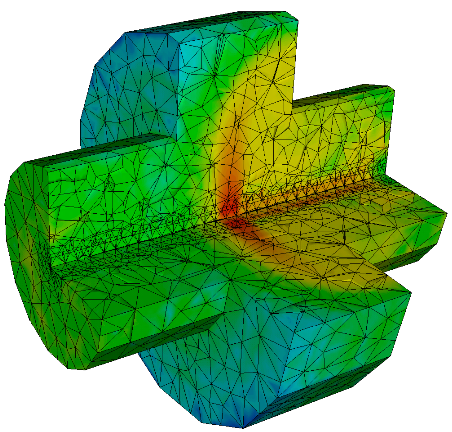
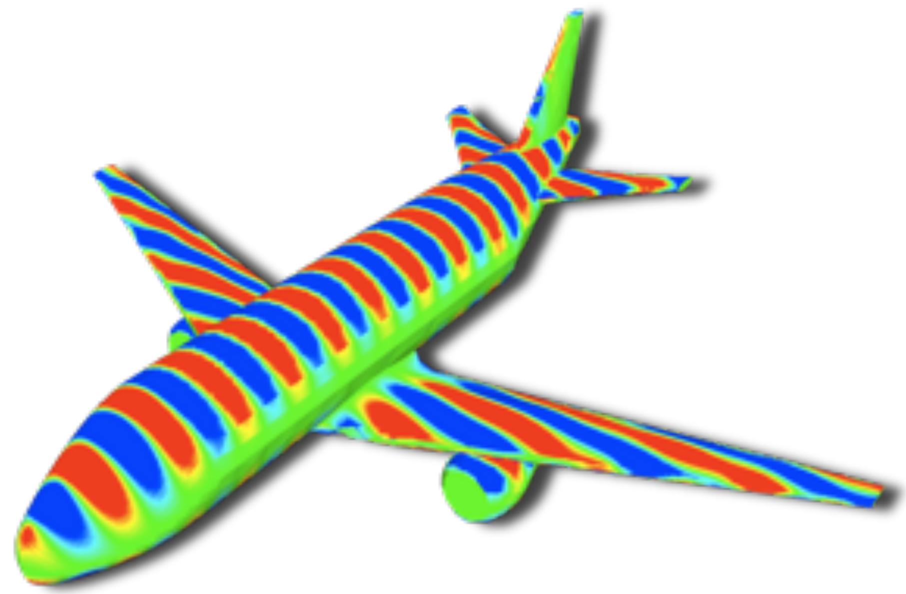
DG-FEM for PDE's
Lecture 3

Jan S Hesthaven
Brown University
Jan.Hesthaven@Brown.edu





A brief overview of what's to come

- Lecture 1: Introduction and DG-FEM in 1D
- Lecture 2: Implementation and numerical aspects
- Lecture 3: Insight through theory
- Lecture 4: Nonlinear problems
- Lecture 5: Extension to two spatial dimensions
- Lecture 6: Introduction to mesh generation
- Lecture 7: Higher order/Global problems
- Lecture 8: 3D and advanced topics

Lecture 3

- ✓ Let's briefly recall what we know
- ✓ Why high order methods ?
- ✓ Part I:
  - ✓ Constructing fluxes for linear systems
  - ✓ Approximation theory on the interval
- ✓ Part II:
  - ✓ Convergence and error estimates
  - ✓ Dispersive properties
  - ✓ Discrete stability and how to overcome

Let us recall

We already know a lot about the basic DG-FEM

- Stability is provided by carefully choosing the numerical flux.
- Accuracy appear to be given by the local solution representation.
- We can utilize major advances on monotone schemes to design fluxes.
- The scheme generalizes with very few changes to very general problems -- multidimensional systems of conservation laws.

Let us recall

We already know a lot about the basic DG-FEM

- Stability is provided by carefully choosing the numerical flux.
- Accuracy appear to be given by the local solution representation.
- We can utilize major advances on monotone schemes to design fluxes.
- The scheme generalizes with very few changes to very general problems -- multidimensional systems of conservation laws.

At least in principle -- but what can we actually prove ?

Why high-order accuracy ?

Let us just make sure we understand why high-order accuracy/methods is a good idea

General concerns/criticism:

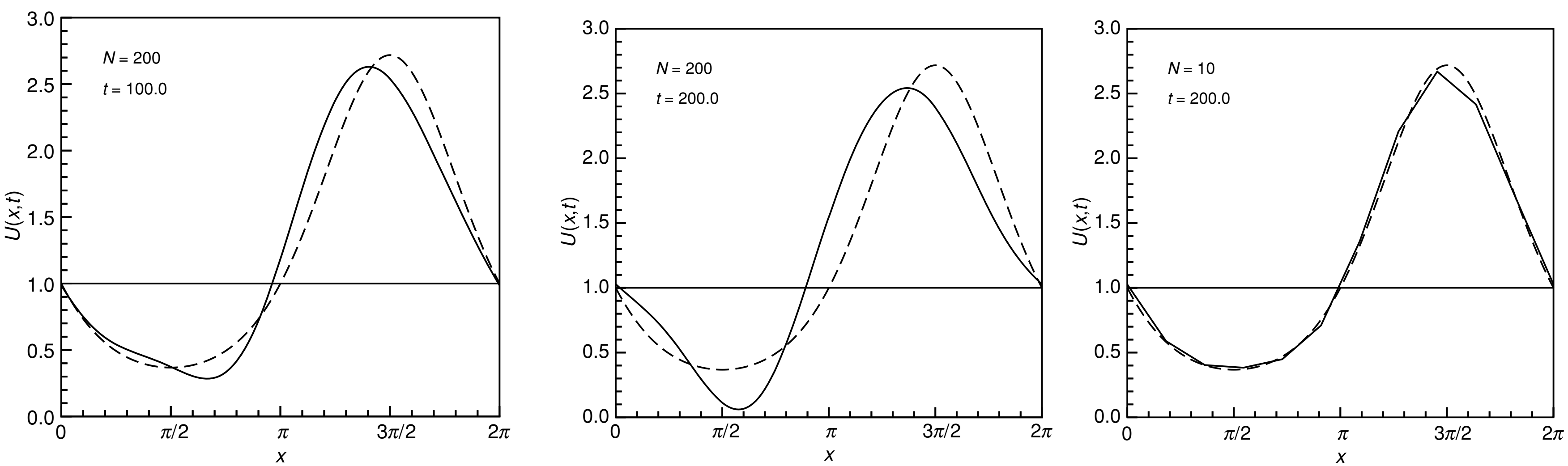
- High-order accuracy is not needed for real appl.
- The methods are not robust/flexible
- They only work for smooth problems
- They are hard to do in complex geometries
- They are too expensive

After having worked on these methods for 15 years, I have heard them all

Why high-order accuracy ?

How do I solve a wave-problem to a given accuracy,
, for a specific period of time, , most efficiently ?

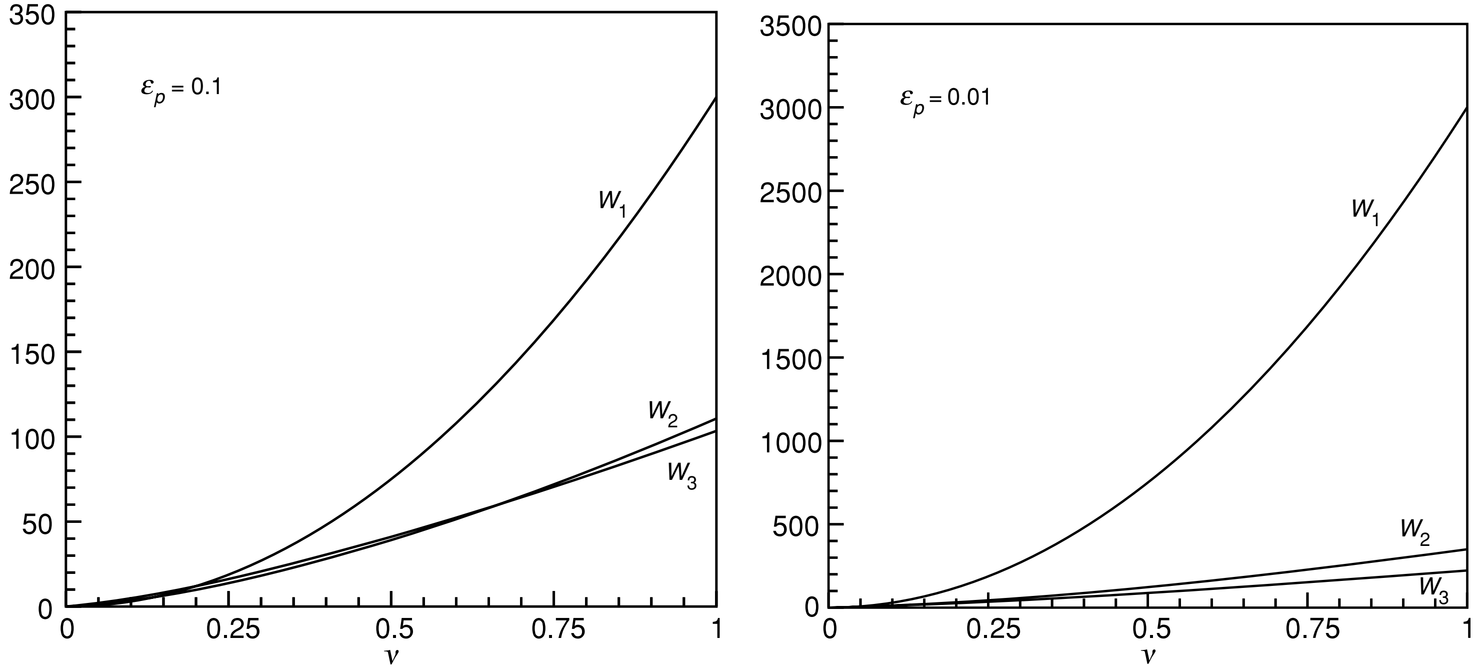
$$\text{Memory} \propto \left(\frac{\nu}{\varepsilon_p}\right)^{\frac{d}{2m}} \quad , \quad \text{Work} \propto (2m)^d \nu \left(\frac{\nu}{\varepsilon_p}\right)^{\frac{d+1}{2m}}$$



