Numerical solution of the stationary convection-diffusion equation in 1D (XIV)

The numerical scheme

$$\begin{aligned} F_{j+1/2} &= \beta_j^0 \varphi_j + \beta_{j+1}^1 \varphi_{j+1}, \quad j = 1, \cdots, J-1 \, , \\ F_{1/2} &= \beta_1^1 \varphi_1 + \gamma_0 \, , \quad F_{J+1/2} = \beta_J^0 \varphi_J + \gamma_1 \, . \end{aligned}$$

$$L_h \varphi_j = F_{j+1/2} - F_{j-1/2} \quad \Rightarrow (2.31) :$$

$$L_h \varphi_j = \alpha_j^{-1} \varphi_{j-1} + \alpha_j^0 \varphi_j + \alpha_j^1 \varphi_{j+1} = \tilde{q}_j, \quad j = 1, \cdots, J \, ,$$

with $\alpha_1^{-1} = \alpha_J^1 = 0$.
This is called the numerical or finite volume scheme.

Numerical solution of the stationary convection-diffusion equation in 1D (XV)

General form of a linear scheme:

$$L_h\varphi_j = \sum_{k\in K} \alpha_j^k \varphi_{j+k} = \tilde{q}_j \ ,$$

with K some index set; in our case

$$K = \{-1,\ 0,\ 1\}$$

Stencil $[L_h]$ of the operator L_h:

$$[L_h]_j = [\ \alpha_j^{-1} \quad \alpha_j^0 \quad \alpha_j^1\ ]\ .$$

Numerical solution of the stationary convection-diffusion equation in 1D (XVI)

$$Ay = b\,, \quad y = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_J \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{q}_1 \\ \vdots \\ \tilde{q}_J \end{bmatrix},$$

A is tridiagonal matrix:

$$A = \begin{bmatrix} \alpha_1^0 & \alpha_1^1 & 0 & \cdots & 0 \\ \alpha_2^{-1} & \alpha_2^0 & \alpha_2^1 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \alpha_{J-1}^{-1} & \alpha_{J-1}^0 & \alpha_{J-1}^1 \\ 0 & \cdots & 0 & \alpha_J^{-1} & \alpha_J^0 \end{bmatrix}.$$

Numerical solution of the stationary convection-diffusion equation in 1D (XVII)

Two important questions

• How well does the numerical solution approximate the exact solution?

• How accurately and efficiently can we solve the linear algebraic system?

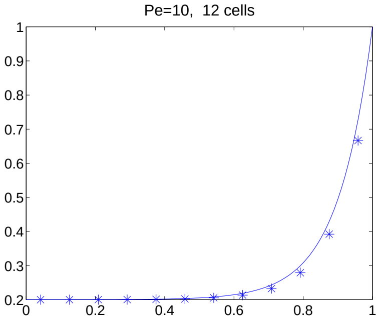
Numerical solution of the stationary convection-diffusion equation in 1D (XIIX)

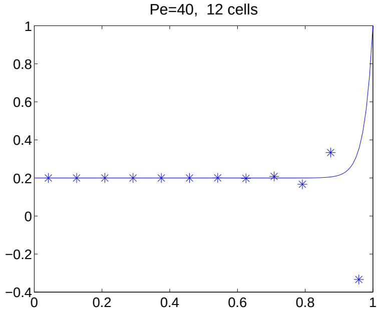
Lets zoom in on the behaviour of the algorithm for different values of the Péclèt number and see where the behaviour changes...

- Two-sided Dirichlet boundary condition
- Uniform mesh
- Convection:Central scheme I (averaging)
- Diffusion:Central discretisation.

Numerical solution of the stationary convection-diffusion equation in 1D (XIX)

Numerical experiments on uniform grid





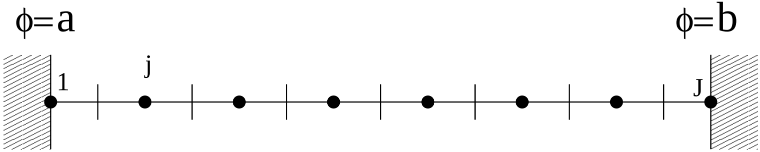
Exact solution (—) and numerical solution (*).

