Asymptotic behavior of splitting schemes involving time-subcycling techniques

Guillaume Dujardin, Pauline Lafitte

Asymptotic behavior of splitting schemes involving time-subcycling techniques

Guillaume Dujardin
INRIA Lille Nord Europe, EPI SIMPAF

Pauline Lafitte,
INRIA Lille Nord Europe, EPI SIMPAF & École Centrale Paris, Lab. MAS

November 12, 2012

Abstract

In order to integrate numerically a well-posed multiscale evolutionary problem such as a Cauchy problem for an ODE system or a PDE system, using time-subcycling techniques consists in splitting the vector field in a fast part and a slow part and take advantage of this decomposition, for example by integrating the fast equation on a much smaller time step than the slow equation (instead of having to integrate the whole system with a very small time step to ensure stability for example). These techniques are designed to improve the computational efficiency and have been been very widely used for schemes, that may have (at least) one component that has to be computed through an explicit scheme thus constrained by a limitation of the time step (CFL). In this paper, we study the long time behavior of such schemes, that are primarily designed to be convergent in short-time to the solution of the original problem. We develop our analysis on ODE toy-models and illustrate our results numerically on more complex PDE systems.

1 Introduction

Time-subcycling is a way to speed up computations for a multiscale problem by splitting the underlying operator and treating the different steps of the resulting numerical scheme with adapted time-steps. Our aim is to determine how appropriately the subcycling techniques capture the right asymptotic state for continuous dynamical systems described by ODEs or PDEs, the solutions of which converge to a steady state as time goes to infinity. In order to save computational time, the subcycling techniques have been very widely used for schemes associated with multiscale systems, that may have (at least) one component that has to be computed through an explicit scheme thus constrained by a limitation of the time step (CFL). Subcycling techniques have been developed extensively for multiscale problems arising in computational fluid and structural dynamics [9, 3]. The applications we have specifically in mind are related to the recent development of the "asymptotic-preserving" schemes in the sense of Jin [6, 7] for kinetic equations. Dividing systems in a suitable timescale was indeed proved efficient for Boltzmann-type and Fokker-Planck equations by way of micro-macro decompositions [4, 8, 2, 1]. However, if subcycling techniques have been

used in several test-cases, up to our knowledge, the asymptotic error to the longtime solution was not precisely computed.

In order to provide the reader with numerical examples and illustrate our results, Sections 2 and 3 are devoted to two different examples of differential systems with two different time scales, in the spirit of the analysis of the Dahlquist equation when studying the asymptotic stability of schemes for stiff ODEs [5] and of the analysis led by Temam [10]. Both systems have exact and explicit solutions so one can do any computation and estimate involving the exact flows.

The first one (analyzed in Section 2) is linear and reads[1]

$$\begin{cases} u' = -Nc(u-v) \\ v' = c(u-v), \end{cases} \tag{1}$$

where $c \geq 0$ and $N \in \mathbb{N}$, with $N \in \mathbb{N}$ being large : it is the stiffness parameter in the problem. The second one (analyzed in Section 3) is nonlinear and reads

$$\begin{cases} u' = -Nc(u-v) - N(u-v)^2 \\ v' = c(u-v) + (u-v)^2. \end{cases} \tag{2}$$

Since for any solution of any such system, one has $u' + Nv' = 0$, the solutions of these two systems are included in the straight lines of slope $-1/N$ in the phase space $\mathbb{R}_u \times \mathbb{R}_v$. In the following, we consider linear splitting schemes between the fast (*i.e.* first) equation of the system and the slow (*i.e.* second) equation. Therefore, for the numerical solutions of the linear system (1), the numerical schemes will always be the composition of matrices of the form

$$M_f(\lambda_f) := \begin{pmatrix} \lambda_f & 1-\lambda_f \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad M_s(\lambda_s) := \begin{pmatrix} 1 & 0 \\ 1-\lambda_s & \lambda_s \end{pmatrix}, \tag{3}$$

Let us denote by $\mathcal{M}_s := \{M_s(\lambda),\, \lambda \in [0,1)\}$ and $\mathcal{M}_f := \{M_f(\lambda),\, \lambda \in [0,1)\}$. These sets are stable under matrix multiplication. The expressions of the parameters λ will depend on the choice of integrator (exact flow or θ-scheme) and the composition of the matrices will depend on the type of splitting one wants to use (Lie or Strang type). In Section 4, we perform the same analysis for a 1D coupled reaction-diffusion system, where the stiffness parameter is independent of the mesh. For this problem, the boundary conditions play a crucial role in the existence of attractive equilibrium points. We focus on several cases of boundary conditions (homogeneous Dirichlet, inhomogeneous Dirichlet as well as homogeneous Neumann conditions), and we discuss the speed of convergence of the numerical splitting methods with subcycling, as well as the A-order issues for the Neumann conditions.

2 Full analysis of the linear system

2.1 The exact solution

Let us compute the exact solution of (1). We consider the matrix

$$A = \begin{pmatrix} -N & N \\ 1 & -1 \end{pmatrix}.$$

[1] From the dimensional point of view, c is homogeneous to the inverse of a characteristic time.

It is diagonalizable and here are its eigenvalues and associated spectral projectors

$$\left(-(N+1),\ P=-\frac{1}{(N+1)}A\right)\ \text{and}\ \left(0,\ Q=\frac{1}{(N+1)}\begin{pmatrix}1 & N\\ 1 & N\end{pmatrix}\right).$$

So the exact solution is, for all $t\in\mathbb{R}$,

$$W(t):=\begin{pmatrix}u(t)\\ v(t)\end{pmatrix}=\left(\mathrm{e}^{-(N+1)ct}P+Q\right)\begin{pmatrix}u^0\\ v^0\end{pmatrix},$$

for the initial values u^0 and v^0 at time $t=0$. In particular, we note that all the solutions converge to the equilibrium state $Q(u^0,v^0)^t$ when t tends to infinity. In the following, we fix $T>0$ and we define

$$F_T=\mathrm{e}^{-(N+1)cT}P+Q, \tag{4}$$

the eigenvalues of which are $\mathrm{e}^{-(N+1)cT}$ and 1.

2.2 General properties of linear splitting schemes

Let $G_{\delta t}$ be defined for $\delta t\in I_N$ as the 2-by-2 matrix of a numerical flow which is a product of matrices of the form (3), where I_N is the intersection, that may depend on N, of the stability intervals of the related schemes (see examples in Section 2.3). In the following, for all $n\in\mathbb{N}$, we will denote by

$$W^n:=\begin{pmatrix}u^n\\ v^n\end{pmatrix}=G_{\delta t}^nW^0$$

the numerical solution at time $n\delta t$ starting from the initial datum $W^0=(u^0,v^0)^t$.

Lemma 2.1 *For all $\delta t\in I_N$, the matrix $G_{\delta t}$ is diagonalizable, with two distinct real eigenvalues. One of these eigenvalues is 1 and the other one lies in $(0,1)$. The vector $(1,1)^t$ is an eigenvector of $G_{\delta t}$ associated to the eigenvalue 1. Hence the matrix $G_{\delta t}$ reads*

$$G_{\delta t}=\begin{pmatrix}1-\alpha(\delta t) & \alpha(\delta t)\\ \beta(\delta t) & 1-\beta(\delta t)\end{pmatrix} \tag{5}$$

for two real-valued functions α and β. Moreover, the spectral decomposition of the matrix $G_{\delta t}$ reads

$$G_{\delta t}=\mu(\delta t)P(\delta t)+Q(\delta t), \tag{6}$$

where $P(\delta t)$ is the matrix of the spectral projector of $G_{\delta t}$ associated to the eigenvalue $\mu(\delta t)=1-\alpha(\delta t)-\beta(\delta t)$ and $Q(\delta t)$ is that associated to the eigenvalue 1. In particular,

$$Q(\delta t)=\frac{1}{\alpha(\delta t)+\beta(\delta t)}\begin{pmatrix}\beta(\delta t) & \alpha(\delta t)\\ \beta(\delta t) & \alpha(\delta t)\end{pmatrix}. \tag{7}$$

Remark 2.2 *We will sometimes use in the following the notation $G[\alpha,\beta]$ in reference to (5).*

Proof. Since all the matrices M_s and M_f have $(1,1)^t$ for eigenvector associated with 1, so does any (finite) product of such matrices and this explains the form of the matrix $G_{\delta t}$ in (5). Moreover, since all the matrices M_s and M_f also have their other real eigenvalue in $(0,1)$, the determinant of a product of such matrices is in $(0,1)$. Hence for all $\delta t \in I_N$, $G_{\delta t}$ is diagonalizable with eigenvalues 1 and $\mu(\delta t) = \mathrm{Tr}(G_{\delta t}) - 1 = \det(G_{\delta t}) \in (0,1)$, so we get the spectral decomposition (6). One can compute the projection matrix $Q(\delta t)$ by computing the left-eigenvector associated with 1 in order to obtain (7). ∎

Let us denote by $S(\delta t)$ the ratio $\alpha(\delta t)/\beta(\delta t)$. We know that for all $n \in \mathbb{N}$, $u^n + S(\delta t)v^n = u^0 + S(\delta t)v^0$, where $(u^n, v^n)^t = G_{\delta t}^n (u^0, v^0)^t$. Hence, the numerical solution evolves on a straight line in the $\mathbb{R}_u \times \mathbb{R}_v$ plane, as the does the exact solution. Moreover, the exact and numerical propagators share an interesting property :

Property 2.3 *For any fixed $\delta t > 0$, $F_{\delta t}^n = F_{n\delta t}$ projects the vector $(u^0, v^0)^t$ onto the line of equation $u = v$ when n tends to infinity and so does $G_{\delta t}^n$ for all $\delta t \in I_N$.*

Proof. The projection property for $F_{n\delta t}$ as $n \to +\infty$ relies on the decomposition (4). Using Lemma 2.1, we get

$$\forall n \in \mathbb{N}, \qquad G_{\delta t}^n = (\mu(\delta t))^n P(\delta t) + Q(\delta t),$$

with $|\mu(\delta t)| < 1$. ∎

Let us denote those limits (which depend on δt)

$$(u_{\mathrm{num}}^\infty, v_{\mathrm{num}}^\infty)^t = \lim_{n\to+\infty} G_{\delta t}^n (u^0, v^0)^t \quad \text{and} \quad (u_{\mathrm{ex}}^\infty, v_{\mathrm{ex}}^\infty)^t = \lim_{n\to+\infty} F_{\delta t}^n (u^0, v^0)^t$$

The asymptotic error $(u_{\mathrm{num}}^\infty, v_{\mathrm{num}}^\infty)^t - (u_{\mathrm{ex}}^\infty, v_{\mathrm{ex}}^\infty)^t = (Q(\delta t) - Q)(u^0, v^0)^t$ is hence measured by the difference of the slopes of the two straight lines $u + Nv = u^0 + Nv^0$ (exact solution) and $u + S(\delta t)v = u^0 + S(\delta t)v^0$ (numerical solution) (see Figure 1).

