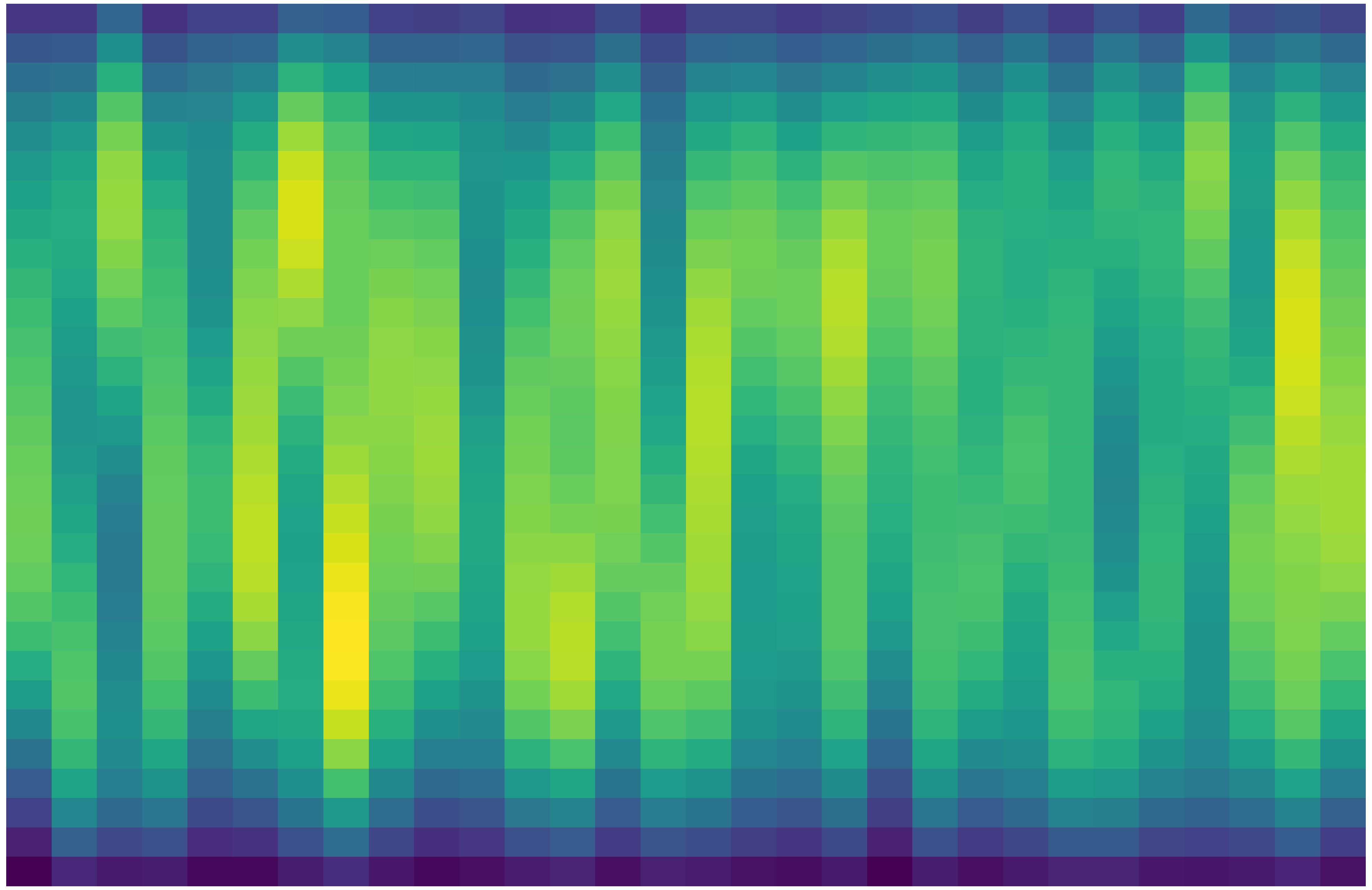
Multigrid Methods — An Overview

Lecture 1: Basics

Luke Olson
Nelder Fellow
Department of Mathematics
Imperial College

home:
Department of Computer Science
University of Illinois at Urbana-Champaign



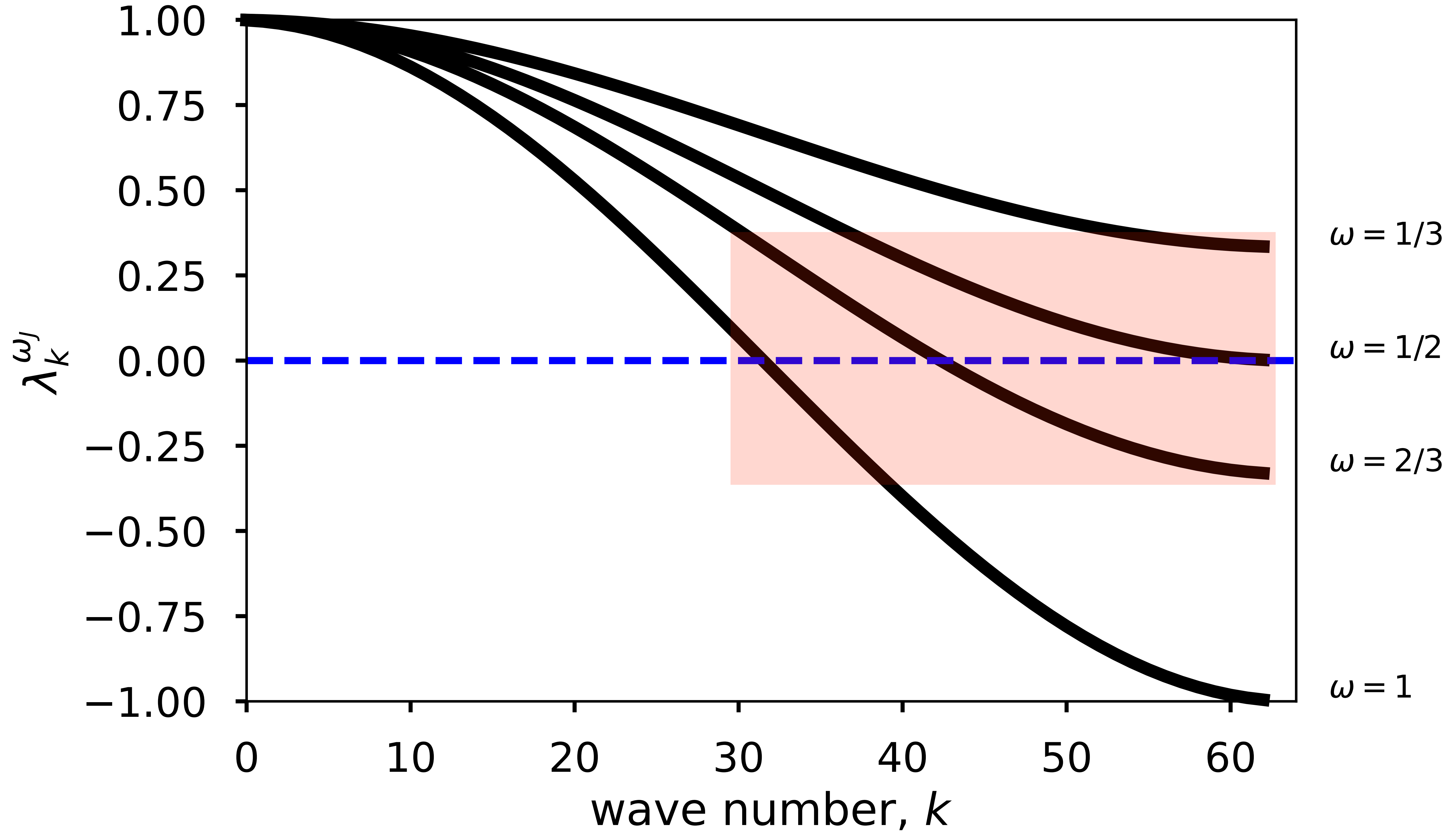
Objectives — high level

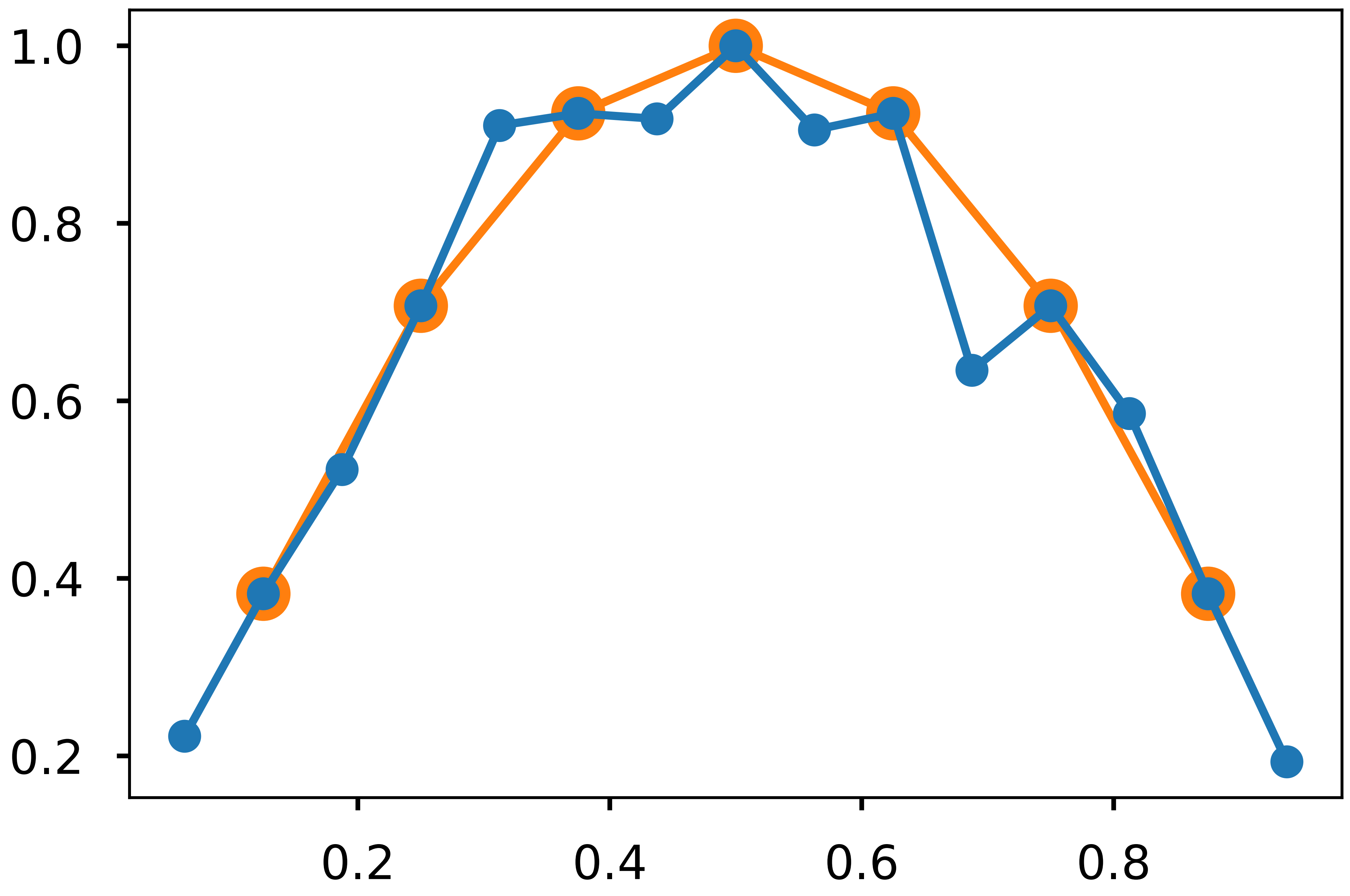
1. **Lecture 1 - Basics**
    - Basic mechanics of a multigrid method
    - 1D, 2D, Poisson
2. **Lecture 2 - Extensions**
    - What can go wrong and how to fix it
3. **Lecture 3 - Algebraic**
    - What to do if we do not have a grid (hierarchy)
4. **Lecture 4 - Some Theory**
    - Why does any of this work

Objectives — today

1. **Lecture 2 - Extensions to multigrid**
    - Illustrate some limitations of multigrid
    - Outline some approaches to fixing multigrid (and why)
    - Identify key pieces in moving to unstructured problems

From last time …





- Smoothing:
  Reduce high frequency error

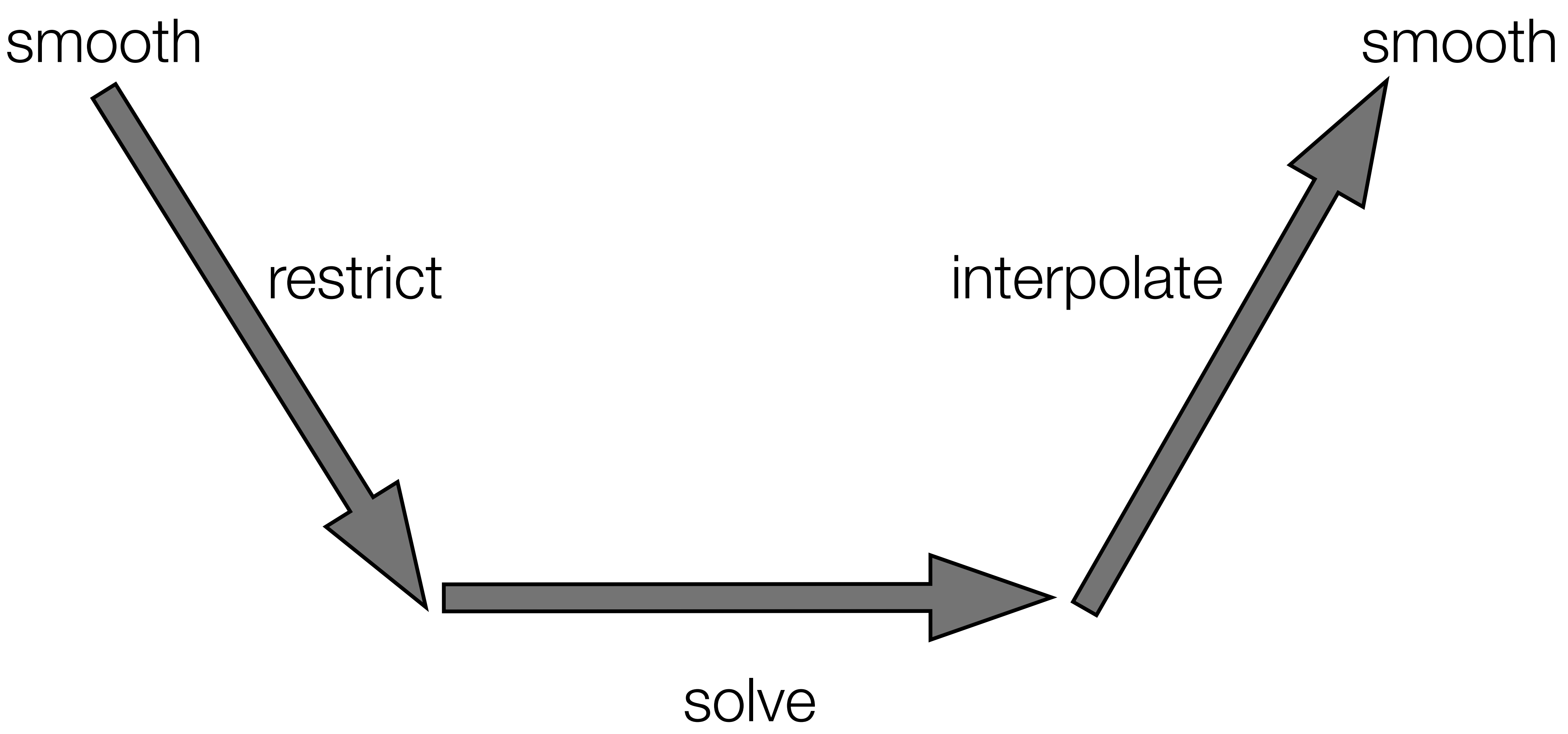
- Coarse-grid correction:
  Reduce smooth-ish things in the range of interpolation