Why the *wiggles* for Pe $= 40$?

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (I)

A sidestep to the vertex-centered case:
For a conservative scheme $[L_h]_j = [\,\alpha_j^{-1} \quad \alpha_j^0 \quad \alpha_j^1]$ on a uniform grid the dependence on j disappears and $\alpha^0 = -(\alpha^{-1} + \alpha^{+1})$

$$\alpha^{-1}\varphi_{j-1} - (\alpha^{-1} + \alpha^{+1}+)\varphi_j + \alpha^{+1}\varphi_{j+1} = 0, \quad j = 1..J-1$$
$$\varphi_0 = a, \quad \varphi_J = b$$



In the uniform grid case the only difference between cell/vertex centered is in treatment of the boundary conditions..

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (II)

The analytical solution of this difference equation with the Dirichlet boundary conditions is given by:

$$\varphi_j = a + (b-a)\frac{1-z^j}{1-z^J}, \quad z = \tfrac{\alpha^{-1}}{\alpha^{+1}}, \quad j = 0, 1, \ldots J$$

Question: Depending on the coefficients $a, b, \alpha^{-1}, \alpha^{+1}$, when will the solution have oscillations?

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (III)

Central scheme I:

$$
\begin{aligned}
\alpha^{j-1} = -\left(\frac{u}{2} + \frac{\varepsilon}{h}\right) = -\frac{\varepsilon}{h}\left(\frac{1}{2}\frac{uh}{\varepsilon} + 1\right) &= \\
-\frac{\varepsilon}{h}\left(\frac{1}{2}Pe_h + 1\right) &= -\frac{\varepsilon}{h}\left(1 + \frac{1}{2}Pe_h\right) \\
\alpha^{j+1} = \left(\frac{u}{2} - \frac{\varepsilon}{h}\right) = -\frac{\varepsilon}{h}\left(-\frac{1}{2}\frac{uh}{\varepsilon} + 1\right) &= \\
-\frac{\varepsilon}{h}\left(-\frac{1}{2}Pe_h + 1\right) &= -\frac{\varepsilon}{h}\left(1 - \frac{1}{2}Pe_h\right)
\end{aligned}
$$

Divide by the common factor $-\frac{\varepsilon}{h}$, Pe_h is the *mesh-* Péclèt number.

$$
\varphi_j = a + (b-a)\frac{1 - \left(\frac{1+\frac{Pe}{2}}{1-\frac{Pe}{2}}\right)^j}{1 - \left(\frac{1+\frac{Pe}{2}}{1-\frac{Pe}{2}}\right)^N}
$$

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (IV)

The condition for an oscillatory solution that:

$$\frac{\alpha^{-1}}{\alpha^{+1}} < 0$$

Corresponds to

$$\frac{1 + \frac{1}{2} Pe_h}{1 - \frac{1}{2} Pe_h} < 0 \quad \Rightarrow \quad 1 - \frac{1}{2} Pe_h < 0 \quad \Rightarrow \quad Pe_h > 2$$

Lets check that...

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (V)

For large values of Pe_h we can even predict how the solution will look. Subsitute $\epsilon = 2/Pe_h$.

$$\varphi_j = a + (b-a)\frac{1-\left(\frac{1+1/\epsilon}{1-1/\epsilon}\right)^j}{1-\left(\frac{1+1/\epsilon}{1-1/\epsilon}\right)^J}$$

Pe_h is *large* and ϵ very small: make an expansion in ϵ.