2nd order FD

Infinite order FD

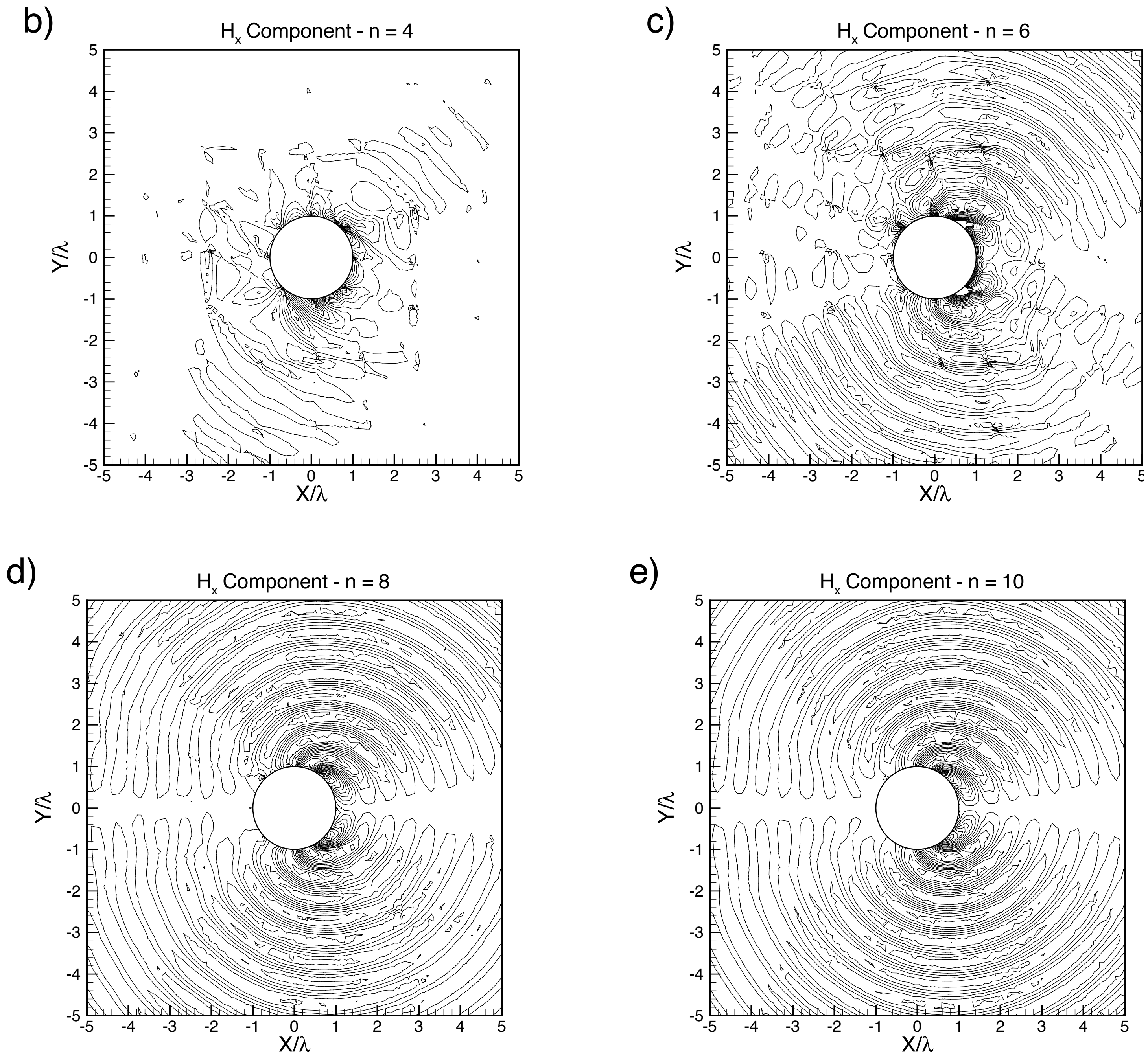
Why high-order accuracy ?



High-order is important if

- High accuracy is required - and it increasingly is !
- Long time integration is needed
- High-dimensional problems (3D) are considered
- Memory restrictions become a bottleneck

Added benefit of high-order support



High-order takes 'some' of the pain out of grid generation

But first a bit more on fluxes

Let us briefly look a little more carefully at linear systems

$$\mathcal{Q}(\boldsymbol{x})\frac{\partial \boldsymbol{u}}{\partial t} + \nabla \cdot \mathcal{F} = \mathcal{Q}(\boldsymbol{x})\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial \boldsymbol{F}_1}{\partial x} + \frac{\partial \boldsymbol{F}_2}{\partial y} = 0,$$

$$\mathcal{F} = [\boldsymbol{F}_1, \boldsymbol{F}_2] = [\mathcal{A}_1(\boldsymbol{x})\boldsymbol{u}, \mathcal{A}_2(\boldsymbol{x})\boldsymbol{u}] \,.$$

Prominent examples are

- Acoustics
- Electromagnetics
- Elasticity

In such cases we can derive *exact upwind* fluxes

Linear systems and fluxes

Assume first that all coefficients vary smoothly

$$\mathcal{Q}(\boldsymbol{x})\frac{\partial \boldsymbol{u}}{\partial t} + \mathcal{A}_1(\boldsymbol{x})\frac{\partial \boldsymbol{u}}{\partial x} + \mathcal{A}_2(\boldsymbol{x})\frac{\partial \boldsymbol{u}}{\partial y} + \mathcal{B}(\boldsymbol{x})\boldsymbol{u} = 0,$$

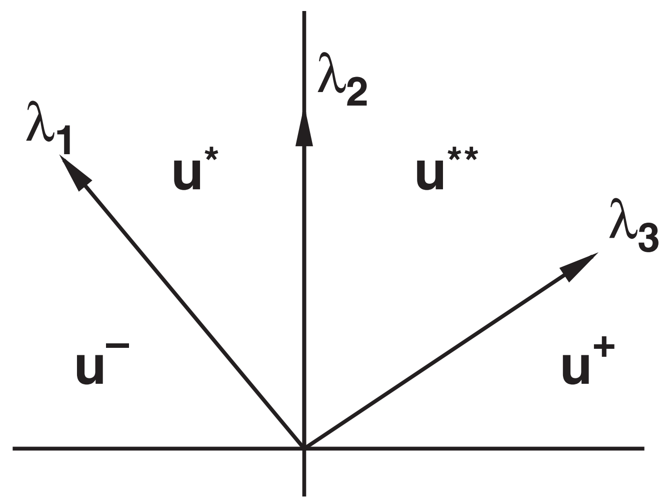
The flux along a normal $\hat{n}$ is then

$$\Pi = \left(\hat{n}_x\mathcal{A}_1(\boldsymbol{x}) + \hat{n}_y\mathcal{A}_2(\boldsymbol{x})\right). \qquad \hat{\boldsymbol{n}}\cdot\mathcal{F} = \Pi\boldsymbol{u}.$$

Now diagonalize this as

$$\mathcal{Q}^{-1}\Pi = \mathcal{S}\Lambda\mathcal{S}^{-1},$$

$$\Lambda = \Lambda^+ + \Lambda^-,$$



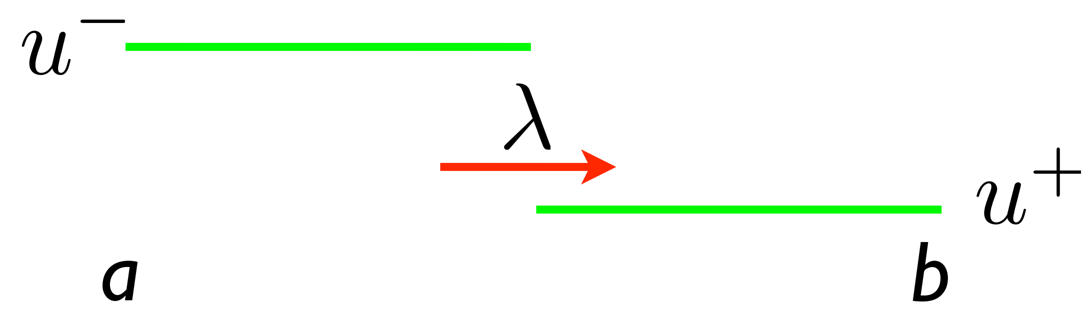
and we obtain

$$(\hat{\boldsymbol{n}}\cdot\mathcal{F})^* = \mathcal{Q}\mathcal{S}\left(\Lambda^+\mathcal{S}^{-1}\boldsymbol{u}^- + \Lambda^-\mathcal{S}^{-1}\boldsymbol{u}^+\right),$$

Linear systems and fluxes

For non-smooth coefficients, it is a little more complex

Consider the problem $$\frac{\partial u}{\partial t} + \lambda \frac{\partial u}{\partial x} = 0, \quad x \in [a,b].$$



Then we clearly have

$$\frac{d}{dt}\int_a^b u\,dx = -\lambda\left(u(b,t) - u(a,t)\right) = f(a,t) - f(b,t),$$

$$\frac{d}{dt}\int_a^b u\,dx = \frac{d}{dt}\left((\lambda t - a)u^- + (b - \lambda t)u^+\right) = \lambda(u^- - u^+).$$

Linear systems and fluxes

Hence, by simple mass conservation, we achieve

$$-\lambda(u^- - u^+) + (f^- - f^+) = 0.$$

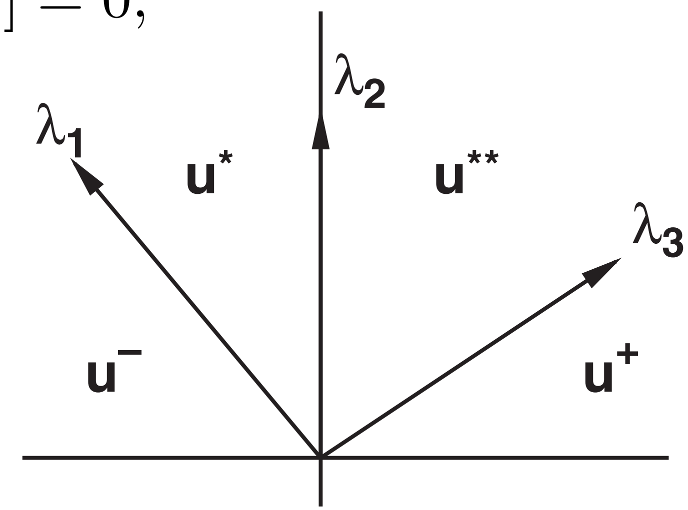
for $a \to x^-, b \to x^+$

These are the Rankine-Hugoniot conditions

For the general system, these are

$$\forall i: \quad -\lambda_i \mathcal{Q}[\boldsymbol{u}^- - \boldsymbol{u}^+] + [(\mathit{\Pi}\boldsymbol{u})^- - (\mathit{\Pi}\boldsymbol{u})^+] = 0,$$

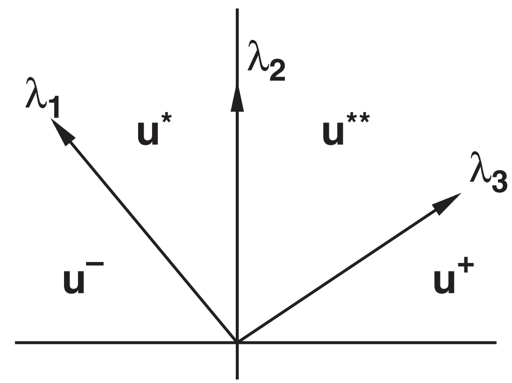
They must hold across each wave and can be used to connect across the interface



Linear systems and fluxes

So for the 3-wave problem we have

$$
\begin{aligned}
&\lambda \mathcal{Q}^-(\boldsymbol{u}^* - \boldsymbol{u}^-) + \left[(\Pi \boldsymbol{u})^* - (\Pi \boldsymbol{u})^-\right] = 0,\\
&\left[(\Pi \boldsymbol{u})^* - (\Pi \boldsymbol{u})^{**}\right] = 0,\\
&-\lambda \mathcal{Q}^+(\boldsymbol{u}^{**} - \boldsymbol{u}^+) + \left[(\Pi \boldsymbol{u})^{**} - (\Pi \boldsymbol{u})^+\right] = 0,
\end{aligned}
$$



and the numerical flux is given as

$$
(\hat{\boldsymbol{n}} \cdot \mathcal{F})^* = (\Pi \boldsymbol{u})^* = (\Pi \boldsymbol{u})^{**},
$$

This approach is general and yields the exact upwind fluxes -- but requires that the system can be solved !

