Stability for implicit-explicit schemes for non-equilibrium kinetic systems in weighted spaces with symmetrization

F. Patricia Medina
Mathematical Sciences Department, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609-2980

Malgorzata Peszynska†
Department of Mathematics, Oregon State University, Corvalis, OR 97331

Abstract

We consider kinetic systems, and prove their stability working in weighted spaces in which the systems are symmetric. We prove stability for various explicit and implicit semi-discrete and fully discrete schemes. The applications include advective and diffusive transport coupled to the accumulation of immobile components governed by non-equilibrium relationships. We also discuss extensions to nonlinear relationships and multiple species.

Keyword: stability for systems, kinetic models, non-equilibrium, adsorption, symmetrization, implicit schemes, explicit schemes

1 Introduction

In this paper we transform and analyze semi-implicit numerical schemes for an evolution system

\[(\phi u)_t + v_t + \nabla \cdot (qu) - \nabla \cdot (\phi d \nabla u) = f, \quad (1a)\]
\[v_t = \alpha(g(u) - v) \quad (1b)\]

which arises in a variety of important applications, e.g., transport in porous media with adsorption. Here \(\alpha > 0\) and \(g(\cdot)\) is monotone, with details below. The positive coefficient \(\phi\) is the porosity.

---

*Research partially supported by NSF DMS-1115827 “Hybrid modeling in porous media”, and NSF DMS-1522734 “Phase transitions in porous media across multiple scales”
†Corresponding author

URL: https://www.wpi.edu/people/faculty/fpmedina (F. Patricia Medina)
http://www.math.oregonstate.edu/ mpesz (Malgorzata Peszynska)
For this system there is no maximum principle, and if \( f = 0 \), there is not even a natural conservation or stability principle in the natural norms of \((u,v)\). Further, the analysis of the simple finite discretization schemes with well known truncation errors, even when \( g \) is linear, has to deal with nonnormality, and is unnecessarily complex, even when \( g(\cdot) \) is linear.

The transformation we propose involves symmetrization, rescaling, and a change of variables. Equivalently, we work in weighted spaces. We exploit the symmetrization to prove strong stability of the problem and of the associated numerical schemes, from which the natural error estimates follow. For fully implicit schemes the framework of m-accretive operators reduces the stability analysis to the verification that the operator is m-accretive. However, for implicit-explicit schemes this is not sufficient, and we draw upon Fourier analysis.

**Overview**  
For the linear case when \( g(u) = cu \), with \( c > 0 \), the abstract form of (1) has the structure of a linear kinetic system

\[
U' + V' + LU = F \quad (2a)
\]
\[
V' + \alpha(V - cU) = 0, \quad (2b)
\]

with the unknowns \( U, V : (0,T) \to H \times H \), where \( H \) is an appropriate Hilbert space to be defined, and the source term \( F : [0,\infty) \to H \) is given. The linear transport operator \( L \) is defined in the sequel, and we will require for \( L \) to be m-accretive to get strong stability.

Our main technical objective is to study the stability of (2) and of one–step implicit and implicit-explicit discrete schemes for (2)

\[
\frac{U^n - U^{n-1}}{\tau} + \frac{V^n - V^{n-1}}{\tau} + LU^n* = F^n \quad (3a)
\]
\[
\frac{V^n - V^{n-1}}{\tau} + \alpha(V^n - cU^n) = 0. \quad (3b)
\]

which is solved at every time step \( n = 1, 2, \ldots \) for the approximations \( U^n, V^n \) to \( u(\cdot, t_n), v(\cdot, t_n) \). Here \( F^n \approx F(t_n) \). This one-step scheme is fully implicit if \( n^* = n \). Other schemes arise for \( n^* \neq n \). The analysis of (3) involves consideration of spatial discretization as well as of time discretization. Our technique of symmetrization allows to demonstrate strong stability of the schemes in a weighted space, even though the original system (2) has nonnormal operators.

Extensions of (2) to nonlinear systems and to systems with multiple components will be also discussed.

**Motivation and context**  
The problem (1) comes from applications in subsurface modeling such as the transport of contaminant undergoing adsorption, or coalbed methane reservoir simulation, but cover also a variety of other applications. In those problems (1) represents the conservation of mass of some
chemical component, with \( u \) denoting the mobile concentration, and \( v \) representing the immobile component, while \( g(\cdot) \) a general monotone (increasing) function. We provide details on the applications in Section 2.

