= 3 cm =

Two–level local mode analysis – an introduction

Kees Oosterlee
(Original lecture: J. Linden)

The basic idea of any multigrid algorithm is to annihilate the high frequency error components efficiently on a fine grid by relaxation (or smoothing) while the low frequencies are approximated from coarser grids (with considerably less work than on the fine grid). The smoothing analysis is used to analyze quantitatively this reduction of the high frequency error components on the fine grid. Assuming the ideal situation in which

- the relaxation does not affect the low frequencies
- all low frequencies are approximated more or less exactly on the coarse grid
- there is no interaction between high and low frequency error components

it already leads to an "estimate" or prediction of the 2–level convergence factor. The local mode analysis can, however, also be used to analyze a **complete 2–level multigrid algorithm** more precisely. In this so–called 2–level analysis not only are the smoothing properties of the relaxation scheme on the fine grid considered but also the transfers between fine and coarse grid, the coarse grid operator and the coupling between high and low frequencies.

This lecture gives an introduction into this 2–level local mode analysis.

Contents:

In this part, the **2–level local mode analysis** is introduced **with the following restrictions:**

- only 2D scalar problems:

$$L^h U^h = F^h$$

- only square meshes:

$$G^h = \{(mh, lh) : k, l \in \mathbb{Z}\}$$

  with meshsize h in both directions;

- only standard coarsening;

$$G^H = G^{2h} = \{(2mh, 2lh) : k, l \in \mathbb{Z}\}$$

  i.e. the meshsizes are doubled in both directions.

Generalisations to higher dimensions, rectangular meshes and non–standard coarsenings are straightforward.

The 2–level operator

- **The 2–level cycle** (Correction Scheme)

  Consider a linear problem

  $$L^h U^h = F^h \quad \text{on grid } G^h$$

  Then: The 2–level cycle starting from the initial approximation u^h consists of the following steps

(1)	ν_1 smoothing steps on the fine grid:	$u^h \Longleftarrow S^{\nu_1}(u^h, F^h)$
(2)	computation of defects on the fine grid:	$d^h := F^h - L^h u^h$
(3)	restriction of defects from fine to coarse:	$d^H := I_h^H d^h$
(4)	solution of the coarse grid problem:	$L^H V^H = d^H$
(5)	prolongation of corrections from coarse to fine:	$v^h := I_H^h V^H$
(6)	adding the corrections to the current fine grid approximation:	$u^h \Longleftarrow u^h + v^h$
(7)	ν_2 smoothing steps on the fine grid:	$u^h \Longleftarrow S^{\nu_2}(u^h, F^h)$

- Steps (1) and (7) are **pre– and post–smoothing**, steps (2)...(6) form the **“coarse grid correction cycle”**.

The 2–level operator continued

- $S(u^h, F^h)$ denotes a relaxation scheme for the linear problem $L^h U^h = F^h$ with
  (i) S is linear in both arguments.
  (ii) $S(u^h, F^h) = u^h \iff u^h = U^h$
  i.e. $\iff u^h$ is the exact discrete solution.
- Instead of $S(u^h, 0)$ we simply write: Su^h.

Let u^h :initial approximation before the 2–level cycle,
$\overline{u}^h$:approximation after the 2–level cycle

Then $v^h = U^h - u^h$: initial error
$\overline{v}^h = U^h - \overline{u}^h$ error after the 2–level cycle

Then:

$$\overline{v}^h = \underbrace{S^{\nu_2} \overbrace{\left(Id - I_H^h \left(L^H\right)^{-1} I_h^H L^h\right)}^{C_h^H} S^{\nu_1}}_{M_h^H} \; v^h$$

- M_h^H is called the **"2–level operator"**.
- It is a (linear) operator which describes how the initial error is "transformed" to the new error by one complete 2–level cycle.
- C_h^H is called the **"coarse grid correction operator"**.

The 2–level local mode analysis

- **Goal:** Not only an analysis of smoothing properties of the relaxation S and prediction of multigrid performance based only on smoothing rates.
- Instead: Quantitative analysis and prediction of error reduction by a **complete 2–level cycle**.