Linear systems and fluxes -- an example

Consider

$$\frac{\partial \boldsymbol{q}}{\partial t} + \mathcal{A}\frac{\partial \boldsymbol{q}}{\partial x} = \frac{\partial}{\partial t}\begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} a(x) & 0 \\ 0 & -a(x) \end{bmatrix} \frac{\partial}{\partial x}\begin{bmatrix} u \\ v \end{bmatrix} = 0,$$

Following the general approach, we have

$$\begin{aligned} a^-(\boldsymbol{q}^* - \boldsymbol{q}^-) + (\Pi \boldsymbol{q})^* - (\Pi \boldsymbol{q})^- &= 0, \\ -a^+(\boldsymbol{q}^* - \boldsymbol{q}^+) + (\Pi \boldsymbol{q})^* - (\Pi \boldsymbol{q})^+ &= 0, \end{aligned}$$

with $$(\Pi \boldsymbol{q})^{\pm} = \hat{\boldsymbol{n}} \cdot (\mathcal{A}\boldsymbol{q})^{\pm} = \hat{\boldsymbol{n}} \cdot \begin{bmatrix} a^{\pm} & 0 \\ 0 & -a^{\pm} \end{bmatrix}\begin{bmatrix} u^{\pm} \\ v^{\pm} \end{bmatrix} = \hat{\boldsymbol{n}} \cdot \begin{bmatrix} a^{\pm}u^{\pm} \\ -a^{\pm}v^{\pm} \end{bmatrix}.$$

Solving this yields

$$(\Pi \boldsymbol{q})^* = \frac{2a^+a^-}{a^+ + a^-}\hat{\boldsymbol{n}} \cdot \left(\begin{bmatrix} \{\{u\}\} \\ -\{\{v\}\} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} [\![u]\!] \\ [\![v]\!] \end{bmatrix} \right),$$

Intermediate velocity

$$a^* = \frac{2a^-a^+}{a^+ + a^-},$$

Linear systems and fluxes -- an example

Consider Maxwell's equations

$$\begin{bmatrix} \varepsilon(x) & 0 \\ 0 & \mu(x) \end{bmatrix} \frac{\partial}{\partial t} \begin{bmatrix} E \\ H \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial}{\partial x} \begin{bmatrix} E \\ H \end{bmatrix} = 0.$$

The exact same approach leads to

$$H^* = \frac{1}{\{\{Z\}\}} \left(\{\{ZH\}\} + \frac{1}{2} [\![E]\!] \right), \quad E^* = \frac{1}{\{\{Y\}\}} \left(\{\{YE\}\} + \frac{1}{2} [\![H]\!] \right),$$

Now assume smooth materials:

$$H^* = \{\{H\}\} + \frac{Y}{2} [\![E]\!], \quad E^* = \{\{E\}\} + \frac{Z}{2} [\![H]\!],$$

$$Z^{\pm} = \sqrt{\frac{\mu^{\pm}}{\varepsilon^{\pm}}} = (Y^{\pm})^{-1},$$

$$\frac{Y}{\varepsilon} = \frac{Z}{\mu} = \frac{1}{\sqrt{\varepsilon\mu}} = c$$

We have recovered the LF flux!

An example

Consider Maxwell's equations

$$\varepsilon(x)\frac{\partial E}{\partial t} = -\frac{\partial H}{\partial x}, \quad \mu(x)\frac{\partial H}{\partial t} = -\frac{\partial E}{\partial x},$$

On the DG form

$$\begin{aligned}\frac{d\boldsymbol{E}_h^k}{dt} + \frac{1}{J^k\varepsilon^k}\mathcal{D}_r\boldsymbol{H}_h^k &= \frac{1}{J^k\varepsilon^k}\mathcal{M}^{-1}\left[\boldsymbol{\ell}^k(x)(H_h^k - H^*)\right]_{x_l^k}^{x_r^k} \\ &= \frac{1}{J^k\varepsilon^k}\mathcal{M}^{-1}\oint_{x_l^k}^{x_r^k}\hat{\boldsymbol{n}}\cdot(H_h^k - H^*)\boldsymbol{\ell}^k(x)\,dx,\end{aligned}$$

with the flux

$$H^- - H^* = \frac{1}{2\{\{Z\}\}}\left(Z^+[\![H]\!] - [\![E]\!]\right),$$

$$E^- - E^* = \frac{1}{2\{\{Y\}\}}\left(Y^+[\![E]\!] - [\![H]\!]\right),$$

An example

MaxwellRHS1D.m

```
function [rhsE, rhsH] = MaxwellRHS1D(E,H,eps,mu)

% function [rhsE, rhsH] = MaxwellRHS1D(E,H,eps,mu)
% Purpose  : Evaluate RHS flux in 1D Maxwell

Globals1D;

% Compute impedance
Zimp = sqrt(mu./eps);

% Define field differences at faces
dE = zeros(Nfp*Nfaces,K); dE(:) = E(vmapM)-E(vmapP);
dH = zeros(Nfp*Nfaces,K); dH(:) = H(vmapM)-H(vmapP);
Zimpm = zeros(Nfp*Nfaces,K); Zimpm(:) = Zimp(vmapM);
Zimpp = zeros(Nfp*Nfaces,K); Zimpp(:) = Zimp(vmapP);
Yimpm = zeros(Nfp*Nfaces,K); Yimpm(:) = 1./Zimpm(:);
Yimpp = zeros(Nfp*Nfaces,K); Yimpp(:) = 1./Zimpp(:);

% Homogeneous boundary conditions, Ez=0
Ebc = -E(vmapB); dE (mapB) = E(vmapB) - Ebc;
Hbc =  H(vmapB); dH (mapB) = H(vmapB) - Hbc;

% evaluate upwind fluxes
fluxE = 1./(Zimpm + Zimpp).*(nx.*Zimpp.*dH - dE);
fluxH = 1./(Yimpm + Yimpp).*(nx.*Yimpp.*dE - dH);

% compute right hand sides of the PDE's
rhsE = (-rx.*(Dr*H) + LIFT*(Fscale.*fluxE))./eps;
rhsH = (-rx.*(Dr*E) + LIFT*(Fscale.*fluxH))./mu;
return
```

Impedance

Compute field jumps

Compute interface impedance

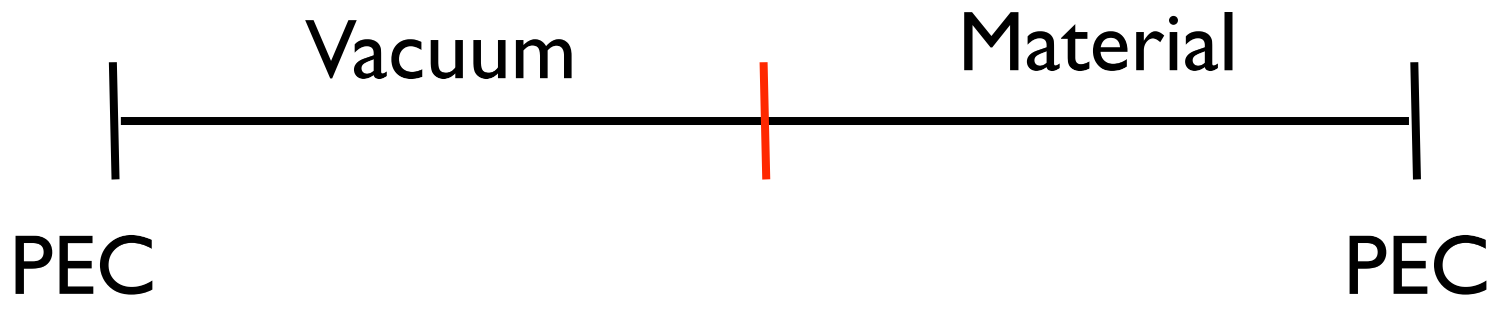
Boundary conditions

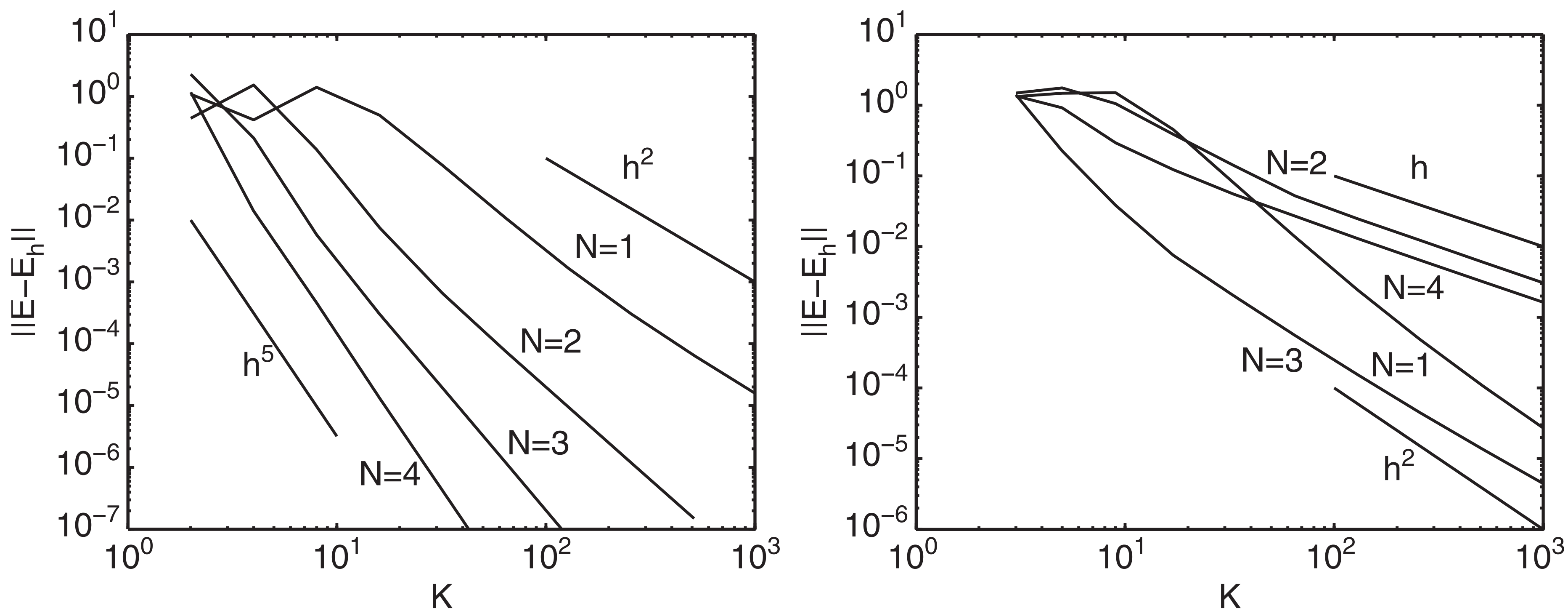
Complete fluxes

Complete computation

An example

Test example is cavity problem





Lets move on

At this point we have a good understanding of stability for linear problems -- through the flux.

Lets now look at accuracy in more detail.

Recall

$$\Omega \simeq \Omega_h = \bigcup_{k=1}^{K} \mathsf{D}^k, \qquad u(x,t) \simeq u_h(x,t) = \bigoplus_{k=1}^{K} u_h^k(x,t),$$

we assume the local solution to be

$$x \in \mathsf{D}^k = [x_l^k, x_r^k]: \ u_h^k(x,t) = \sum_{n=1}^{N_p} \hat{u}_n^k(t)\psi_n(x) = \sum_{i=1}^{N_p} u_h^k(x_i^k,t)\ell_i^k(x).$$

modal basis   nodal basis

Local approximation

To simplify matters, introduce local affine mapping

$$x \in \mathsf{D}^k : \; x(r) = x_l^k + \frac{1+r}{2} h^k, \;\; h^k = x_r^k - x_l^k, \qquad r \in [-1, 1]$$

We have already introduced the Legendre polynomials

$$u(r) \simeq u_h(r) = \sum_{n=1}^{N_p} \hat{u}_n \tilde{P}_{n-1}(r) = \sum_{i=1}^{N_p} u(r_i) \ell_i(r),$$

$$\boldsymbol{u} = \mathcal{V}\hat{\boldsymbol{u}}, \;\; \mathcal{V}^T \boldsymbol{\ell}(r) = \tilde{\boldsymbol{P}}(r), \;\; \mathcal{V}_{ij} = \tilde{P}_j(r_i).$$

and r_i are the Legendre Gauss Lobatto points:

It is robust -- but is it accurate ?