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (VI)

$$\varphi_j = a + (b-a)\frac{1-(-1)^j(1+2j\epsilon)}{1-(-1)^J(1+2J\epsilon)} + \mathcal{O}(\epsilon^2)$$

Case 1: J odd:

$$\varphi_j \approx a + (b-a)\frac{1-(-1)^j - 2(-1)^j j\epsilon}{2(1+J\epsilon)}$$

j even $\phi_j \approx a-(b-a)j\epsilon$; $\quad$ j odd $\varphi_j \approx b-(b-a)(J-j)\epsilon$

Analytical solution of the *discretised* stationary convection-diffusion equation in 1D (VII)

$$\varphi_j = a + (b-a)\frac{1-(-1)^j(1+2j\epsilon)}{1-(-1)^J(1+2J\epsilon)} + \mathcal{O}(\epsilon^{\in})$$

Case 2: J even:

$$\varphi_j \approx a + (b-a)\frac{1-(-1)^j(1+2j\epsilon)}{-2J\epsilon}$$

j even: $\phi_j \approx a-(b-a)\frac{j}{J}$; $\qquad$ j odd: $\varphi_j \approx a-(b-a)\left(\frac{-1}{J\epsilon}-\frac{j}{J}\right)$

The discrete maximum principle (I)

We can also explain the behavior for $P_h > 2$ from the properties of the operator: no wiggles means no local extrema: the discrete counterpart of the maximum principle.

$$c_n, x_n \in \mathbb{R} \quad n = 1, 2, \ldots N :$$
$$\left\{ \sum_{n=1}^{N} c_n = 0, \quad c_n < 0 | n > 1, \quad \sum_{n=1}^{N} c_n x_n \leq 0 \right\} \Rightarrow$$
$$x_n = x_1, \quad n = 1, \ldots, N \quad \vee \quad x_1 < \max\{x_n : n > 1\}$$

x_j is the average of the neighbouring x_j's...

The discrete maximum principle (II)

Scheme in stencil notation:

$$L_h\varphi_j = \sum_{k\in K} \alpha_j^k \varphi_{j+k} = \tilde{q}_j \ , \quad j = 1, \cdots, J \ .$$

The operator L_h is *of positive type* if

$$\sum_{k\in K} \alpha_j^k = 0, \quad j = 2, \cdots, J-1$$

and

$$\alpha_j^k < 0, \quad k \neq 0, \quad j = 2, \cdots, J-1$$

The discrete maximum principle (III)

Discrete maximum principle.
If L_h is of positive type and

$$L_h\varphi_j \leq 0, \quad j = 2, \cdots, J-1 \ ,$$

then $\varphi_j \leq \max\{\varphi_1, \varphi_J\}$.
Extrema can only occur on the boundary of the domain, just like in the continuous case!

The discrete maximum principle (IV)

When is our scheme (Central Scheme I) of positive type?

$$[L_h] = \left[\begin{array}{ccc} -\frac{1}{2}u - \frac{\varepsilon}{h} & 2\frac{\varepsilon}{h} & \frac{1}{2}u - \frac{\varepsilon}{h} \end{array}\right]$$

Of pos. type iff $\alpha^{+1} < 0$:

$$\frac{1}{2}u - \frac{\varepsilon}{h} < 0 \Rightarrow \frac{uh}{\varepsilon} < 2 \Rightarrow Pe_h < 2$$

So to avoid wiggles in the solution we should reduce the meshwidth such that $h < 2/Pe$. This is totally impossible in the limit $Pe \Rightarrow \infty$ Alternatives? How'bout this upwind scheme?

The discrete maximum principle (V)

When is the upwind scheme of positive type?

$$u > 0 : \quad [L_h] = \left[-u - \frac{\varepsilon}{h} \qquad u + 2\frac{\varepsilon}{h} \qquad -\frac{\varepsilon}{h} \right] .$$

$$u < 0 : \quad [L_h] = \left[-\frac{\varepsilon}{h} \qquad -u + 2\frac{\varepsilon}{h} \qquad u - \frac{\varepsilon}{h} \right] .$$

The *Upwind Scheme* is of positive type for all Pe_h ☺,
but heavily overpredicts the thickness of the boundary layer in the solution ☹
The *Upwind Scheme* corresponds to *Central Scheme I*, with ε replaced by $\varepsilon' = \varepsilon + \varepsilon_a, \quad \varepsilon_a = uh/2$, which dominates the solution for $Pe >> 1$.