Numerical analysis of (1) with non-equilibrium kinetics was given in [1] for diffusion only, with focus on non-Lipschitz \( g(\cdot) \) important for liquid adsorption. In [2] Lagrangian techniques for advection with non-equilibrium adsorption and in [3] the Lagrangian transport combined with Galerkin approximation to diffusion were analyzed. In addition, in a sequence of papers devoted to the scalar conservation laws with relaxation terms [4] a problem similar to (4a) but without diffusion is studied, and convergence order of \( O(\sqrt{h}) \) is established. In turn, in [5] we studied the stability of schemes for a single equation analogue of (4a) without diffusion and where \( v \) was eliminated, and in [6] we extended the analysis to cover the linear case with diffusion. Furthermore, previous results on stability of schemes of (2) for the case of initial equilibrium were shown in [7, 5, 6].

Our approach in this paper provides a unified framework for the analysis of a variety of explicit and implicit finite difference schemes for the non-equilibrium advection-diffusion problems. In particular, it establishes strong stability as well as optimal error estimates of order \( O(h) \) or \( O(h^2) \).

Outline In Section 2, we motivate the study of (2), provide examples of \( L \), and provide literature review. In Section 3, we describe the main idea of symmetrization in the abstract setting leading to the stability of the numerical schemes. In Section 4, we provide concrete examples of fully discrete schemes for (2) and evaluate their stability, and in Section 5, we illustrate the theory with numerical examples, and convergence studies. We close in Section 6, where we outline extensions to the nonlinear and multi-species case, and discuss future work.

Notation and assumptions Throughout the paper we assume that \( c > 0, \alpha > 0 \); otherwise, the system is decoupled and trivial. \( I \) always denotes the identity operator or matrix, as is clear from the content.

With the original variables in (1) denoted by \( u(x, t) \) and \( v(x, t) \), we consider the vectors \( U(t) = u(\cdot, t) = (u(x, t))_x \in H \). Each \( U(t), V(t) \) lives in a Hilbert space \( H \), with the inner product denoted by \( \langle \cdot, \cdot \rangle_H \); we drop the subscript \( H \) when it does not lead to a confusion. The domain of an operator \( L \) is denoted by \( D(L) \), and the time derivative \( U'(t) \) or \( \frac{dU}{dt}(t) \) generalizes the partial derivative \( \frac{\partial}{\partial t} \), and is defined in an appropriate abstract setting, such as that developed in [8].

The vector \( W = W(t) = [U(t), V(t)]^T \) lives in \( H \times H \), which is endowed either with the natural or weighted inner product, with details below. We also consider new variables \( \tilde{W} \) in appropriate spaces. The (matrices of) operators on \( W \) or \( \tilde{W} \) in the product space \( H \times H \) are denoted similarly to those on \( H \). In particular, for \( L : D(L) \subset H \to H \) we define \( L = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} \) on \( D(L) = D(L) \times H \subset H \times H \). We define \( A, D \) analogously.
In discrete schemes, we consider uniform time stepping \( t_n = n\tau \) with \( n = 0, 1, \ldots \), and time step \( \tau \). For fully discrete schemes, we denote the spatial grid parameter by \( h \) and consider finite dimensional analogues of the operators such as \( L_h \) for \( L \). For the unknown \( u(x, t) \) we denote by \( u^n(x) \approx u(x, t_n) \) its semi-discrete in time approximation, by \( U^n \in H \) the collection \( (u^n(x))_x \in H \). In turn, \( U_h(t) = (u_j(t))_j \in H_h \) where \( u_j(t) \approx u(x_j, t) \) is the semi-discrete in space approximation, in a discrete (usually finite dimensional) subspace \( H_h \subseteq H \) of dimension dependent on \( h \). Finally, fully discrete approximations \( U^n_h \approx U_h(t_n) \).

Most of our results are formulated on \( H \), and are shown to apply on \( H_h \).

2 Motivation and literature review

In this section we develop the applications which motivate \( 2 \) and provide details on its abstract setup, in particular on the properties of \( L \) as they follow for the special cases of \( 1 \) under assumed boundary conditions. Our presentation of the model follows the literature on coalbed methane adsorption \([9, 10, 11]\) where \( 1 \) arises directly, and finite volumes are used; see also our expository work in \([12]\). In Section 2.3 we discuss a particular direction in which \( 2 \) is reduced to a single equation; this is discussed in \([5]\) under initial equilibrium assumption.

2.1 Applications

In porous media, the model of transport with adsorption \( 1 \) describes the evolution of the concentration of a chemical. Consider an open bounded region of flow \( \Omega \subseteq \mathbb{R}^k, k = 1, 2, 3 \), in which the volumetric flux \( q \), with \( \nabla \cdot q = 0 \), is given; assume also the porosity \( \phi(x) > 0 \) and the (uniformly positive definite) diffusion coefficient \( d \) are known. If the chemical is adsorbed in the porous medium, the mass conservation must include also the rate of change of the adsorbed immobile amount denoted by \( v \). The mass conservation of the chemical being transported by advection \( Au = \nabla \cdot (qu) \) and diffusion–dispersoon \( Du = -\nabla \cdot (d\nabla u) \), with adsorption term, is

\[
(\phi u)_t + v + Au + Du = f, \quad x \in \Omega, \quad t > 0 \tag{4a}
\]

and it remains to specify the relation of \( v \) to \( u \).

