LMECA 2660 : Numerical simulation of 1D convection - diffusion equation

P. Billuart, P. Parmentier and G. Winckelmans

February 2017

We consider the 1-D convection - diffusion linear equation,

$$\frac{\partial u}{\partial t} + c\,\frac{\partial u}{\partial x} = \nu\,\frac{\partial^2 u}{\partial x^2}\ . \tag{1}$$

We consider the case of a Gaussian initial condition :

$$u(x,0) = \frac{Q}{\sqrt{\pi\,\sigma_0^2}}\ \exp\left(-\frac{x^2}{\sigma_0^2}\right) \tag{2}$$

with $Q = \int_{-\infty}^{\infty} u(x,0)\,dx$.
It is straightforward to show that the analytical solution to this problem is a diffusing Gaussian function moving at constant velocity :

$$u(x,t) = \frac{Q}{\sqrt{\pi\,(\sigma_0^2 + 4\nu t)}}\ \exp\left(-\frac{(x-ct)^2}{(\sigma_0^2 + 4\nu t)}\right)\ , \tag{3}$$

and whose spatial Fourier transform, $\widehat{u}(k,t)$, is also a Gaussian and is given by (cfr. Reminder about Fourier transforms and Fourier series theory) :

$$\widehat{u}(k,t) = Q\ \exp\left(-\frac{k^2(\sigma_0^2 + 4\,\nu t)}{4}\right)\ \exp\left(-ikct\right)\ . \tag{4}$$

The integral of the solution, Q, is conserved, while the energy decreases :

$$\begin{aligned} E(t) &= \int_{-\infty}^{\infty} \frac{u(x,t)^2}{2}\,dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{2}\,\widehat{u}(k,t)\ \left(\widehat{u}(k,t)\right)^*\ dk\ \ \text{(Parseval equality)} \\ &= \frac{Q^2}{\sqrt{8\pi\,(\sigma_0^2 + 4\nu t)}}\ . \end{aligned}$$

Note : Even though the domain of definition for the present problem is unbounded, we will here use a large periodic numerical domain ($L = 32\,\sigma_0$). As a consequence, we will use discrete Fourier series instead of Fourier transforms. For further details regarding the definitions, the links between those, etc. : see the Reminder.

- **Discrete Fourier series** :
  1. Taking the FFT (e.g., using Matlab), obtain the coefficients of the discrete Fourier series of the initial Gaussian for the meshes that will be used numerically later in this homework : $\frac{h}{\sigma_0} = \frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{8}$, thus corresponding to $N = \frac{L}{h} = 64$, 128 and 256.
  2. Compare them with the coefficients of the Fourier transform of the Gaussian function in unbounded domain, Eq. (4), by plotting the logarithm of their modulus as a function of j. Comment.
- **Partially decentered scheme** :
  1. Using Taylor series, obtain the partially decentered discretization of highest possible order for the convective term, $c\frac{\partial u}{\partial x}|_i$, using explicit finite differences involving u_{i-2}, u_{i-1}, u_i and u_{i+1}.
  2. What are the order and truncation error of this scheme ?
  3. Using a modal analysis with $u_i(t) = \sum_j \widehat{U}_j(t)e^{ik_jx_i}$ and $\frac{d\widehat{U}_j}{dt} = \lambda_j\,\widehat{U}_j$, obtain the expression for λ_k as a function of the dimensionless wavenumber kh. Compare it to the exact case and quantify the phase and/or amplitude error(s) as a function of kh.
  4. Compare the λ_k with those obtained when using a second, fourth and sixth order centered scheme (E2, E4 and E6 : see Lecture notes). Plot, in the complex plane, the $\frac{\lambda_k\,h}{c}$ corresponding to those 4 schemes with $0 \leq \frac{kh}{\pi} \leq 1$ and mark the points corresponding to $\frac{kh}{\pi} = 0$, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ and 1.
- **Produce a C code** :
  1. whose temporal integration is performed using a classical Runge-Kutta 4 scheme,
  2. whose spatial discretization for the convective term can be performed using the following finite differences schemes : second order explicit (E2), fourth order explicit (E4), the decentered scheme of the previous question, fourth order implicit (I4), and, finally, sixth order implicit (I6),
  3. whose spatial discretization of the diffusive term is performed using the second order explicit finite differences scheme (E2),
  4. that provides, at each time step, the following global diagnostics : $Q_h^n = h\sum_i u_i^n$ and $E_h^n = h\sum_i \frac{(u_i^n)^2}{2}$,
  5. that provides, at each time step, the global error in the least squares sense : $R_h^n = \sqrt{h\sum_i(u_i^n - u(x_i, t^n))^2}$.