A second look at approximation

We will need a little more notation

Regular energy norms

$$\|u\|_{\Omega}^2 = \int_{\Omega} u^2 \, d\boldsymbol{x} \qquad \|u\|_{\Omega,h}^2 = \sum_{k=1}^{K} \|u\|_{\mathsf{D}^k}^2 \,, \quad \|u\|_{\mathsf{D}^k}^2 = \int_{\mathsf{D}^k} u^2 \, d\boldsymbol{x}.$$

Sobolev norms

$$\|u\|_{\Omega,q}^2 = \sum_{|\alpha|=0}^{q} \|u^{(\alpha)}\|_{\Omega}^2, \quad \|u\|_{\Omega,q,h}^2 = \sum_{k=1}^{K} \|u\|_{\mathsf{D}^k,q}^2 \,, \quad \|u\|_{\mathsf{D}^k,q}^2 = \sum_{|\alpha|=0}^{q} \|u^{(\alpha)}\|_{\mathsf{D}^k}^2,$$

Semi-norms

$$|u|_{\Omega,q,h}^2 = \sum_{k=1}^{K} |u|_{\mathsf{D}^k,q}^2, \quad |u|_{\mathsf{D}^k,q}^2 = \sum_{|\alpha|=q} \|u^{(\alpha)}\|_{\mathsf{D}^k}^2.$$

Approximation theory

Recall

$$\Omega \simeq \Omega_h = \bigcup_{k=1}^{K} \mathsf{D}^k, \qquad u(x,t) \simeq u_h(x,t) = \bigoplus_{k=1}^{K} u_h^k(x,t),$$

we assume the local solution to be

$$x \in \mathsf{D}^k = [x_l^k, x_r^k] : \; u_h^k(x,t) = \sum_{n=1}^{N_p} \hat{u}_n^k(t)\psi_n(x) = \sum_{i=1}^{N_p} u_h^k(x_i^k,t)\ell_i^k(x).$$

The question is in what sense is $u(x,t) \simeq u_h(x,t)$

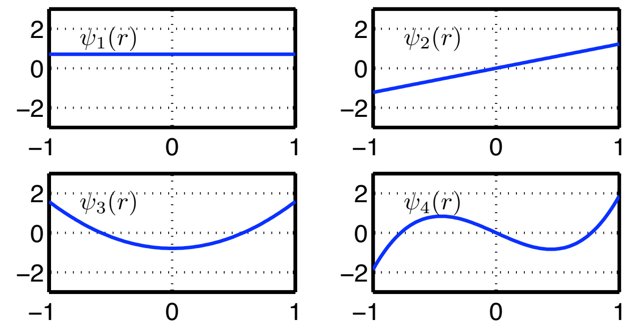
We have observed improved accuracy in two ways

- Increase K/decrease h
- Increase N

Approximation theory

Let us assume all elements have size h and consider

$$v(r) = u(hr) = u(x);$$



We consider expansions as

$$v_h(r) = \sum_{n=0}^{N} \hat{v}_n \tilde{P}_n(r), \quad \tilde{P}_n(r) = \frac{P_n(r)}{\sqrt{\gamma_n}}, \quad \gamma_n = \frac{2}{2n+1}. \quad \tilde{v}_n = \int_{\mathsf{I}} v(r)\tilde{P}_n(r)\, dr.$$

Theorem 4.1. *Assume that $v \in H^p(\mathsf{I})$ and that v_h represents a polynomial projection of order N. Then*

$$\|v - v_h\|_{\mathsf{I},q} \le N^{\rho - p} |v|_{\mathsf{I},p},$$

where

$$\rho = \begin{cases} \frac{3}{2}q, & 0 \le q \le 1 \\ 2q - \frac{1}{2}, & q \ge 1 \end{cases}$$

and $0 \le q \le p$.

Approximation theory

A sharper result can be obtained by using

Lemma 4.4. *If $v \in H^p(\mathsf{I})$, $p \geq 1$ then*

$$\|v^{(q)} - v_h^{(q)}\|_{\mathsf{I},0} \leq \left[\frac{(N+1-\sigma)!}{(N+1+\sigma-4q)!}\right]^{1/2} |v|_{\mathsf{I},\sigma},$$

where $\sigma = \min(N+1, p)$ and $q \leq p$.

Note that in the limit of N>>p we recover

$$\|v^{(q)} - v_h^{(q)}\|_{\mathsf{I},0} \leq N^{2q-p}|v|_{\mathsf{I},p},$$

A minor issues arises -- these results are based on projections and we are using interpolations ?

Approximation theory

We consider

interpolation projection

$$v_h(r) = \sum_{n=0}^{N} \hat{v}_n \tilde{P}_n(r), \quad \tilde{v}_h(r) = \sum_{n=0}^{N} \tilde{v}_n \tilde{P}_n(r), \qquad \boldsymbol{v} = \mathcal{V}\hat{\boldsymbol{v}},$$

Compare the two

$$(\mathcal{V}\hat{\boldsymbol{v}})_i = v_h(r_i) = \sum_{n=0}^{\infty} \tilde{v}_n \tilde{P}_n(r_i) = \sum_{n=0}^{N} \tilde{v}_n \tilde{P}_n(r_i) + \sum_{n=N+1}^{\infty} \tilde{v}_n \tilde{P}_n(r_i),$$

$$\mathcal{V}\hat{\boldsymbol{v}} = \mathcal{V}\tilde{\boldsymbol{v}} + \sum_{n=N+1}^{\infty} \tilde{v}_n \tilde{P}_n(\boldsymbol{r}),$$

$$v_h(r) = \tilde{v}_h(r) \boxed{+ \tilde{\boldsymbol{P}}^T(r)\mathcal{V}^{-1} \sum_{n=N+1}^{\infty} \tilde{v}_n \tilde{P}_n(\boldsymbol{r}).}$$

Approximation theory

Consider this term

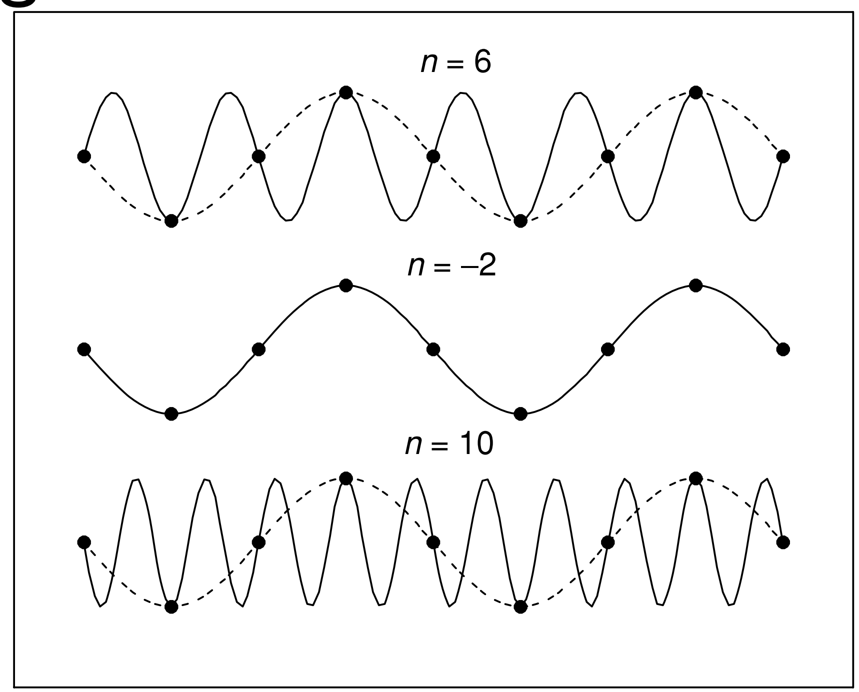
$$\tilde{\boldsymbol{P}}^T(r)\mathcal{V}^{-1}\sum_{n=N+1}^{\infty}\tilde{v}_n\tilde{P}_n(\boldsymbol{r}) = \sum_{n=N+1}^{\infty}\tilde{v}_n\left(\tilde{\boldsymbol{P}}^T(r)\mathcal{V}^{-1}\tilde{P}_n(\boldsymbol{r})\right),$$

$$\tilde{\boldsymbol{P}}^T(r)\mathcal{V}^{-1}\tilde{P}_n(\boldsymbol{r}) = \sum_{l=0}^{N}\tilde{p}_l\tilde{P}_l(r), \quad \mathcal{V}\tilde{\boldsymbol{p}} = \tilde{P}_n(\boldsymbol{r}),$$

Caused by interpolation of high-frequency unresolved modes

Aliasing

Caused by the grid



Approximation theory

This has a some impact on the accuracy

Theorem 4.5. *Assume that* $v \in H^p(\mathsf{I})$, $p > \frac{1}{2}$, *and that* v_h *represents a polynomial interpolation of order* N. *Then*

$$\|v - v_h\|_{\mathsf{I},q} \leq N^{2q-p+1/2}|v|_{\mathsf{I},p},$$

where $0 \leq q \leq p$.

To also account for the cell size we have

Theorem 4.7. *Assume that* $u \in H^p(\mathsf{D}^k)$ *and that* u_h *represents a piecewise polynomial approximation of order* N. *Then*

$$\|u - u_h\|_{\Omega,q,h} \leq Ch^{\sigma-q}|u|_{\Omega,\sigma,h},$$

for $0 \leq q \leq \sigma$, *and* $\sigma = \min(N+1, p)$.

Approximation theory

Combining everything, we have the general result

Theorem 4.8. *Assume that* $u \in H^p(\mathsf{D}^k)$, $p > 1/2$, *and that* u_h *represents a piecewise polynomial interpolation of order* N. *Then*

$$\|u - u_h\|_{\Omega,q,h} \leq C \frac{h^{\sigma-q}}{N^{p-2q-1/2}} |u|_{\Omega,\sigma,h},$$

for $0 \leq q \leq \sigma$, *and* $\sigma = \min(N+1, p)$.

with $h = \max_k h^k$

Lets summarize Part I

Fluxes:

✓ For linear systems, we can derive exact upwind fluxes using Rankine-Hugoniot conditions.

Accuracy:

✓ Legendre polynomials are the right basis
✓ Local accuracy depends on elementwise smoothness
✓ Aliasing appears due to the grid but is under control
✓ For smooth problems, we have a spectral method
✓ Convergence can be recovered in two ways
  ✓ Increase N
  ✓ Decrease h

Convergence of the solution at all times ?