Therefore, we set the following

Definition 2.4 *The relative asymptotic error is the difference*

$$\varepsilon^\infty = \frac{|S(\delta t) - N|}{N},$$

and we say that the asymptotic order (A-order) is at least $p \in \mathbb{N}^\star$ if when δt tends to 0, we have

$$\varepsilon^\infty = \mathcal{O}(\delta t^p).$$

Of course, as usual, the A-order is the supremum of the set of such p.

Our first result is the following

Theorem 2.5 *Let $G_{\delta t}$ be defined for $\delta t \in I_N$ and assume that it is a product of matrices of the form (3). If the linear scheme $G_{\delta t}$ has local order at least $p+1$ (hence its global order is at least p) when solving the numerical ODE system (1), then its A-order is at least p.*

Proof. Since the numerical flow $G_{\delta t}$ has local order $p+1$, its difference with the exact flow $F_{\delta t}$ reads

$$G_{\delta t} - F_{\delta t} = \begin{pmatrix} 1-\alpha(\delta t) & \alpha(\delta t) \\ \beta(\delta t) & 1-\beta(\delta t) \end{pmatrix} - \mathrm{e}^{-(N+1)c\delta t} P - Q = \mathcal{O}(\delta t^{p+1}).$$

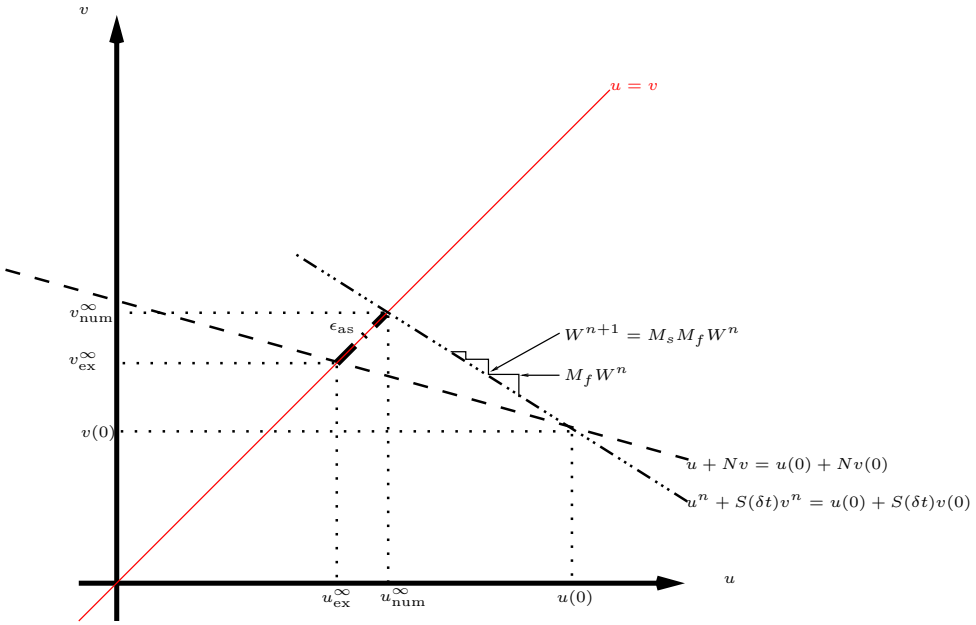


Figure 1: Evolution of the exact and numerical solution in the phase space $\mathbb{R}_u \times \mathbb{R}_v$. We note $W^n = (u^n, v^n)^t$.

This implies the following Taylor expansions for α and β:

$$\alpha(\delta t) = \frac{N}{N+1}(1 - \mathrm{e}^{-c(N+1)\delta t}) + \mathcal{O}(\delta t^{p+1}) \quad \text{and} \quad \beta(\delta t) = \frac{1}{N+1}(1 - \mathrm{e}^{-c(N+1)\delta t}) + \mathcal{O}(\delta t^{p+1}).$$

We infer that the slope of the equilibrium state is

$$S(\delta t) = \frac{\alpha(\delta t)}{\beta(\delta t)} = N + \mathcal{O}(\delta t^p).$$

■

Now, we define linear splitting schemes for the linear differential system (1), based on the composition of exact flows of the split vector fields and on θ-schemes solving the split equations approximately. We focus on their asymptotic behavior. We know from Property 2.3 and Theorem 2.5 that for all initial datum $(u^0, v^0) \in \mathbb{R}^2$, the numerical solutions provided by such splitting schemes (provided they are consistent with equation (1)) converge to an asymptotic state when the numerical time $n\delta t$ tends to infinity (and δt is fixed). The typical questions of interest are the following: What is the size of this relative asymptotic error with respect to the numerical time step δt ? Can we do better than the estimate on the relative A-order provided by Theorem 2.5 ?

2.3 Lie and Strang splitting schemes

Denoting by δt the numerical time step related to the "slow" equation, the time step associated to the "fast" equation is then $\delta t/N$. When dealing with slow/fast Lie-splitting methods, one has to choose which equation will be integrated first: either the slow equation first, and then the fast one

(which we denote by FS)[2], or the fast equation and then the slow one (which we denote by SF). Note that, in our very simple linear setting, the eigenvalues, eigenvectors, spectral projectors, etc of any FS splitting method can be deduced from those of a SF splitting formula in a way explained in Appendix A and the analysis extends straightforwardly. Therefore, we restrict ourselves to the study of SF Lie-splitting schemes.

The (exact or numerical) integration of the fast equation of (2) over a time step δt yields the flow

$$\Phi_{f,\delta t} \qquad \text{with matrix} \qquad M_f(\lambda_f(\delta t)).$$

In the same way, for the slow equation, we define

$$\Phi_{s,\delta t} \qquad \text{with matrix} \qquad M_s(\lambda_s(\delta t)),$$

with

$$\lambda_s(\delta t) = \lambda_f\left(\frac{\delta t}{N}\right).$$

In the following, we may use the superscript θ, which is is either a parameter within $[0,1]$ of a θ-scheme or $\theta = \text{ex}$ for the exact solution. This means that

$$\begin{cases} \lambda_f^{\theta_f}(\delta t) = \dfrac{1 - Nc\theta_f\delta t}{1 + (1-\theta_f)Nc\delta t}; & \lambda_s^{\theta_s}(\delta t) = \dfrac{1 - c\theta_s\delta t}{1 + (1-\theta_s)c\delta t}; \\ \lambda_f^{\text{ex}}(\delta t) = e^{-Nc\delta t}; & \lambda_s^{\text{ex}}(\delta t) = e^{-c\delta t}. \end{cases}$$

In case $\theta_f \in (1/2, 1]$ (resp. $\theta_s \in (1/2, 1]$), we assume that $(2\theta_f - 1)cN\delta t/N < 2$ (resp. $(2\theta_s - 1)c\delta t < 2$) so that $\lambda_f^{\theta_f}(\delta t/N) \in (0,1)$ (resp. $\lambda_s^{\theta_s}(\delta t) \in (0,1)$) (the associated schemes are A-stable). The stability interval I_N is the intersection of these domains in δt.

Remark 2.6 *Recall that* $\lambda_f^{\theta_f}(\delta t) = 1 - Nc\delta t + N^2c^2(1-\theta_f)\delta t^2 + \mathcal{O}(\delta t^3)$.

For any functions of δt λ_f, λ_s, we consider the following four schemes : given $W^n \in \mathbb{R}^2$, we set

- **Scheme #1** : (Lie type - slow time - subcycled) $W^{n+1} = G_1(\delta t)W^n$ where

$$G_1(\delta t) = M_s(\lambda_s(\delta t))M_f(\lambda_f(\delta t/N))^N$$

- **Scheme #2** : (Lie type - fast time - no subcycling) $W^{n+1} = G_2(\delta t)W^n$ where

$$G_2(\delta t) = (M_s(\lambda_s(\delta t/N))M_f(\lambda_f(\delta t/N)))^N$$

- **Scheme #3** : (Strang type - slow time - subcycled) $W^{n+1} = G_3(\delta t)W^n$ where

$$G_3(\delta t) = M_s(\lambda_s(\delta t/2))\ M_f(\lambda_f(\delta t/N))^N\ M_s(\lambda_s(\delta t/2))$$

- **Scheme #4** : (Strang type - fast time - no subcycling) $W^{n+1} = G_4(\delta t)W^n$ where

$$G_4(\delta t) = (M_s(\lambda_s(\delta t/(2N)))\ M_f(\lambda_f(\delta t/N))\ M_s(\lambda_s(\delta t/(2N))))^N$$

[2]We chose this notation because of the usual convention on the composition of flows : the first to be applied is written on the right-hand side of the others.