$$e_1 = e_0 - P(P^T A P)^{-1} P^T A G e$$

Algorithm: two-level multigrid

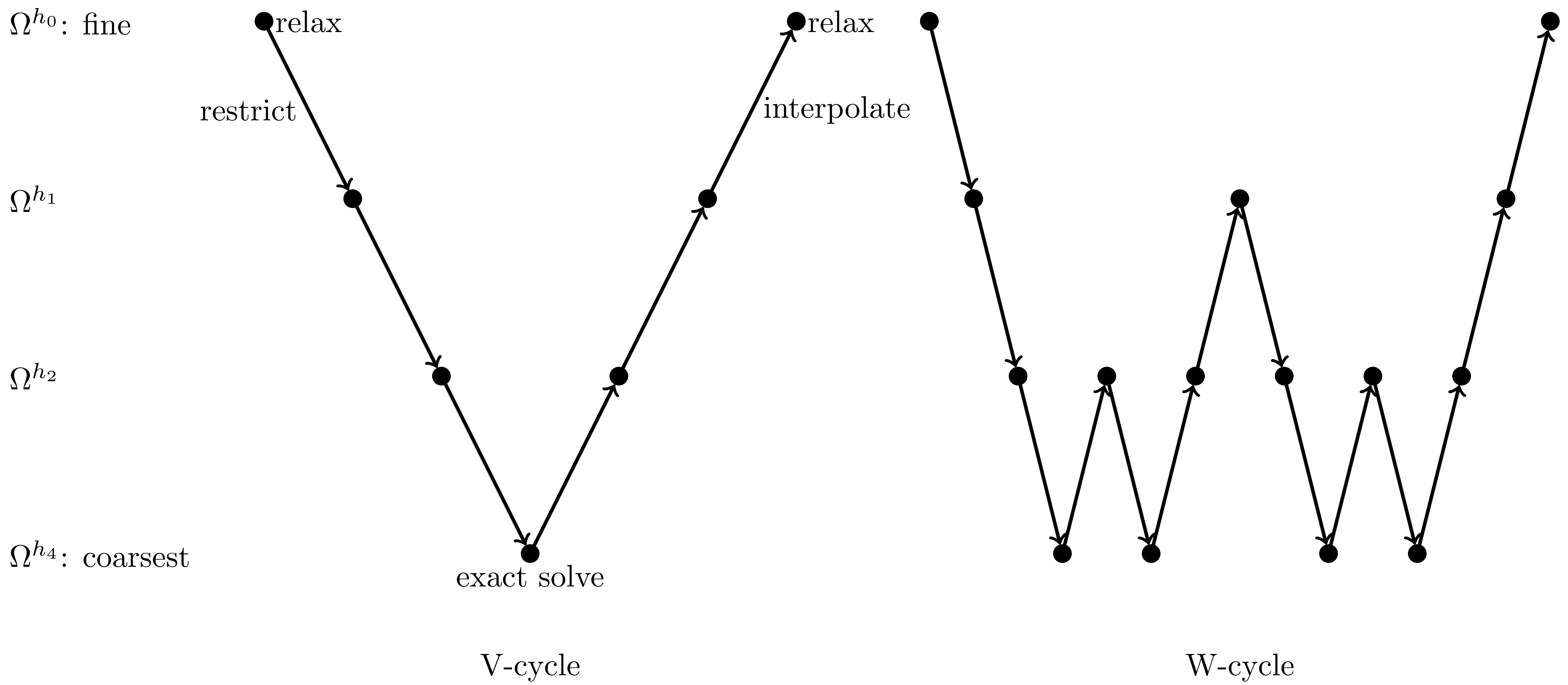
Input: initial guess

1. Smooth ν_{pre} times on $Au = f$
2. Compute $r = f - Au$
3. Compute $r_c = Rr$
4. Solve $A_c e_c = r_c$
5. Interpolate $\hat{e} = Pe_c$
6. Correct $u \leftarrow u + \hat{e}$
7. Smooth ν_{post} times on $Au = f$



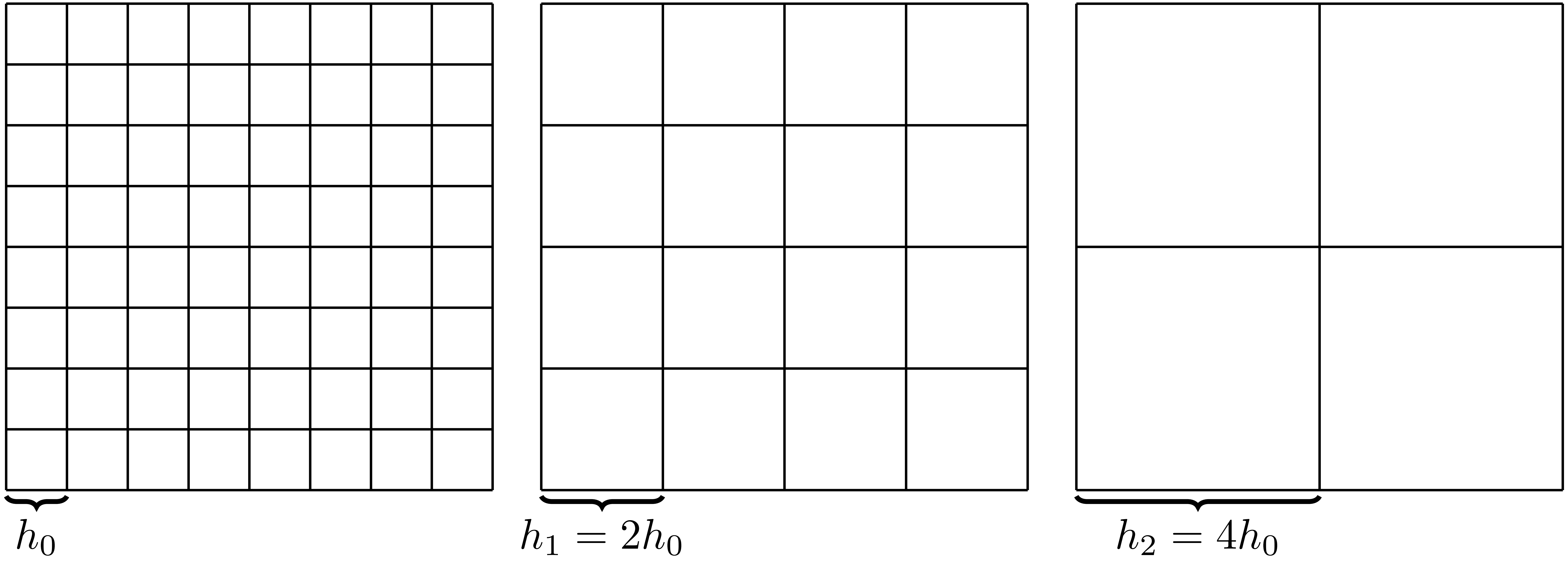
A two-level "V" cycle

The Multigrid V-Cycle and W-Cycle



- Two-grid cycle can expose issues with coarser interpolation
- W-Cycle can account for inadequate coarser level solves
- **Exact solve?** Usually a pseudo-inverse

Multigrid in 2D



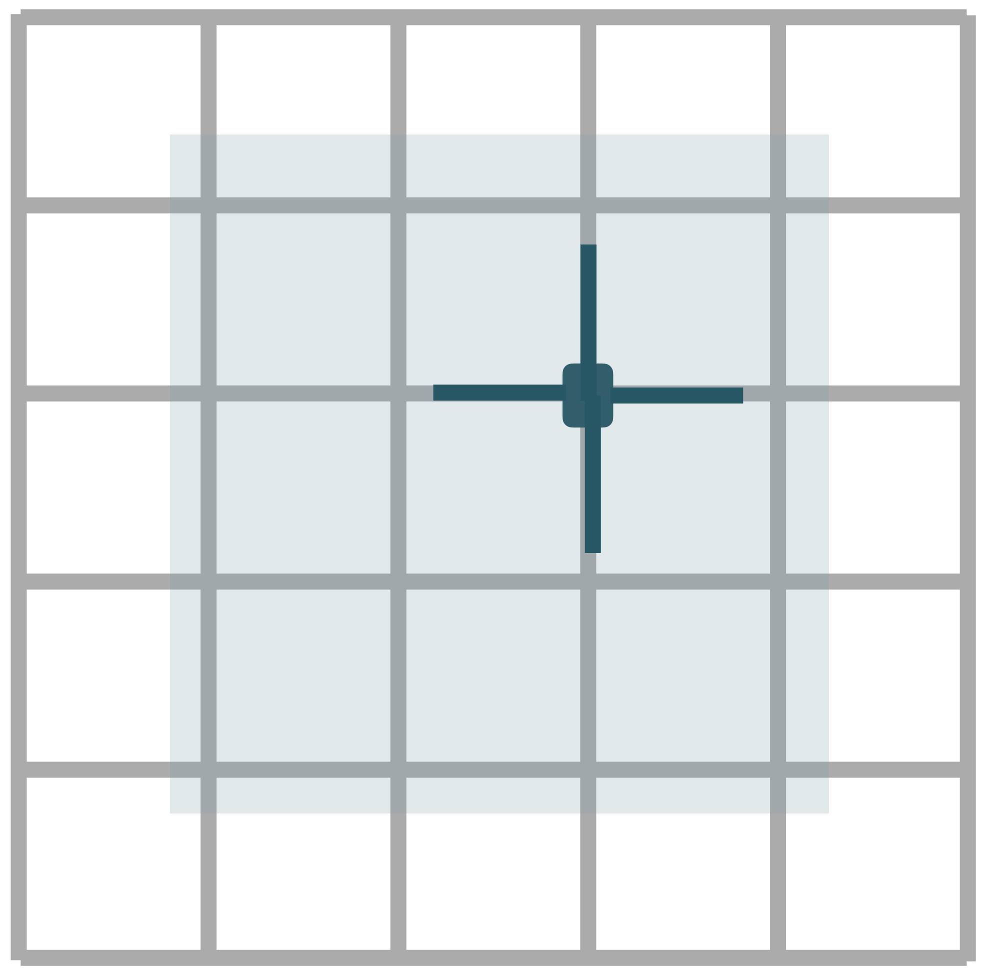
- Again, assume we have a series of uniform grids
- Relaxation remains the same (what is ω?)

Multigrid in 2D

- Model problem

$$
\begin{aligned}
-u_{xx} - u_{yy} &= f \\
u &= 0 \qquad \text{on boundary}
\end{aligned}
$$

- Results in the stencil / matrix

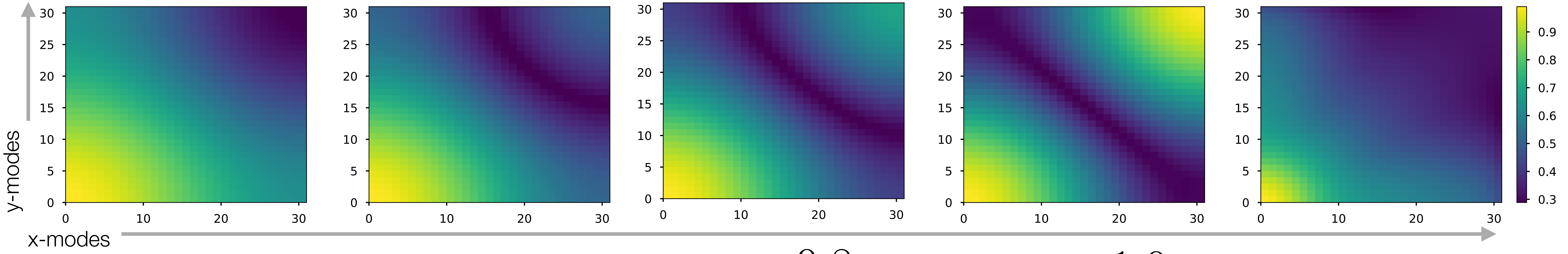


$$
\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix} \longrightarrow \frac{1}{h^2}
\begin{bmatrix}
4 & -1 & & & -1 & & & & & & & & & & & \\
-1 & 4 & -1 & & & -1 & & & & & & & & & & \\
 & -1 & 4 & -1 & & & -1 & & & & & & & & & \\
 & & -1 & 4 & & & & -1 & & & & & & & & \\
-1 & & & & 4 & -1 & & & -1 & & & & & & & \\
 & -1 & & & -1 & 4 & -1 & & & -1 & & & & & & \\
 & & -1 & & & -1 & 4 & -1 & & & -1 & & & & & \\
 & & & -1 & & & -1 & 4 & & & & -1 & & & & \\
 & & & & -1 & & & & 4 & -1 & & & -1 & & & \\
 & & & & & -1 & & & -1 & 4 & -1 & & & -1 & & \\
 & & & & & & -1 & & & -1 & 4 & -1 & & & -1 & \\
 & & & & & & & -1 & & & -1 & 4 & & & & -1 \\
 & & & & & & & & -1 & & & & 4 & -1 & & \\
 & & & & & & & & & -1 & & & -1 & 4 & -1 & \\
 & & & & & & & & & & -1 & & & -1 & 4 & -1 \\
 & & & & & & & & & & & -1 & & & -1 & 4
\end{bmatrix}
$$

Relaxation in 2D

$$\sin\left(\frac{k_x i\pi}{n+1}\right)\sin\left(\frac{k_y j\pi}{n+1}\right)$$

Convergence factor over 10 sweeps



$\omega = 0.5$ $\omega = 0.667$ $\omega = 0.8$ $\omega = 1.0$ GS

$$I - \omega D^{-1}A \qquad\qquad I - (D-E)^{-1}A$$

- weighted Jacobi: Same issue — need to select a parameter
- Gauss-Seidel improved
- Red-Black Gauss-Seidel and other schemes even more effective

Interpolation in 2D

- Bilinear interpolation, tensor of 1D interpolation

$$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

- Example 3x3 -> 7x7

$$\begin{bmatrix} 0.5 & & \\ 1 & & \\ 0.5 & 0.5 & \\ & 1 & \\ & 0.5 & 0.5 \\ & & 1 \\ & & 0.5 \end{bmatrix} \otimes \begin{bmatrix} 0.5 & & \\ 1 & & \\ 0.5 & 0.5 & \\ & 1 & \\ & 0.5 & 0.5 \\ & & 1 \\ & & 0.5 \end{bmatrix} \longrightarrow \frac{1}{4}\begin{bmatrix} 1 & & & & & \\ 2 & & & & & \\ 1 & 1 & & & & \\ & 2 & & & & \\ & 1 & 1 & & & \\ & & 2 & & & \\ & & 1 & & & \\ 2 & & & & & \\ 4 & & & & & \\ 2 & 2 & & & & \\ & 4 & & & & \\ & 2 & 2 & & & \\ & & 4 & & & \\ & & 2 & & & \\ 1 & & & 1 & & \\ 2 & & & 2 & & \\ 1 & 1 & & 1 & 1 & \\ & 2 & & & 2 & \\ & 1 & 1 & & 1 & 1 \\ & & 2 & & & 2 \\ & & 1 & & & 1 \\ & & & 2 & & \\ & & & 4 & & \\ & & & 2 & 2 & \\ & & & & 4 & \\ & & & & 2 & 2 \\ & & & & & 4 \\ & & & & & 2 \\ & \vdots & & \vdots & & \end{bmatrix}$$

Luke Olson
http://lukeo.cs.illinois.edu/

Notebook

- `9-multigrid-2d.ipynb`

A few observations so far: **One**

- Let's consider a V(1,1) cycle — weighted Jacobi, etc.

- The **error** propagation for this looks like

$$e_1 = \underbrace{G(I - P(P^T A P)^{-1} P^T A)G}_{M} e_0$$

$$G = I - \omega D^{-1} A$$

$$M e_k \leftarrow 0?$$

- One thing we can do, is consider bounds on each operation.