- **Remember:** For any iteration of the form

$$w^{i+1} = Mw^i$$

  with a linear operator M the spectral radius

$$\rho(M)$$

  gives the asymptotic speed of convergence.
- **Therefore:** We must "compute"

$$\rho\left(M_h^H\right)$$

- Applying Fourier analysis, however, is a bit more complicated than in the case of smoothing (or one–level) analysis.

  This is because on an infinite grid G^h the Fourier components are not eigenfunctions of the 2–level operator M_h^H. M_h^H actually couples groups of Fourier components.

Local mode analysis of the 2–level operator:

2–level operator:

$$\overline{v}^h = \underbrace{S^{\nu_2} \overbrace{\left(Id - I_H^h \left(L^H\right)^{-1} I_h^H L^h\right)}^{C_h^H} S^{\nu_1}}_{M_h^H} \; v^h$$

Let $\quad v^h = e^{i\theta \underline{x}/h}, \; \theta = (\theta_1, \theta_2)$

with $\quad |\theta| := \max\left(|\theta_1|, \theta_2|\right) \leq \pi$

- From smoothing analysis, we already know:

$$S^{\nu_1} e^{i\theta \underline{x}/h} = \mu^{\nu_1}(\theta) \cdot e^{i\theta \underline{x}/h}$$

  where $\mu(\theta)$ is the smoothing factor of the relaxation S.
  (e.g. for relaxations based on regular splittings
  $L^h = A^h + B^h$, i.e. $S = \left(B^h\right)^{-1} A^h$).

- We already know

$$L^h e^{i\theta \underline{x}/h} = \widetilde{L}^h(\theta) \cdot e^{i\theta \underline{x}/h}$$

  where $\widetilde{L}^h(\theta) \in C$ is the Fourier symbol of L^h.

- How does the fine to coarse transfer I_h^H act on Fourier components?

1. $I_h^H =$ **injection**

$$I_h^H e^{i\theta \underline{x}/h} = e^{i2\theta \underline{x}/2h} = e^{i2\theta^{00}\underline{x}/h}$$

where $\theta^{00} = (\theta_1^{00}, \theta_2^{00})$ is defined by

$$\theta_i^{00} := \left\{ \begin{array}{cccccc} \theta_i & \text{if} & |\theta_i| \leq \frac{\pi}{2} & & & \\ \theta_i - \pi & \text{if} & \pi \geq & \theta_i & > \frac{\pi}{2} \\ \theta_i + \pi & \text{if} & -\pi \leq & \theta_i & < -\frac{\pi}{2} \end{array} \right\} \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$$

i.e. on G^{2h}, the frequency 2θ coincides with the $2\theta^{00}$.
note: $\theta^{00} = \theta$ if $|\theta| \leq \frac{\pi}{2}$.

2. $I_h^H = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ **Full–Weighting**

Then:

$$I_h^H e^{i\theta \underline{x}/h} = \underbrace{\frac{1}{4}(1 + \cos\theta_1)(1 + \cos\theta_2)}_{\tilde{I}_h^H(\theta)} e^{i2\theta^{00}\underline{x}/2h}$$

$\tilde{I}_h^H(\theta)$: Fourier symbol of Full-Weighting operator;
Note: For injection, the Fourier symbol is 1 for all frequencies.

Harmonic frequencies: – 2D –

- The fine to coarse transfer I_h^H maps four fine grid Fourier components onto one coarse grid Fourier component (standard coarsening $H = 2h$).

More precisely:

Let $\theta = (\theta_1, \theta_2) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^2$ be a low frequency.

Define:

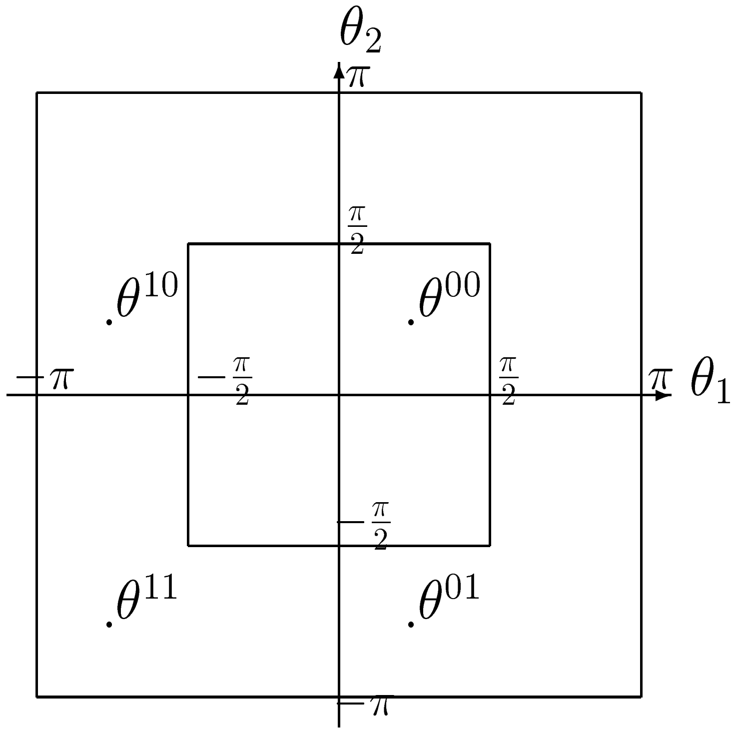
$$\boxed{i = 1,2: \quad \overline{\theta}_i := \left\{ \begin{array}{lll} \theta_i + \pi & if & \theta_i \leq 0 \\ \theta_i - \pi & if & \theta_i > 0 \end{array} \right\}}$$

Then:

$$\left\{ \begin{array}{lll} \theta^{00} & := & (\theta_1, \theta_2) \\ \theta^{10} & := & (\overline{\theta}_1, \theta_2) \\ \theta^{01} & := & (\theta_1, \overline{\theta}_2) \\ \theta^{11} & := & (\overline{\theta}_1, \overline{\theta}_2) \end{array} \right\} \stackrel{I_h^H}{\Longrightarrow} 2\theta^{00}$$

- These four fine grid frequencies are called **harmonic frequencies** or simply **harmonics**.

Harmonic frequencies: – 2D –



Standard coarsening:

- 3 high frequencies
- 1 low frequency
- the harmonic components

$$e^{i\theta^{kl}\underline{x}/h}, \quad k, l = 0, 1$$

coincide on the coarser grid $2h$ with the coarse grid component

$$e^{i2\theta^{00}\underline{x}/2h}$$

Local mode analysis of the 2–level operator continued:

2–level operator: $M_H^h = S^{\nu_2} \left(Id - I_H^h \left(L^H \right)^{-1} I_h^H L^h \right) S^{\nu_1}$

1. The fine to coarse operator I_h^H maps the four harmonic frequencies onto the same coarse grid component.
2. The coarse to fine operator I_H^h maps a single coarse grid component back onto a linear combination of the four corresponding fine grid harmonics.
3. All the other operators in M_h^H are "single grid grid operators" which map a Fourier component onto a multiple of itself. (note: this is not true for checker–board relaxations.)

$\Longrightarrow$ If we take the initial error v^h as a linear combination of the 4 harmonics

$$v^h(x) = A^{00} e^{i\theta^{00}\underline{x}/h} + A^{10} e^{i\theta^{10}\underline{x}/h} + A^{01} e^{i\theta^{01}\underline{x}/h} + A^{11} e^{i\theta^{11}\underline{x}/h}$$

then the error after a complete 2–level cycle is again such a linear combination:

$$\begin{aligned} \overline{v}^h(x) &:= M_h^H v^h \\ &= \overline{A}^{00} e^{i\theta^{00}\underline{x}/h} + \overline{A}^{10} e^{i\theta^{10}\underline{x}/h} + \overline{A}^{01} e^{i\theta^{01}\underline{x}/h} + \overline{A}^{11} e^{i\theta^{11}\underline{x}/h} \end{aligned}$$

and we can compute the new amplitudes $\overline{A}^{kl}$ from the old ones A^{kl}, $(k, l = 0, 1)$.