Using the notations of Lemma 2.1, we obtain that for Scheme #1,

$$\alpha_1(\delta t) = 1 - (\lambda_f(\delta t/N))^N \qquad \text{and} \qquad \beta_1(\delta t) = (1 - \lambda_s(\delta t))(\lambda_f(\delta t/N))^N,$$

for Scheme #2,

$$\alpha_2(\delta t) = 1 - \lambda_f(\delta t/N) \qquad \text{and} \qquad \beta_2(\delta t) = (1 - \lambda_s(\delta t/N))\lambda_f(\delta t/N),$$

for Scheme #3,

$$\alpha_3(\delta t) = (1 - \lambda_f(\delta t/N)^N)[\lambda_s(\delta t/2)]^N \quad \text{and} \quad \beta_3(\delta t) = (1 - \lambda_s(\delta t/2))(1 + [\lambda_f(\delta t/N)^N]\lambda_s(\delta t/2)),$$

and for Scheme #4,

$$\alpha_4(\delta t) = (1 - \lambda_f(\delta t/N))\lambda_s(\delta t/2) \quad \text{and} \quad \beta_4(\delta t) = (1 - \lambda_s(\delta t/2))(1 + \lambda_f(\delta t/N)\lambda_s(\delta t/2)).$$

Asymptotic order The above computations enable us to prove the following

Property 2.7 *A linear Lie-splitting method such as Scheme #1 and #2 has an A-order of at least 1. Moreover, if it involves to schemes of order at least 2, then its A-order is at most 1. However, it is possible to build linear Lie-splitting methods of A-order at least 2 involving schemes of order 1.*

Proof. The fact that Scheme #1 and Scheme #2 have order at least 1 follows from Theorem 2.5. Let us consider Scheme #1 and assume that we have the following Taylor expansion for $\lambda_s(\delta t)$ and $\lambda_f(\delta t/N)$:

$$\lambda_f(\delta t/N) = 1 - c\delta t + c^2 A_f \delta t^2 + \mathcal{O}(\delta t^3) \quad \text{and} \quad \lambda_s(\delta t) = 1 - c\delta t + c^2 A_s \delta t^2 + \mathcal{O}(\delta t^3).$$

We derive that

$$S_1(\delta t) = \frac{\alpha_1(\delta t)}{\beta_1(\delta t)} = N + cN(A_s - A_f + (N+1)/2)\delta t + \mathcal{O}(\delta t^2).$$

When the two schemes are of order at least 2, we have $A_f = A_s = 1/2$, so that the A-order is exactly 1. If one wants to build a Lie-splitting scheme such that its A-order is at least 2, then one just has to solve the equation $A_f - A_s = (N+1)/2$ for A_f and A_s. A similar computation yields

$$S_2(\delta t) = \frac{\alpha_2(\delta t)}{\beta_2(\delta t)} = N + c((1 - A_f)N + A_s)\delta t + \mathcal{O}(\delta t^2).$$

Hence, the choice $(A_f, A_s) = (1, 0)$ leads to a Lie-splitting method of A-order at least 2 with two underlying methods of order 1. ∎

Remark 2.8 *The crucial point lies in the fact that the linear combination of the derivatives A_f and A_s involves N in both cases, so that the slow and fast schemes have to be specifically designed with the knowledge of N if one wants to achieve the second A-order. Let us examine the θ-schemes case. One infers from Remark 2.6 that $(A_f, A_s) = (1 - \theta_f, 1 - \theta_s)$. So, as soon as $N > 1$, one cannot build schemes of type #1 or #2 of order at least 2 with θ-schemes, unless, in Scheme #2, the slow scheme is fully explicit and the fast scheme is fully implicit. In this very particular case, the A-order is infinite because $\alpha_2 = N\beta_2$. Note that, if a fully implicit scheme is at hand for the fast equation, it seems unwise to use a subcycling technique anyway, since there is no stability constraint on δt from the fast scheme part.*

Property 2.9 *A linear Strang-splitting method such as Scheme #3 and #4 involving only schemes of order at least 2 has an A-order of at least 2. Moreover, it is possible to build a Strang-splitting scheme of A-order at least 2 involving two schemes of order only 1.*

Proof. The fact that a Strang-splitting method involving two methods of order 2 is of A-order 2 comes from Theorem 2.5. Assume we have the same Taylor expansion as in the proof of Property 2.7. For Scheme #3, we have

$$S_3(\delta t) = \frac{\alpha_3(\delta t)}{\beta_3(\delta t)} = N + Nc(2A_s - 1 + 2 - 4A_f)\delta t/4 + \mathcal{O}(\delta t^2),$$

and for Scheme #4

$$S_4(\delta t) = \frac{\alpha_4(\delta t)}{\beta_4(\delta t)} = N + c(N(2A_f - 1) + 2 - 4A_s)\delta t/4 + \mathcal{O}(\delta t^2).$$

For example, one can choose $(A_f, A_s) = (1/4, 0)$ to have a Scheme #3 of A-order at least 2 involving two schemes of order 1. ∎

Remark 2.10 *In contrast to what occurs in the Lie case, the dependence upon N in the Strang subcycled scheme #3 is decoupled from the combination of A_f and A_s. In particular, in case the fast and slow schemes are θ-schemes, the above condition $1 - 2A_s + 4A_f - 2 = 0$ for Scheme #3 reads $4\theta_f - 2 + 1 - 2\theta_s = 0$ so that we have a fairly natural one-parameter family of couples of schemes of order 1, not depending on N, that lead to a subcycled scheme (#3) of A-order at least 2. In particular, on can choose to use an Euler explicit scheme for the slow equation ($\theta_f = 1$) and a semi-implicit scheme for the fast equation ($\theta_s = 1/4$) so that the subcycled Strang-splitting scheme #3 is at least of A-order 2. It is also possible to build a second A-order #4 scheme with θ-schemes provided one solves $2N(1-2\theta_s)+2\theta_f-1 = 0$. One sees in that case that the influence of the choice of θ_f weakens as N increases.*

Remark 2.11 *We can exchange the influence of the choices of A_s and A_f in the A-order by Strang-splitting with the order FSF, that is, by introducing*

$$\widetilde{G_3}(\delta t) = M_f(\lambda_f(\delta t/(2N)))^N \; M_s(\lambda_s(\delta t)) \; M_f(\lambda_f(\delta t/(2N)))^N,$$
$$\widetilde{G_4}(\delta t) = (M_f(\lambda_f(\delta t/(2N))) \; M_s(\lambda_s(\delta t)) \; M_f(\lambda_f(\delta t/(2N))))^N,$$

thanks to the computations detailed in Appendix A. The coefficient in front of δt^2 is then $4\theta_s - 2 + 1 - 2\theta_f = 0$ (resp. $2(1 - 2\theta_s) + N(2\theta_f - 1) = 0$) for Scheme $\widetilde{\#3}$ (resp. $\widetilde{\#4}$). One concludes easily that it is then possible to build a $\widetilde{\#3}$ scheme of A-order 2 with an explicit fast scheme ($\theta_f = 1$) and a semi-implicit slow scheme ($\theta_s = 3/4$). For $\widetilde{\#4}$ schemes, one notes that the choice of the fast scheme is now the more important.

Convergence rate Let us perform the same analysis on the convergence rate to equilibrium, *i.e.* the eigenvalues μ_i, $i \in \{1, \dots, 4\}$. We get the following Taylor expansions of $\rho_i(\delta t) = \mu_i(\delta t) - \mathrm{e}^{-(N+1)\delta t}$, that we summarize in the following table in the (A_f, A_s) form :

i	(A_f, A_s)
$\rho_1(\delta t)$	$c^2(N(2A_f-1)+2A_s-1)\delta t^2/2+\mathcal{O}(\delta t^3)$
$\rho_2(\delta t)$	$c^2(N^2(2A_f-1)+2A_s-1)\delta t^2/(2N)+\mathcal{O}(\delta t^3)$
$\rho_3(\delta t)$	$c^2(2N(2A_f-1)+2A_s-1)\delta t^2/4+\mathcal{O}(\delta t^3)$
$\rho_4(\delta t)$	$c^2(2N^2(2A_f-1)+2A_s-1)\delta t^2/(4N)+\mathcal{O}(\delta t^3)$

One notes at once that second order fast and slow schemes generate a second order approximation of the convergence rate, (as well as an A-order of 2 for Schemes #3 and #4). Besides, one can manage to construct a second order approximated rate choosing at least one of the fast and slow schemes to be of order 1, but the A-order will be exactly 1. The only combination of θ-schemes leading to a second order approximated rate and of A-order 2 consists in taking the Crank-Nicolson scheme for both the fast and slow schemes, using the Strang splitting (Schemes #3 and #4). In any case, the choice of the fast scheme plays a greater role than that of the slow scheme for the approximated rate. As predicted in Appendix A, schemes $\widetilde{\#3}$ and $\widetilde{\#4}$ have similar rates of convergence as schemes #3 and #4 : the choice of the fast scheme is always predominant.

2.4 Conclusion

Let us remind the reader that the applications we have in mind involve a fast equation for which an implicit scheme is hard to solve, thus implying the use of an explicit scheme, inducing a stability constraint on the numerical time-step δt. In that case, the subcycling techniques are computationally less costly, thus relevant.

We proved in this section that, in view of the aforementioned goal, we can indeed build a scheme of type $\widetilde{\#3}$ with $\theta_f = 1$ (explicit), $\theta_s = 1/4$ (semi-implicit), which will be of second A-order, even though it is (locally) consistent of order 1 with (1) and has a rate of convergence which approximates the exact rate at order 1. It is the only scheme, among the four types described above and involving an explicit resolution of the fast equation, that achieves a second A-order.

3 Analysis of the nonlinear system

3.1 Analysis of the exact solutions

In this section, we investigate the long time behavior of the two-scale nonlinear system (2). Let us first write this system in the form

$$\begin{cases} u' &= -N(u-v)[c+(u-v)] \\ v' &= (u-v)[c+(u-v)]. \end{cases} \tag{8}$$

This way, we are able to derive the following

Property 3.1 *Let $(u^0, v^0) \in \mathbb{R}^2$ be given. The maximal solution starting at (u^0, v^0) lies on the straight line of equation $u+Nv = u^0+Nv^0$. It is defined for all non-negative time if $u^0+c \geq v^0$ and it dies in finite positive time if $u^0 + c < v^0$. Moreover, if $u_0 + c = v_0$ then the solution is constant, and if $u^0 + c > v^0$ then the solution tends to the intersection of the two straight lines of equations $u + Nv = u^0 + Nv^0$ and $u = v$, i.e. to the point of coordinates $(u^0 + Nv^0)/(N + 1) \times (1, 1)$.*

Proof. The linear change of variable $(X, Y) = (u + Nv, u - v)$ yields the following equivalent differential system

$$\begin{cases} X' &= & 0 \\ Y' &= & -(N+1)Y(c-Y). \end{cases}$$

The second equation of this system has for maximal solution starting at $t = 0$ in $Y^0 \in \mathbb{R}$ the function

$$Y(t) = Y^0 \frac{\mathrm{e}^{-c(N+1)t}}{1 + \frac{Y^0}{c}(1 - \mathrm{e}^{-c(N+1)t})}$$

defined as long as $-c < Y^0(1 - \mathrm{e}^{-c(N+1)t})$. The result follows from this observation and the fact that the system reads (8). ∎

3.2 Splitting schemes with or without subcycling for the nonlinear problem (2)

Let us recall this result providing an estimate of the order of a splitting scheme (with or without subcycling) as a function of the order of the underlying schemes and the order of the splitting method.