A few observations so far: **One**

- Take the operator

$$M = G(I - P(P^T AP)^{-1} P^T A)G$$

- And makes some bounds

$$\underbrace{\|I - P(P^T AP)^{-1} P^T A\|}_{\text{approximation property}} \underbrace{\|G\|^2}_{\text{smoothing property}}$$

- General $\|G\| \leq 1$

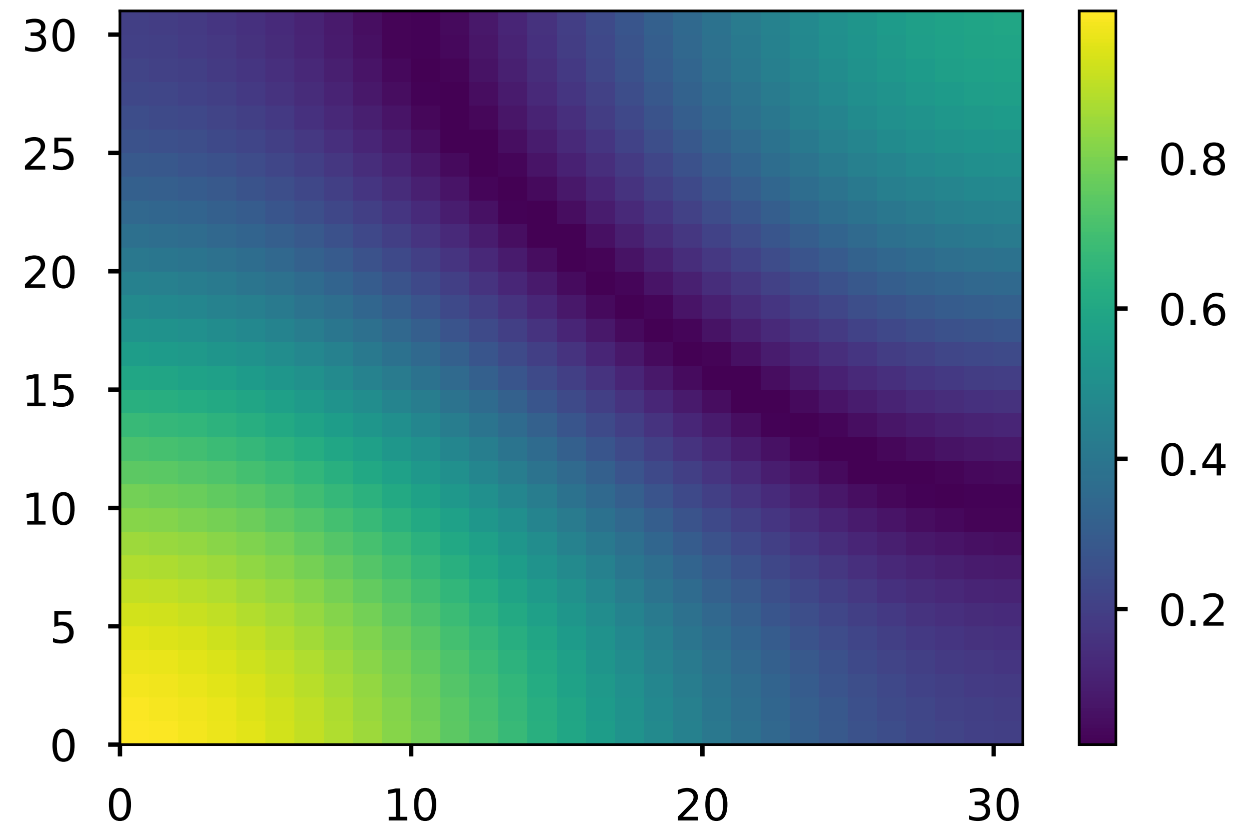
1D over [n/2,n]: $\|G\| \leq \dfrac{1}{3}$

2D over [n/2,n]: $\|G\| \leq \dfrac{3}{5}$

Luke Olson
http://lukeo.cs.illinois.edu/

A few observations so far: **One**

$$\underbrace{\|I - P(P^TAP)^{-1}P^TA\|}_{\text{approximation property}}\underbrace{\|G\|^2}_{\text{smoothing property}}$$



- General $\|G\| \leq 1$

1D over [n/2,n]: $\|G\| \leq \frac{1}{3}$

2D over [n/2,n]: $\|G\| \leq \frac{3}{5}$

- Also, if $w \quad \text{s.t.} \quad Aw \in \mathcal{N}(P^T)$

then $(I - P(P^TAP)^{-1}P^TA)w = w$

$$\|\cdot\| \geq 1$$

Luke Olson
http://lukeo.cs.illinois.edu/

A few observations so far: **Two**

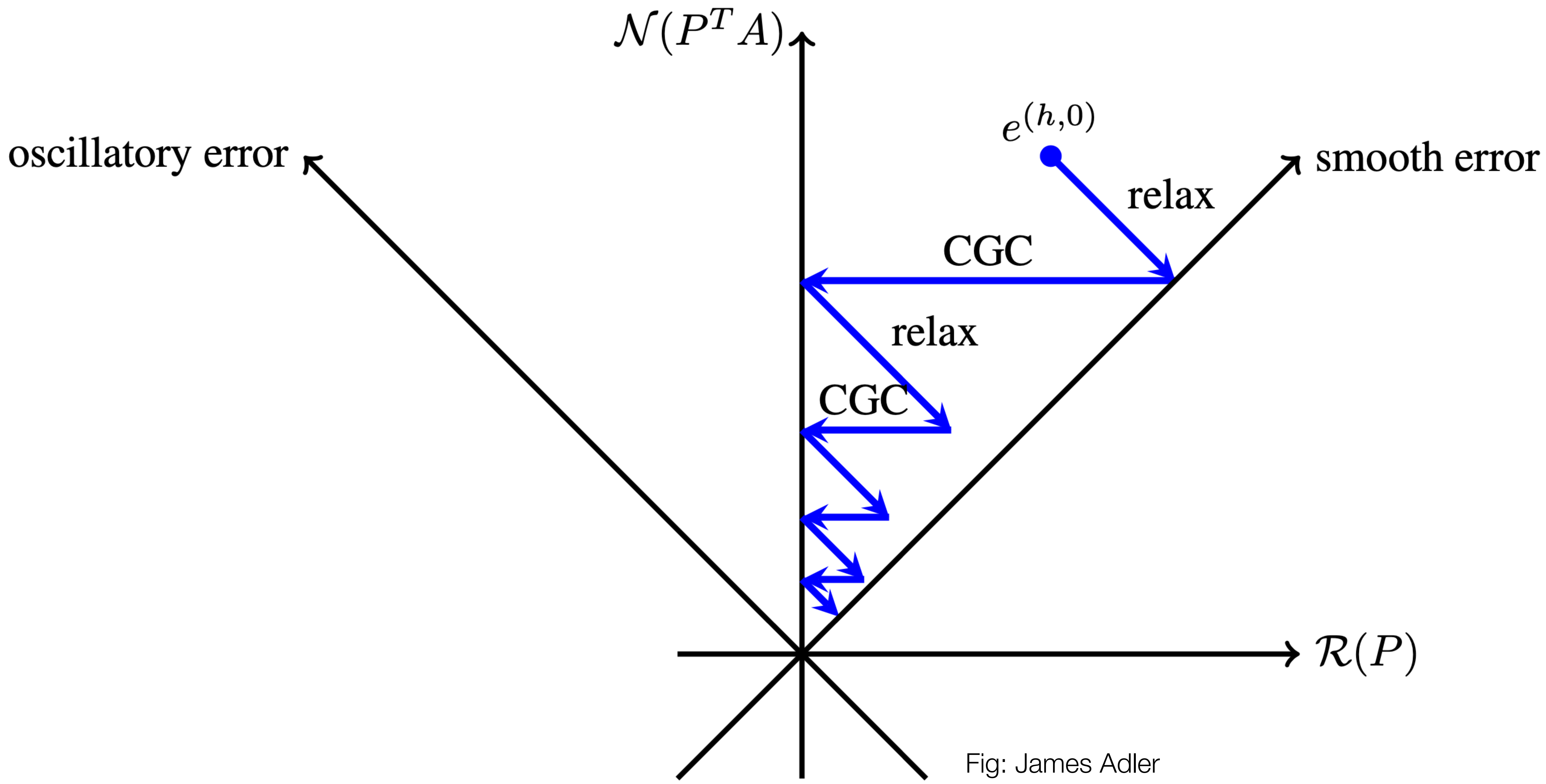


Fig: James Adler

- Complementary processes:
  - **relaxation:** targets for (Fourier) smoothing
  - **coarse grid correction:** targets things in the range of interpolation

A few observations so far: **Three**

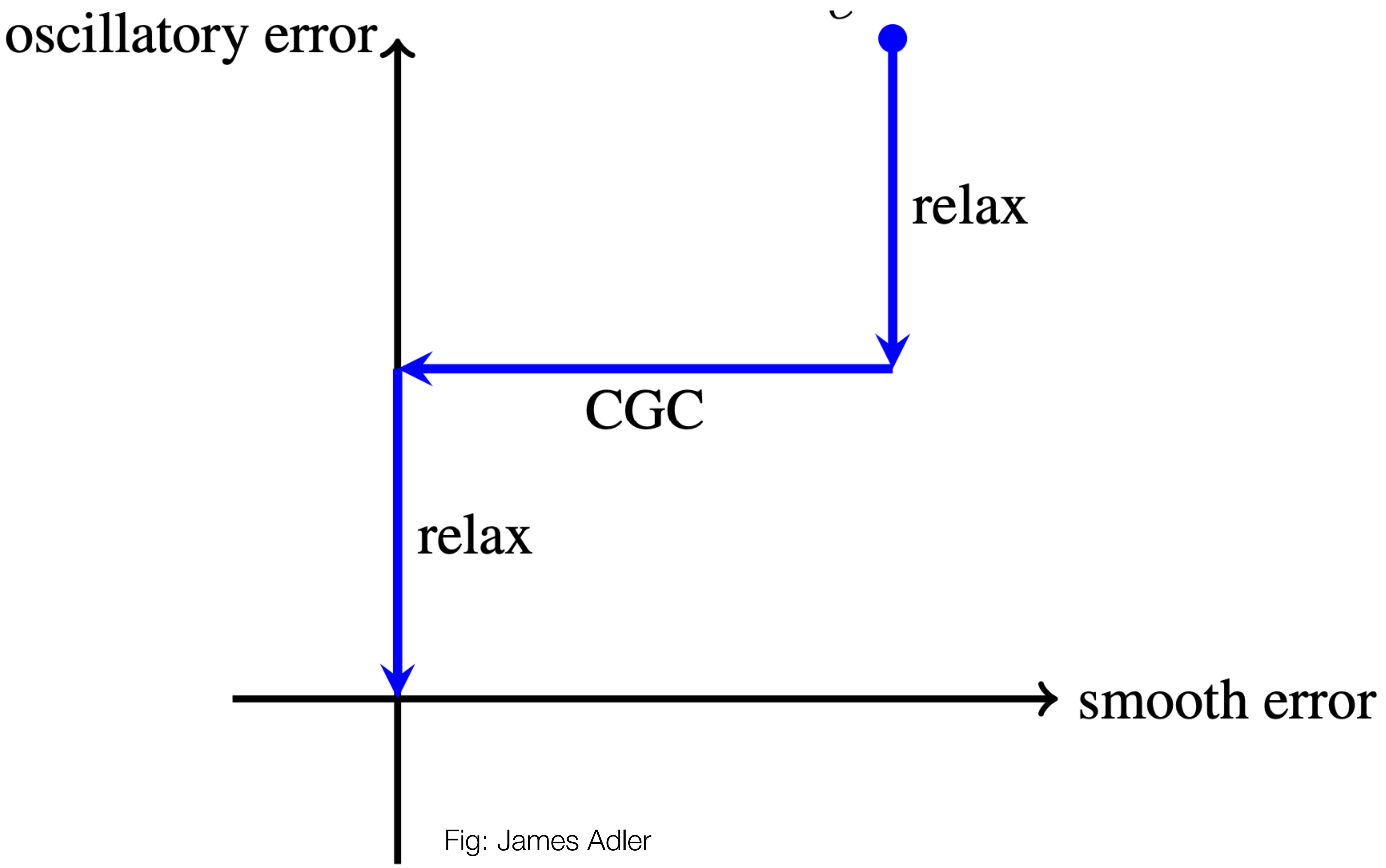


Fig: James Adler

$$e_1 \leftarrow (I - P(P^T A P)^{-1} P^T A) G e_0$$

$$G e_0 \in \mathcal{R}(P) \quad \Rightarrow \quad e_1 = 0$$

interpolation should capture what relaxation misses

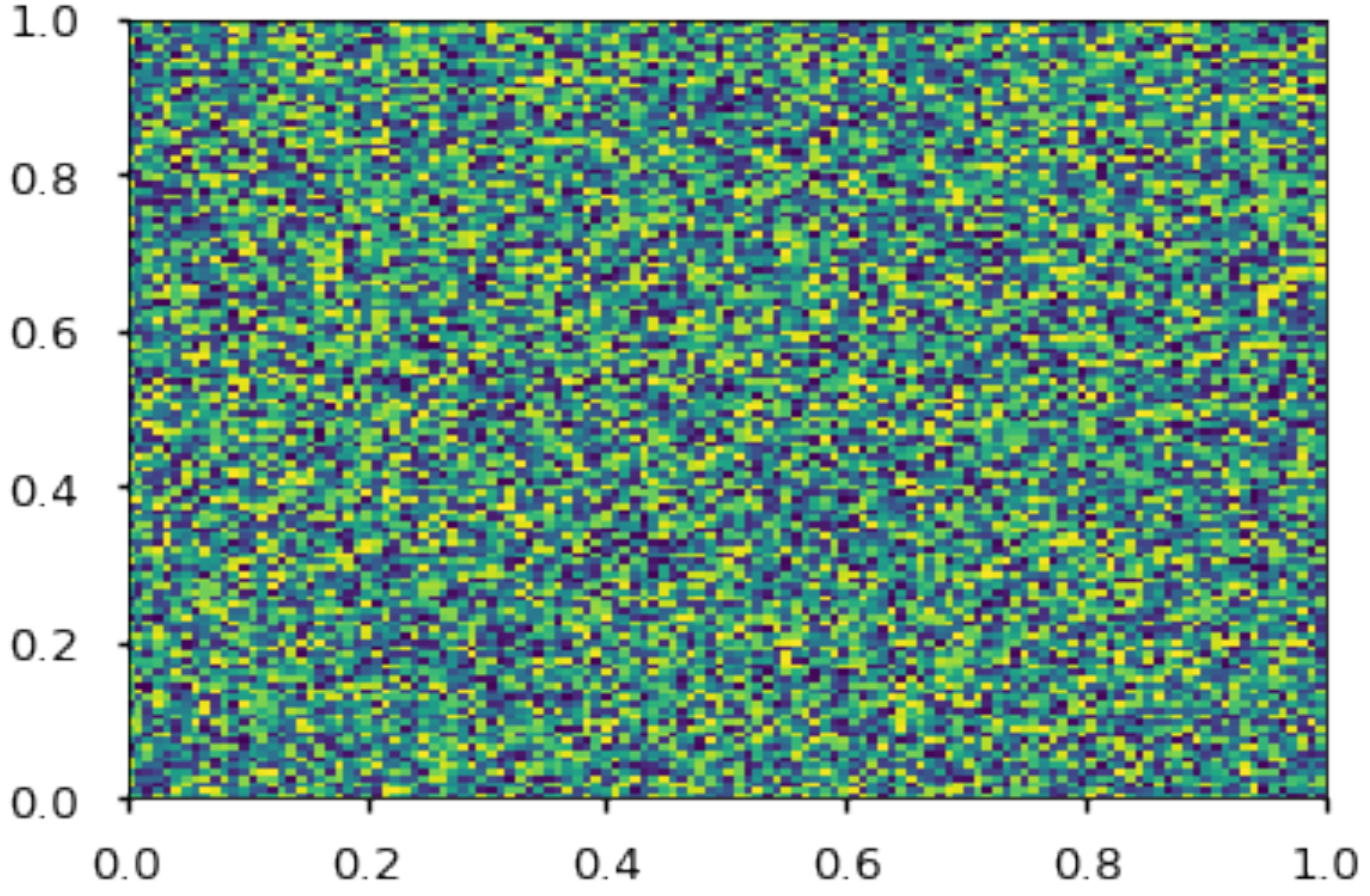
What can go wrong?!