- **"The linear space of the four harmonics is invariant under the 2–level operator"**

Fourier representation of the complete 2–level cycle

$$M_h^H(\theta^{00}) = S^{\nu_2}(\theta^{00}) \cdot C_h^H(\theta^{00}) \cdot S^{\nu_1}(\theta^{00})$$

- This complex 4×4 matrix transforms the amplitudes of an initial error of the form

  $$v^h(x) = A^{00} e^{i\theta^{00}\underline{x}/h} + A^{10} e^{i\theta^{10}\underline{x}/h} + A^{01} e^{i\theta^{01}\underline{x}/h} + A^{11} e^{i\theta^{11}\underline{x}/h}$$

  onto the amplitudes of error

  $$\overline{v}^h(x) = \overline{A}^{00} e^{i\theta^{00}\underline{x}/h} + \overline{A}^{10} e^{i\theta^{10}\underline{x}/h} + \overline{A}^{01} e^{i\theta^{01}\underline{x}/h} + \overline{A}^{11} e^{i\theta^{11}\underline{x}/h}$$

  after a complete 2–level cycle:

  $$\begin{pmatrix} \overline{A}^{00} \\ \overline{A}^{10} \\ \overline{A}^{01} \\ \overline{A}^{11} \end{pmatrix} = M_h^H(\theta^{00}) \begin{pmatrix} A^{00} \\ A^{10} \\ A^{01} \\ A^{11} \end{pmatrix}$$

- The **spectral radius**

  $$\rho_h^H(\theta^{00}) = \rho\left(M_h^H(\theta^{00})\right)$$

  is the asymptotic reduction factor of these amplitudes.

- **Therefore:**

  $$\rho^\star := \max\left\{\rho_h^H(\theta^{00}) : \left|\theta^{00}\right| \leq \frac{\pi}{2}\right\}$$

  is the worst asymptotic reduction factor of all Fourier components. It, therefore, measures the speed of convergence of the complete 2–level cycle.

  $\rho^\star$ is called the **2–level factor**.

Remarks:

- In practice:
  - work out the Fourier representations of all operators which are involved M_h^H
  - use a computer to combine them to M_h^H and to find $\rho\left(M_h^H\right)$ and the maximum $\rho^\star$.
- **Note:** The spectral radius $\rho\left(M_h^H\right)$ does only depend on $\nu := \nu_1 + \nu_2$.

  For the spectral radius it doesn't matter how many smoothing steps are performed before and after the coarse grid correction.

  But: This is true only for 2–level cycles.

- Generalisations to systems of equations are straightforward but technically complicated.

  E.g.: Instead of 4×4 matrices, for a $q \times q$ system one gets a Fourier representation of M_h^H in form of complex $4q \times 4q$ matrices.

Example: Results for the 5–point Laplacian

- $L^h = \frac{1}{h^2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}$
- standard coarsening
- lexicographic Gauss–Seidel smoothing $\Longrightarrow S$
  Remember: smoothing rate = 0.5 !
- full weighting $\Longrightarrow I_h^H$
- bilinear interpolation $\Longrightarrow I_H^h$

			W-cycle
ν_1, ν_2	$\mu^{\nu_1+\nu_2}$	$\rho^\star$	N=128
1,0	0.5	0.400	0.395
1,1	0.25	0.193	0.190
2,1	0.125	0.119	0.117

$\rho^\star$ is the prediction of multigrid performance obtained by the 2–level local mode analysis.
$\mu^{\nu_1+\nu_2}$ is prediction by smoothing analysis.

Simplified 2–level reasonings, an example:

Consider the linear **convection–diffusion problem**:

$$LU = -\varepsilon \Delta U + aU_x + bU_y = \ldots$$

with $1 \gg \varepsilon > 0$.

Discretization by **central differences** and **explicit artificial viscosity** on a square cartesian mesh with meshsize h:

$$L^h = -\varepsilon^h \Delta^h + \frac{a}{2h}\begin{bmatrix} -1 & 0 & 1 \end{bmatrix} + \frac{b}{2h}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

with

$$\varepsilon^h := \max\left(\varepsilon, \beta\frac{|a|h}{2}, \beta\frac{|b|h}{2}\right)$$

where $\beta \geq 1$.

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

- isotropic artificial viscosity scheme
- $\beta \geq 1 \Longrightarrow$ **resulting difference operator is diagonally dominant** $\Longrightarrow$ **stability.**

- If $\varepsilon \ll 1$ usual multigrid cycles converge rather slowly. In fact, by 2-level cycles, the errors are asymptotically reduced by a factor of only 0.5. Multigrid cycles with more than 2 levels converge even slower.