Theorem 3.2 *Let us consider a differential system of the form*

$$\begin{cases} u' &= & Nf(u,v) \\ v' &= & g(u,v), \end{cases}$$

where f and g are smooth functions from $\mathbb{R}^2$ to $\mathbb{R}$. We denote by $\varphi_{e,\delta t}$ the exact flow of this equation. Let us denote by $\varphi_f(\delta t)$ (respectively) $\varphi_s(\delta t)$ the propagators at time δt of the two split equations:

$$\begin{cases} u' &= & Nf(u,v) \\ v' &= & 0 \end{cases} \qquad \text{(resp.)} \qquad \begin{cases} u' &= & 0 \\ v' &= & g(u,v). \end{cases}$$

Assume that $S_{f,\delta t}$ and $S_{s,\delta t}$ are numerical methods of respective orders p and q. Assume that a splitting method is defined for $a_1, \ldots, a_n, b_1, \ldots, b_2 \in \mathbb{C}$ by the formula

$$\Phi_{\delta t} = \Pi_{i=1}^n (S_{s,b_i\delta t} \circ S_{f,a_i\delta t}), \tag{9}$$

so that this method with the exact flows has order r. Then the order of the method $\Phi_{\delta t}$ is at least $\min(p, q, r)$, and so is the order of the method with subcycling

$$\Phi_{\delta t}^{sc} = \Pi_{i=1}^n (S_{s,b_i\delta t} \circ (S_{f,a_i\delta t/N})^N). \tag{10}$$

Proof. Since the method $S_{s,\delta t}$ has order p, we may write, when $\delta t \to 0$,

$$S_{s,\delta t} = \varphi_{s,\delta t} + \mathcal{O}(\delta t^{p+1}),$$

and similarly

$$S_{f,\delta t/N} = \varphi_{f,\delta t/N} + \mathcal{O}(\delta t^{q+1}).$$

The smoothness of the propagators implies that all $p \in \mathbb{N}^\star$,

$$S_{f,\delta t/N}^p = \varphi_{f,\delta t/N}^p + \mathcal{O}(\delta t^{q+1}),$$

where the constant in the Landau symbol depends on p. In particular, for $p = N$, using the semi-group property of the exact flow, we have

$$S^N_{f,\delta t/N} = \varphi_{f,\delta t} + \mathcal{O}(\delta t^{q+1}).$$

This implies that

$$\begin{aligned}
\Phi^{sc}_{\delta t} = \Pi_{i=1}^n (S_{s,b_i\delta t} \circ (S_{f,a_i\delta t/N})^N) &= \Pi_{i=1}^n (\varphi_{s,b_i\delta t} + \mathcal{O}(\delta t^{p+1})) \circ (\varphi_{f,a_i\delta t} + \mathcal{O}(\delta t^{q+1})) && (11)\\
&= \Pi_{i=1}^n (\varphi_{s,b_i\delta t} \circ \varphi_{f,a_i\delta t}) + \mathcal{O}(\delta t^{\min(p,q)+1}) && (12)\\
&= \varphi_{e,\delta t} + \mathcal{O}(\delta t^{\min(p,q,r)+1}), && (13)
\end{aligned}$$

since the splitting method is assumed to have order r when used with the exact flows. This proves the result for $\Phi^{sc}_{\delta t}$. The result for $\Phi_{\delta t}$ is even simpler. ■

In the following, we consider numerical splitting methods for the nonlinear problem (2) in the same way as for the linear problem (1) in Section 2.3: Scheme #1 is a SF Lie-splitting method with subcycling, Scheme #2 is a SF Lie-splitting method without subcycling, Scheme #3 is a FSF Strang-splitting method with subcycling, and Scheme #4 is a FSF Strang-splitting method without subcycling.

Once again, we consider numerical flows for the integration of the split equations described by θ-schemes, *i.e.* for the fast equation, the first component of $\Phi^{\theta_f}_{f,\delta t}(u^n, v^n)$ solves the equation in X

$$X - u^n = N\delta t\theta_f\big(c(v^n - u^n) - (u^n - v^n)^2\big) + N\delta t(1-\theta_f)\big(c(v^n - X) - (X - v^n)^2\big),$$

while its second one is its second argument and, for the slow equation, the second component of $\Phi^{\theta_f}_{f,\delta t}(u^n, v^n)$ solves the equation in X

$$X - v^n = \delta t\theta_s\big(c(u^n - v^n) + (u^n - v^n)^2\big) + \delta t(1-\theta_s)\big(c(u^n - X) + (u^n - X)^2\big),$$

while its first component is its first argument.

3.3 Numerical examples of splitting methods for problem (2)

We run the four schemes with four different values of the couple (θ_f, θ_s). We sum up the results on the asymptotic order in Table 1 and provide numerical results in Figure 2 and Figure 3. These results were obtained with final time $T = 2.0$, speed $c = 1$, factor $N = 50$, initial datum $(u^0, v^0) = (5, 1)$, so that, using the analysis carried out in the proof of Property 3.1, the exact solution at final time is within a distance smaller than 10^{-40} of its asymptotic limit $55/51 \times (1, 1)$.

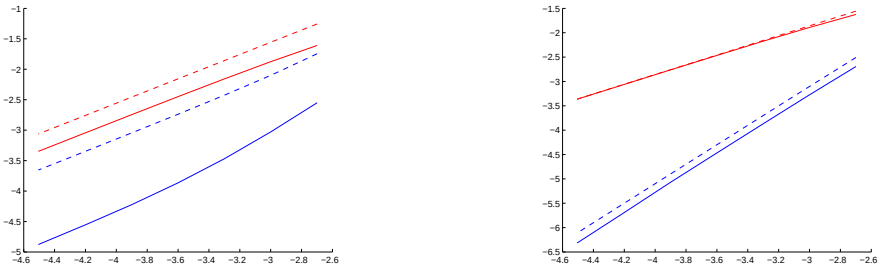
By Theorem 3.2, we know that the Lie-splitting schemes (Scheme #1 and Scheme #2) have classical order 1 for any possible choice of (θ_f, θ_s). The first two columns of Table 1 show that the asymptotic order if also 1 in these cases. Theorem 3.2 also implies that the Strang-splitting scheme #3 has at least order 1 with the choice $(\theta_f, \theta_s) = (1, 0)$ and the asymptotic orders at the end of the first line of Table 1 show that the asymptotic order is also 1 in this case. The same theorem also ensures that Scheme #3 has order 2 when applied with $(\theta_f, \theta_s) = (1/2, 1/2)$. The asymptotic orders displayed at the end of the second line of Table 1 show that the asymptotic order is also 2 in this case. The end of the 2 last lines is surely the interesting part of this section: for $(\theta_f, \theta_s) = (1, 3/4)$ and $(\theta_f, \theta_s) = ((N-1)/(2N), 1/4)$, the classical order of the splitting method is, by Theorem 3.2 at least 1. In the first case $(\theta_f, \theta_s) = (1, 3/4))$, the numerical results show that the subcycled scheme #3 has A-order 2 while the Strang-splitting scheme #4 has A-order 1. We

(θ_f, θ_s)	Scheme #1	Scheme #2	Scheme #3	Scheme #4
$(1.0, 0.0)$	0.9671	1.0000	1.2808	1.0499
$(0.5, 0.5)$	0.9677	1.0000	2.0071	1.9952
$(1.0, 0.75)$	0.9686	1.0021	1.9889	1.0548
$(\frac{N-1}{2N}, 0.25)$	0.9672	0.9991	1.5041	1.9932

Table 1: Asymptotic error for the 4 schemes for some values of (θ_f, θ_s).

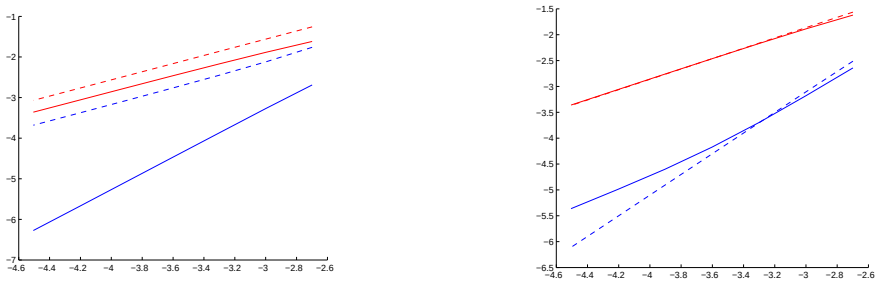
recall that, for these parameters, the Scheme #3 was of A-order 2 in the linear setting (see Remark 2.10). In the second case $(\theta_f, \theta_s) = ((N-1)/(2N), 1/4)$, the same phenomenon occurs : Scheme #3 has A-order 1 while Scheme #4 has asymptotic A-order 2. We recall that these values of the parameters were chosen in the linear setting in such a way that the (linear) Scheme #4 has A-order 2.

These examples suggest that, in this context, the A-order of a scheme applied to the linear problem is the same to the A-order of the scheme applied to the nonlinear problem. This can be explained by the fact that the two problems (1) and (2) have the same set of attractive equilibrium points (the straight line $u = v$), they project the initial datum (u_0, v_0) (chosen in an appropriate subset of the phase plane $(u^0 + c < v^0)$) on the same equilibrium point $(u^0 + Nv^0)/(N+1) \times (1, 1)$, and in the neighborhood of this equilibrium point, $(u - v)^2 << |u - v|$. In particular, these examples show that it is possible to build in the nonlinear setting, as well in the linear setting, splitting methods with asymptotic order greater than the classical order of the schemes used for solving the split-equations.



(a) (b)

Figure 2: Logarithm of the asymptotic error as a function of the time step: Scheme #1 (full red line), Scheme #2 (dotted red line), Scheme #3 (full blue line), Scheme #4 (dotted blue line). $(\theta_f, \theta_s) = (1.0, 0.0)$ (on the left) and $(\theta_f, \theta_s) = (0.5, 0.5)$ (on the right).