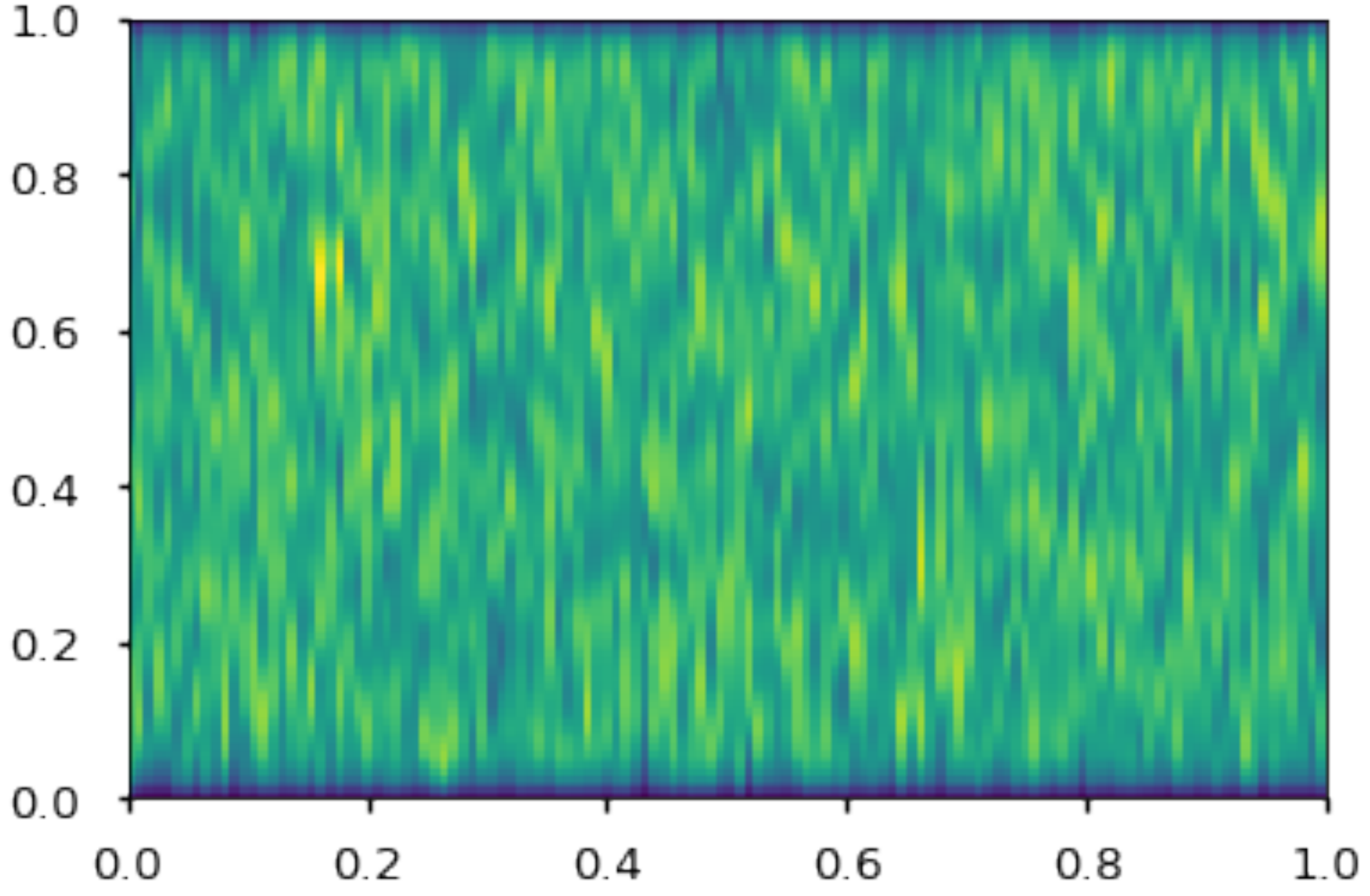
- Demo: `10-multigrid-anisotropy.ipynb`

Consider an *anisotropic problem*

$$-\varepsilon u_{xx} - u_{yy} = f$$
$$u = 0 \qquad \text{on boundary}$$

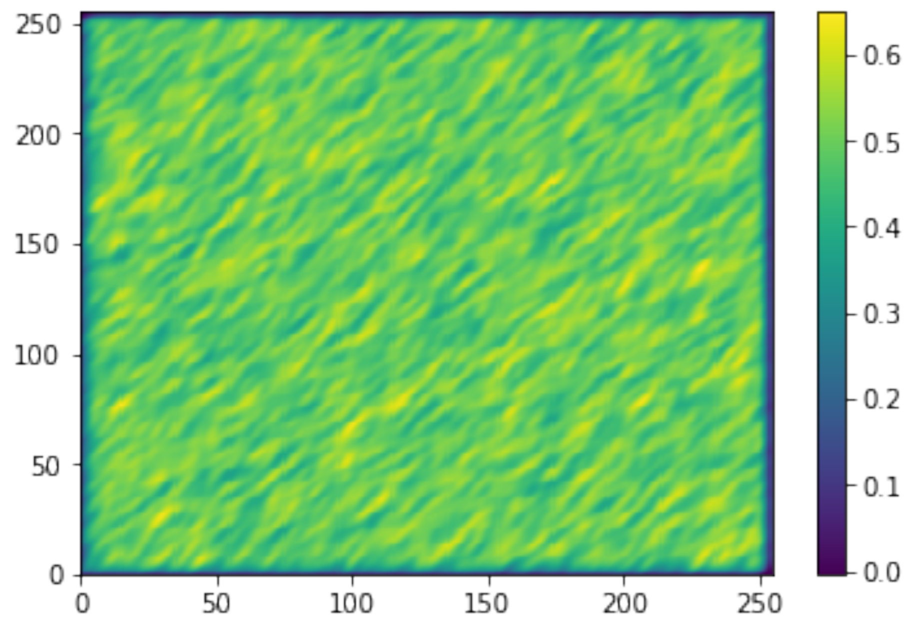
$$\begin{bmatrix} & -1 & \\ -\varepsilon & 2+2\varepsilon & -\varepsilon \\ & -1 & \end{bmatrix}$$

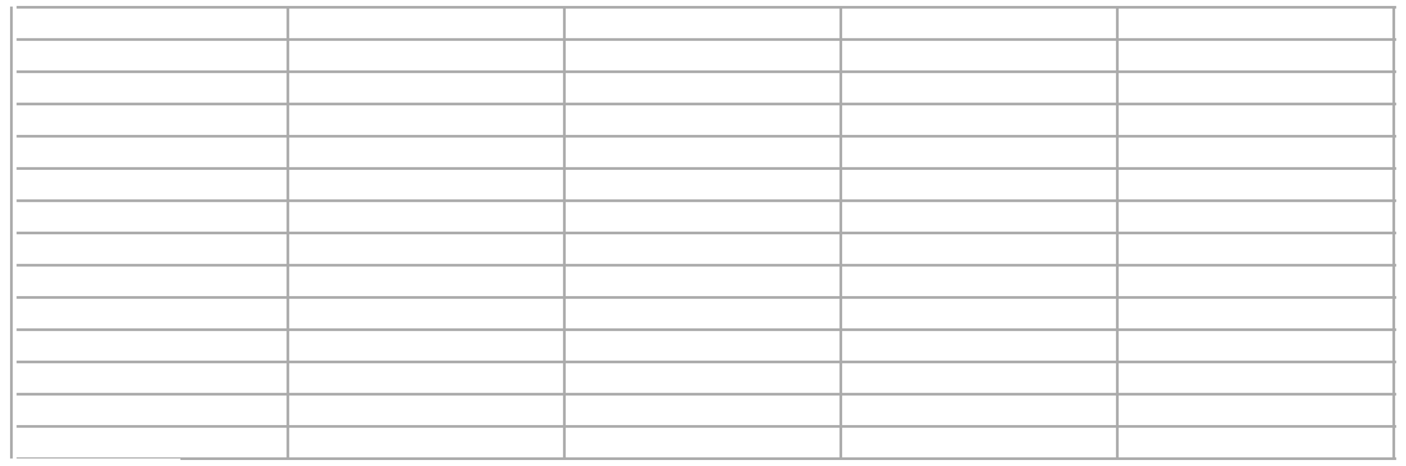


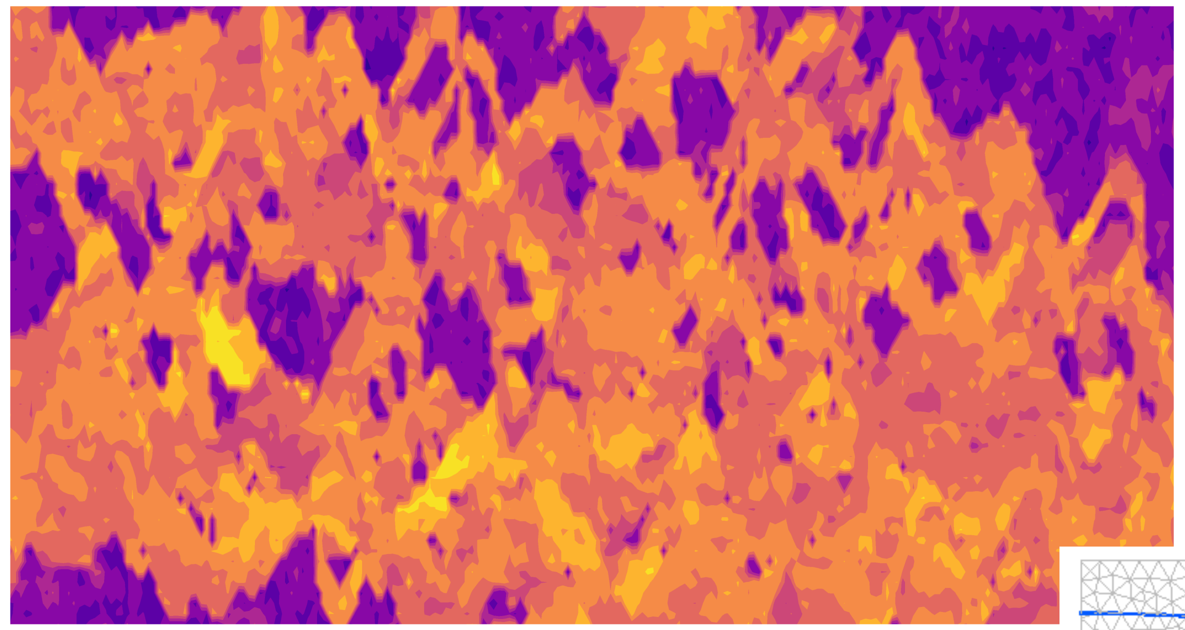


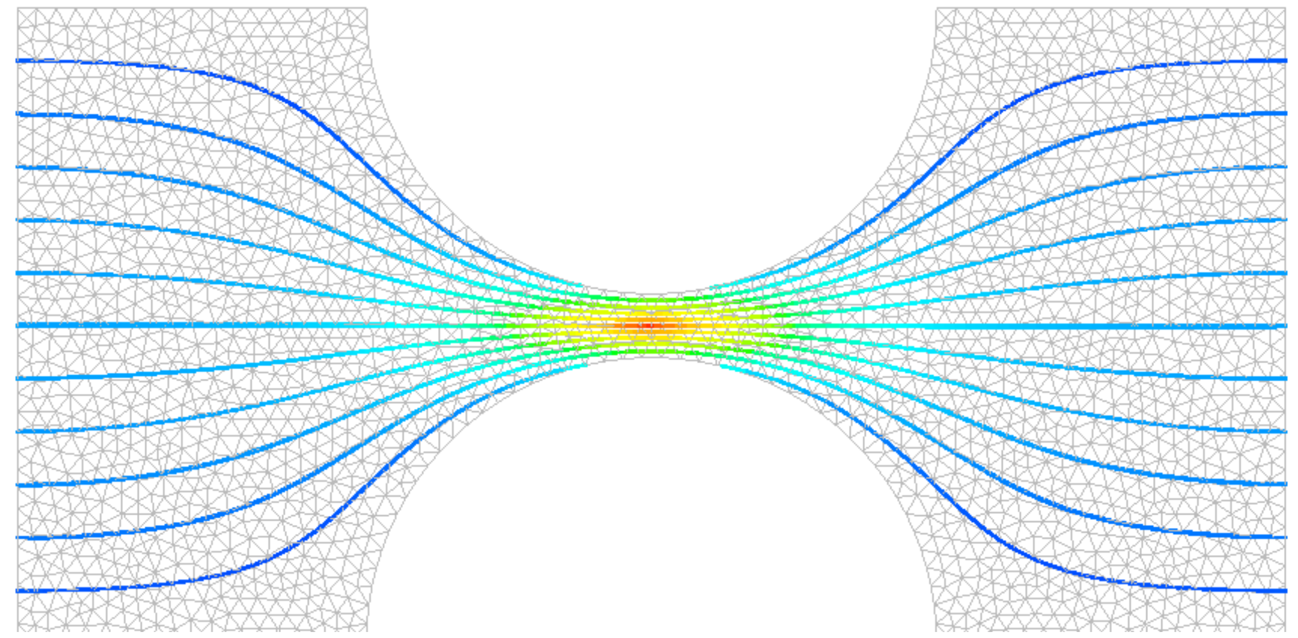
What can go wrong?!

- Anisotropy
- Mesh stretching
- Jumping coefficients
- Non-elliptic



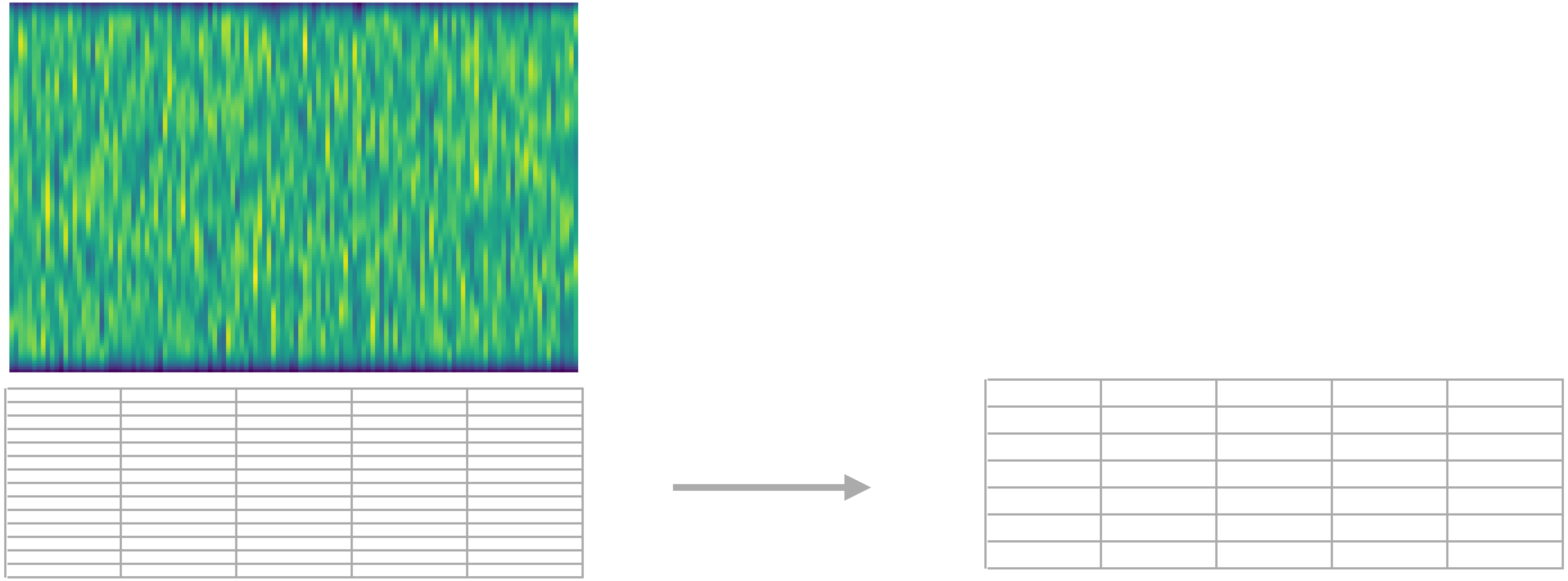






Options for more robust Multigrid

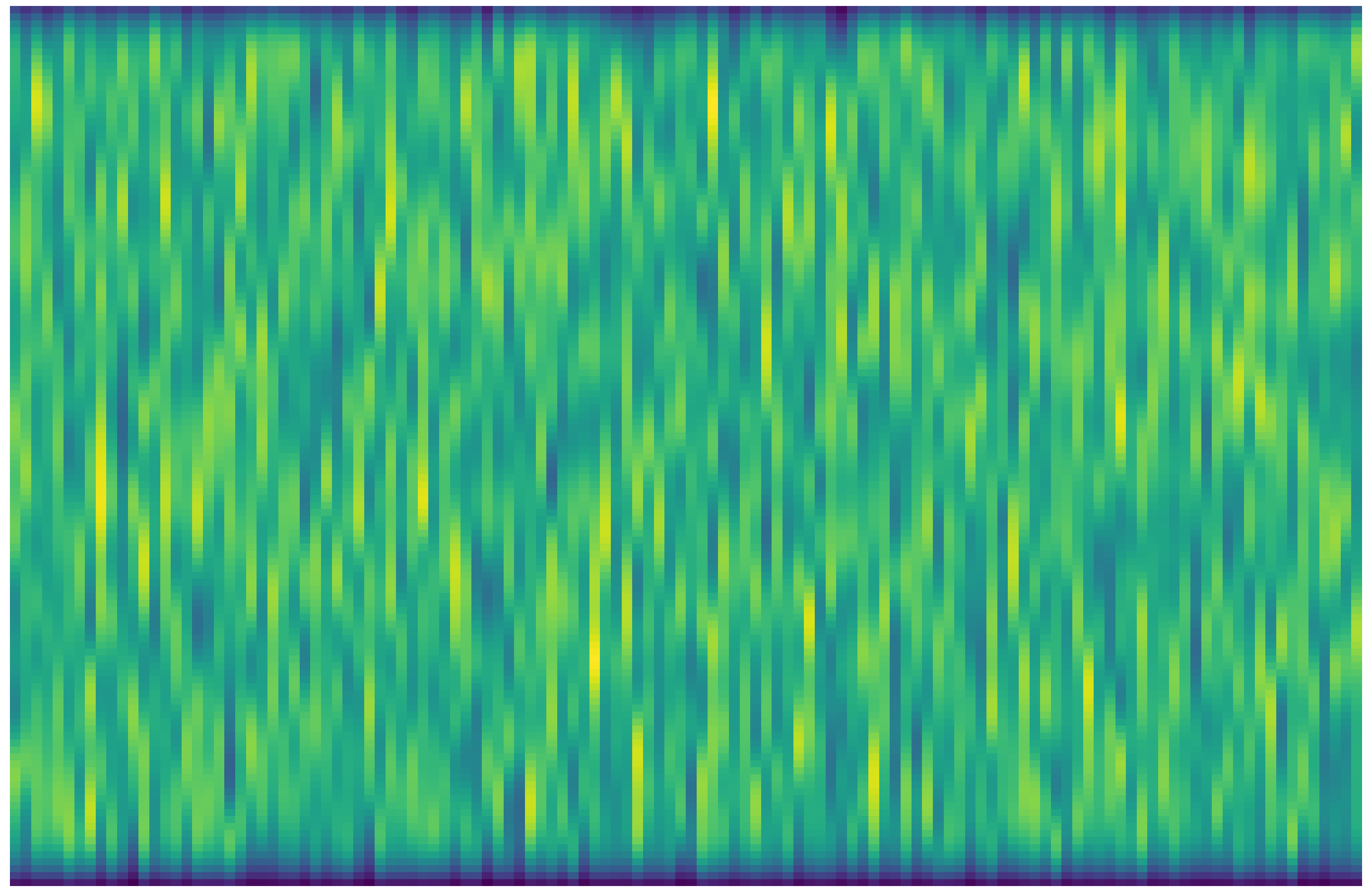
- **Semicoarsening:** Coarsen in the direction of smoothness