The equilibrium relationship \( v = g(u) \) which can be used to complete \( 4a \) assumes that the time scale of transport is much slower than that of the adsorption. In turn, the non-equilibrium or kinetic model

\[
v_t + \alpha(v - g(u)) = 0 \quad \tag{4b}
\]

allows to treat the time scales of adsorption and of transport on par with each other, with \( \alpha > 0 \) denoting the rate of the process. As \( \alpha \to \infty \), it is expected that \( 4b \) has solutions close to the equilibrium. The linear relationship \( g(u) = cu \), with \( c > 0 \) is what is assumed throughout most of this paper.

For well-posedness, we require appropriate boundary conditions on \( u \) as well as initial conditions for both \( u \) and \( v \).
2.1.1 Abstract setting

In the abstract form, the model (4a) and (4b), upon absorbing nonessential constants in the definitions of $u$, are written as (2), in which $L = \mathcal{D} + A$ is the abstract diffusion-advection transport operator, and where (2) is posed as a Cauchy problem in an appropriate function space.

Consider $L : D(L) \to H$ in a Hilbert space $H$, with domain $D(L)$. We recall that $L$ is accretive if $\langle LU, U \rangle \geq 0$ for any $U \in D(L)$. Additionally, $L$ is m-accretive if $I + L$ is onto $H$, that is, for any $F \in H$ the problem $U + LU = F$ is solvable (from accretiveness there follows the uniqueness of the solution).

For an m-accretive $L$, the following results are well known; see, e.g., [8], (Sec.I.4). The dynamics of $U'(t) + LU = 0, U(0) \in H$ is governed by a linear contraction semigroup, so that, in particular, $U(t) \in D(L)$. If $L$ is self-adjoint, additional regularity and convergence properties follow. The nonhomogeneous case of $U'(t) + LU = F$ requires that $F \in (C^1[0, \infty), H)$. See, e.g. [8], (Prop.4.1) and [13], (Cor. 3.B).

For the applications of (1) described in Sec. 2.1 we consider $H = L^2(\Omega)$ with the inner product $\langle \psi, \xi \rangle = \int_\Omega \psi\xi$. In the case of periodic boundary conditions, without loss of generality, one can use $\Omega = (0, 1)^k$, but we only analyze $k = 1$ case. We will recall the standard abstract results for $Au = -d\nabla^2 u + \nabla \cdot (qu)$, with weak rather than the classical (partial) derivatives. The definition of $L$ and $D(L)$ accounts for the boundary conditions. For details on this abstract setup see [13], (Ex IV.2., p108) and [8], (Prop. I. 4.2, p21). For periodic case, see [13], (Ex IV.1, p107)), and for advection see [13], (Example IV.1).

Remark 1 (i) Let $L = D$, with $d > 0$, and with homogeneous Dirichlet boundary conditions imposed. We have $D(L) = H^1_0(\Omega) \cap H^2(\Omega)$, and $L$ is m-accretive self-adjoint. (ii) As in (i), but with homogeneous Neumann conditions, $D(L) = H^2(\Omega)$, $L$ is m-accretive, and self-adjoint. (iii) As in (i), with periodic boundary conditions, e.g., when $\Omega = (0, 1)$ we have $D(L) = \{\psi \in H^2(\Omega), \psi(0) = \psi(1), \psi'(0) = \psi'(1)\}$. The operator $L$ is m-accretive and self-adjoint. (iv) Case $L = A$, with properly posed conditions on the inflow boundary, or with periodic boundary conditions, e.g., for $\Omega = (0, 1)$, $D(L) = \{\psi \in H^1(\Omega) : \psi(0) = \psi(1)\}$. The operator $D$ is m-accretive but not selfadjoint. (v) Case $L = A + D$, and $d > 0, q \neq 0$. With periodic b.c., $D(L)$ is as in (iii), and $L$ is m-accretive but not selfadjoint.