Lecture 3

- ✓ Let’s briefly recall what we know
- ✓ Why high-order methods ?
- ✓ Part I:
  - ✓ Constructing fluxes for linear systems
  - ✓ Approximation theory on the interval
- ✓ Part II:
  - ✓ Convergence and error estimates
  - ✓ Dispersive properties
  - ✓ Discrete stability and how to overcome

Lets recall convergence etc

We consider the system

$$\frac{\partial \boldsymbol{u}}{\partial t} + \mathcal{A}\frac{\partial \boldsymbol{u}}{\partial x} = 0,$$

which we assume is wellposed in the sense

$$\|\boldsymbol{u}(t)\|_{\Omega} \leq C \exp(\alpha t)\|\boldsymbol{u}(0)\|_{\Omega}.$$

The semi-discrete scheme is given as

$$\frac{d\boldsymbol{u}_h}{dt} + \mathcal{L}_h \boldsymbol{u}_h = 0.$$

Inserting the exact solution u into the scheme yields

$$\frac{d\boldsymbol{u}}{dt} + \mathcal{L}_h \boldsymbol{u} = \mathcal{T}(\boldsymbol{u}(x,t)),$$

truncation error

Convergence and all that

Let us introduce the error

$$\boldsymbol{\varepsilon}(\boldsymbol{x}, t) = \boldsymbol{u}(\boldsymbol{x}, t) - \boldsymbol{u}_h(\boldsymbol{x}, t),$$

What we really seek is convergence

$$\forall t \in [0, T] : \lim_{\mathrm{dof} \to \infty} \|\boldsymbol{\varepsilon}(t)\|_{\Omega,h} \to 0.$$

This is often a little complicated to get to due to the requirement for all t.

Let us get to it in a different way.

Convergence and all that

Let us consider the error equation

$$\frac{d}{dt}\boldsymbol{\varepsilon} + \mathcal{L}_h\boldsymbol{\varepsilon} = \mathcal{T}(\boldsymbol{u}(\boldsymbol{x},t)),$$

The solution is given as

$$\boldsymbol{\varepsilon}(t) - \exp\left(-\mathcal{L}_h t\right)\boldsymbol{\varepsilon}(0) = \int_0^t \exp\left(\mathcal{L}_h(s-t)\right)\mathcal{T}(\boldsymbol{u}(s))\,ds,$$

Now consider

$$\|\boldsymbol{\varepsilon}(t)\|_{\Omega,h} \le \|\exp\left(-\mathcal{L}_h t\right)\boldsymbol{\varepsilon}(0)\|_{\Omega,h} + \left\|\int_0^t \exp\left(\mathcal{L}_h(s-t)\right)\mathcal{T}(\boldsymbol{u}(s))\,ds\right\|_{\Omega,h}$$

$$\left\|\int_0^t \exp\left(\mathcal{L}_h(s-t)\right)\mathcal{T}(\boldsymbol{u}(s))\,ds\right\|_{\Omega,h} \le \int_0^t \|\exp\left(\mathcal{L}_h(s-t)\right)\|_{\Omega,h}\|\mathcal{T}(\boldsymbol{u}(s))\|_{\Omega,h}\,ds,$$

Convergence and all that

So if we require consistency

$$\begin{cases} \lim_{\mathrm{dof}\to\infty} \|\boldsymbol{\varepsilon}(0)\|_{\Omega,h} = 0, \\ \lim_{\mathrm{dof}\to\infty} \|\mathcal{T}(\boldsymbol{u}(t))\|_{\Omega,h} = 0 \end{cases}$$

and stability

$$\lim_{\mathrm{dof}\to\infty} \|\exp\left(-\mathcal{L}_h t\right)\|_{\Omega,h} \le C_h \exp(\alpha_h t), \quad t \ge 0,$$

we obtain convergence

$$\forall t \in [0, T] : \lim_{\mathrm{dof}\to\infty} \|\boldsymbol{\varepsilon}(t)\|_{\Omega,h} \to 0.$$

This is of course part of the celebrated Lax-Richtmyer equivalence theorem

Convergence and all that

Recall

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0,$$

for which we proved stability as

$$\frac{1}{2}\frac{d}{dt}\|u_h\|^2_{\Omega,h} \leq c\|u_h\|^2_{\Omega,h},$$

This generalizes easily to systems when upwinding is used on the characteristic variables.

Combining this with the accuracy analysis yields

$$\|u - u_h\|_{\Omega,h} \leq \frac{h^N}{N^{p-5/2}}|u|_{\Omega,p,h},$$

but we observed

$$\|u(T) - u_h(T)\|_{\Omega,h} \leq h^{N+1}(C_1 + TC_2).$$

Error estimates

To get closer to the observed behavior, we need to be a little more careful.

Define $$\mathcal{B}(u,\phi) = (u_t,\phi)_\Omega + a(u_x,v)_\Omega = 0$$

we have $$\mathcal{B}(u,u) = 0 = \frac{1}{2}\frac{d}{dt}\|u\|^2_\Omega;$$

For two different solutions we have

$$\varepsilon(t) = u_1(t) - u_2(t)$$

$$\frac{1}{2}\frac{d}{dt}\|\varepsilon\|^2_\Omega = 0, \quad \Longrightarrow \quad \|\varepsilon(T)\|_\Omega = \|u_1(0) - u_2(0)\|_\Omega,$$

Error estimates

We will now mimic this for the semi-discrete problem

$$\mathcal{B}_h(u_h,\phi_h) = ((u_h)_t,\phi_h)_{\Omega,h} + a((u_h)_x,\phi_h)_{\Omega,h} - (\hat{\boldsymbol{n}}\cdot(au_h-(au)^*),\phi_h)_{\partial\Omega,h} = 0,$$

Let us use a central flux

$$(au)^* = \{\{au\}\},$$

to obtain

$$\mathcal{B}_h(u_h,\phi_h) = ((u_h)_t,\phi_h)_{\Omega,h} + a((u_h)_x,\phi_h)_{\Omega,h} - \frac{1}{2}([\![au_h]\!],\phi_h)_{\partial\Omega,h} = 0.$$

Observe

$$\mathcal{B}_h(u,\phi_h) = 0, \quad \Longrightarrow \quad \mathcal{B}_h(\varepsilon,\phi_h) = 0, \quad \varepsilon = u - u_h.$$

Using

$$\boxed{\mathcal{B}_h(\varepsilon_h,\varepsilon_h) = \frac{1}{2}\frac{d}{dt}\|\varepsilon_h\|^2_{\Omega,h}.}$$

Error estimates

Now consider

$$\frac{1}{2}\frac{d}{dt}\|\varepsilon_N\|^2_{\Omega,h} = \mathcal{B}_h(\mathcal{P}_N u - u, \varepsilon_h),$$

one proves (with some work)

$$\begin{aligned}|\mathcal{B}_h(u - \mathcal{P}_N u, \varepsilon_h)| &\leq \frac{1}{2}\left((\{\{aq\}\}, \{\{aq\}\})_{\partial\Omega,h} + (\varepsilon_h, \varepsilon_h)_{\partial\Omega,h}\right)\\ &\leq C|a|h^{2\sigma-1}\|u\|^2_{\Omega,h,\sigma+1},\end{aligned}$$

$$\Longrightarrow \qquad \frac{d}{dt}\|\varepsilon_h\|^2_{\Omega,h} \leq C|a|h^{2\sigma-1}\|u\|^2_{\Omega,h,\sigma+1},$$

$$\Longrightarrow \qquad \|\varepsilon_h(T)\| \leq (C_1 + C_2 T)h^{N+1/2},$$

Better -- but not quite there

Error estimates

The observe full order

$$\|u(T) - u_h(T)\|_{\Omega,h} \leq h^{N+1}(C_1 + TC_2).$$

is in fact a special case !

It only works when

- ✓ When full upwinding on all characteristic variables are used
- ✓ Proof is only valid for the linear case
- ✓ Proof relies on 1D superconvergence results

In spite of this, optimal convergence is observed in many problems - why ?

Why often optimal anyway ?

Assume stability

$$\lim_{\mathrm{dof}\to\infty} \| \exp\left(-\mathcal{L}_h t\right) \|_{\Omega,h} \le C_h \exp(\alpha_h t), \quad t \ge 0,$$

Recall

$$\|\boldsymbol{\varepsilon}(t)\|_{\Omega,h} \le \| \exp\left(-\mathcal{L}_h t\right) \boldsymbol{\varepsilon}(0)\|_{\Omega,h} + \left\| \int_0^t \exp\left(\mathcal{L}_h (s-t)\right) \mathcal{T}(\boldsymbol{u}(s))\, ds \right\|_{\Omega,h}$$

Error in I.C.

Error accumulation

$$\|u - u_h\|_{\Omega,q,h} \le C \frac{h^{\sigma-q}}{N^{p-2q-1/2}} |u|_{\Omega,\sigma,h},$$

$$\sigma = \min(N+1, p).$$

Dispersive properties

Consider again

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, \quad u(x,0) = \exp(ilx), \quad \Rightarrow \quad u(x,t) = \exp(i(lx-\omega t)),$$

The scheme is given as

$$\frac{h}{2}\mathcal{M}\frac{d\boldsymbol{u}_h^k}{dt} + a\mathcal{S}\boldsymbol{u}^k = \boldsymbol{e}_N\left[(au_h^k) - (au_h^k)^*\right]_{x_r^k} - \boldsymbol{e}_0\left[(au_h^k) - (au_h^k)^*\right]_{x_l^k},$$

$$(au)^* = \{\{au\}\} + |a|\frac{1-\alpha}{2}[\![u]\!].$$

Look for solutions of the form

$$\boldsymbol{u}_h^k(x^k,t) = \boldsymbol{U}_h^k\exp[i(lx^k - \omega t)],$$

Dispersive properties

We recover

$$\left[2\mathcal{S} - \alpha \boldsymbol{e}_N \left(\boldsymbol{e}_N^T - \exp(iL(N+1))\boldsymbol{e}_0^T\right)\right.$$
$$\left. + (2-\alpha)\boldsymbol{e}_0 \left(\boldsymbol{e}_0^T - \exp(-iL(N+1))\boldsymbol{e}_N^T\right)\right] \boldsymbol{U}_h^k = i\Omega\mathcal{M}\boldsymbol{U}_h^k.$$

Where

$$L = \frac{lh}{N+1} = \frac{2\pi}{\lambda}\frac{h}{N+1} = 2\pi p^{-1}, \quad \Omega = \frac{\omega h}{a},$$

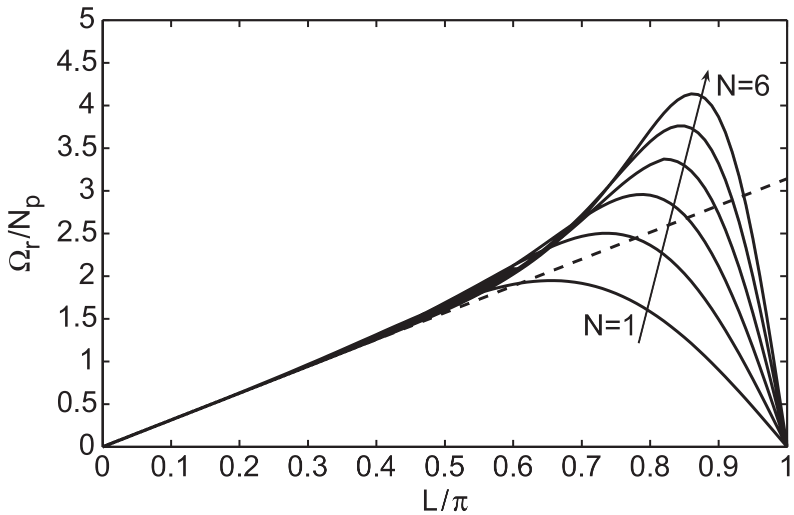
$$p = \frac{\lambda}{h/(N+1)}$$ = DoF per wavelength