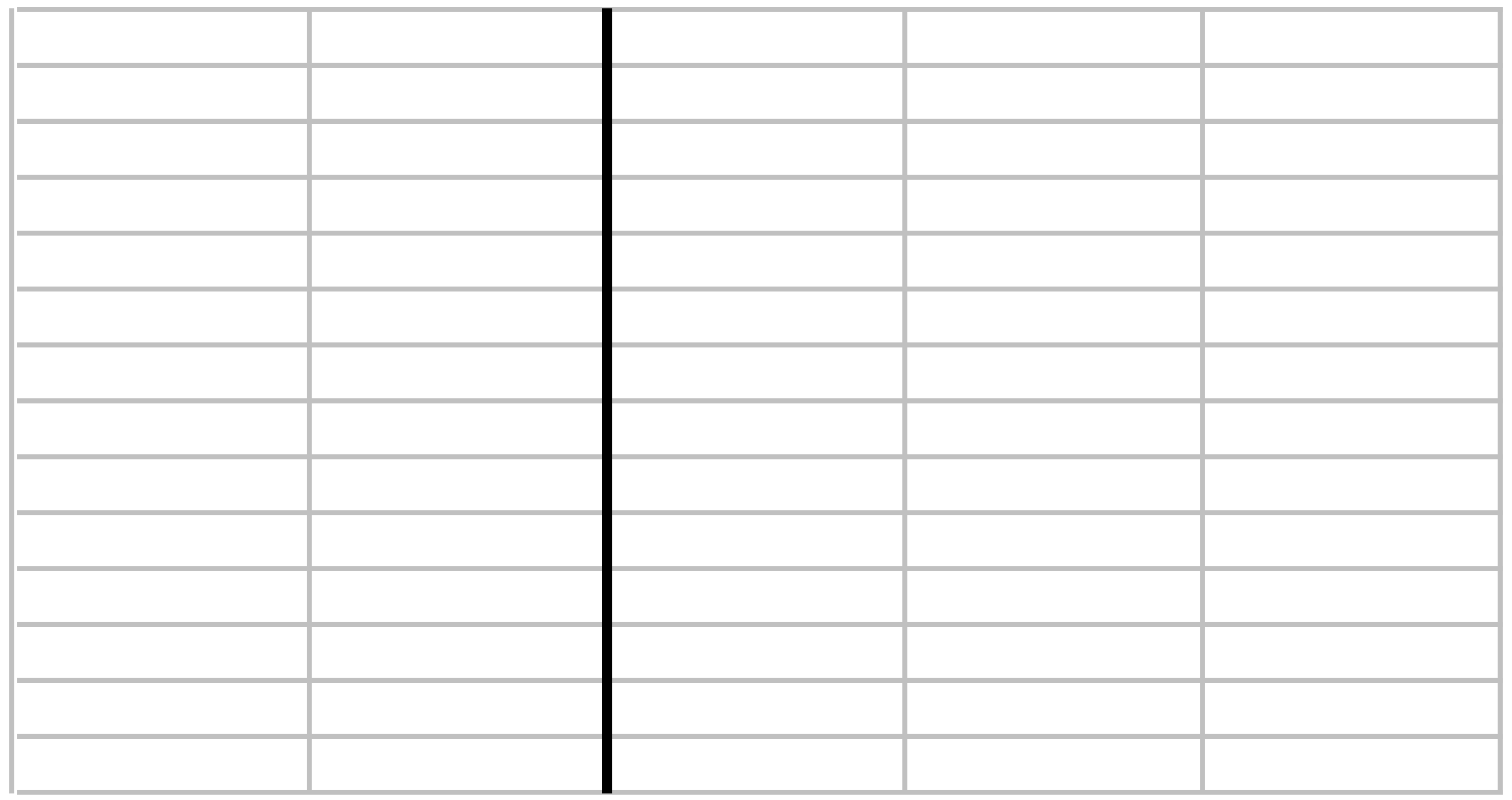


- Downside: if anisotropy varies in another direction, we need a different grid

Options for more robust Multigrid

- **Line/plane relaxation**
  Perform relaxation in groups (in a line)

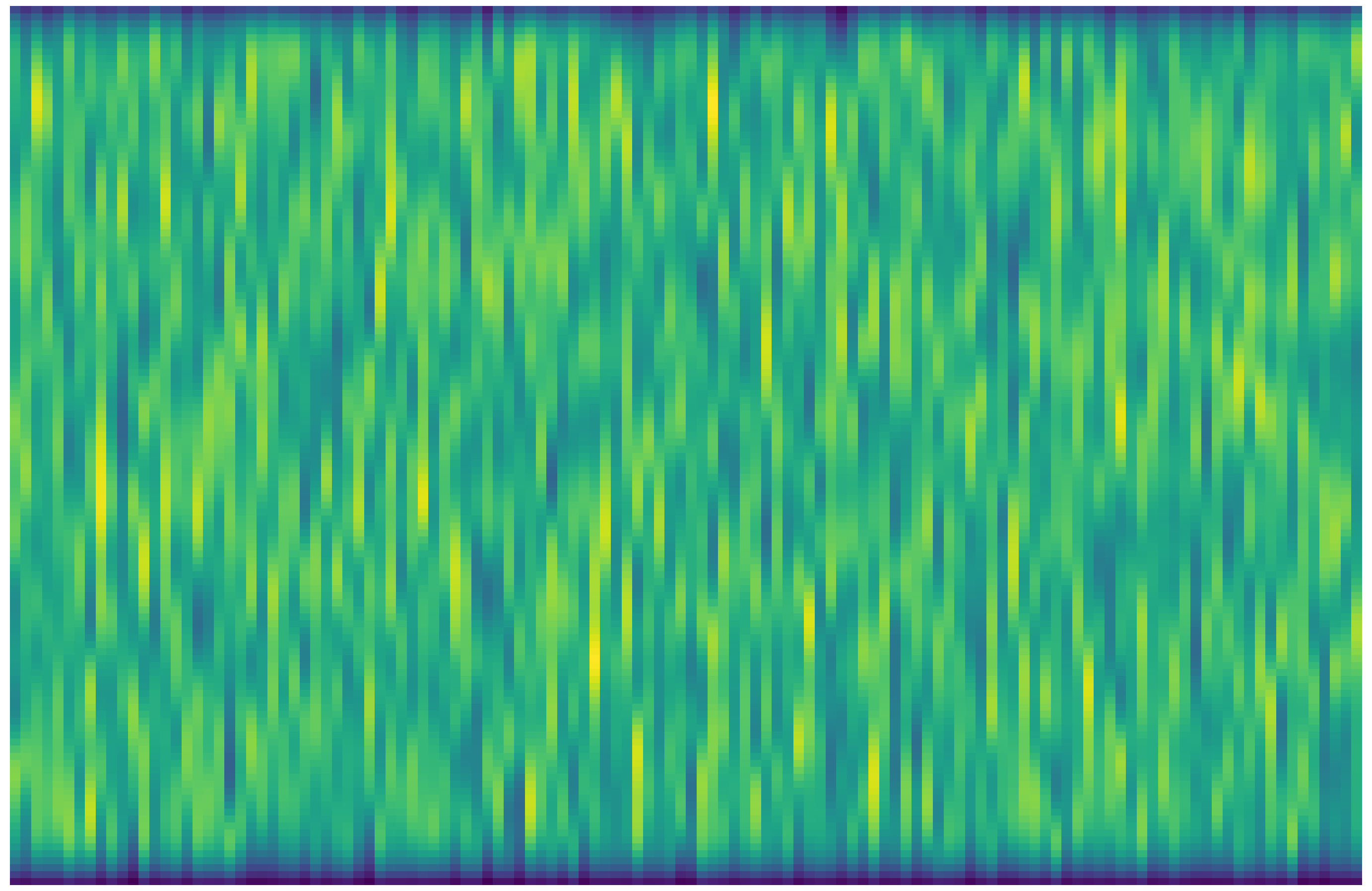


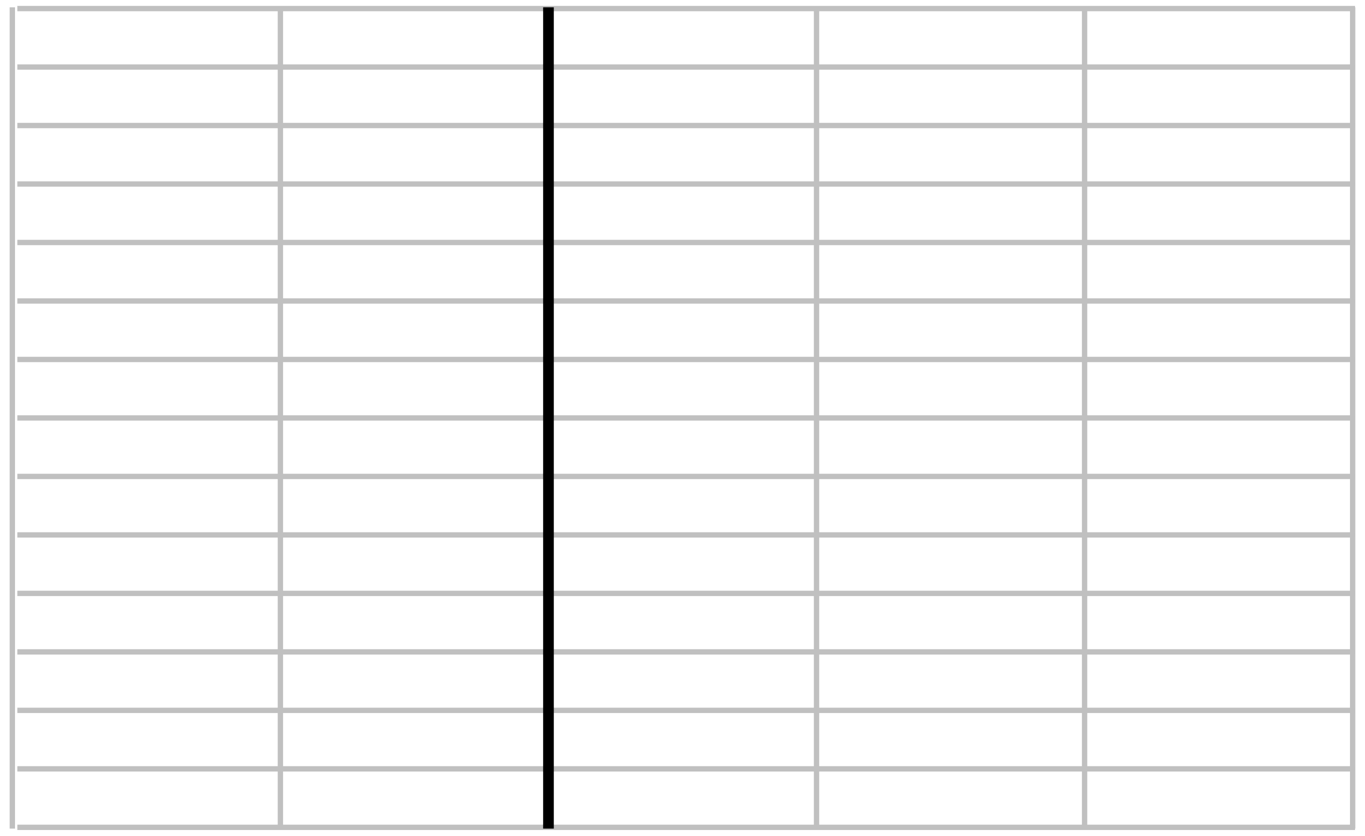


- Example:
    - smooth error in the y-direction (x-lines)
    - no smoothing in the x-direction (y-lines)

Options for more robust Multigrid

- **Line/plane relaxation**
- For each x-line (lines of strong anisotropy):
  - Eliminate the residual on the entire line
  - (Gauss-Siedel, by lines)



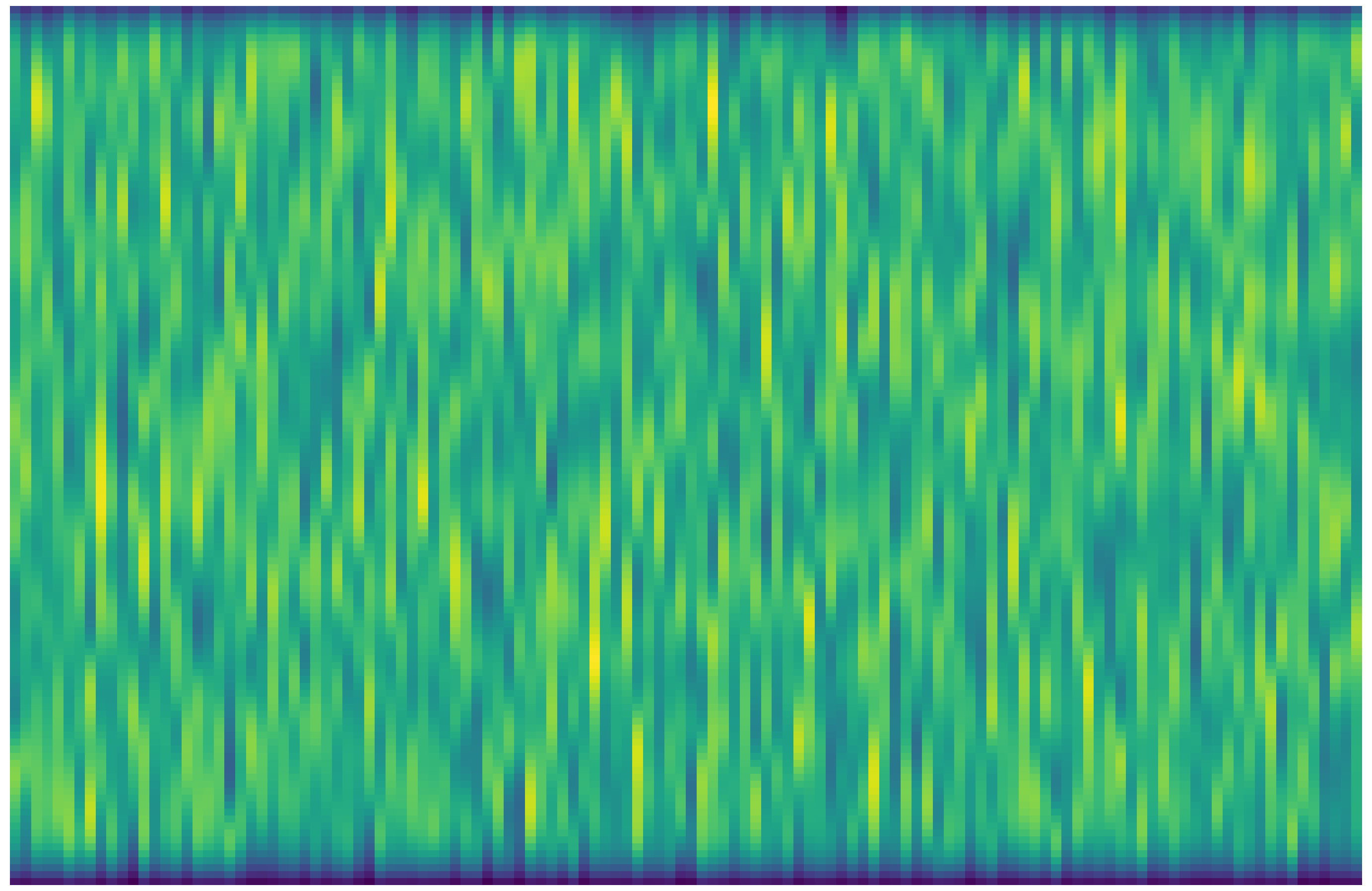


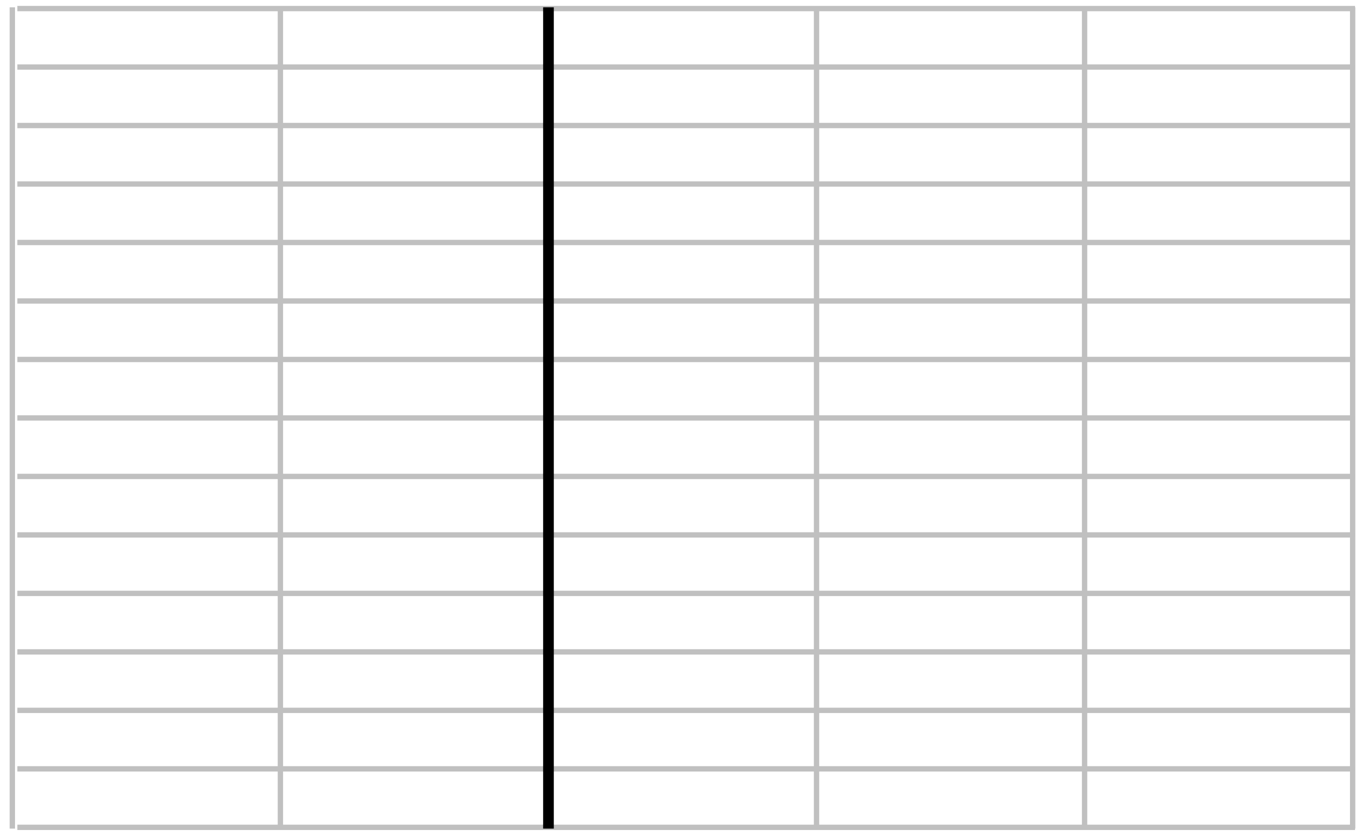
Options for more robust Multigrid

- **Line/plane relaxation**
-

$$A = (n+1)^2 \begin{bmatrix} A_{1D} & -\varepsilon I & & & \\ -\varepsilon I & A_{1D} & -\varepsilon I & & \\ & -\varepsilon I & A_{1D} & -\varepsilon I & \\ & & \ddots & & \\ & & & -\varepsilon I & A_{1D} \end{bmatrix}$$