2.2 Related models and previous work

In models of transport with gas adsorption, (4b) allows to account for sub-scale diffusion accompanying the overall transport; see [14, 15]. More general models in which α is a monotone operator can be used, e.g., to model hysteresis in adsorption [16] or non-equilibrium phase transitions [17]. Further, non-equilibrium relation (4b) is used to model transport in media with multiscale character, such as in the classical Warren-Root and Barenblatt models of double porosity [18, 19]; see also modeling and analysis in [20, 21, 22], and numerical analysis in [7, 23].
In previous work for nonlinear $g(\cdot)$, proved a-priori error estimates for Lagrangian-Galerkin methods for (1), and in the analysis is for diffusion only. Our paper handles the advection and diffusion problem together for linear $g(\cdot)$, and handles the analysis as well as implicit and explicit numerical schemes in the same framework.

2.3 Nonlocal formulation in $u$ under the initial equilibrium assumption

One can reformulate the coupled kinetic system (2a) as a single equation with nonlocal in time terms. Recall Volterra convolution integral term defined by $\int_0^t \beta(t-s)U(s)ds$. We solve (2b) for $V(t)$ in terms of $U(t)$ and substitute to (2a) to give the following Volterra integro-differential equation

$$U' + \alpha U' * \beta + AU = F + \beta(t)(V(0) - cU(0)),$$  
(5a)

solved for $U$, where $\beta(t) = \alpha e^{-\alpha t}$. The variable $V$ can be recovered from $U$ by

$$V(t) = e^{-\alpha t}V(0) + \int_0^t \alpha cU(s)e^{-\alpha(t-s)}ds.$$  
(5b)

Now we see that the second term on the right hand side of (5a) acts like a source/sink term decreasing with $t$, and it vanishes under the assumption of initial equilibrium

$$V(0) - cU(0) = 0.$$  
(6)

The one-way coupling in (5) focuses the attention on $U$ while keeping track of the memory effects expressed by $U' * \beta$.

The effect of memory terms isolated from the source term can be studied if (6) is assumed. This approach was followed in [7, 5, 6]. Strong stability of $u$ for the numerical schemes was proven for $L = D$ in [7], and for linear or nonlinear advection operator $L = A$ in [5], where we exploited the positivity of the kernel $\beta$. Even though we did not prove it, the numerical results suggested that a maximum principle holds for $U$.

If (6) cannot be assumed, the positive source term in (5a) can be expected to disturb the maximum principle and/or stability. Indeed, a simple example in Sec. 5.1 readily demonstrates it.

As concerns numerical schemes, in [7] the kernel was also allowed to be weakly singular, e.g., $\beta = O(t^{-1/2})$, which corresponds to subscale diffusion, i.e., the case when there is diffusion in (2b). A more general case of nonlocal terms and of operators $L$ was studied in [5] in which the multiscale model derived in [24], but initial data was assumed in equilibrium. In [5] we conjectured experimentally that the presence of the memory term $u' * \beta$ would lead to an increased regularity of the solution $u$ when $\beta$ was weakly singular, but this effect appears weaker for the bounded $\beta$.

Beyond porous media, the effect of non-equilibrium (relaxation) such as in (4b), was studied, e.g., in [25], and is an important component of pseudo-parabolic models [26].
3 Stability for the abstract symmetrized evolution system

We start by motivating the symmetrization and discussing the properties and well-posedness of the symmetrized system in the abstract form on some general Hilbert spaces $H$.

First we re-arrange (2) in an equivalent form

\[
\begin{align*}
U' - \alpha(V - cU) + LU &= F, \quad (7a) \\
V' + \alpha(V - cU) &= 0. \quad (7b)
\end{align*}
\]

This arrangement is similar to those used in multiscale models such as the Warren-Root or Barenblatt models [18, 19]. For these models however $c = 1$; their analysis and numerics can be found, e.g., in [23]. When $c \neq 1$, the analysis requires additional work.

In a vector-matrix form with $w = [u, v]^T$ we write (7)

\[
\begin{align*}
W' + BW &= W' + CW + LW = [F, 0]^T, \quad (8)
\end{align*}
\]

with

\[
\mathcal{C} = \alpha \left[ \begin{array}{cc} cI & -I \\ -cI & I \end{array} \right], \quad L = \left[ \begin{array}{cc} L & 0 \\ 0 & 0 \end{array} \right].
\]

This system of evolution equations is solved for $W(t) = [U(t), V(t)]^T \in H \times H$. The space $H \times H$ is endowed with the natural inner product $\langle \cdot, \cdot \rangle_{H \times H}$ on the product space, where $\langle [U, V]_T, [\phi, \psi]_T \rangle_W = \langle U, \phi \rangle_H + \langle V, \psi \rangle_H$.

**Challenge** The system (8) is linear, and thus is trivially well-posed, e.g., if $H = \mathbb{R}^P$, $P \in \mathbb{N}$. However, in Section 5.1 we show with a simple example on $H = \mathbb{R}$, that the system (8) is not stable in $\| \cdot \|_{H \times H}$, even though the solutions to the homogeneous problem eventually decay to 0.