Local grid refinement (I)

A second alternative is to use *local grid refinement*
Reduce the grid size **only** in the business part of the domain:
In the **boundary layer** at the outflow boundary.

$$\text{Exact sol.} \quad \Rightarrow \quad \delta = \mathcal{O}(\varepsilon) = \mathcal{O}(\mathrm{Pe}^{-1}).$$

Of course we locally need $h < \delta$ to capture the boundary layer
Take $\delta \equiv 6\varepsilon$.
Put 6 equal cells in the boundary layer region $\Omega_1 = [1 - \delta, 1]$
and 6 equal cells in the rest: $\Omega_2 = [0, 1 - \delta]$, with $\Omega = \Omega_1 \cup \Omega_2$.

Local grid refinement (II)

Lets compare the performance of the different schemes:

- Central Scheme I

$$\varphi_{j+1/2} = \frac{\varphi_{j+1} + \varphi_j}{2}$$

- Central Scheme II

$$\varphi_{j+1/2} = \frac{h_j \varphi_{j+1} + h_{j+1} \varphi_j}{h_j + h_{j+1}}$$

- Upwind Scheme

$$\varphi_{j+1/2} = \varphi_j, \quad (u > 0);$$

The coefficients of the stencil are now dependent on j...
Central Scheme II will be more accurate, right?

Local grid refinement (III)

Summary of results:

Scheme	Pe=10	Pe=40	Pe=400	Pe=4000
Central Scheme I	.0607	.0852	.0852	.0852
Upwind	.0785	.0882	.0882	.0882
Central Scheme II	.0607	.0852	.0856	.3657

Maximum error norm; 12 cells.

Local grid refinement (IV)

Central Scheme I

$$[L_h] = \left[\begin{array}{rl} - & \dfrac{1}{2} - \dfrac{\varepsilon}{H} \\ & \dfrac{\varepsilon}{H} + \dfrac{2\varepsilon}{h+H} \\ & \dfrac{1}{2} - \dfrac{2\varepsilon}{h+H} \end{array}\right].$$

Central Scheme II

$$[L_h] = \left[\begin{array}{rl} - & \dfrac{1}{2} - \dfrac{\varepsilon}{H} \\ - & \dfrac{1}{2} + \dfrac{h}{h+H} + \dfrac{\varepsilon}{H} + \dfrac{2\varepsilon}{h+H} \\ & \dfrac{H}{h+H} - \dfrac{2\varepsilon}{h+H} \end{array}\right].$$

Global and local truncation error (I)

Definition The *global truncation error* is defined as

$$e_j \equiv \varphi(x_j) - \varphi_j \ , \quad j = 1, \cdots , J \ ,$$

with $\varphi(x)$ the exact solution.
Only in very few cases we can determine the global error directly, so we use another quantity to estimate the global error.

Global and local truncation error (II)

Definition The *local truncation error* of the discrete operator L_h is defined as

$$\tau_j \equiv L_h\left(\varphi(x_j) - \varphi_j\right) \equiv L_h e_j \ , \quad j = 1, \cdots, J \ .$$

$\varphi(x)$ is an *exact* solution of the continuous problem $L\varphi(x) = q(x)$
φ_j is an *exact* solution of the discretized problem $L_h\varphi_j = q_j$
But not the otherway around...

Global and local truncation error (III)

Let $e = (e_1, \cdots e_J)^T, \quad \tau = (\tau_1, \cdots \tau_j)^T$.

$$e = L_h^{-1}\tau\,.$$

$$\|e\| \leq \|L_h^{-1}\|\|\tau\|\,.$$

This means that a scheme with smaller $\|\tau\|$ *will have necessarily a smaller* $\|e\|$, *because* $\|L_h^{-1}\|$ *may be larger!!!*
This explains why it is possible that Central Scheme II (linear interpolation for convection) may be worse than averaging ☹

Global and local truncation error (IV)