$$A_{1D} = (n+1)^2 \begin{bmatrix} 2+2\varepsilon & -1 & & \\ -1 & 2+2\varepsilon & -1 & \\ & & \ddots & \\ & & -1 & 2+2\varepsilon \end{bmatrix}$$





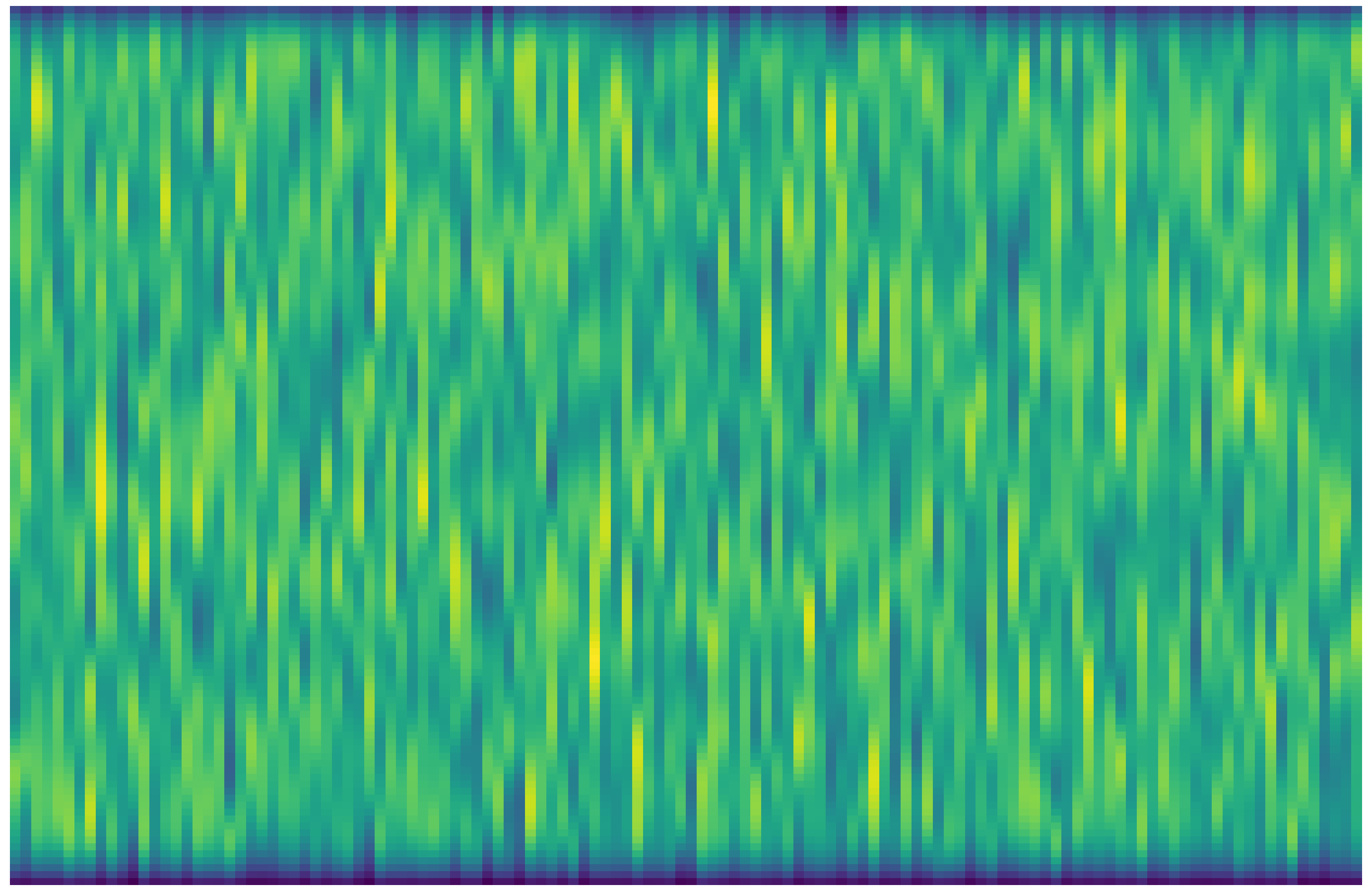
Options for more robust Multigrid

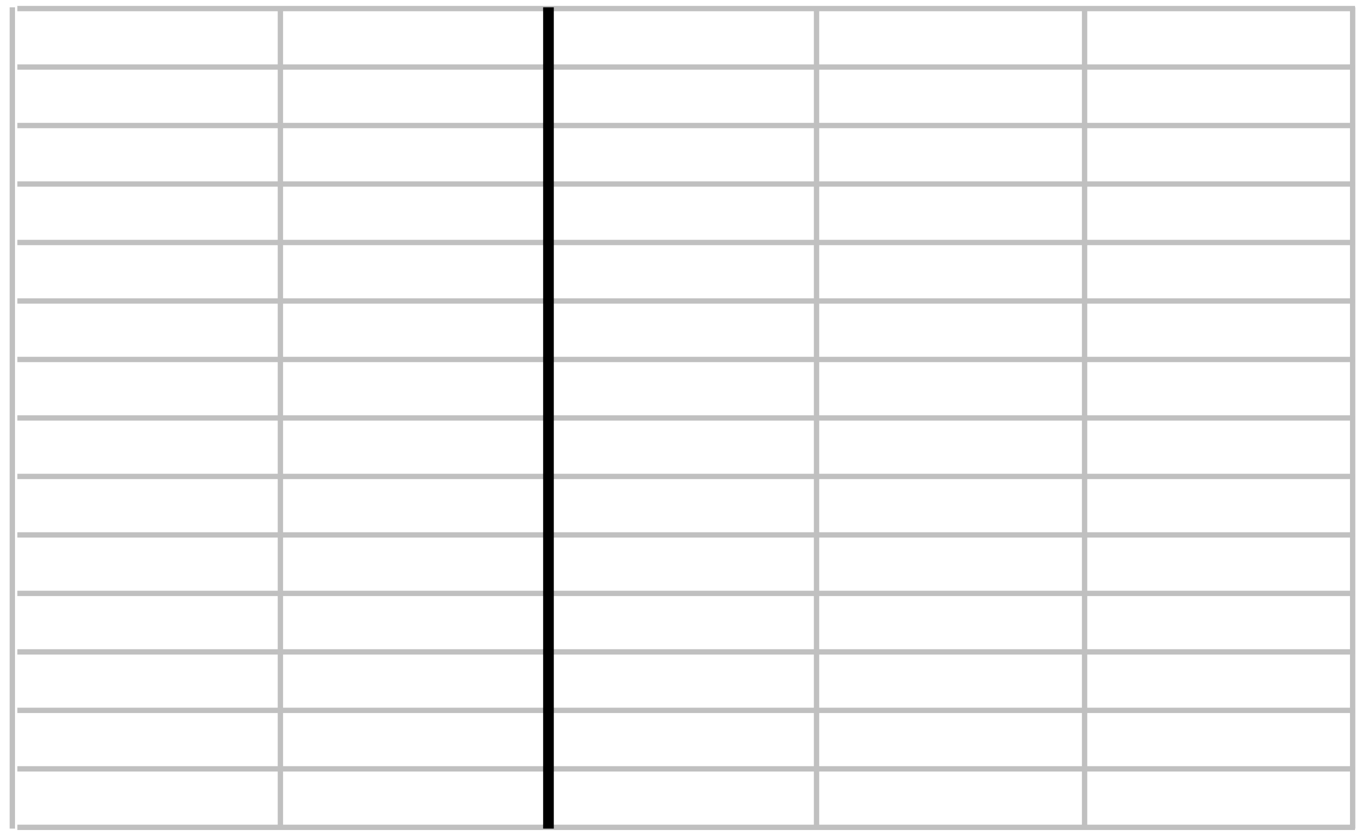
- **Line/plane relaxation**
- For each line, solve

$$A_{1D} v_k = g_k$$

$$v_k = \begin{bmatrix} \vdots \\ v_{k,j-1} \\ v_{k,j} \\ v_{k,j+1} \\ \vdots \end{bmatrix} \quad g_k = \begin{bmatrix} \vdots \\ f_{k,j-1} + \varepsilon\left(v_{k-1,j-1} + v_{k+1,j-1}\right) \\ f_{k,j} + \varepsilon\left(v_{k-1,j} + v_{k+1,j}\right) \\ f_{k,j+1} + \varepsilon\left(v_{k-1,j+1} + v_{k+1,j+1}\right) \\ \vdots \end{bmatrix}$$

- In 3D, lines become planes…



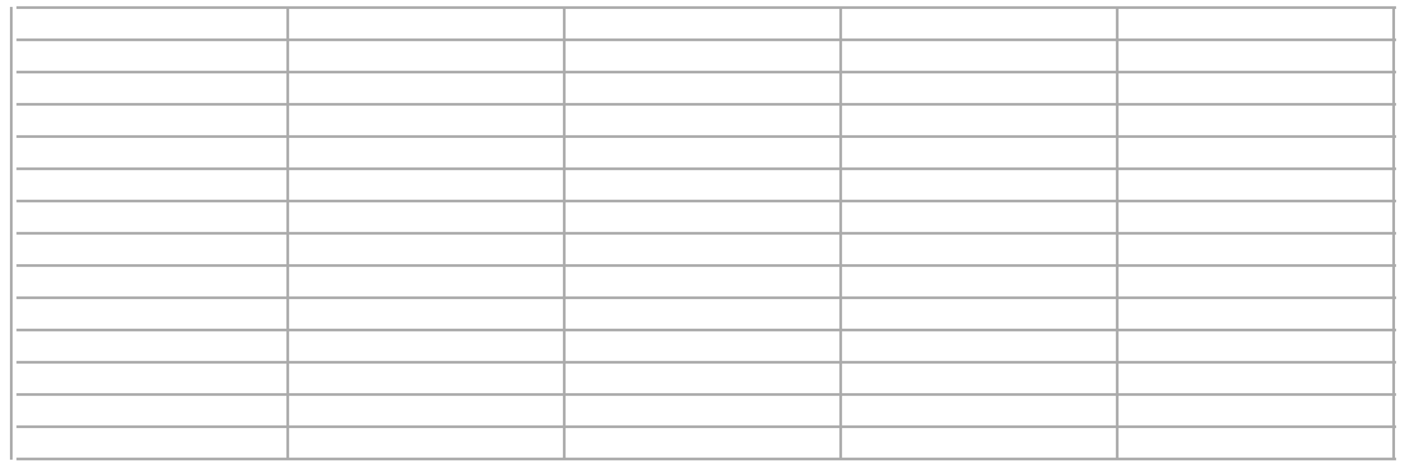
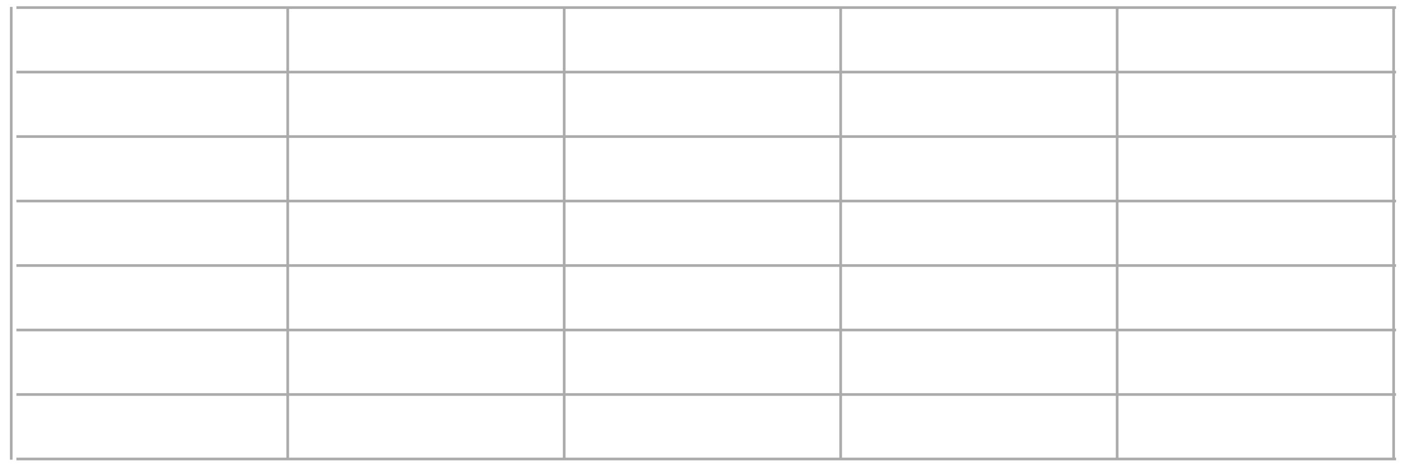


Notebook

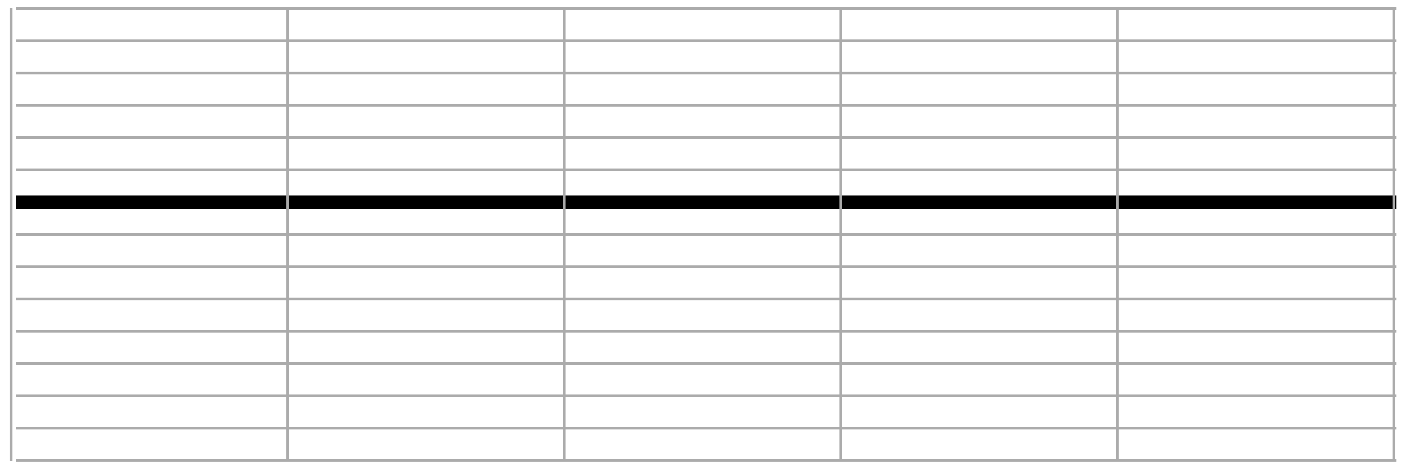
- `11-multigrid-2d-line-relaxation.ipynb`