(a) (b)

Figure 3: Logarithm of the asymptotic error as a function of the time step: Scheme #1 (full red line), Scheme #2 (dotted red line), Scheme #3 (full blue line), Scheme #4 (dotted blue line). $(\theta_f, \theta_s) = (1.0, 0.75)$ (on the left) and $(\theta_f, \theta_s) = (N-1)/(2N), 0.75)$ (on the right).

4 A coupled reaction-diffusion system

4.1 The homogeneous Dirichlet problem

This section aims at studying the behavior of time-splitting schemes involving subcyling techniques for solving the following system of partial differential equations

$$\left\{ \begin{array}{rcll} \partial_t u &=& \nu_1 \Delta u + c_1(v-u) & t>0, x\in(0,L) \\ \partial_t v &=& \nu_2 \Delta v + c_2(u-v) & t>0, x\in(0,L) \end{array} \right. , \tag{14}$$

with homogeneous Dirichlet boundary conditions at $x=0$ and $x=L$, and given initial data u^0 and v^0 in an appropriate function space. Of course, $\Delta = \partial_x^2$ is the Laplace operator and $L>0$ is given. We focus on the case where one of the equations in System (14) is "fast" and the other is "slow". Moreover, we assume the "speed" ratios allow us to actually do subcycling. This means that

$$\frac{\nu_1}{\nu_2} = \frac{c_1}{c_2} = N \in \mathbb{N}^\star, \tag{15}$$

and $N >> 1$[3]. Consequently, in accordance with Section 2, we will use the notation $\nu = \nu_2$ and $c = c_2$.

In that case, the first equation in (14) is the "fast" one, so u is the "fast" unknown and the second one (on v) is the slow one.

Let us recall that we have the following

Theorem 4.1 *For all initial data $(u^0, v^0) \in L^2(0,L)$, System (14) has a unique solution $t \mapsto (u(t), v(t))$ in $C^0([0,+\infty), L^2(0,L)) \cap C^\infty((0,+\infty)\times[0,L])$, satisfying $(u,v)(0) = (u^0, v^0)$.*

Proof. If one looks for solutions of the form

$$u(t,x) = \sum_{k=1}^{+\infty} \alpha_k(t) \sin\left(\frac{k\pi}{L}x\right) \quad \text{and} \quad v(t,x) = \sum_{k=1}^{+\infty} \beta_k(t) \sin\left(\frac{k\pi}{L}x\right),$$

then the coefficients satisfy the differential systems

$$\begin{cases} \dot\alpha_k(t) = -N(c+\nu\frac{k^2\pi^2}{L^2})\alpha_k(t) + Nc\beta_k(t) \\ \dot\beta_k(t) = +c\alpha_k(t) - (c+\nu\frac{k^2\pi^2}{L^2})\beta_k(t) \end{cases}$$

and the eigenvalues λ_k and μ_k of the matrices $M_k = \begin{pmatrix} -N(c+\nu\frac{k^2\pi^2}{L^2}) & +Nc \\ +c & -(c+\nu\frac{k^2\pi^2}{L^2}) \end{pmatrix}$ are both real, negative and satisfy

$$\lambda_k \sim -N\frac{k^2\pi^2}{L^2} \quad \text{and} \quad \mu_k \sim -\frac{k^2\pi^2}{L^2}.$$

■

The following theorem deals with the asymptotic behavior of the solutions of System (14):

Theorem 4.2 *For all solutions (u,v) of System (14) and all $t \geq 0$, we have*

$$\int_0^L (|u^2| + N|v|^2)(t)\mathrm{d}x \leq \Big(\int_0^L (|u|^2 + N|v|^2)(0)\mathrm{d}x \Big) \mathrm{e}^{-\frac{2\pi^2\nu}{L^2}t}$$

[3] Yet, we are not interested in the limit $N \to +\infty$.

Proof. Let (u, v) be a smooth solution of (14). We compute

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int_0^L (|u|^2 + N|v|^2)\mathrm{d}x\right)(t) &= N\nu\int_0^L u(t)\Delta u(t) + N\nu\int_0^L v(t)\Delta v(t) + Nc\int_0^L (u(v-u)+v(u-v))(t)\\ &= -N\nu\int_0^L |\nabla u(t)|^2 - \nu\int_0^L N|\nabla v(t)|^2 - Nc\int_0^L |u(t)-v(t)|^2\\ &\leq -\frac{2\pi^2\nu}{L^2}\frac{1}{2}\int_0^L (|u(t)|^2 + N|v(t)|^2)\mathrm{d}x,\end{aligned}$$

using that $N > 1$, $\int_0^L |u(t) - v(t)|^2 \geq 0$ and the Poincaré inequality

$$\int_0^L |u(t)|^2 \leq \frac{L^2}{\pi^2}\int_0^L |\nabla u(t)|^2.$$

■

Space discretization In the following, we will use the classical finite-difference discretization of minus the Laplace operator, using the symmetric tridiagonal $M \times M$ matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & \dots & 0\\ -1 & 2 & -1 & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & 0\\ \vdots & \ddots & -1 & 2 & -1\\ 0 & \dots & 0 & -1 & 2\end{pmatrix},$$

where $M \in \mathbb{N}^\star$ and $\delta x = L/(M+1)$.

We note for all $i \in \{0, \dots, M+1\}$, $x_i = i.\delta x$. For any function u on $(0, L)$, we write $U = (u_1, \dots, u_M)$ any vector of approximations of $u(x_1), \dots, u(x_M)$.

Let us recall that the eigenvalues of A are

$$\lambda_p = 4\sin^2\Big(\frac{p\pi}{2(M+1)}\Big), \qquad 1 \leq p \leq M, \tag{16}$$

and that for all $p \in \{1, \dots, M\}$, the vector

$$(\sin(1p\pi/(M+1)), \sin(2p\pi/(M+1)), \dots, \sin(Mp\pi/(M+1))),$$

is an eigenvector of A associated to λ_p. In the following, we denote by

$$A = PDP^{-1}, \tag{17}$$

the corresponding diagonalisation of A.

Numerical analysis of the rate of convergence for several time-splitting schemes Assume $\delta t > 0$ is given. The methods we have in mind all share the same basic idea : we discretize in time separately the spatially-discretized versions of both equations of System (14). The "slow" one is discretized on an interval of length δt and we denote by $\Phi_{slow,\delta t}$ its numerical flow. The "fast" one is discretized on an interval of length $\delta t/N$ and we denote by $\Phi_{fast,\delta t/N}$ its numerical

flow. Then, we compute numerical flows using splitting methods and subcycling by considering numerical flows such as

$$\Psi_{Lie,\delta t} = \Phi_{slow,\delta t} \circ \Phi^N_{fast,\delta t/N},$$

and

$$\Psi_{Strang,\delta t} = \Phi_{slow,\delta t/2} \circ \Phi^N_{fast,\delta t/N} \circ \Phi_{slow,\delta t/2}{}^4.$$

As we did in Section 2 and in Section 3, we consider θ-schemes for the solution of the slow and fast equations. We choose two parameters $(\theta_f, \theta_s) \in [0,1]^2$. The numerical integrators involved in the splitting scheme therefore read:

$$\Phi_{fast,\delta t/N}(u^n, v^n) = \Big[\Big(I - \theta_f \delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A\Big)\Big(I + (1-\theta_f)\delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A\Big)^{-1} u^n + c\delta t v^n \, , \, v^n\Big], \quad (18)$$

and

$$\Phi_{slow,\delta t}(u^n, v^n) = \Big[u^n \, , \, \Big(I - \theta_s \delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A\Big)\Big(I + (1-\theta_s)\delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A\Big)^{-1} v^n + c\delta t u^n\Big]. \quad (19)$$

This way, a stability condition reads

$$\delta t \leq \frac{1}{2} \frac{1}{c + 4\nu/(\delta x)^2}. \quad (20)$$

Note that the $1/2$ in the above relation is restrictive, since the two schemes are stable if one replaces this $1/2$ by 1, but we will use this restriction later. Note also that, more importantly, the stability condition on the scheme (20) is actually independent on N, and this is a very interesting feature of splitting schemes involving subcycling.

Let us define for $i \in \{s, f\}$,

$$B_i = I - \theta_i \delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A) \qquad \text{and} \qquad C_i = \Big(I + (1-\theta_i)\delta t(cI + \nu \frac{\delta t}{(\delta x)^2} A)\Big).$$

Since they are polynomials in A, the matrices $I, C_s, C_f, B_s, B_f, C_s^{-1}, C_f^{-1}$ and A do commute. The matrices of the linear mappings $\Phi_{slow,\delta t}$ and $\Phi_{fast,\delta t/N}$ in the canonical basis of $\mathbb{R}^{2M}$ read respectively

$$M_s = \begin{pmatrix} I & 0 \\ c\delta t C_s^{-1} & B_s C_s^{-1} \end{pmatrix} \qquad \text{and} \qquad M_f = \begin{pmatrix} B_f C_f^{-1} & c\delta t C_f^{-1} \\ 0 & I \end{pmatrix} \quad (21)$$

Therefore, the matrix of $\Phi^N_{fast,\delta t/N}$ reads

$$M_f^N = \begin{pmatrix} (B_f C_f^{-1})^N & c\delta t C_f^{-1}(I + (B_f C_f^{-1})^1 + \cdots + (B_f C_f^{-1})^{N-1}) \\ 0 & I \end{pmatrix}.$$

The corresponding matrix of $\Psi_{Lie,\delta t}$ is hence

$$\begin{pmatrix} (B_f C_f^{-1})^N & c\delta t C_f^{-1}(I + (B_f C_f^{-1})^1 + \cdots + (B_f C_f^{-1})^{N-1}) \\ c\delta t C_s^{-1}(B_f C_f^{-1})^N & B_s C_s^{-1} + c^2\delta t^2(I + B_f C_f^{-1} + \cdots + (B_f C_f^{-1})^{N-1}) \end{pmatrix}. \quad (22)$$

[4]Since this method is conjugated to the method $\Phi_{Lie,\delta t}$, the results on the latter method extend to $\Phi_{Strang,\delta t}$