Unless $c = 1$, the operator $\mathcal{C}$ and $B$ are not self-adjoint and nonnormal with respect to $\langle \cdot, \cdot \rangle_{H \times H}$. In consequence, the analysis of the numerical schemes for (8) is quite complicated. Therefore, we consider a weighted inner product on $H \times H$ or, equivalently, a change of variables. This idea which we explain in Sections 3.1 and 3.3 makes the subsequent analysis of numerical schemes fairly straightforward.

3.1 Symmetrization and rescaling

We propose to consider a weighted (scaled) inner product $\langle \cdot, \cdot \rangle_H$ on $H \times H$, and an associated norm $\| \cdot \|_c$

\[
\langle [U, V]_T, [\phi, \psi]_T \rangle_c = c(U, \phi)_H + (V, \psi)_H, \quad \| [U, V]_T \|_c^2 = c \| U \|_H^2 + \| V \|_H^2. \quad (10)
\]

In other words, instead of $H \times H$ we consider the new (Hilbert) space $W_c$ which is $H \times H$ endowed with $\langle \cdot, \cdot \rangle_c$. In the new space $W_c$ we are able to prove the
stability of the evolution system, and of the appropriate numerical schemes for
the diffusion-advection examples.

The use of weighted inner product space can be interpreted as changing
variables from $U$ to $\tilde{U} = \sqrt{c}U$, since $\| [U, V]^T \|_c^2 = \| [\sqrt{c}U, V]^T \|_{H \times H}^2$. We also
denote the change of variables from $w$ to $\tilde{w}$ with

$$\tilde{W} = [\tilde{U}, V]^T = [\sqrt{c}U, V]^T. \quad (11)$$

To show how we exploit the space $W_c$, we rewrite (7) by scaling the first
component equation of (7) by $\sqrt{c}$, and re-distributing the appropriate constants,
by linearity of $L$. We see that (7) is equivalent to

$$\sqrt{c}U' - \alpha(\sqrt{c}V - c \sqrt{c}U) + L \sqrt{c}U = \sqrt{c}F \quad \text{(12a)}$$

$$V' + \alpha(\sqrt{c}V - \sqrt{c} \sqrt{c}U) = 0. \quad \text{(12b)}$$

where we have also used $cU = \sqrt{c} \sqrt{c}U$ in (12b). Rewriting in the new variables

$$\tilde{W}' + \tilde{B} \tilde{W} = \tilde{W}' + \tilde{C} \tilde{W} + \tilde{L} \tilde{W} = [\tilde{F}, 0]^T \quad (13)$$

with $\tilde{F} = \sqrt{c}F$ and the operators defined as

$$\tilde{B} = \tilde{C} + \tilde{L}, \quad \tilde{C} = \alpha \begin{bmatrix} cI & -\sqrt{c}I \\ -\sqrt{c}I & I \end{bmatrix}. \quad (14)$$

3.2 Well-posedness

Now we complete the formal discussion of the well-posedness of (8). We see
that $B : D(B) \to H \times H$ with $D(B) = D(L) \times H$ is dense in $H \times H$. Similarly,
$\tilde{B} : D(\tilde{B}) \to W_c$, and simply $D(\tilde{B}) = D(B) = D(L) \times H$.

**Proposition 1** Let $L$ be $m$-accretive on $H$. Then the operator $B$ is $m$-accretive
on $W_c$. Equivalently, $\tilde{B}$ is $m$-accretive on $H \times H$.

**Proof:** (i) The proof follows from (14) and (11) by the calculation

$$\langle \tilde{B} \tilde{W}, \tilde{W} \rangle_{H \times H}$$

$$= \langle ([L + c\alpha I] \tilde{U} - \alpha \sqrt{c} V - \alpha \sqrt{c} \tilde{U} + \alpha V]_t, [\tilde{U}, V]^T \rangle_{H \times H}$$

$$= \langle L \tilde{U} + c\alpha \tilde{U} - \alpha \sqrt{c} V, \tilde{U} \rangle_H + \langle -\alpha \sqrt{c} \tilde{U} + \alpha V, V \rangle_H$$

$$= \langle L \tilde{U}, \tilde{U} \rangle_H + c\alpha \langle \tilde{U}, \tilde{U} \rangle_H - 2\alpha \sqrt{c} \langle V, \tilde{U} \rangle_H + \alpha \langle V, V \rangle_H$$

$$= \langle L \tilde{U}, \tilde{U} \rangle_H + \alpha \| \sqrt{c} \tilde{U} - V \|_{H \times H}^2 \geq \langle L \tilde{U}, \tilde{U} \rangle_H \geq 0 \quad (15)$$