Options for more robust Multigrid

- **Semicoarsening**
  Coarsen in the direction of smoothness

- **Line/plane relaxation**
  Perform relaxation in groups (in a line)



- **Operator based Interpolation** (e.g. BoxMG)

$$Ae = 0$$

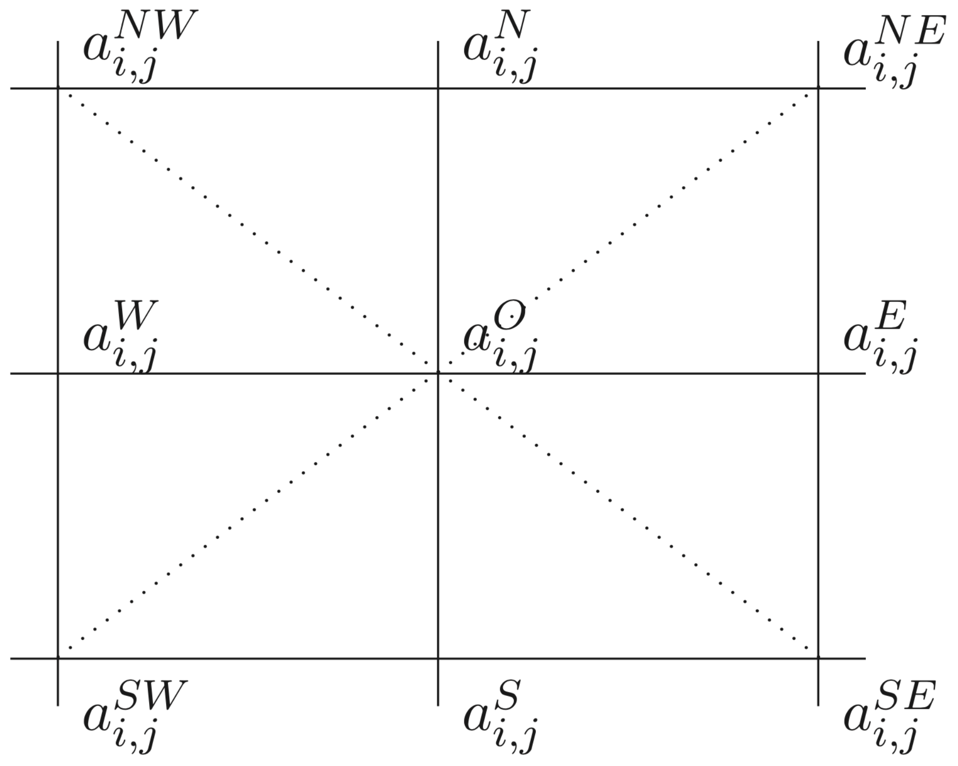
J. E. Dendy, Black box multigrid, J. Comput. Phys., 1982

J. E. Dendy and J. D. Moulton, Black box multigrid with coarsening by a factor of three, J. Numer. Lin. Alg. App., 2010

BoxMG

Math and figures from:
Robust and adaptive multigrid methods: comparing structured and algebraic approaches, MacLachlan, Moulton, Chartier

- Node (i,j) stencil:

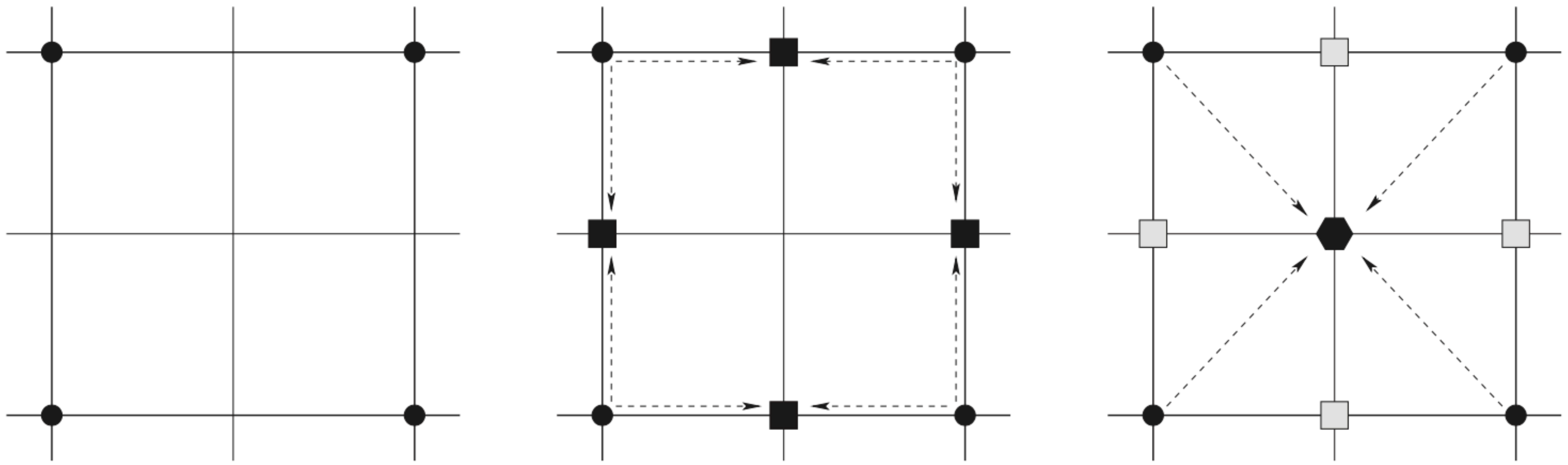


$$Ax = b$$

$$a_{i,j}^{SW} x_{i-1,j-1} + a_{i,j}^{S} x_{i,j-1} + a_{i,j}^{SE} x_{i+1,j-1} + a_{i,j}^{W} x_{i-1,j} + a_{i,j}^{O} x_{i,j} + a_{i,j}^{E} x_{i+1,j} + a_{i,j}^{NW} x_{i-1,j+1} + a_{i,j}^{N} x_{i,j+1} + a_{i,j}^{NE} x_{i+1,j+1} = b_{i,j}$$

BoxMG

Math and figures from:
Robust and adaptive multigrid methods: comparing structured and algebraic approaches, MacLachlan, Moulton, Chartier



1. Inject coarse points (left)
2. Assume the error is constant along x-lines (and y-lines)
3. Infer interpolation from the edges (right)

Example: Assuming Ae=0 and contant y-lines:

$$(a^S + a^O + a^N)e = -(a^{SW} + a^W + a^{NW})e^c$$
$$- (a^{SE} + a^E + a^{NW})e^c$$

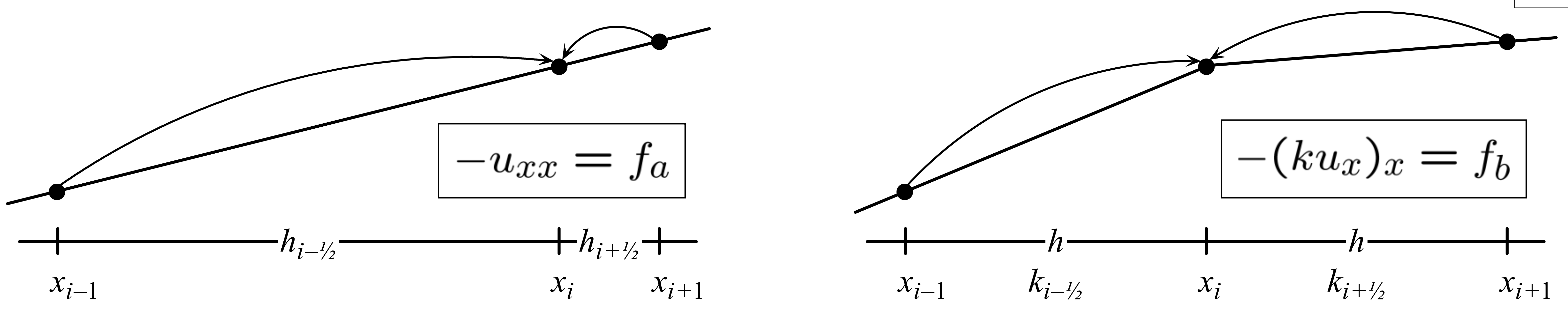
Interpolation based on entries in A

Algebraic Multigrid (AMG) uses matrix coefficients

- Geometric information alone is not sufficient

Linear Interpolation Operator-Dependent Interpolation

Slide credit: R. Falgout, 2/1/2016

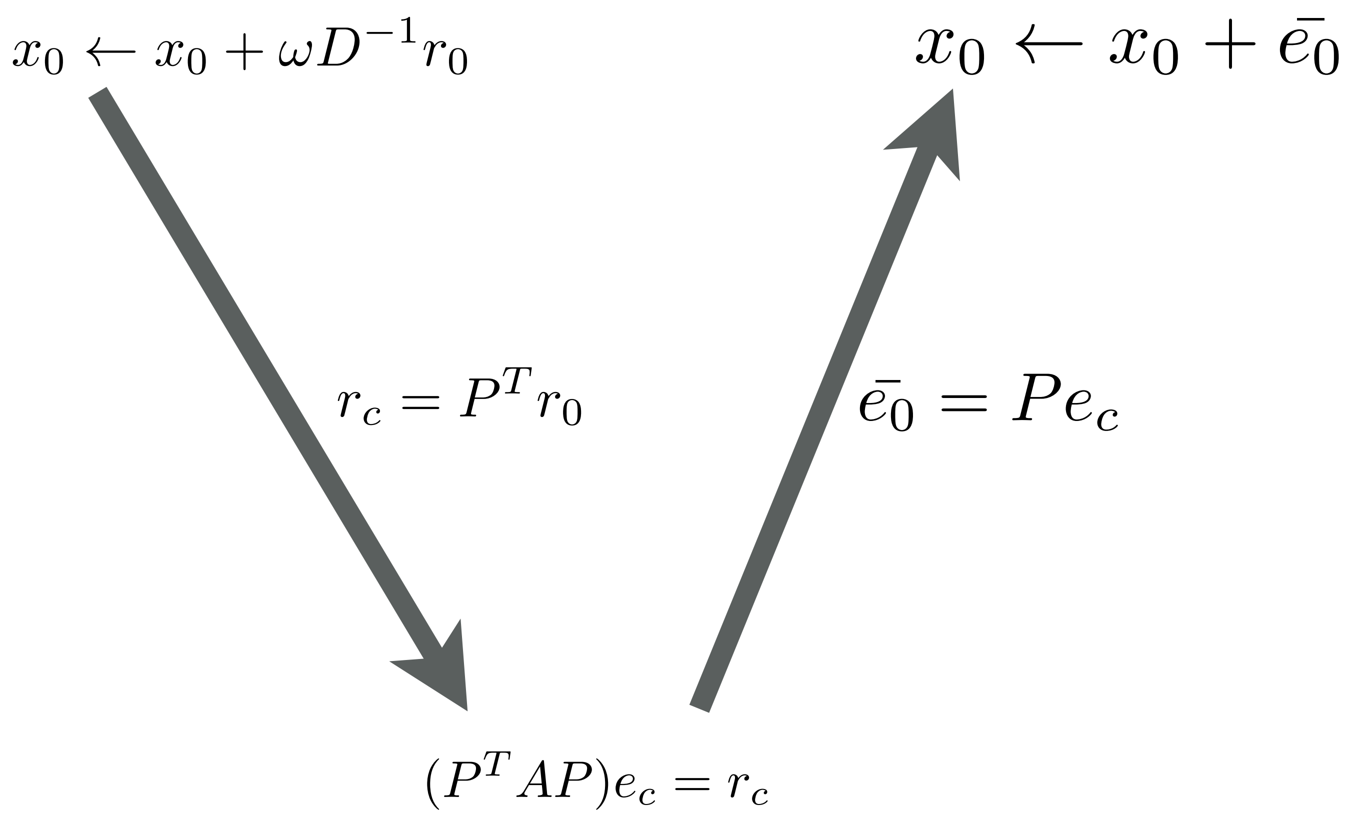


- AMG does not use geometric information, but captures both linear & operator-dep interpolation

$$(A\mathbf{u})_i = a_{i,i-1}u_{i-1} + a_{i,i}u_i + a_{i,i+1}u_{i+1}$$

$$u_i = \left(-\frac{a_{i,i-1}}{a_{i,i}}\right)u_{i-1} + \left(-\frac{a_{i,i+1}}{a_{i,i}}\right)u_{i+1}$$

Algebraic Observation



$$e_1 \leftarrow (I - P(P^T A P)^{-1} P^T A) G e_0$$

coarse grid correction

relax

relax

residual

restrict

coarse solve

interpolate

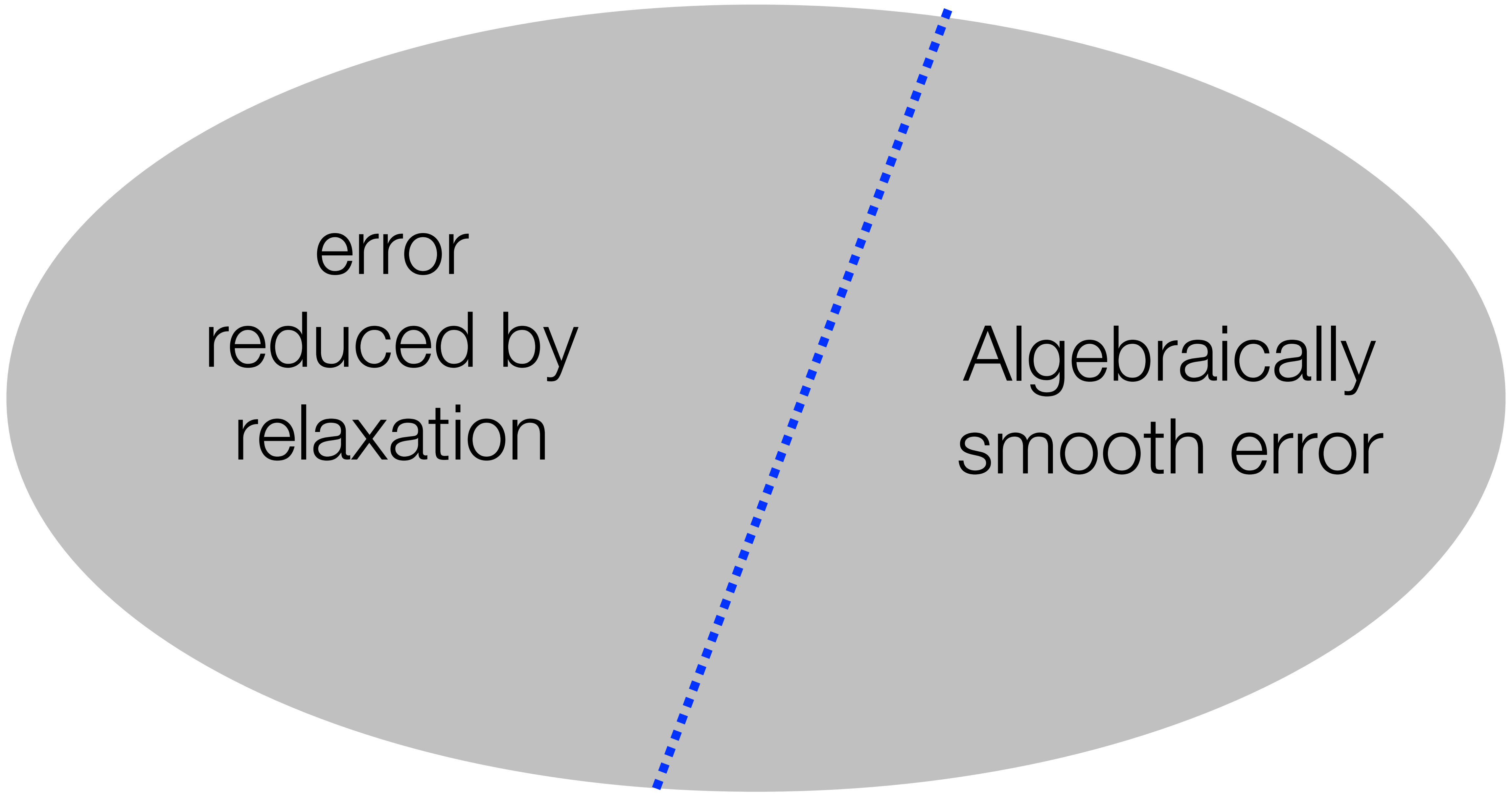
correct

$$G e_0 \in \mathcal{R}(P) \quad \Rightarrow \quad e_1 = 0$$

interpolation should capture what relaxation misses

Algebraically Smooth Error

- **“Algebraically smooth” error may not be geometrically smooth**



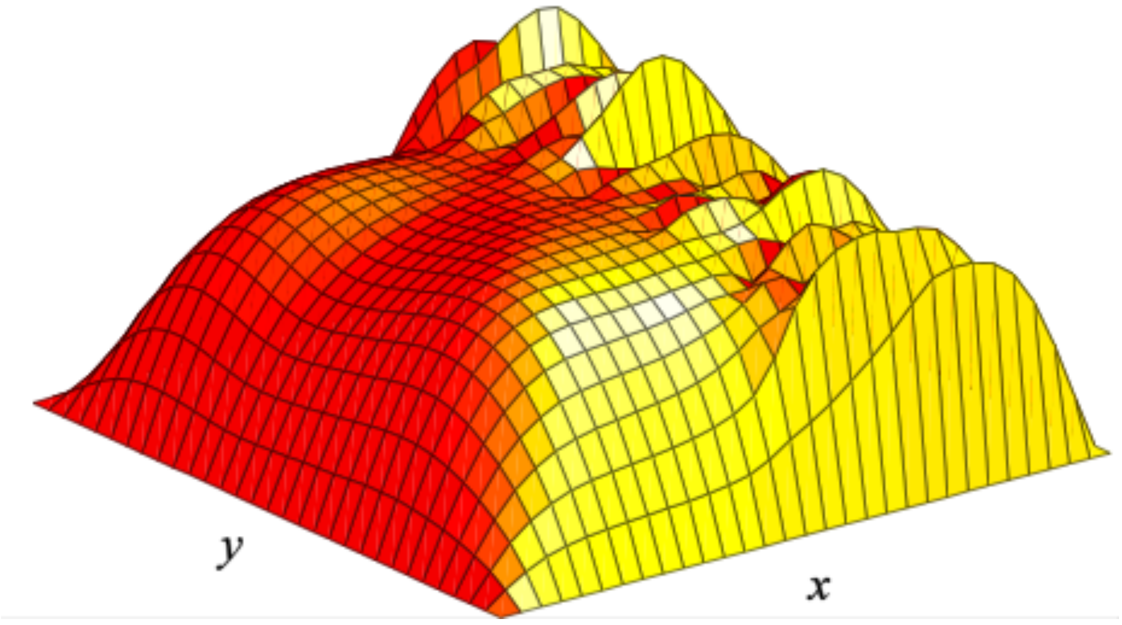
Error left by relaxation can be geometrically oscillatory