Hint : To solve the periodic tridiagonal system required by the implicit schemes, use the *Thomas algorithm* that will be briefly described hereunder (very small complexity, $\mathcal{O}(N)$ operations). A C code implementation of this algorithm will be provided.

- **Pure convection case ($\nu = 0$) - Numerical simulation and analysis** :
  1. Perform the numerical simulation using each scheme and for three different meshes : $\frac{h}{\sigma_0} = \frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{8}$, thus corresponding to $N = \frac{L}{h} = 64$, 128 and 256.
  2. Compare the obtained solutions to one another, and also to the analytical solution at $\frac{ct}{L} = \frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ and 1. Comment.
  3. Plot and compare the evolution of the global diagnostics as a function of time : $\frac{Q_h^n}{Q}$, $\frac{E_h^n}{E(0)}$ and $\frac{R_h^n}{\sqrt{E(0)}}$. Comment.
  4. Plot the global error achieved at $\frac{ct}{L} = 1.0$ as a function of $\frac{h}{\sigma_0}$, and in a log-log diagram. Determine the order of convergence of the different numerical methods used and comment.

  Note : It is important to satisfy the stability constraints of the temporal integration scheme (here the RK4 scheme) that imply a limitation on the CFL number : $CFL = \frac{c\Delta t}{h}$. In order to make fair comparisons, perform all your simulations with the same value : $CFL = 1$. Based on your findings in the previous question (i.e. the plot, in the complex plane, of the $\frac{\lambda_k h}{c}$ related to the different discretizations of the convective term with respect to kh), and the stability curve of the RK4 scheme (see Lecture notes), verify that this choice of CFL indeed respects the stability constraints of the RK4 scheme in each case.

  Remark : It is important to distinguish the stability and the accuracy of a scheme. Indeed, a stable scheme is not necessarily accurate. In the pure convection case with the RK4 scheme, a $CFL = 1$ should be a good choice ; it satisfies the stability criteria and it guarantees a proper accuracy of the temporal integration, while providing a reasonable simulation duration. In short, there is no point to have neither a too small time step (too long simulation time), nor a too large one (loss of accuracy and stability problems).

- **Convection-diffusion ($Re = \frac{c\sigma_0}{\nu} = 40$) - Numerical simulation and analysis**
  1. Perform the numerical simulation using the second order schemes for the discretization of the convective and diffusive terms, the mesh $\frac{h}{\sigma_0} = \frac{1}{4}$ and $CFL = 1$.
  2. Plot the numerical solution at the times $\frac{ct}{L} = \frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ and 1, and compare them to the analytical solution. Comment.
  3. Plot, compare and comment the global diagnostics as a function of time.
  4. At $\frac{ct}{L} = 1.0$, compare your results and diagnostics to those obtained in the pure convection case (cfr. question 3), using the same discretization for the convective term, mesh and CFL.

Note : Regarding the stability in the convection-diffusion case, one should also comply with the limitations on the Fourier number, $r = \frac{\nu \Delta t}{h^2}$, in addition to the limitations on the CFL number. We here have $CFL = \frac{h}{\sigma_0}\ Re\ r$: we thus obtain $r = \frac{1}{10}$, which comfortably guarantees stability for the diffusive term with a RK4 scheme, at least for the centered schemes.

Additional information about periodic domains

In this problem, we wish to study a function in an unbounded domain. As it is numerically impossible to build an infinite computational domain, we will rather use a periodic numerical domain of length L. This approximation is only valid if the characteristic length of the problem, σ_0, is significantly smaller than L. We here choose $L = 32\sigma_0$.

To illustrate the implementation of such condition, let's consider a centered second order scheme as an example. To obtain the solution at the point i, we need to use the information contained at the points $i-1$, i and $i+1$. The domain is discretized using N points numbered from 0 to $N-1$. In a periodic domain, to compute the solution at the point $N-1$, we then use the information at the points $N-2$, $N-1$ and 0. Likewise, to obtain the solution at the point 0, we use the points $N-1$, 0 and 1.

Remarks and instruction

1. This homework is a one man/woman job.
2. The program will be written in C.
3. The visualisations and FFTs can be performed using the software or language of your choice (Matlab, Python,...).
4. Provide clean and clear results. Provide nice and large plots with legends, titles, axes, ...
5. A printed version (recto-verso) must be handed over (a box will be placed in front of the A.026 office in the Stévin building).
6. A copy of your program and of your report shall also be uploaded on **Moodle**.
7. The homework is due on the **16th of March 2017 at 6pm**.
8. **Note of warning** : A zero-tolerance policy will be applied regarding plagiarism. Systematic and automatic testing will be carried out (internet, other present or former students, etc.). It is not forbidden to assist one another, but it must then be explicitly specified. You will however be evaluated on the basis of your personal contribution.

