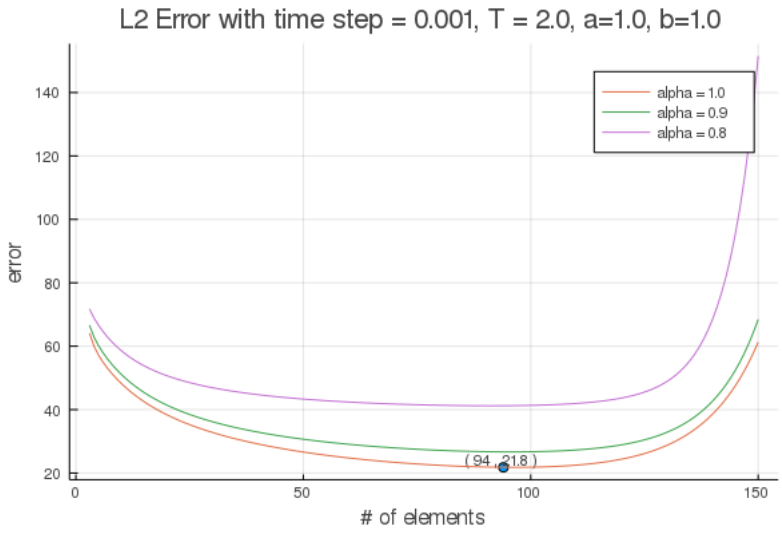
Since each end point of an element is multiply defined, I took the L2 error at each element boundary twice, once for the left element and once for the right. Or as an equation:

$$\text{error} = \sum_{t=0}^{T}\sum_{i=0}^{N}(u_{il}^{t} - u(ih,tk))^2 + (u_{ir}^{t} - u(ih,tk))^2$$

Here are three plots (where the divergence with many elements should come from the CFL condition):



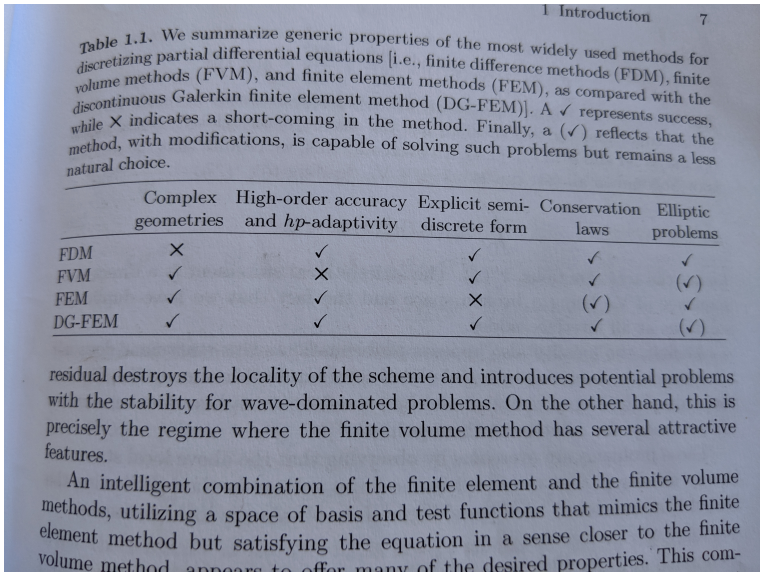

L2 Error with time step = 0.001, T = 2.0, a=1.0, b=0.0
alpha = 1.0
alpha = 0.9
alpha = 0.8
error
# of elements
(45 , 1.01)
30
25
20
15
10
5
0
50
100
150


L2 Error with time step = 0.001, T = 2.0, a=1.0, b=1.0
alpha = 1.0
alpha = 0.9
alpha = 0.8
error
# of elements
(94 , 21.8)
140
120
100
80
60
40
20
0
50
100
150

7 Properties of DG

Here are a few properties of and thoughts on DG that make it potentially interesting:

- Like CG, higher order accuracy can be achieved by increasing number of elements, or by increasing the order of basis functions, but we have the freedom to refine acccuracy element wise with the order of basis functions.
- Capable of modeling physical discontinuities.
- Parellelizable
- Can be formulated with only reference to local elements.
- Why is this better than CG?



1 Introduction 7

Table 1.1. We summarize generic properties of the most widely used methods for discretizing partial differential equations [i.e., finite difference methods (FDM), finite volume methods (FVM), and finite element methods (FEM), as compared with the discontinuous Galerkin finite element method (DG-FEM)]. A ✓ represents success, while × indicates a short-coming in the method. Finally, a (✓) reflects that the method, with modifications, is capable of solving such problems but remains a less natural choice.

	Complex geometries	High-order accuracy and *hp*-adaptivity	Explicit semi-discrete form	Conservation laws	Elliptic problems
FDM	×	✓	✓	✓	✓
FVM	✓	×	✓	✓	(✓)
FEM	✓	✓	×	(✓)	✓
DG-FEM	✓	✓	✓	✓	(✓)

residual destroys the locality of the scheme and introduces potential problems with the stability for wave-dominated problems. On the other hand, this is precisely the regime where the finite volume method has several attractive features.

An intelligent combination of the finite element and the finite volume methods, utilizing a space of basis and test functions that mimics the finite element method but satisfying the equation in a sense closer to the finite volume method appears to offer many of the desired properties. This com-