where we have exploited the symmetry $\langle V, \tilde{U} \rangle_H = \langle \tilde{U}, V \rangle_H$ and completed the
square. The last step followed since $L$ is accretive on $H$. 8
Similarly, we have that
\[
\langle BW, W \rangle = \langle B[U, v]^T, [U, V]^T \rangle_c
\]
\[
= \langle [(L + c\alpha I)U - \alpha V, -\alpha U + \alpha V]^T, [U, V]^T \rangle_c
\]
\[
= c\langle LU, U \rangle_H + c^2\alpha\langle U, U \rangle_H - 2\alpha\langle U, V \rangle_H + \alpha\langle V, V \rangle_H
\]
\[
= c\langle LU, U \rangle_H + \alpha\langle cU - V, cU - V \rangle_H \geq c\langle LU, U \rangle_H \geq 0. \tag{16}
\]

(ii) To show that \( B \) is \( m \)-accretive, i.e., that \( I + B \) is onto \( W_c \), we show how to solve the system \((I + B)W = F\) for any \( F = [F, G] \in H \times H \). To this end, we consider the solution of the stationary counterpart of (7)
\[
U - \alpha(V - cU) + LU = F,
\]
\[
V + \alpha(V - cU) = G.
\]
(In our problem (2) we have \( G = 0 \) but it is easy to consider the general case.) Solving the second equation for \( V \) in terms of \( U \), back-substituting to the first equation, and \( \alpha - \frac{\alpha^2}{1 + \alpha} = \frac{\alpha}{1 + \alpha} \), we see that \( U \) satisfies
\[
\left( (1 + \frac{c\alpha}{1 + \alpha})I + L \right) U = F + \frac{\alpha}{1 + \alpha}G
\]
which can be solved for any \( F \in H \), since \( L \) is \( m \)-accretive.

**Corollary 1** Assume \( \alpha, c > 0 \) and \( L : D(L) \to H \) is \( m \)-accretive, and \( \tilde{W}_{\text{init}} \in W_c \). By Hille-Yosida Theorem as quoted in [8] Prop. 4.2, p21 and [8] Thm I.5.1, p25, we conclude that there exists a unique solution to the Cauchy problem, with \( \tilde{w}(t) \in D(\tilde{B}) \)
\[
\tilde{W}' + \tilde{B}\tilde{W} = [\tilde{F}, 0]^T, \quad \tilde{W}(0) = \tilde{W}_{\text{init}} \in H.
\]
The evolution of \( \tilde{W} \) is governed by the linear contraction semigroup. When \( \tilde{F} = 0 \), we have the stability
\[
\frac{d}{dt} \| \tilde{W} \|^2 \leq 0. \tag{18}
\]

### 3.3 Alternative motivation for symmetrization

We provide here another way to motivate the symmetrization and rescaling proposed in Section 3.1. We consider the homogeneous case of (7), and take the inner product of each component equation with \( u \) and \( v \), respectively. We obtain
\[
\langle U', U \rangle - \alpha\langle V, U \rangle + \alpha\langle cU, U \rangle + \langle LU, U \rangle = 0
\]
\[
\langle V', V \rangle + \alpha\langle V, V \rangle - \alpha\langle cU, V \rangle = 0.
\]
Adding these identities directly does not produce useful results for stability in \( \| (U, V) \| \), because the cross-terms do not cancel. However, up to the scaling, the
second term in the first identity is similar to the third one in the second identity. Multiplying the first equation with $c$, and adding the resulting equations, we obtain
\[ c\langle U', U \rangle - \alpha \langle V, cU \rangle + c(LU, U) + \langle V', V \rangle + \alpha \langle V, V \rangle - \alpha \langle cU, V \rangle = 0 \]
Rearranging the terms, by symmetry of the inner product, we get
\[ c\langle U', U \rangle + \langle V', V \rangle + \alpha \langle V, V \rangle - 2\alpha \langle V, cU \rangle + \alpha \langle cU, cU \rangle + c\langle LU, U \rangle + \langle V, V \rangle - 2\alpha \langle V, cU \rangle = 0 \quad (19) \]
Next, for the first two terms in (19) we write
\[ c\langle U', U \rangle + \langle V', V \rangle = \frac{1}{2} \frac{d}{dt} \| U \|^2 + \frac{1}{2} \frac{d}{dt} \| V \|^2 = \frac{1}{2} \frac{d}{dt} \| [\sqrt{c}U, V]^T \|^2. \]
The next three terms in (19) are easily combined to give $\alpha \langle V - cU, V - cU \rangle = \alpha \| V - cU \|^2 \geq 0$. Since $L$ is accretive, upon $c(LU, U) \geq 0$ we obtain from (19)
\[ \frac{d}{dt} \| [\sqrt{c}U, V]^T \|^2 \leq 0. \quad (20) \]
In other words, we see that the system (7) is stable in the quantity of interest $\| [\sqrt{c}U, V]^T \|$, or in $[U, V]_{W_c}$.