Resolution of a tridiagonal and periodic system : extension of the Thomas algorithm

We here present a summary of the Thomas algorithm. For further details, Google should be useful. The goal of the Thomas algorithm is to solve the system $Ax = q$ with :

$$A = \begin{bmatrix} a & c & & & & \\ b & a & c & & & \\ & b & a & c & & \\ & & \ddots & \ddots & \ddots & \\ & & & b & a & c \\ & & & & b & a \end{bmatrix} \tag{5}$$

This algorithm works in 2 steps : the *forward pass* whose aim is to simplify the equations, and the *backward pass* that effectively solves the system.

1. The forward pass simplifies the equations and sets them under the form
$$x_i + \tilde{a}_i x_{i+1} = \tilde{q}_i.$$
The first equation becomes :
$$\tilde{a}_1 = \frac{c}{a} \qquad \tilde{q}_1 = \frac{q_1}{a} \tag{6}$$
For the others, equation (i) becomes $(i) - (i-1)b$:
$$\tilde{a}_i = \frac{c}{a - b\tilde{a}_{i-1}} \qquad \tilde{q}_i = \frac{q_i - b\tilde{q}_{i-1}}{a - b\tilde{a}_{i-1}} \tag{7}$$
Indeed, by recurrence, one obtains
$$\left.\begin{array}{llllllll} (i-1) & x_{i-1} & + & \tilde{a}_{i-1}x_i & & & = & \tilde{q}_{i-1} \\ (i) & bx_{i-1} & + & ax_i & + & cx_{i+1} & = & q_i \end{array}\right\} \tag{8}$$
$$(i) - (i-1)\,b \quad \Rightarrow \quad x_i + \frac{c}{a - b\tilde{a}_{i-1}} x_{i+1} = \frac{q_i - b\tilde{q}_{i-1}}{a - b\tilde{a}_{i-1}} \tag{9}$$
2. Backward pass. Now that the system is upper triangular, we are able to solve it very efficiently. For the last equation :
$$x_n = \tilde{q}_n \tag{10}$$
For the other equations, we proceed from bottom to top :
$$x_i = \tilde{q}_i - \tilde{a}_i x_{i+1} \tag{11}$$

However, the system we here wish to solve is a periodic problem. The matrix is then :

$$A = \begin{bmatrix} a & c & & & & b \\ b & a & c & & & \\ & b & a & c & & \\ & & \ddots & \ddots & \ddots & \\ & & & b & a & c \\ c & & & & b & a \end{bmatrix}, \tag{12}$$

Therefore, we are not able to directly use *Thomas Algorithm* as detailed above. However, let's consider the matrix $A_c = A\,(1:n-1, 1:n-1)$, i.e. the A matrix without the last row and the last column. Such a matrix satisfies the structure required by the algorithm. The problem then becomes :

$$A_c x_c = q_c - \begin{bmatrix} b \\ \vdots \\ c \end{bmatrix} x_n \tag{13}$$

$$b x_{n-1} + a x_n + c x_1 = q_n \tag{14}$$

To solve the first equation, let's assume that x_c has the form :

$$x_c = x^{(1)} + x^{(2)} x_n \tag{15}$$

In that case, x_1, x_2 are given by (linearity of the system) :

$$A_c x^{(1)} = q_c \qquad A_c x^{(2)} = - \begin{bmatrix} b \\ \vdots \\ c \end{bmatrix} . \tag{16}$$

This allows to solve the last equation :

$$x_n = \frac{q_n - c x_1^{(1)} - b x_{n-1}^{(1)}}{a + c x_1^{(2)} + b x_{n-1}^{(2)}} \ . \tag{17}$$

Reminder about Fourier transforms and Fourier series

The Fourier transform $\widehat{f}(k)$ of a function $f(x)$ is defined as :

$$\widehat{f}(k) = \mathcal{F}\left(f(x)\right) \; = \; \int_{-\infty}^{\infty} f(x) \, \exp\left(-\imath k x\right) dx \; ,$$

while the inverse transform is defined as :

$$f(x) = \mathcal{F}^{-1}\left(\widehat{f}(k)\right) \; = \; \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(k) \, \exp\left(\imath k x\right) dk \; .$$

We easily verify that :

$$\mathcal{F}\left(f(x - x_0)\right) \; = \; \widehat{f}(k) \; \exp\left(-\imath k x_0\right) \; .$$