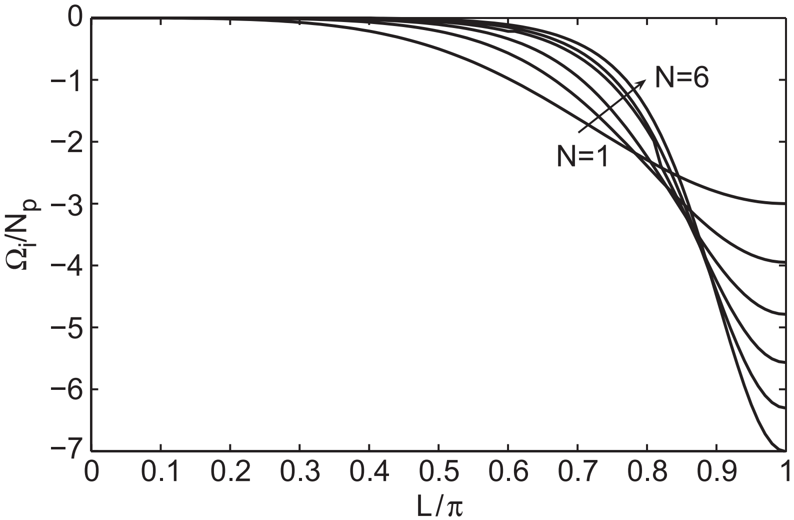
So for a fixed L we solve the eigenvalue problem

.. and the eigenvalue will tell us how the wave propagates

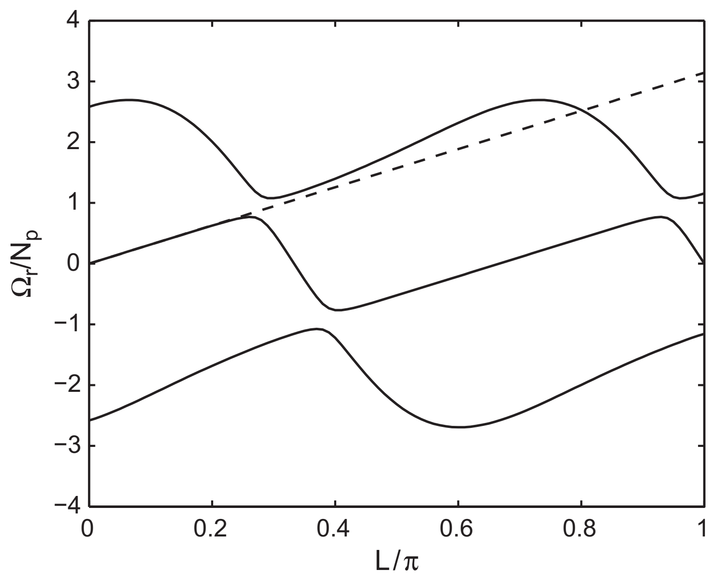
Dispersive properties

Upwind fluxes





Central fluxes



N=2

Dispersive properties

There are some analytic results available (upwind)

$$\left|\mathcal{R}(\tilde{l}h) - \mathcal{R}(lh)\right| \simeq \frac{1}{2}\left[\frac{N!}{(2N+1)!}\right]^2 (lh)^{2N+3},$$

$$\left|\mathcal{I}(\tilde{l}h)\right| \simeq \frac{1}{2}\left[\frac{N!}{(2N+1)!}\right]^2 (1-\alpha)^{(-1)^N} (lh)^{2N+2},$$

The dispersive accuracy is excellent!

Define the relative phase error $\quad \rho_N = \left|\dfrac{\exp(ilh) - \exp(i\tilde{l}h)}{\exp(ilh)}\right|,$

$$\rho_N \simeq \begin{cases} 2N+1 < lh - C(lh)^{1/3}, & \text{no convergence} \\ lh - o(lh)^{1/3} < 2N+1 < lh + o(lh)^{1/3}, & \mathcal{O}(N^{-1/3}) \text{ convergence} \\ 2N+1 \gg lh, & \mathcal{O}(hl/(2N+1))^{2N+2} \text{ convergence} \end{cases}$$

Convergence for $\quad 2 \simeq \dfrac{lh}{N+1} = 2\pi p^{-1}; \quad \Rightarrow \quad \boxed{p \geq \pi}$

Discrete stability

So far we have not done anything to discretize time.

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \quad \Rightarrow \quad \frac{d\boldsymbol{u}_h}{dt} + \mathcal{L}_h \boldsymbol{u}_h = 0.$$

We shall consider the use of ERK methods

$$\begin{aligned}
\boldsymbol{k}^{(1)} &= \mathcal{L}_h\left(\boldsymbol{u}_h^n, t^n\right), \\
\boldsymbol{k}^{(2)} &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \frac{1}{2}\Delta t \boldsymbol{k}^{(1)}, t^n + \frac{1}{2}\Delta t\right), \\
\boldsymbol{k}^{(3)} &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \frac{1}{2}\Delta t \boldsymbol{k}^{(2)}, t^n + \frac{1}{2}\Delta t\right), \\
\boldsymbol{k}^{(4)} &= \mathcal{L}_h\left(\boldsymbol{u}_h^n + \Delta t \boldsymbol{k}^{(3)}, t^n + \Delta t\right), \\
\boldsymbol{u}_h^{n+1} &= \boldsymbol{u}_h^n + \frac{1}{6}\Delta t\left(\boldsymbol{k}^{(1)} + 2\boldsymbol{k}^{(2)} + 2\boldsymbol{k}^{(3)} + \boldsymbol{k}^{(4)}\right),
\end{aligned}$$

Discrete stability

and also a Low Storage form

$$\boldsymbol{p}^{(0)} = \boldsymbol{u}^n,$$

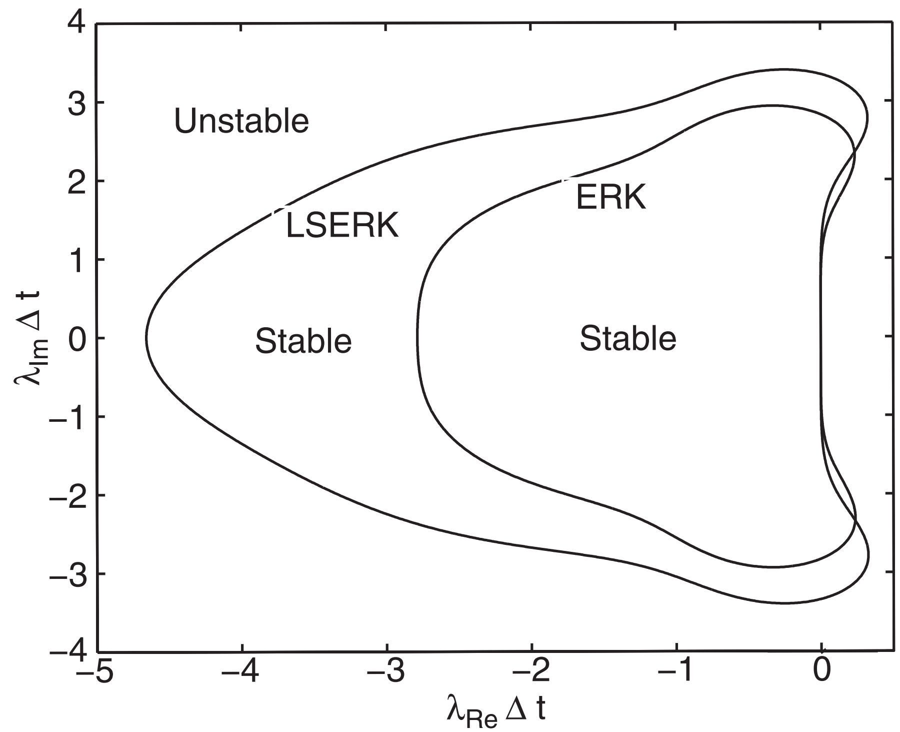
$$i \in [1, \dots, 5] : \begin{cases} \boldsymbol{k}^{(i)} = a_i \boldsymbol{k}^{(i-1)} + \Delta t \mathcal{L}_h \left(\boldsymbol{p}^{(i-1)}, t^n + c_i \Delta t \right), \\ \boldsymbol{p}^{(i)} = \boldsymbol{p}^{(i-1)} + b_i \boldsymbol{k}^{(i)}, \end{cases}$$

$$\boldsymbol{u}_h^{n+1} = \boldsymbol{p}^{(5)}.$$

Consider

$$u_t = \lambda u, \quad \mathrm{Real}(\lambda) \leq 0,$$

The *stability region* defines the timestep that gives stability.



Discrete stability

Consider

$$\mathcal{L}_h = \frac{2a}{h}\mathcal{M}^{-1}\left[\mathcal{S} - \mathcal{E}\right],$$

We have

$$\begin{aligned}
\frac{h^2}{4a^2}\|\mathcal{L}_h\|_{\mathsf{I}}^2 &= \frac{h^2}{4a^2}\sup_{\|u_h\|=1}\|\mathcal{L}_h u_h\|_{\mathsf{I}}^2 \\
&\leq \|\mathcal{D}_r\|_{\mathsf{I}}^2 + \|\mathcal{M}^{-1}\mathcal{E}\|_{\mathsf{I}}^2 + 2\sup_{\|u_h\|=1}\left(\mathcal{D}_r u_h, \mathcal{M}^{-1}\mathcal{E}u_h\right)_{\mathsf{I}} \\
&\leq C_1 N^4 + C_2 N^2 + C_3 N^3 \leq C N^4,
\end{aligned}$$

So we should expect

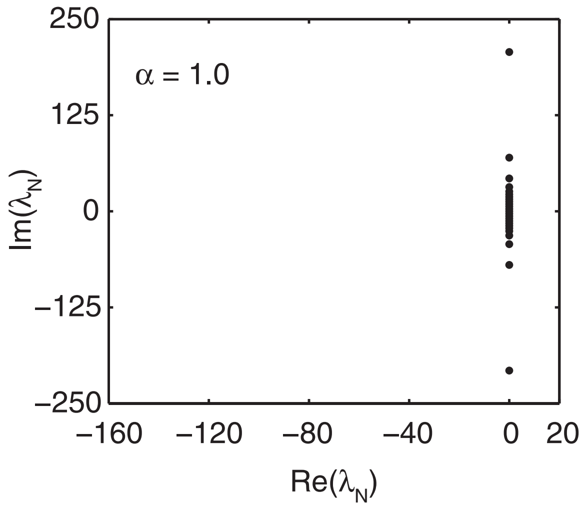
$$\|\mathcal{L}_h\|_{\mathsf{D}^k} \leq C\frac{a}{h^k}N^2$$

Which would indicate

$$\boxed{\Delta t \leq C\frac{h}{aN^2}}$$

Discrete stability

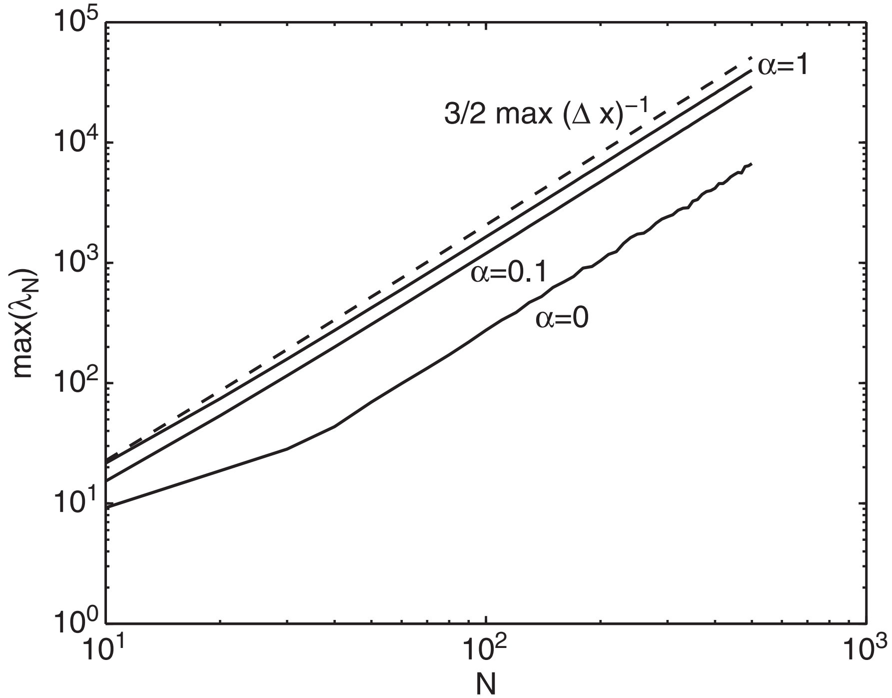
The structure also matters



The estimate

$$\Delta t \leq C\,\frac{h}{aN^2}$$

is sharp !