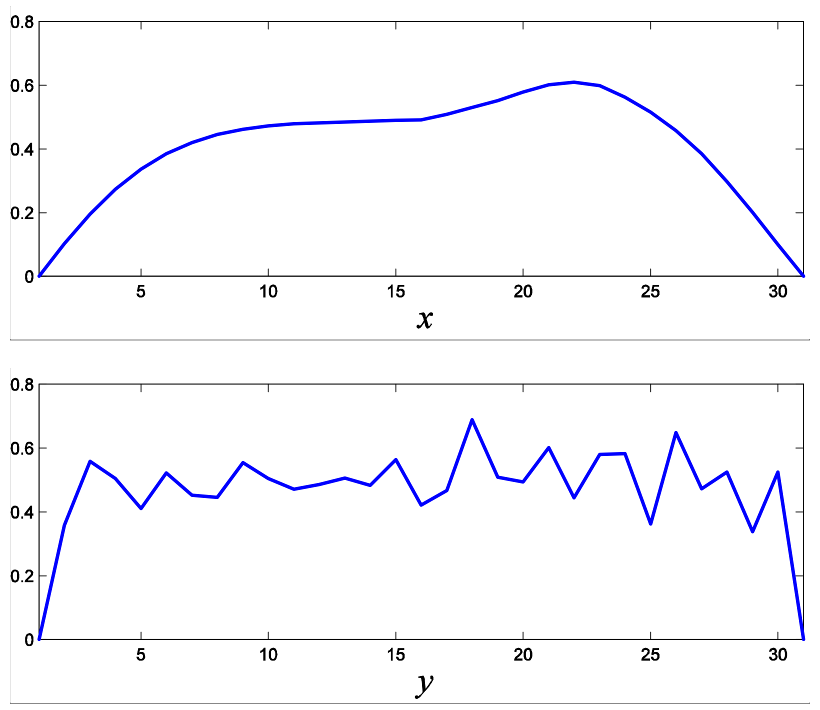
- 7 GS sweeps on

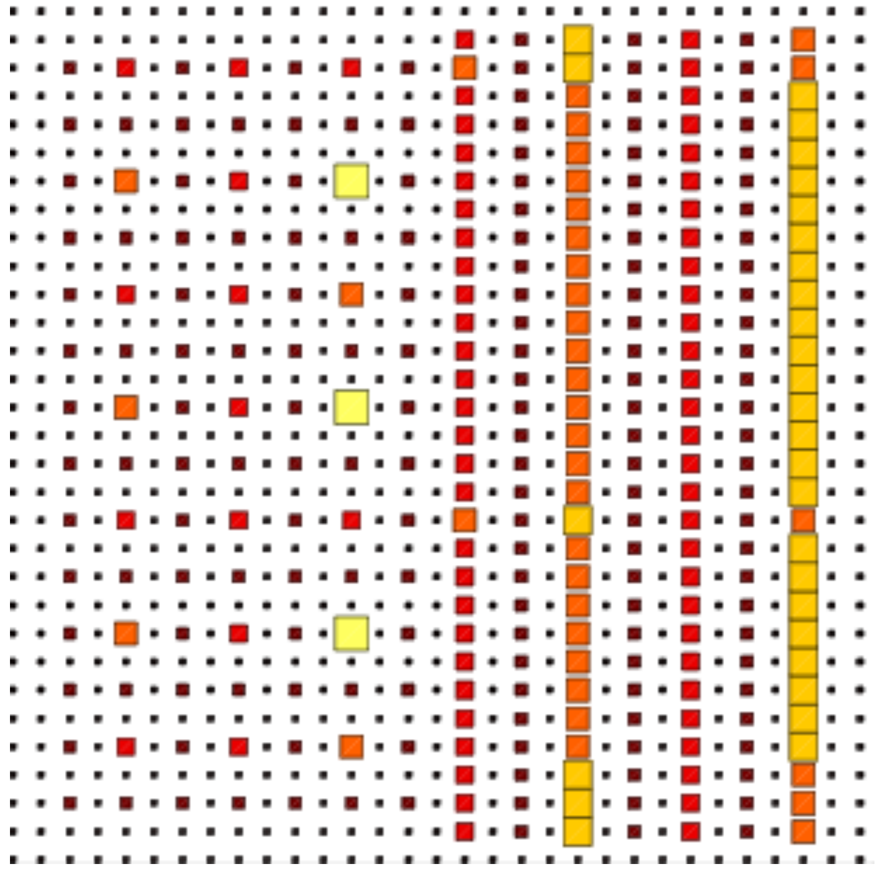
$$-au_{xx} - bu_{yy} = f$$

$a = b$	$a \gg b$

- Caution: this example
  - targets geometric smoothness
  - uses pointwise smoothers







Slide credit: R. Falgout, 2/1/2016

AMG coarsens grids in the direction of geometric smoothness

Main idea: Algebraically smooth error

- Take a relaxation scheme such as w-Jacobi

$$e \leftarrow (I - M^{-1}A)e$$

- If relaxation stagnates, then the remaining error exhibits poor convergence, so

- Formally (characterized by small eigenvalues)

$$(I - M^{-1}A)e \approx e \Rightarrow M^{-1}Ae \approx 0 \Rightarrow r \approx 0$$

$$\langle Ae, e \rangle << 1$$

Main idea: Algebraically smooth error

- We then have

$$
\begin{aligned}
\langle Ae, e\rangle &= \sum_i e_i (A_{ii} e_i + \sum_{j\neq i} A_{ij} e_j) \\
&= \sum_i e_i \left(\sum_{j \neq i} -A_{ij}(e_i - e_j) \right) \\
&= \sum_{i<j} -A_{ij} \cdot e_i \cdot (e_i - e_j) + \sum_{i>j} -A_{ij} \cdot e_i \cdot (e_i - e_j) \\
&= \sum_{i<j} -A_{ij} \cdot e_i \cdot (e_i - e_j) - \sum_{i<j} -A_{ij} \cdot e_i \cdot (e_i - e_j) \\
&= \sum_{i<j} -A_{ij} \cdot (e_i - e_j)^2
\end{aligned}
$$

assume zero row sum

swap i, j

- Ok, so smooth error varies **slowly** in the direction of large matrix coefficients

Briggs, William L. and Henson, Van Emden and McCormick, Steve F., A Multigrid Tutorial (2Nd Ed., 20000

Main idea: Algebraically smooth error

- We have assumed **geometric** smoothness to show

$$\mathbf{e}^T A\mathbf{e} = \sum_{i<j} (-a_{ij})(e_i - e_j)^2 \ll 1$$

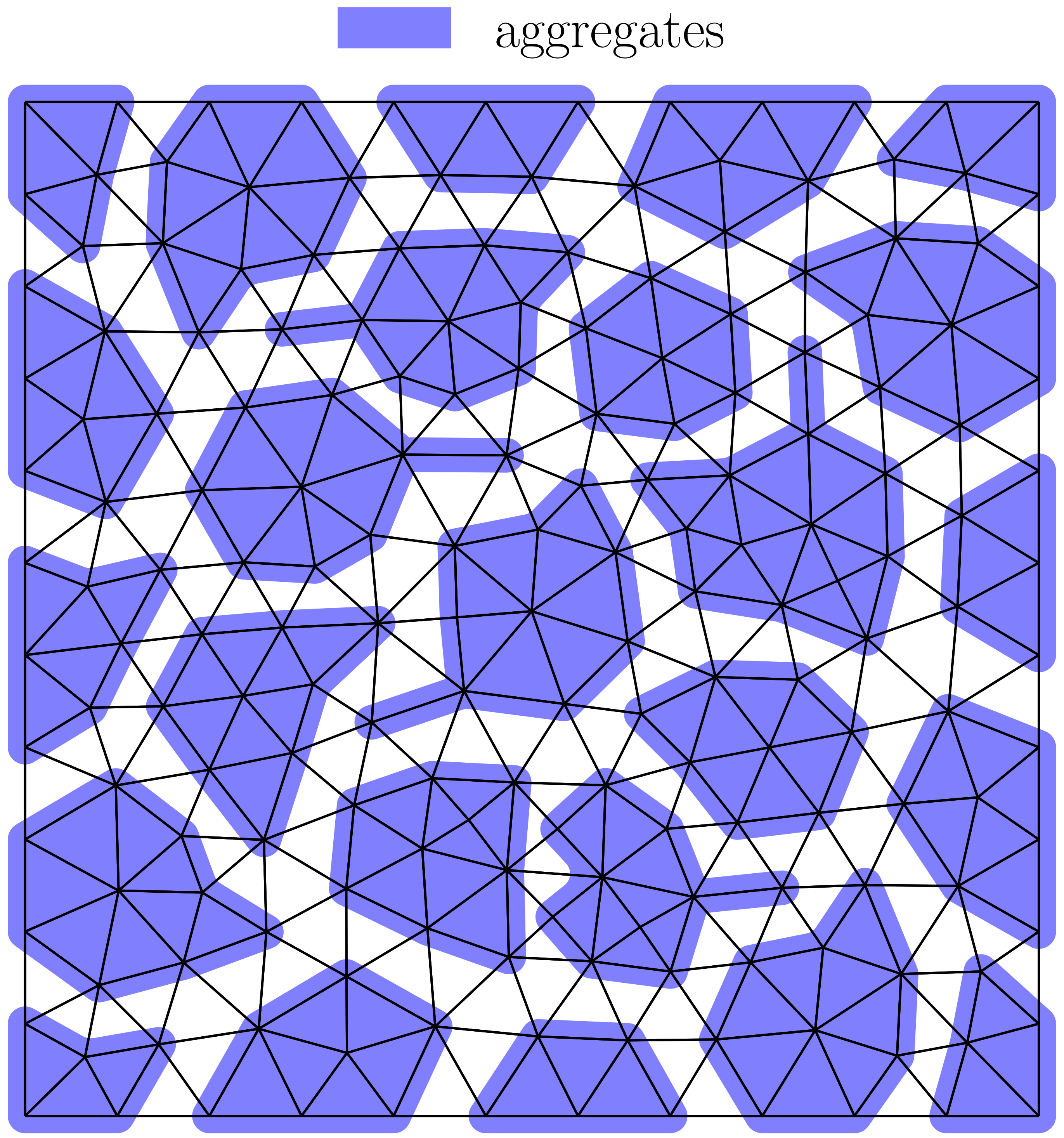
- **CF AMG**: Smooth error varies slowly in the direction of "large" matrix coefficients
- **Strength of connection**: Given a threshold $0 < \theta \leq 1$, we say that variable u_i strongly depends on variable u_j if

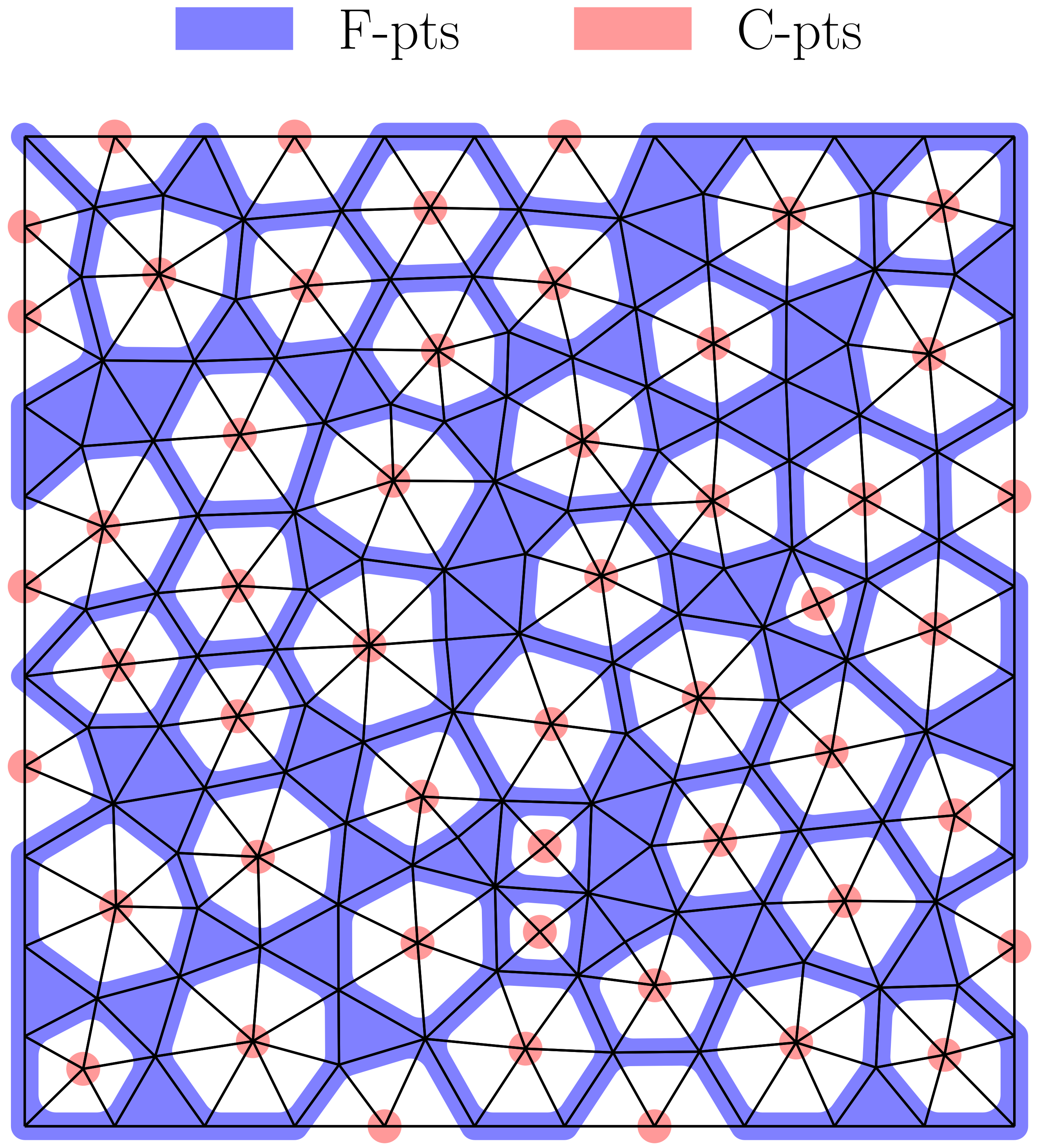
$$-a_{ij} \geq \theta \max_{k \neq i} \{-a_{ik}\}$$

- Often positive off-diagonals are treated as weak
- This definition of strength of connection is not symmetric

Demo: `12-AMG-coarse-mesh.ipynb`

Two (general) forms of AMG





- Smoothed Aggregation AMG (SA-AMG)
- Interpolation constructed from candidate vectors
- Clear approach to *optimize* interpolation

- Coarse-Fine AMG (CF-AMG) or Ruge-Stüben
- Coarse grid points are a subset of the fine grid points
- Edge-wise construction of interpolation, allowing straightforward control of sparsity
- Incorporating near-nullspace is not straightforward

Up next…

- What happens when we drop the notion of a *grid ? AMG* (Monday)
- What does this work at all? Theory. (next Friday)