We also demonstrate that :

$$\mathcal{F}\left(\exp\left(-\frac{x^2}{\sigma^2}\right)\right) \; = \; \sqrt{\pi \, \sigma^2} \; \exp\left(-\frac{k^2 \, \sigma^2}{4}\right) \; .$$

A periodic function $f(x)$, of period L, may be represented by a Fourier series :

$$\begin{aligned} f(x) \; &= \; \sum_{j=-\infty}^{\infty} \widehat{F}(k_j) \exp(\imath k_j x) \; , \\ \text{where} \quad \widehat{F}(k_j) \; &= \; \frac{1}{L} \int_{-L/2}^{L/2} f(x) \, \exp(-\imath k_j x) \, dx \end{aligned}$$

and with $k_j = \frac{2\pi}{L} j$. We easily verify that the Fourier series of the function shifted in space by x_0, $f(x - x_0)$, is simply $\widehat{F}(k_j) \; \exp\left(-\imath k_j x_0\right)$.

Finally, a discrete and periodic function of period L, with $x_i = -L/2 + i\,h$ ($i = 0, 1, 2, \ldots, N-1$, where N is even, and $h = L/N$) and $f_i = f(x_i)$, can be represented by a discrete Fourier series :

$$\begin{aligned} f_i = f(x_i) \; &= \; \sum_{j=-N/2}^{N/2} \widehat{F}(k_j) \exp(\imath k_j x_i) \; , \\ \text{where} \quad \widehat{F}_j = \widehat{F}(k_j) \; &= \; \frac{1}{N} \sum_{i=0}^{N-1} f(x_i) \, \exp(-\imath k_j x_i) \; . \end{aligned}$$

Here again, the discrete Fourier series of the function shifted in space by x_0, $f(x_i - x_0)$, is simply : $\widehat{F}(k_j) \; \exp\left(-\imath k_j x_0\right)$.

The coefficients $\widehat{F}(k_j)$ are typically obtained using the Fast Fourier Transform (FFT). If the function $f(x_i)$ is real-valued, the coefficients are complex conjugates : $\widehat{F}(-k_j) = \widehat{F}_r(-k_j) + \imath\, \widehat{F}_\imath(-k_j) = \left(\widehat{F}(k_j)\right)^* = \widehat{F}_r(k_j) - \imath\, \widehat{F}_\imath(k_j)$. We then also have :

$$f(x_i) \;\; = \;\; \widehat{F}_r(k_0) + \sum_{j=1}^{N/2-1} 2\left[\widehat{F}_r(k_j)\cos(k_j x_i) - \widehat{F}_\imath(k_j)\sin(k_j x_i)\right] + 2\widehat{F}_r(k_{N/2})\cos(k_{N/2}x_i) \; .$$

The first term (i.e., the "zero mode") corresponds to the mean value of the function. The last term (i.e. the "flip-flop mode") corresponds to the highest wavenumber mode (i.e., $k_{N/2}h = \pi$) and is taken as purely real (its imaginary part is set to zero). We thus have N discrete and real values of the function $(f_0, f_1, \ldots, f_{N-1})$, and N discrete and real values for the reconstruction of the Fourier coefficients : $1 + 2 * (N/2 - 1) + 1 = N$

By extension, a non periodic function, but evaluated on an interval L which is large enough for the function to be considered periodic, may be represented approximately on this interval L by using the Fourier transform $\widehat{f}(k)$ obtained in an unbounded domain :

$$f(x) \;\; \simeq \;\; \frac{1}{L}\sum_{j=-\infty}^{\infty} \widehat{f}(k_j)\exp(\imath k_j x) \; .$$

For example, for the Gaussian function of Eq. (3), considered on a periodic domain of period L, and with $L \gg \sigma_0$, we obtain :

$$u(x,t) \;\; \simeq \;\; \frac{Q}{L}\sum_{j=-\infty}^{\infty} \exp\left(-\frac{k_j^2(\sigma_0^2 + 4\,\nu t)}{4}\right)\,\exp(\imath k_j(x - ct)) \; .$$

The larger L/σ, the better the approximation.

For the discrete version, we then get, as an approximation :

$$f_i = f(x_i) \;\; \simeq \;\; \frac{1}{L}\sum_{j=-\infty}^{\infty} \widehat{f}(k_j)\exp(\imath k_j x_i) \; ,$$

and thus, for the Gaussian function of Eq. (3) :

$$u_i(t) = u(x_i,t) \;\; \simeq \;\; \frac{Q}{L}\sum_{j=-\infty}^{\infty} \exp\left(-\frac{k_j^2(\sigma_0^2 + 4\,\nu t)}{4}\right)\,\exp(\imath k_j(x_i - ct)) \; .$$