Estimate of local truncation error for the following scheme $L_h\varphi_j$

$$\begin{aligned} L_h\varphi_j \equiv \quad & -\left(\frac{u}{2}+\frac{\varepsilon}{h_{j-1/2}}\right)\varphi_{j-1}+\varepsilon\left(\frac{1}{h_{j-1/2}}+\frac{1}{h_{j+1/2}}\right)\varphi_j \\ & +\left(\frac{u}{2}-\frac{\varepsilon}{h_{j+1/2}}\right)\varphi_{j+1}=h_jq_j, \quad j=2,\cdots,J-1\ , \end{aligned}$$

Although the exact sol. is unknown, nevertheless τ can be estimated.

Global and local truncation error (V)

Use **Taylor's formula** to approximate $\varphi_{j\pm1}$ to express $L_h\varphi(x_j)$ in φ and $\frac{d^n f(\xi)}{dx^n}(x_j)$

$$f(x) = f(x_0) + \sum_{k=0}^{n-1} \frac{1}{k!}\frac{d^k f(x_0)}{dx^k}(x - x_0)^k + \frac{1}{n!}\frac{d^n f(\xi)}{dx^n}(x - x_0)^n$$

Global and local truncation error (VI)

$$\varphi(x_{j\pm1}) = \varphi(x_j) \pm h_{j\pm1/2}\varphi^{(1)}(x_j) + \frac{1}{2}h_{j\pm1/2}^2\varphi^{(2)}(x_j)$$
$$\pm\frac{1}{6}h_{j\pm1/2}^3\varphi^{(3)}(x_j) + \frac{1}{24}h_{j\pm1/2}^4\varphi^{(4)}(x_j) + \mathcal{O}(h_{j\pm1/2}^5) .$$

Here $\mathcal{O}$ is *Landau's order symbol*:
A function $f(h) = \mathcal{O}(h^p)$ if there exist a constant M independent of h and a constant $h_0 > 0$ such that

$$\frac{|f(h)|}{h^p} < M, \quad \forall h \in (0, h_0) .$$

Global and local truncation error (VII)

Combining Taylor with the continuous operator:

$$L\varphi(x) = u\varphi_{,x} - \varepsilon\varphi_{,xx} = q(x)s \tag{1}$$

We can show that up to $\mathcal{O}(\Delta^4)$ the exact solution of L satisfies $\tilde{L}_h$

$$L_h\varphi(x_j) = \tilde{L}_h\varphi(x_j) + \mathcal{O}(\Delta^4)\ ,$$

$\varphi(x)$ is the exact sol.

$$\begin{aligned}
\tilde{L}_h\varphi(x_j) &= \frac{1}{2}q_j(h_{j-1/2} + h_{j+1/2}) \\
&+ \left(\frac{1}{4}u\varphi^{(2)} - \frac{1}{6}\varepsilon\varphi^{(3)}\right)(h_{j+1/2}^2 - h_{j-1/2}^2) \\
&+ \left(\frac{1}{12}u\varphi^{(3)} - \frac{1}{24}\varepsilon\varphi^{(4)}\right)(h_{j+1/2}^3 + h_{j-1/2}^3) + \mathcal{O}(\Delta^4)\ .
\end{aligned}$$

Global and local truncation error (IIX)

The local truncation error is given as

$$\tau_j = L_h e_j = L_h[\varphi(x_j) - \varphi_j] = \tilde{L}\varphi(x_j) - h_j q_j + \mathcal{O}(\Delta^4) \; .$$

$$\begin{aligned}
\tau_j &= \frac{1}{2} q_j (h_{j-1/2} - 2h_j + h_{j+1/2}) \\
&+ \left(\frac{1}{4} u \varphi^{(2)} - \frac{1}{6} \varepsilon \varphi^{(3)} \right) (h_{j+1/2}^2 - h_{j-1/2}^2) \\
&+ \left(\frac{1}{12} u \varphi^{(3)} - \frac{1}{24} \varepsilon \varphi^{(4)} \right) (h_{j+1/2}^3 + h_{j-1/2}^3) + \mathcal{O}(\Delta^4) \; .
\end{aligned}$$