Discrete stability

General guidelines

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, \quad \Longrightarrow \quad \Delta t \leq C\frac{1}{|a|}\min_{k,i}\frac{h^k}{2}(\Delta_i r),$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \mathcal{A}\frac{\partial \boldsymbol{u}}{\partial x} = 0, \quad \Longrightarrow \quad \Delta t \leq C\frac{1}{\max(|\lambda(\mathcal{A})|)}\min_{k,i}\frac{h^k}{2}(\Delta_i r),$$

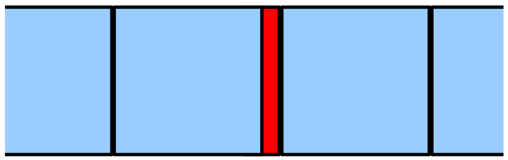
There are tricks to play to improve on this

✓ Mappings to improve the scaling
✓ Covolume filtering techniques
✓ Local time-stepping

See text for a discussion of other methods

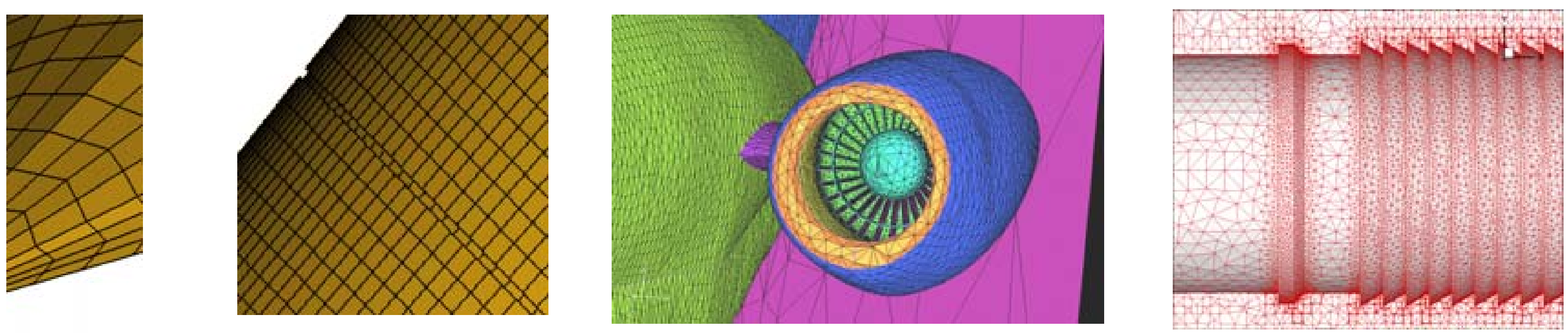
Local time-stepping

Problem: Small cells, even just one, cause a very small global time-step in an explicit scheme.



$$\Delta t \leq C\sqrt{\varepsilon\mu}\Delta x \simeq C_1\sqrt{\varepsilon\mu}\frac{N^2}{h}$$

A significant problem for large scale complex applications



<u>Old idea</u>: take only time-steps required by local restrictions.

<u>Old problems</u>: accuracy and stability

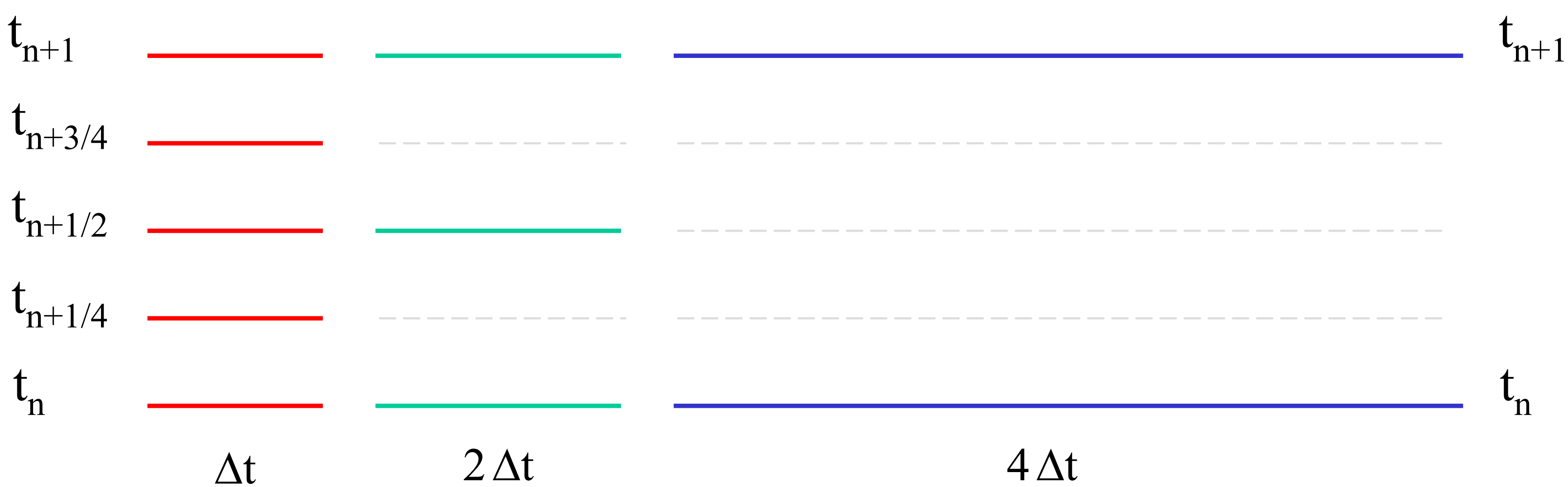
Local time-stepping

Substantial recent work by

Cohen, Grote, Lanteri, Piperno, Gassner, Munz etc

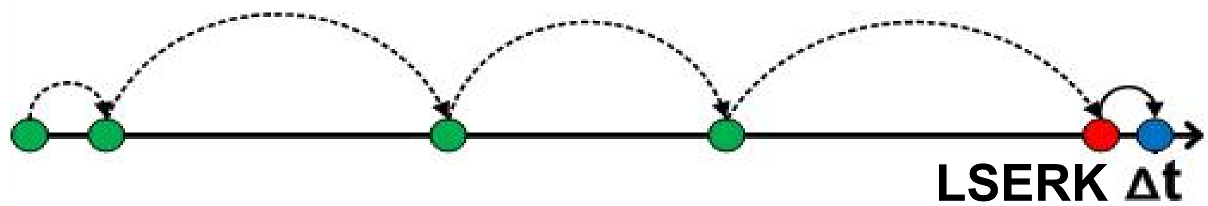
Most of the recent work is based on LF-like schemes, restricted to 2nd order in time.

Layout for **multi-rate** local time-stepping

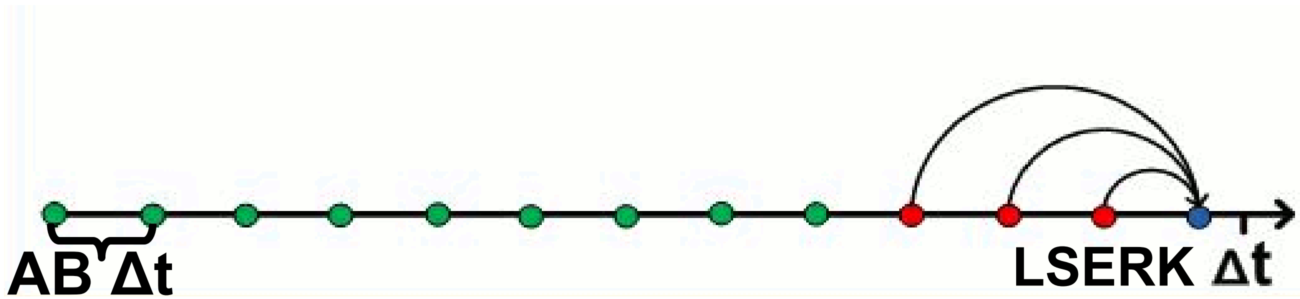


Local time-stepping

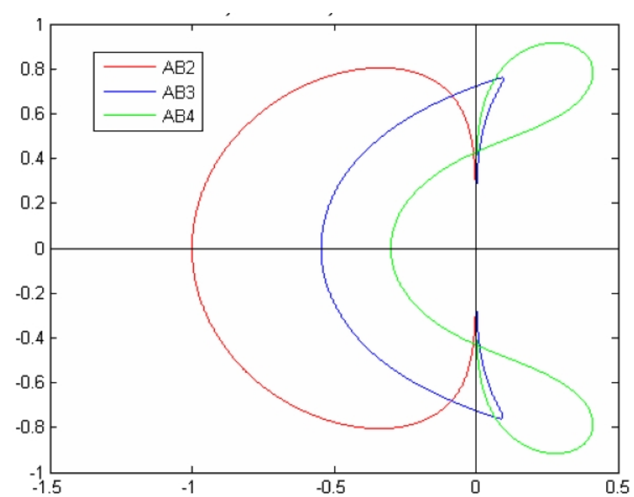
Recall the ERK scheme



We consider a multi-step scheme

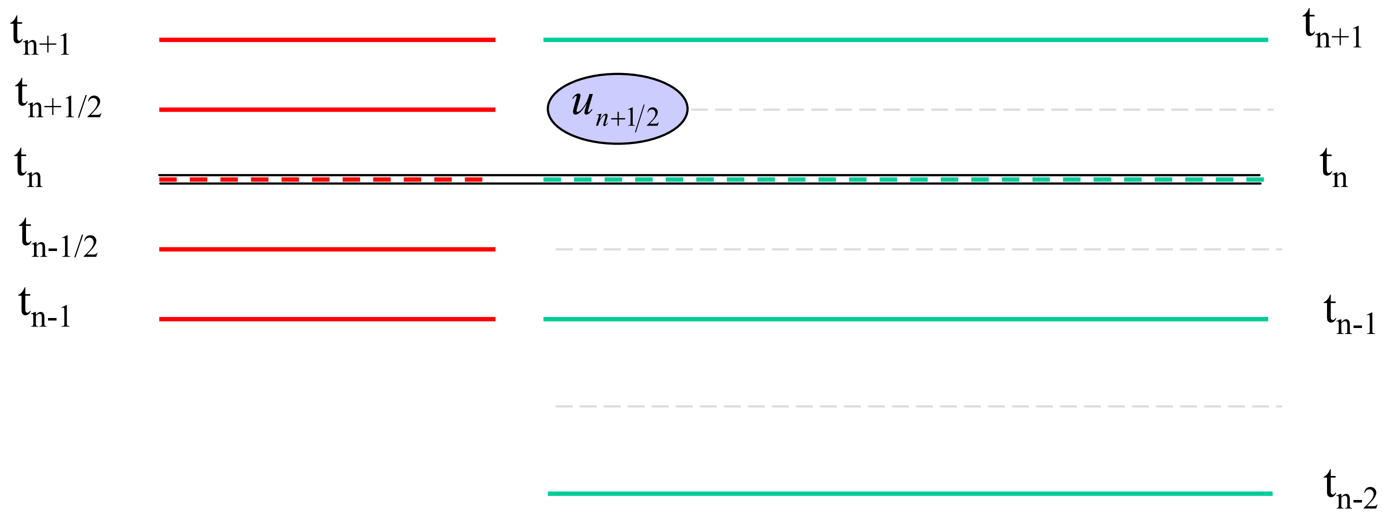


$$u_{n+1} = u_n + \frac{\Delta t}{12}\left[23F(u_n) - 16F(u_{n-1}) + 5F(u_{n-2})\right]$$