Denoting by $\mathcal{P}$ the matrix (see (17))

$$\mathcal{P} = \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix},$$

we obtain that the matrix $\mathcal{D} := \mathcal{P}^{-1}\Psi_{Lie,\delta t}\mathcal{P}$ is exactly the same as that of (22) where A is replaced with D in the definition of the matrices B_f, B_s, C_f and C_s. In particular, it consists in four squared blocks of size $2M \times 2M$, each of which is diagonal. We infer that the eigenvalues of $\Psi_{Lie,\delta t}$ are the roots of the M polynomials

$$\lambda^2 - \Big((\phi_{f,p}^{-1}\psi_{f,p})^N + \psi_{s,p}^{-1}\phi_{s,p} + c^2\delta t^2\phi_{f,p}^{-1}\phi_{s,p}^{-1}\sum_{k=0}^{N-1}(\phi_{f,p}^{-1}\psi_{f,p})^k\Big)\lambda + (\phi_{f,p}^{-1}\psi_{f,p})^N\psi_{s,p}^{-1}\phi_{s,p}, \tag{23}$$

where[5] for all $i \in \{s, f\}$,

$$\forall p \in \{1, \dots, M\}, \quad \psi_{i,p} = (1 - \theta_i\delta t(c + \nu\frac{\lambda_p}{(\delta x)^2})) \quad \text{and} \quad \phi_{i,p} = (1 + (1-\theta_i)\delta t(c + \nu\frac{\lambda_p}{(\delta x)^2})).$$

We extend the functions $p \mapsto \lambda_p$, $p \mapsto \psi_{i,p}$ and $p \mapsto \psi_{i,p}$ to the continuous interval $(0, M+1)$. The stability condition (20) ensures that for all p, and all $i \in \{s, f\}$, $\psi_{i,p} \in [1/2, 1] \subset (0, 1]$ and $\phi_{i,p}^{-1} \in [2/3, 1] \subset (0, 1]$. The functions $p \mapsto \psi_{i,p}$ and $p \mapsto \psi_{i,p}$ are smooth, decreasing on $(0, M+1)$ with values in $(0, 1)$. Hence, any finite product of such functions and any finite sum is smooth and decreasing on $(0, M+1)$. For example,

$$P(p) = (\phi_{f,p}^{-1}\psi_{f,p})^N, \quad Q(p) = \psi_{s,p}^{-1}\phi_{s,p} \quad \text{and} \quad \Sigma(p) = c^2\delta t^2\phi_{f,p}^{-1}\phi_{s,p}^{-1}\sum_{k=0}^{N-1}(\phi_{f,p}^{-1}\psi_{f,p})^k,$$

are positive decreasing functions of p on $(0, M+1)$. Note that the discriminant of the polynomial (23) is

$$d_p := \Big(P(p) + Q(p) + \Sigma(p)\Big)^2 - 4Q(p)P(p) = \Big(Q(p) - P(p) + \Sigma(p)\Big)^2 + 4P(p)\Sigma(p) > 0, \tag{24}$$

so that the eigenvalues of $\Psi_{Lie,\delta t}$ are the real numbers

$$\lambda_p^- = \frac{P(p) + Q(p) + \Sigma(p) - \sqrt{d_p}}{2} \qquad \text{and} \qquad \lambda_p^+ = \frac{P(p) + Q(p) + \Sigma(p) + \sqrt{d_p}}{2}$$

for $p \in \{1, \dots, M\}$. Note that, with the stability condition (20), we have for all p, $0 < \lambda_p^- < \lambda_p^+$. Moreover, we have the following monotonicity property:

Lemma 4.3 *Assume $N \geq 2$. The map $p \mapsto \lambda_p^+$ is decreasing.*[6].

Proof. Note that for all $i \in \{s, f\}$ and all $p \in (0, M+1)$, the function $\theta_i \mapsto \phi_{i,p}^{-1}\psi_{i,p}$ is decreasing. Hence,

$$1 - \delta t(c + \nu\lambda_p) \leq \phi_{s,p}^{-1}\psi_{s,p} \qquad \text{and} \qquad (\phi_{f,p}^{-1}\psi_{f,p})^N \leq \Big(\frac{1}{1 + \delta t(c + \nu\lambda_p)}\Big)^N.$$

[5] Recall that the λ_p are the eigenvalues of the matrix A defined in (16).

[6] Note that d_p is not a decreasing function of p in general.

Since $N \geq 2$, we have $(1/(1+\delta t(c+\nu\lambda_p)))^N \leq (1/(1+\delta t(c+\nu\lambda_p)))^2$ and hence for all $p \in (0, M+1)$,

$$\phi_{s,p}^{-1}\psi_{s,p} - (\phi_{f,p}^{-1}\psi_{f,p})^N \geq \frac{(1-\delta t(c+\nu\lambda_p))\Big(1+\delta t(c+\nu\lambda_p)\Big)^2 - 1}{\Big(1+\delta t(c+\nu\lambda_p)\Big)^2}.$$

The stability condition (20) ensures that $\delta t(c+\nu\lambda_p) \in (0, 1/2]$. Since the function $x \mapsto (1-x)(1+x)^2 - 1 = x(1-x-x^2)$ is positive on $(0, 1/2]$, we finally get that for all $p \in (0, M+1)$,

$$Q(p) - P(p) = \phi_{s,p}^{-1}\psi_{s,p} - (\phi_{f,p}^{-1}\psi_{f,p})^N > 0. \tag{25}$$

Recall that, with (24),

$$\sqrt{d_p} > \sqrt{(Q(p) - P(p) + \Sigma(p))^2}.$$

Since $Q(p) - P(p) > 0$ and $\Sigma(p) > 0$, we infer

$$\sqrt{d_p} > Q(p) - P(p). \tag{26}$$

We can use this inequality to derive an estimate from above on the quantity Differentiating the function $p \mapsto \lambda_p^+$ with respect to p yields (we omit the p as argument of the variables to keep notations short)

$$\begin{aligned}
2\sqrt{d_p}\frac{\mathrm{d}}{\mathrm{d}p}\lambda_p^+ &= (\underbrace{P'+Q'+\Sigma'}_{<0})\sqrt{d_p} + (P+Q+\underbrace{\Sigma}_{>0})(\underbrace{P'+Q'+\Sigma'}_{<0}) - 2(PQ)' \\
&< (P'+Q'+\Sigma')(Q-P) + (P+Q)(P'+Q'+\Sigma') - 2P'Q - 2PQ' \\
&< (P'+Q'+\underbrace{\Sigma'}_{<0})(\underbrace{2Q}_{>0}) - 2P'Q - 2PQ' \\
&< 2(P'+Q')Q - 2P'Q - 2PQ' \\
&< \underbrace{2Q'}_{<0}(\underbrace{Q-P}_{>0}) \\
&< 0,
\end{aligned}$$

using (25) again. This implies that the derivative of $p \mapsto \lambda_p^+$ is negative on $(0, M+1)$ and proves the lemma. ∎

Hence the biggest eigenvalue of $\Psi_{Lie,\delta t}$ is λ_1^+. Of course, an asymptotic expansion of that biggest eigenvalue as $\delta t \to 0^+$ helps us controlling the exponential decay of the L^2 norm of the numerical solution provided by $\Psi_{Lie,\delta t}$. This allows us to prove the following

Theorem 4.4 *Let $c, \nu > 0$, $N \geq 2$ and $\Phi_{Lie,\delta t}$ be defined as above. Assume $M \in \mathbb{N}^\star$ is given. There exists[7] $C, \gamma, h > 0$ such that for all $T > 0$, all $U^0, V^0 \in \mathbb{R}^M$, all $\delta t \in (0, h)$ and all $n \in \mathbb{N}$ with $n\delta t \leq T$, we have*

$$\|\Psi_{Lie,\delta t}^n(U^0, V^0)\|_2 \leq C\mathrm{e}^{-\gamma n\delta t}\|(U^0, V^0)\|_2. \tag{27}$$

Remark 4.5 *Note that one can impose $\gamma \geq N\nu\lambda_1/((N+1)(\delta x)^2)$ in this case (provided h is small enough). Since $N\nu\lambda_1/(\delta x)^2 \to N\nu\frac{\pi^2}{L^2}$ as $\delta x \to 0^+$ (or equivalently as $M \to +\infty$), we have, at least asymptotically with respect to δx, a numerical decay rate of the appropriate order with respect to the parameters ν and L: we compare the exact decay rate $\nu\pi^2/L^2$ from Theorem 4.2 with the asymptotic numerical one $N\nu\pi^2/(L^2(N+1))$ (recall that N is large).*

[7] The reason for the constant C is the lack of symmetry of the matrix $\mathcal{D}$.

Proof. Let $M \in \mathbb{N}^\star$ be fixed and $p \in \{1, \dots, M\}$ be given. Since

$$\phi_{f,p}^{-1}\psi_{f,p} = \frac{1 - \theta_f \delta t(c + \nu \frac{\lambda_p}{(\delta x)^2})}{1 + (1-\theta_f)\delta t(c + \nu \frac{\lambda_p}{(\delta x)^2})},$$

we may write for $k \in \{0, \dots, N\}$

$$(\phi_{f,p}^{-1}\psi_{f,p})^k = 1 - kf(p)\ \delta t + \mathcal{O}(\delta t^2),$$

where $f(p) = c + \nu\lambda_p/(\delta x)^2$ is a bounded quantity when $\delta t \to 0^+$ (recall that M is fixed). We infer that

$$\sum_{k=0}^{N-1} (\phi_{f,p}^{-1}\psi_{f,p})^k = N - f(p)\frac{N(N-1)}{2}\delta t + \mathcal{O}(\delta t^2).$$

Following the same way, we obtain Taylor expansions for $P(p)$, $Q(p)$, $\Sigma(p)$ and then d_p and eventually λ_p^+ when δt tends to 0:

$$\lambda_p^+ = 1 - \frac{f(p)(N+1) - \sqrt{f(p)^2(N-1)^2 + 4Nc^2}}{2}\delta t + \mathcal{O}(\delta t^2),$$

and therefore,

$$\frac{T}{\delta t}\ln(\lambda_1^+) = -\frac{T}{2}\Big(f(1)(N+1) - \sqrt{f(1)^2(N-1)^2 + 4Nc^2}\Big) + \mathcal{O}(\delta t).$$