4 Stability of numerical schemes

In this section we discuss numerical schemes for (8) and their stability and convergence properties. We focus on one-step time-discrete schemes for (8) solved for $W^n \approx w(\cdot, t_n) \in H$
\[ \frac{1}{\tau}(W^n - W^{n-1}) + CW^n + LW^n* = F^n, \quad n \geq 1. \quad (21) \]
If $n* = n$, the scheme is fully implicit, and if $n* = n-1$, we have implicit-explicit schemes. Note that our treatment of the (stiff) coupling term CW is always implicit. Here $F^n$ is some appropriately defined time-discrete approximation to $F(\cdot, t_n)$, and $W^0$ is known from the initial conditions.
We also note, upon (12), (14), that (21) is equivalent to
\[ \frac{1}{\tau}(\tilde{W}^n - \tilde{W}^{n-1}) + \tilde{C}\tilde{W}^n + \tilde{L}\tilde{W}^{n*} = \tilde{F}^n, \quad n \geq 1, \quad (22) \]
where $\tilde{C}$ is symmetric.

In fully discrete schemes, the abstract operators $L, C$ are replaced by their finite dimensional analogues $L_h, C_h$ depending on the spatial discretization parameter $h$, and they are solved for the vectors of spatial unknowns $W_h^n = (w^n_j)_j$ where $w^n_j \approx w(x_j, t_n)$ as in
\[ \frac{1}{\tau}(W_h^n - W_h^{n-1}) + C_h W_h^n + L_h W_h^{n*} = F_h^n, \quad (23) \]
with an analogous version for (22), which we skip. Here \( F^n_h \) is an appropriate discretization of \( F \).

In addition, we recall that finite element formulations lead, instead of (23), to

\[
\frac{1}{\tau} M_h (W^n_h - W^{n-1}_h) + C M_h W^n_h + L_h W^{n*}_h = F^n_h, \quad n \geq 1.
\]  

(24)

where \( M_h = \begin{bmatrix} M_h & 0 \\ 0 & M_h \end{bmatrix} \), and \( M_h \) is the symmetric positive definite mass (Gram) matrix. For generality, we adopt (24) as the general fully discrete formulation, since (23) is its special case upon setting \( M_h = I \).

4.1 Fully implicit schemes for m-accretive \( L \) and \( L_h \)

We consider here \( n = n^* \) in (21) or (24).

First observation is somewhat surprising. One might expect when \( F = 0 \), that \( \| W^n \|_{H \times H} \leq \| W^{n-1} \|_{H \times H} \), but this does not hold, e.g., if \( \bar{B} \) is nonnormal. (See example in Sec. 5.1).

However, based on the discussion in Sec. 3, stability can be shown easily in weighted spaces.

**Lemma 1** Let \( L \) be m-accretive. Then the fully implicit scheme (22) is strongly stable in the weighted spaces. We have, when \( F = 0 \), that

\[
\| W^n \|_c \leq \| W^{n-1} \|_c \iff \| \tilde{W}^n \| \leq \| \tilde{W}^{n-1} \|,
\]  

(25)
i.e., the operator \( (I + \tau \bar{B})^{-1} \) is a contraction. For \( F \neq 0 \), we have

\[
\| \tilde{W}^n \| \leq \| \tilde{W}^{n-1} \| + \| \tilde{F}^n \|.
\]  

(26)

For (24) we have

\[
\| M^{1/2}_h W^n \|_c \leq \| M^{1/2}_h W^{n-1} \|_c \iff \| M^{1/2}_h \tilde{W}^n \| \leq \| M^{1/2}_h \tilde{W}^{n-1} \|
\]  

(27)

**Proof:** The proof is immediate when we rewrite (22) as

\[
\frac{1}{\tau} (\tilde{W}^n - \tilde{W}^{n-1}) + \bar{B} W^n = \tilde{F}^n, \quad n \geq 1.
\]  

(28)

Rearranging, and taking the inner product with \( \tilde{W}^n \) we obtain

\[
\langle (I + \tau \bar{B}) \tilde{W}^n, \tilde{W}^n \rangle = \langle \tilde{W}^{n-1}, \tilde{W}^n \rangle + \tau \langle \tilde{F}^n, \tilde{W}^n \rangle.
\]