Local time-stepping

Challenge: Achieving this at high-order accuracy

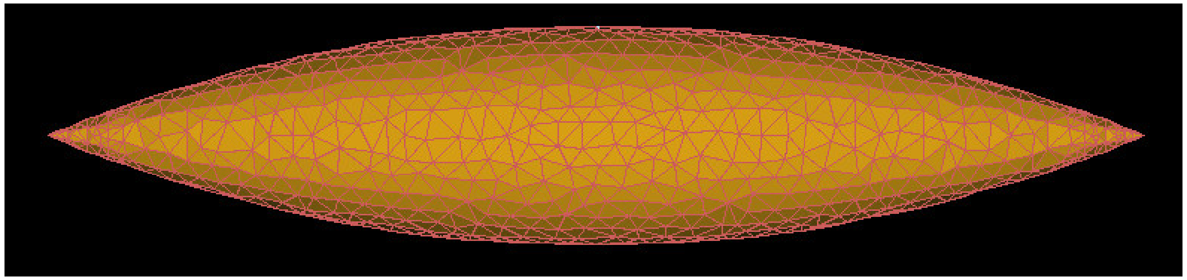


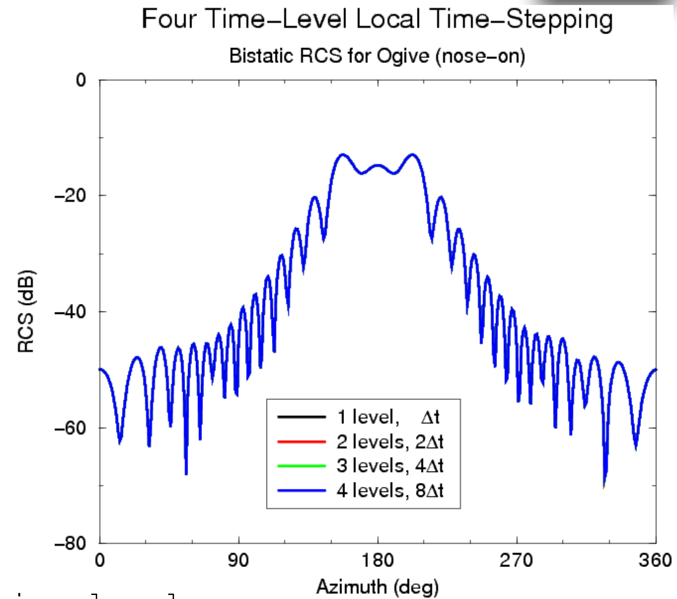
For all interior cells $\quad u_{n+1} = u_n + \frac{\Delta t}{12}\left[23F(u_n) - 16F(u_{n-1}) + 5F(u_{n-2})\right]$

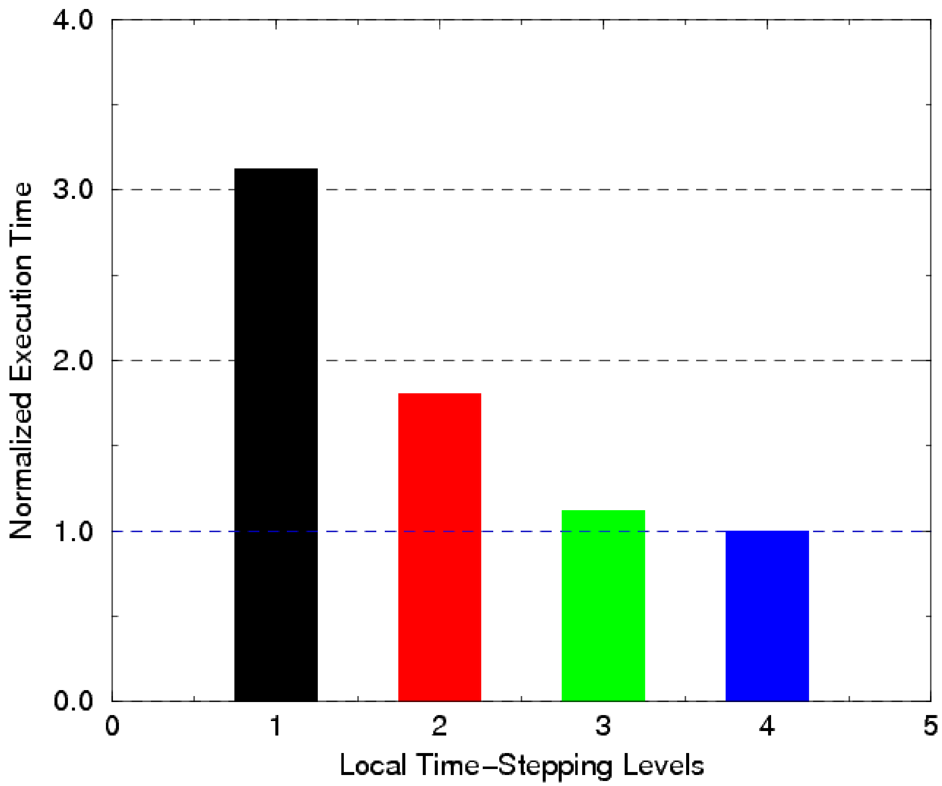
At interface cells $u_{n+1/2}$ $\quad u_{n+1/2} = u_n + \frac{\Delta t}{12}\left[17F(u_n) - 7F(u_{n-1}) + 2F(u_{n-2})\right]$

This generalizes to many levels and arbitrary time-step fractions

Local time-stepping





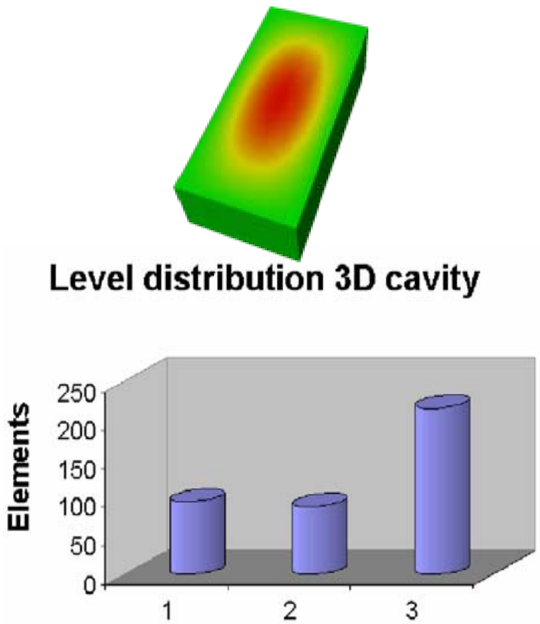


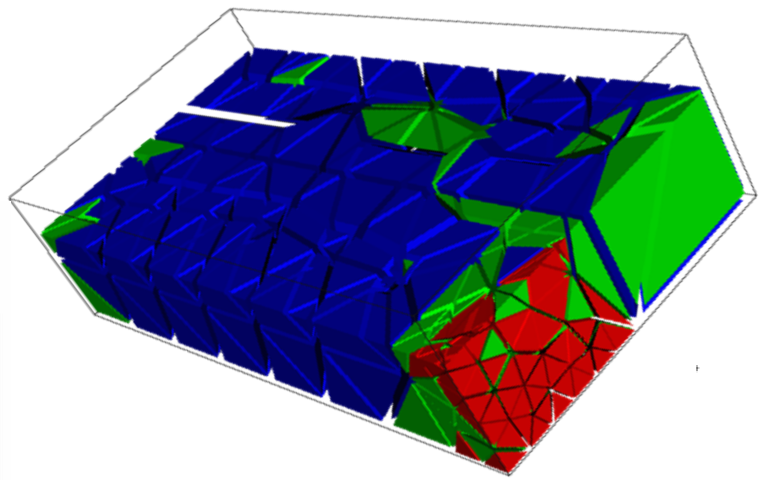
- One time level:
  - N_o = 23742
- Two time levels:
  - N_o = 151 (<1%)
  - N_1 = 23591 (99%)
- Three time levels:
  - N_o = 151 (<1%)
  - N_1 = 1959 (8%)
  - N_2 = 21632 (91%)
- Four time levels:
  - N_o = 151 (<1%)
  - N_1 = 1959 (8%)
  - N_2 = 12622 (53%)
  - N_3 = 9010 (38%)

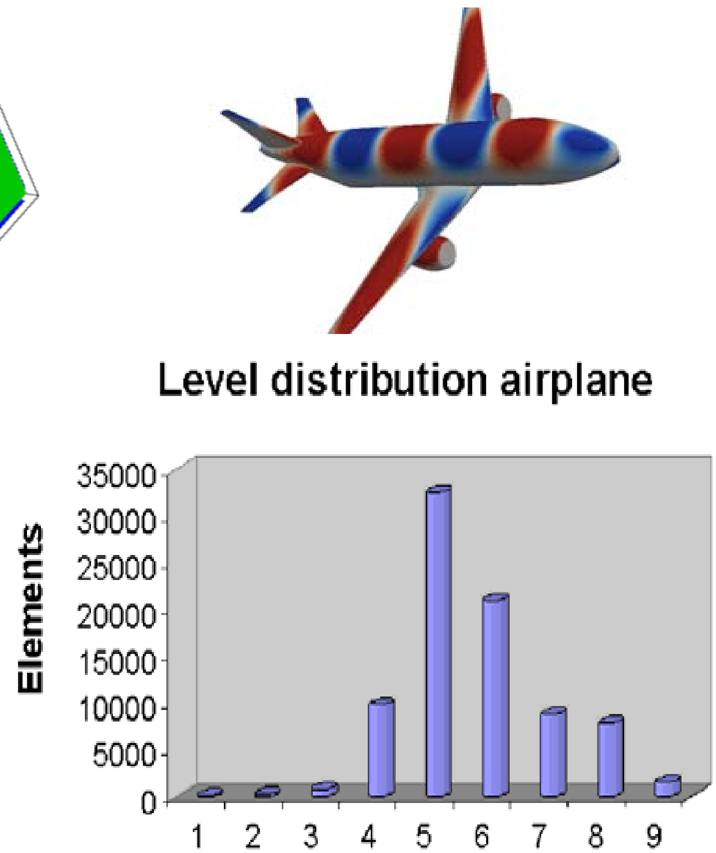
Computations by HyperComp Inc

Local time-stepping

Segmentation is done in preprocessing







Ideally suited for <u>local</u> DG scheme

<u>Known problems</u>:

No known stability proof

Time-step is not optimal (about 80%)

Local time-stepping

The potential speed up is considerable -- and the more complex the better !

Example	Simulation time with		
	Adams-Bashford (global time step)	Adams-Bashford (local time step)	LSERK (global time step)
Resonator	100%	59%	45%
3dB-Coupler	100%	29%	45%
Airplane	100%	15%	45%

Computations by Nico Godel, Hamburg

A brief summary

We now have a good understanding all key aspects of the DG-FEM scheme for linear first order problems

- We understand both accuracy and stability and what we can expect.
- The dispersive properties are excellent.
- The discrete stability is a little less encouraging. A scaling like

$$\Delta t \leq C \frac{h}{aN^2}$$

is the Achilles Heel -- but there are ways!

... but what about nonlinear problems ?