  This happens also in cases where the smoothing properties of the relaxation are excellent!

- The reason for this (surprising) fact is the different amount of artificial viscosity on fine and coarse grids. Note that $\varepsilon^{2h} = 2\varepsilon^h$, i.e. the amount of artificial viscosity on the coarse grid $2h$ is twice the amount of artificial viscosity on the fine grid h.

  The result from that is that certain low frequency error components are not "correctly" approximated on the coarse grid: "They loose a factor 2 in amplitude".

- This can be studied and quantified in detail by local mode analysis. It turns out that only very special low frequency error components show the above behaviour, the so–called characteristic components which are constant along the characteristic of the advection operator.

Simplified 2–level analysis

Consider the **coarse grid correction operator**: $(H = 2h)$

$$C_h^{2h} = Id - I_{2h}^{h} \left(L^{2h}\right)^{-1} I_h^{2h} L^h$$

What happens when C_h^{2h} is applied to very low frequencies ?

- neglect the coupling of the harmonics
- assume that the transfer operators act like identities (which is nearly correct for very low frequencies).

Then:

$$C_h^{2h} \approx Id - \left(L^{2h}\right)^{-1} L^h$$

on very low frequencies.

Let $\theta = (\theta_1, \theta_2)$ be such low frequency:

Then:

$$\begin{aligned} C_h^{2h}(\theta) &\approx 1 - \frac{\widetilde{L}^h(\theta)}{\widetilde{L}^{2h}(2\theta)} \\ &= \quad 1 - 4\frac{\varepsilon^h(4 - 2\cos\theta_1 - 2\cos\theta_2) + ih(a\sin\theta_1 + b\sin\theta_2)}{\varepsilon^{2h}(4 - 2\cos 2\theta_1 - 2\cos 2\theta_2) + i2h(a\sin 2\theta_1 + b\sin 2\theta_2)} \end{aligned}$$

Therefore: Let $\theta_2 = \gamma\theta_1$.

Then compute:

$$\lim_{\theta_1 \to 0} \left(1 - \frac{\widetilde{L}^h(\theta_1, \theta_2)}{\widetilde{L}^{2h}(2\theta_1, 2\theta_2)} \right)$$

- This is an "estimate" of the coarse grid approximation for the lowest frequencies. It should be very small !

Result:

$$\lim_{\theta_1 \to 0} \left(1 - \frac{\widetilde{L}^h(\theta_1, \theta_2)}{\widetilde{L}^{2h}(2\theta_1, 2\theta_2)} \right) = \left\{ \begin{array}{cl} 0 & \text{if} \quad \gamma \neq -\frac{a}{b} \\ 1 - \frac{\varepsilon^h}{\varepsilon^{2h}} & \text{if} \quad \gamma = -\frac{a}{b} \end{array} \right\}$$

- $\gamma = -\frac{a}{b} \implies a\theta_1 + b\theta_2 = 0$
  $\implies$ "characteristic components"
  constant along the characteristic
  of the advection operator;

- if $\varepsilon \ll 1 \implies \varepsilon^{2h} = 2\varepsilon^h$
  $\implies 1 - \frac{\varepsilon^h}{\varepsilon^{2h}} = 0.5$

i.e.: A 2–level cycle reduces the (very) low frequency characteristic error components only by a **factor 0.5**.

Note: The (very) low frequencies are not damped by relaxation.

Remarks:

- One has always to be (at least) aware of this factor 0.5 when dealing with singular perturbed problems (e.g. in fluid dynamics, with upwind or artificial viscosity schemes). A 2–level factor of 0.5 very often leads to very unsatisfactory multilevel convergence (see below).
- The cure against it:
  - invest more work on coarser grids;
    W-cycles ($\gamma = 2, 3$), adaptive cycling;
  - find relaxations which also damp the low characteristic error frequencies (e.g. downstream relaxation);
  - modify the coarse grid operator in a proper way
    ($\approx$ less artificial viscosity on coarser grids)
  - use FMG to only solve your problem up to truncation error.

  Note: the above factor 0.5 limits the **algebraic** convergence of 2–level cycles (i.e. convergence of residuals to zero). “Differential convergence” is (sometimes) not seriously affected.