Note that, since $0 < c < f(1)$, we have $0 < 4Nc^2 < 4Nf(1)$ and hence

$$f(1)^2(N+1)^2 - f(1)^1(N-1)^2 = 4Nf(1)^2 > 4Nc^2$$

and therefore

$$f(1)(N+1) - \sqrt{f(1)^2(N-1)^2 + 4Nc^2} > 0.$$

Since λ_1^+ is the biggest eigenvalue of $\Psi_{Lie,\delta t}$, this proves the result. Note also that the constant γ can be taken arbitrary close to

$$\frac{1}{2}f(1)(N+1) - \sqrt{f(1)^2(N-1)^2 + 4Nc^2} = \frac{1}{2}\Big(f(1)(N+1) - \sqrt{f(1)^2(N+1)^2 - 4N(f(1)^2 - c^2)}\Big).$$

Using the mean value theorem, for some $c_\theta \in (0, 4N(f(1)^2 - c^2))$, the latter quantity is equal to

$$\frac{1}{2}\frac{1}{2}\frac{4N(f(1)^2 - c^2)}{\sqrt{f(1)^2(N+1)^2 - c_\theta}} > N\frac{f(1)^2 - c^2}{f(1)(N+1)} = \frac{N}{N+1}\frac{(f(1)+c)}{f(1)}(f(1)-c) \geq \frac{N}{N+1}\nu\frac{\lambda_1}{(\delta x)^2}.$$

■

The non homogeneous Dirichlet problem In this section we consider System (14) equipped with inhomogeneous Dirichlet boundary conditions, namely

$$u(t,0) = u_l, \qquad u(t,L) = u_r, \qquad v(t,0) = v_l, \qquad v(t,L) = v_r, \tag{28}$$

where u_l, v_l, u_r and v_r are four given real numbers. As in the homogeneous case above, there is a unique stationary solution to the boundary value problem :

Property 4.6 *The PDE system* (14) *with non homogeneous Dirichlet boundary conditions has a unique stationary solution given by*

$$\begin{cases} u: \quad x \mapsto \frac{u_l+v_l}{2} + \frac{(u_r+v_r-u_l-v_l)x}{2L} + \frac{(u_l-v_l)[\cosh(x/\alpha)-\cosh(L/\alpha)\sinh(x/\alpha)\sinh(L/\alpha)]}{2} + \frac{(u_r-v_r)\sinh(x/\alpha)\sinh(L/\alpha)}{2} \\ v: \quad x \mapsto \frac{u_l+v_l}{2} + \frac{(u_r+v_r-u_l-v_l)x}{2L} - \frac{(u_l-v_l)[\cosh(x/\alpha)-\cosh(L/\alpha)\sinh(x/\alpha)\sinh(L/\alpha)]}{2} - \frac{(u_r-v_r)\sinh(x/\alpha)\sinh(L/\alpha)}{2} \end{cases} \tag{29}$$

where $\alpha = \sqrt{\nu/(2c)}$.

The solution of (14)-(28) is then obtained from the unique solution to (14) with homogeneous Dirichlet conditions described in Theorem 4.1 by adding (29).

Let us consider two θ-schemes for the discretization as described in (18)-(19) : the sequence $((U^n, V^n)^t)_{n\in\mathbb{N}}$ is defined by an arithmetic-geometric recursion : given U^0, V^0, for all $n \geq 0$,

$$\begin{pmatrix} U^{n+1} \\ V^{n+1} \end{pmatrix} = \mathcal{M} \begin{pmatrix} U^n \\ V^n \end{pmatrix} + \mathcal{M}_u \begin{pmatrix} u_l \\ \vdots \\ u_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \mathcal{M}_v \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_l \\ \vdots \\ v_r \end{pmatrix} \tag{30}$$

where $\mathcal{M}$ is defined as a product of matrices (21) and $\mathcal{M}_u$ and $\mathcal{M}_v$ are $2M$-by-$2M$ matrices.

Let us list the numerical experiments we conducted :

- Scheme #1 (Lie - SF - slow time - subcycled) : $M_s := M_s(\delta t)$ and $M_f := M_f(\delta t/N)$

$$\mathcal{M} = M_s M_f^N, \;\; \mathcal{M}_u = \frac{\delta t}{N} M_s \sum_{k=0}^{N-1} M_f^k \begin{pmatrix} B_f^{-1} & 0 \\ 0 & 0 \end{pmatrix} \;\text{ and }\; \mathcal{M}_v = \delta t \begin{pmatrix} 0 & 0 \\ 0 & B_s^{-1} \end{pmatrix}$$

- Scheme #2 (Lie - SF - fast time - no subcycling) : $M_s := M_s(\delta t)$ and $M_f := M_f(\delta t/N)$

$$\mathcal{M} = (M_s M_f)^N, \;\; \mathcal{M}_u = \frac{\delta t}{N} \sum_{k=0}^{N-1} (M_s M_f)^k \begin{pmatrix} B_f^{-1} & 0 \\ 0 & 0 \end{pmatrix} \;\text{ and }\; \mathcal{M}_v = \frac{\delta t}{N} \sum_{k=0}^{N-1} (M_s M_f)^k \begin{pmatrix} 0 & 0 \\ 0 & B_s^{-1} \end{pmatrix}$$

- Scheme #3 (Strang - SFS - slow time - subcycled) : $M_s := M_s(\delta t/2)$ and $M_f := M_f(\delta t/N)$

$$\mathcal{M} = M_s M_f^N M_s, \;\; \mathcal{M}_u = \frac{\delta t}{N} M_s \sum_{k=0}^{N-1} M_f^k \begin{pmatrix} B_f^{-1} & 0 \\ 0 & 0 \end{pmatrix} \;\text{ and }\; \mathcal{M}_v = \delta t (I_{2M} + M_s M_f^N) \begin{pmatrix} 0 & 0 \\ 0 & B_s^{-1} \end{pmatrix}$$

- Scheme #4 (Strang - SFS - fast time - no subcycling) : $M_s := M_s(\delta t/(2N))$ and $M_f := M_f(\delta t/N)$

$$\mathcal{M} = (M_s M_f M_s)^N, \;\; \mathcal{M}_u = \frac{\delta t}{N} \sum_{k=0}^{N-1} (M_s M_f M_s)^k \begin{pmatrix} B_f^{-1} & 0 \\ 0 & 0 \end{pmatrix} \;\text{ and }\; \mathcal{M}_v = \frac{\delta t}{2N} \sum_{k=0}^{N-1} (M_s M_f M_s)^k (I_{2M} + M_s M_f)$$

The rate of convergence is given by Theorem 4.4. To compute the asymptotic numerical solution of a given method of type (32), we need to solve the $2M$-by-$2M$ linear system

$$(I_{2M} - \mathcal{M})W = \mathcal{M}_u \begin{pmatrix} u_l \\ \vdots \\ u_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \mathcal{M}_v \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_l \\ \vdots \\ v_r \end{pmatrix}.$$

Let us denote

$$P_u := \begin{pmatrix} I_M & 0 \\ 0 & 0 \end{pmatrix} \text{ and } P_v := \begin{pmatrix} 0 & 0 \\ 0 & B_s^{-1} I_M \end{pmatrix}.$$

According to Theorem 4.4, 1 is not an eigenvalue of $\mathcal{M}$ in the Lie case, so that $\mathcal{M}$ is invertible for Schemes #1 and #2. The numerical tests we conducted for several values of θ_f, θ_s and N showed that the matrix $I_{2M} - \mathcal{M}$ is also invertible for Schemes #3 and #4 and that, for a given set of parameters $(\theta_f, \theta_s, \delta t, \delta x)$ $(I_{2M} - \mathcal{M})^{-1}(\mathcal{M}_u P_u + \mathcal{M}_v P_v)$ are equal for all schemes #1 to #4, independently of N. To this day, we have not proved this property. We show here the graph obtained with Scheme #1 for the following sets of parameters , $N = 10$ being fixed :

- $N = 10$,
- $(u_l, u_r, v_l, v_r) = (1, 2, -1, 4)$, $M = 20, 40, 80, 160$, $\delta x = L/(M+1)$, $\delta t = \delta x^2/\nu_1/2$, $(\theta_f, \theta_s) = (1, 1)$ [explicit,explicit]
- $(u_l, u_r, v_l, v_r) = (2, 4, -1, 4)$, $M = 20, 40, 80, 160$, $\delta x = L/(M+1)$, $\delta t = 0.01$, $(\theta_f, \theta_s) = (1/2, 1/2)$ [Crank-Nicolson,Crank-Nicolson]

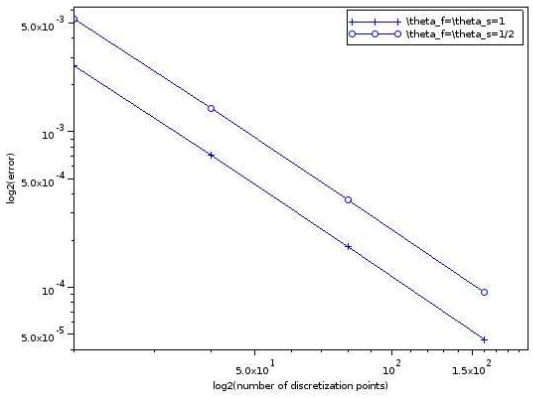


Figure 4: L^∞-error of the asymptotic numerical and exact states for explicit/explicit and Crank-Nicolson/Crank-Nicolson schemes. The numerical order is 1.94.

From 4, we conclude that the asymptotic state depends only on the spatial discretization through δx and does not depend on the time discretization δt or the values of θ. Moreover, the numerical order is close to 2 is δx.

4.2 The homogeneous Neumann problem

We consider the following system of partial differential equations

$$\left\{ \begin{array}{rcll} \partial_t u &=& N\nu\Delta u + Nc(v-u) & t>0, x\in(0,L) \\ \partial_t v &=& \nu\Delta v + c(u-v) & t>0, x\in(0,L) \end{array} \right. , \tag{31}$$

equipped with homogeneous Neumann boundary conditions at $x=0$ and $x=L$, and given initial data u^0 and v^0 in an appropriate function space. In that case, the first equation in (31) is the "fast" one, so u is the "fast" unknown and the second one (on v) is the slow one.