Global and local truncation error (IX)

Smooth grid:

$$|h_{j+1/2} - h_{j-1/2}| = \mathcal{O}(\Delta^2) \quad \text{and}$$
$$|h_{j-1/2} - 2h_j + h_{j+1/2}| = \mathcal{O}(\Delta^3)\ .$$

Any grid which does not fullfill the conditions for a Smooth grid is defined to be a *Rough* grid.

$$\begin{aligned}
\tau_j &= \frac{1}{2} q_j (h_{j-1/2} - 2h_j + h_{j+1/2}) \\
&+ \left(\frac{1}{4} u\varphi^{(2)} - \frac{1}{6}\varepsilon\varphi^{(3)}\right)(h^2_{j+1/2} - h^2_{j-1/2}) \\
&+ \left(\frac{1}{12} u\varphi^{(3)} - \frac{1}{24}\varepsilon\varphi^{(4)}\right)(h^3_{j+1/2} + h^3_{j-1/2}) + \mathcal{O}(\Delta^4)\ .
\end{aligned}$$

What will be the order of the local truncation error on a *Smooth* grid?

Global and local truncation error (X)

$$\text{Smooth grid:} \quad \tau_j = \mathcal{O}(\Delta^3).$$

$$\text{Rough grid:} \quad \tau_j = \mathcal{O}(\Delta).$$

This is cause for common misunderstanding that smooth grids are required for good accuracy. *But this is not necessary in general.* Locally refined grid used in the preceding numerical experiments is rough, but nevertheless the accuracy was satisfactory.
Again: local truncation error is *not* a good indicator of the accuracy of the solution.

Global and local truncation error (XI)

Estimate of local truncation error at the boundaries (assume locally uniform grid)

$$\tau_1 = \frac{h}{4}\varepsilon\varphi^{(2)} + \mathcal{O}(h^2)$$

Interior, uniform grid: $\tau_j = \mathcal{O}(h^3)$.
However, it is not necessary to improve the local accuracy near a Dirichlet boundary!
Neumann at $x = 1$:

$$\tau_J = \frac{h^2}{24}\varepsilon\varphi^{(3)} + \mathcal{O}(h^3)\ .$$

This also does not require improvement!
This follows from estimate of *global* truncation error.

Estimating the global error (I)

The *global* error is directly related to the *local truncation* error:

$$L_h e = \tau$$

Try to find a *Barrier function* E such that:

$$L_h E \geq |\tau|$$

We will show that this implies $|e| \leq E$.

Estimating the global error (II)

$$L_h(\pm e - E) \leq 0 \, .$$

We want to use the discrete maximum principle:
Therefore the scheme should be of positive type or $Pe < 2$.
The maximum occurs at Dirichlet boundary $\Rightarrow$

$$\pm e_j - E_j \leq \pm e_1 - E_1, \quad j = 2, \cdots, J \, . \tag{2.58}$$

Assume locally uniform grid: $h_1 = h_2 = h_{\frac{3}{2}}$:

$$L\varphi_1 = \left(\frac{u}{2} + \frac{3\varepsilon}{h_1}\right)\varphi_1 + \left(\frac{u}{2} - \frac{\varepsilon}{h_1}\right)\varphi_2 = a\varphi_1 - b\varphi_2 \leq 0$$

$$0 < b < a$$

Estimating the global error (III)

This means:

$$b\,(\pm e_2 - E_2) \leq b\,(\pm e_1 - E_1) \quad \text{Using max principle}$$

$$a\,(\pm e_1 - E_1) \leq b\,(\pm e_2 - E_2) \quad \text{Using } L_h \leq 0$$

Because $0 < b < a$ this implies:

$$\pm e_1 - E_1 \leq 0 \Rightarrow |e_1| \leq E_1$$

And using the discrete maximum principle:

$$|e_j| \leq E_j, \quad j = 1, \cdots, J.$$

Choosing/Finding a suitable $E(x)$ is an art.