Let us recall that we have the following

Theorem 4.7 *For all initial data $(u^0,v^0)\in L^2(0,L)$, System (31) has a unique solution $t\mapsto (u(t),v(t))$ in $C^0([0,+\infty),L^2(0,L))\cap C^\infty((0,+\infty)\times[0,L])$, satisfying $(u,v)(0)=(u^0,v^0)$.*

Proof. It follows the same steps as the proof of Theorem 14, if we look for solutions of the form

$$u(t,x)=\sum_{k=1}^{+\infty}\alpha_k(t)\cos(\frac{k\pi}{L}x) \quad \text{and} \quad v(t,x)=\sum_{k=1}^{+\infty}\beta_k(t)\cos(\frac{k\pi}{L}x).$$

■

The following theorem deals with the asymptotic behavior of the solutions of System (31):

Theorem 4.8 *For all smooth solutions (u,v) of System (31), we define their mean values as the functions M_u and M_v through the formulae*

$$M_u(t)=\frac{1}{L}\int_0^L u(t,x)\mathrm{d}x \quad \textit{and} \quad M_v(t)=\frac{1}{L}\int_0^L v(t,x)\mathrm{d}x.$$

One has the following decay estimates for the functions $u-M_u$ and $v-M_v$:

$$\int_0^L (|u-M_u|^2+N|v-M_v|^2)(t)\mathrm{d}x \leq \Big(\int_0^L (|u-M_u|^2+N|v-M_v|^2)(0)\mathrm{d}x\Big)\mathrm{e}^{-\frac{2\pi^2\nu}{L^2}t}.$$

Moreover, both functions u and v converge in mean square exponentially fast to the constant function $(M_u(0)+NM_v(0))/(N+1)$.

Proof. Let (u,v) be a smooth solution of (31). Consider the functions $\tilde{u}(t,x)=u(t,x)-M_u(t)$ and $\tilde{v}(t,x)=v(t,x)-M_v(t)$. These functions solve the differential system

$$\left\{ \begin{array}{rcl} M_u'(t) &=& -Nc(M_u(t)-M_v(t)) \\ M_v'(t) &=& c(M_u(t)-M_v(t)) \end{array} \right. .$$

From the study of (1), we infer that the mean values functions M_u and M_v both converge when t tends to infinity to $(M_u(0)+NM_v(0))/(N+1)$ with the exponential rate of $-c(N+1)$. Moreover, the functions $\tilde{u}$ and $\tilde{v}$ solve the system

$$\left\{ \begin{array}{rcll} \partial_t \tilde{u} &=& N\nu\Delta\tilde{u} + Nc(\tilde{v}-\tilde{u}) & t>0, x\in(0,L) \\ \partial_t \tilde{v} &=& \nu\Delta\tilde{v} + c(\tilde{u}-\tilde{v}) & t>0, x\in(0,L) \end{array} \right. .$$

Hence we can compute as in the proof of Theorem 4.2:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int_0^L(|\tilde{u}(t)|^2+N|\tilde{v}(t)|^2)\mathrm{d}x &= N\nu\int_0^L\tilde{u}\Delta\tilde{u}+N\nu\int_0^L\tilde{v}\Delta\tilde{v}+Nc\int_0^L(\tilde{u}(\tilde{v}-\tilde{u})+\tilde{v}(\tilde{u}-\tilde{v}))\\ &= -N\nu\int_0^L|\nabla\tilde{u}|^2-N\nu\int_0^L|\nabla\tilde{v}|^2-Nc\int_0^L|\tilde{u}-\tilde{v}|^2\\ &\leq -\frac{2\pi^2\nu}{L^2}\frac{1}{2}\int_0^L(|\tilde{u}(t)|^2+N|\tilde{v}(t)|^2)\mathrm{d}x,\end{aligned}$$

using that $\int_0^L|\tilde{u}-\tilde{v}|^2\geq 0$ and the Poincaré-Wirtinger inequality

$$\int_0^L|\tilde{u}(t)|^2\leq\frac{L^2}{\pi^2}\int_0^L|\nabla\tilde{u}(t)|^2.$$

■

The main change with respect to the non homogeneous Dirichlet case is the fact that we now consider $N-1$ intervals so that the Laplacian matrix that now reads

$$A=\begin{pmatrix}1 & -1 & 0 & \dots & 0\\ -1 & 2 & -1 & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & 0\\ \vdots & \ddots & -1 & 2 & -1\\ 0 & \dots & 0 & -1 & 1\end{pmatrix},$$

and is not invertible. It is common knowledge that 0 is a simple eigenvalue, associated with $(1,\dots,1)^t$. Generically, the above mentioned schemes read : given U^0, V^0, for all $n\geq 0$,

$$\begin{pmatrix}U^n\\ V^n\end{pmatrix}=\mathcal{M}^n\begin{pmatrix}U^0\\ V^0\end{pmatrix} \tag{32}$$

where $\mathcal{M}$ is defined as a product of matrices (21). Ensuring that the sufficient CFL condition is fulfilled for all the matrices composing $\mathcal{M}$ also ensures that they are each non negative, that is, all their coefficients are non negative amd that their L^∞-norm is less than 1, so that $\mathcal{M}$ is non negative and its L^∞-norm is less than 1. Moreover, since the equations are coupled, $\mathcal{M}$ is primitive, that is, there exists an integer $p\geq 1$ such that $\mathcal{M}^p$ is positive. We can therefore apply Perron-Frobenius's theorem that allows us to conclude that 1 is a simple eigenvalue of $\mathcal{M}$ associated with $(1,\dots 1,1,\dots,1)^t$. Indeed, we have the following decomposition :

$$I_{2M}=\mathcal{P}+\mathcal{Q}$$

where $\mathcal{Q}=\frac{1}{M(1+S(\delta t,\delta x))}(1,\dots,1,1,\dots,1)^t(1,\dots 1,S(\delta t,\delta x)(1,\dots,1))$ is the spectral projector associated to 1. The definition of A-order is then easily adapted from Definition 2.4. As for the rate of convergence, the computations that were led in the Dirichlet case can be adapted by restricting to the $\mathcal{M}$-stable supplementary subspace.

At the moment, the numerical simulations, unfortunately, show that for all Schemes from #1 to #4, the A-order is 0 due to the approximations in computing a large power of the matrix $\mathcal{M}^t$ to recover the matrix $\mathcal{Q}$. This shows that, in the Neumann case, the subcycled and not subcycled methods, Lie or Strang, are all ill-convincing.

A FS to SF computations

Let us define the matrix

$$\Pi := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Let A be a 2-by-2 matrix. Then ΠA exchanges the lines of A and $A\Pi$ exchanges the columns. Thus, if $\lambda \in \mathbb{R}$,

$$\Pi M_s(\lambda) \Pi = M_f(\lambda),$$

and, if $\alpha, \beta \in (0,1)$,

$$\Pi G[\alpha, \beta] \Pi = G[\beta, \alpha].$$

Since $\Pi^2 = Id$, it means that $M_s(\lambda)$ and $M_f(\lambda)$ are similar, thus share the same spectrum. In Section 2, we computed the A-orders and rates of convergence of SF (fast, then slow) and FSF (fast, then slow, then fast) type schemes. We show here that the results we obtained can easily be applied to FS and SFS schemes.

Lie-splitting schemes Consider $\lambda_s, \lambda_f \in (0,1)$. According to Lemma 2.1 and Remark 2.2, we define $\alpha(\lambda_s, \lambda_f)$ and $\beta(\lambda_s, \lambda_f)$ as

$$M_s(\lambda_s) M_f(\lambda_f) = G[\alpha(\lambda_s, \lambda_f), \beta(\lambda_s, \lambda_f)].$$

Since

$$M_f(\lambda_f) M_s(\lambda_s) = \Pi M_s(\lambda_f) M_f(\lambda_s) \Pi,$$

we infer that

$$M_f(\lambda_f) M_s(\lambda_s) = \Pi G[\beta(\lambda_f, \lambda_s), \alpha(\lambda_f, \lambda_s)] \Pi$$

Consequently, we can deduce the convergence rate and the A-order of the FS methods at once from the results we obtained for the SF methods.

Strang-splitting methods In the same way, knowing $M_f(\lambda_f) M_s(\lambda_s) M_f(\lambda_f)$, one can deduce the convergence rate and the A-order of $M_f(\lambda_f) M_s(\lambda_s) M_f(\lambda_f)$ by noting that

$$M_s(\lambda_s) M_f(\lambda_f) M_s(\lambda_s) = \Pi M_f(\lambda_s) M_s(\lambda_f) M_f(\lambda_s) \Pi.$$

References

[1] José Antonio Carrillo, Thierry Goudon, and Pauline Lafitte. Simulation of fluid and particles flows: asymptotic preserving schemes for bubbling and flowing regimes. *J. Comput. Phys.*, 227(16):7929–7951, 2008.

[2] José Antonio Carrillo, Thierry Goudon, Pauline Lafitte, and Francesco Vecil. Numerical schemes of diffusion asymptotics and moment closures for kinetic equations. *J. Sci. Comput.*, 36(1):113–149, 2008.

[3] W.J.T. Daniel. A partial velocity approach to subcycling structural dynamics. *Computer Methods in Applied Mechanics and Engineering*, 192:375 – 394, 2003.

[4] Pauline Godillon-Lafitte and Thierry Goudon. A coupled model for radiative transfer: Doppler effects, equilibrium, and nonequilibrium diffusion asymptotics. *Multiscale Model. Simul.*, 4(4):1245–1279 (electronic), 2005.

[5] E. Hairer and G. Wanner. *Solving ordinary differential equations II: Stiff and differential-algebraic problems*, volume 2. Springer, 2004.

[6] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.

[7] S. Jin. Asymptotic preserving (ap) schemes for multiscale kinetic and hyperbolic equations: a review. *Lecture Notes for Summer School onâĂİ Methods and Models of Kinetic TheoryâĂİ(M&MKT), Porto Ercole (Grosseto, Italy)*, 2010.

[8] M. Lemou and L. Mieussens. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM Journal on Scientific Computing*, 31(1):334–368, 2008.

[9] Serge Piperno. Explicit/implicit fluid/structure staggered procedures with a structural predictor and fluid subcycling for 2d inviscid aeroelastic simulations. *International Journal for Numerical Methods in Fluids*, 25(10):1207–1226, 1997.

[10] Roger Temam. Multilevel methods for the simulation of turbulence. A simple model. *J. Comput. Phys.*, 127(2):309–315, 1996.