Since \( L \) is accretive, by Proposition 1 so is \( \bar{B} \) on \( H \times H \). Applying this property and Cauchy-Schwartz inequality we get

\[
\| \tilde{W}^n \|^2 = \langle \tilde{W}^n, \tilde{W}^n \rangle \leq \langle \tilde{W}^n, \tilde{W}^n \rangle + \tau \langle \bar{B} \tilde{W}^n, \tilde{W}^n \rangle
\]

\[
= \langle (I + \tau \bar{B}) \tilde{W}^n, \tilde{W}^n \rangle = \langle \tilde{W}^{n-1} + \tau \tilde{F}^n, \tilde{W}^n \rangle \leq (\| \tilde{W}^{n-1} \| + \tau \| \tilde{F}^n \|) \| \tilde{W}^n \|.
\]
where we have also used (28). Upon dividing by \( \| \hat{W}^{n-1} \| \) we get (25). Alternatively, we start from (21) and work in weighted spaces in which \( B \) is accretive.

To prove (27), we proceed analogously using the properties of \( \hat{B}_h \) on \( H_h \) for (24). Additionally we carry out an easy calculation similar to (15) involving \( M_h \), which takes advantage of positive definiteness and symmetry of \( M_h \).

**Remark 2** Lemma 1 reduces the stability analysis of implicit schemes to the verification whether \( L \) (or \( L_h \)) is accretive. In addition, for finite difference schemes the result (26) applies directly to the error analysis, since (21) can be interpreted as the error equation, in which the right-hand-side represent the truncation error. We see, in particular, that the error accumulates linearly.

We collect several results for \( L = D \) in Sec. 4.2; these are in the framework of the method of lines (MOL). However, for some \( L \) even those covered in Rem. 1, and some schemes, \( L_h \) is non-symmetric, and it is hard to verify if it is m-accretive even if \( L \) is. In particular, for \( L = A \) or \( L = D + A \), and non-implicit schemes, we proceed by von-Neumann analysis.

### 4.2 FD discretization for diffusion with Dirichlet boundary conditions

First we consider FD discretization. For the sake of exposition, we provide details for \( k = 1 \) and \( \Omega = (0, 1) \). We seek the interior values \( u^n_j, j = 1, \ldots, N_h \), where \( h = \frac{1}{N_h+1} \) is the spatial grid parameter. We also seek \( v^n_j \) on the same grid of interior points.

After symmetrization and rescaling, at every time step, one solves the problem (23), rewritten as

\[
\frac{\hat{U}^n_h - \hat{U}^{n-1}_h}{\tau} - \alpha (\sqrt{cV^n_h} - c\hat{U}^n_h) + L_h\hat{U}^{n*}_h = \hat{F}^n_h \quad (29a)
\]

\[
\frac{V^n_h - V^{n-1}_h}{\tau} + \alpha (V^n_h - \sqrt{c\hat{U}^n_h}) = 0. \quad (29b)
\]

where the well known Dirichlet matrix is

\[
L_h := \frac{d}{h^2} \begin{bmatrix}
2 & -1 & 0 & \ldots & 0 & 0 \\
-1 & 2 & -1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & -1 & 2 & -1 \\
0 & 0 & \ldots & 0 & -1 & 2 \\
\end{bmatrix} \quad (29c)
\]

In addition, we note that \( M_h \) in (23) is the identity matrix. Also, we recall that \( L_h \) is symmetric positive definite on \( \mathbb{R}^{N_h} \), thus m-accretive. In fact, for this chosen domain \( \Omega \), the eigenvalues of \( L_h \) are \( \frac{d}{h^2} 2(1 - \cos(p\pi h)) > 0, p = 1, \ldots, N_h \).

When \( n^* = n \), (29) has the form of (24), with \( M_h = 1 \). We can thus apply Lemma 1. We obtain the following result which holds for other domains \( \Omega \), and Dirichlet boundary conditions, as long as \( L_h \) is symmetric and positive definite.
Corollary 2  The implicit in time scheme (29) for \( n^* = n \) is strongly stable in the sense of (27).

When \( n^* = n - 1 \), we consider the homogeneous version of (29) in the form

\[
H_1 \tilde{W}^n = H_0 \tilde{W}^{n-1},
\]

where \( H_1 \) and \( H_0 \) are the block matrices, with \( b = \alpha \tau \),

\[
H_1 = \begin{bmatrix}
(1 + bc)I & -b\sqrt{c}I \\
-b\sqrt{c}I & (1 + b)I
\end{bmatrix}, \quad H_0 = \begin{bmatrix}
I - L_h & O \\
O & I
\end{bmatrix}.
\]

Observe that the matrix \( I - L_h \) is symmetric, thus so is \( H_0 \). If \( X_p, p = 1, \ldots, P \) are the eigenvectors for \( L_h \), it is easy to show that each \( [X_p, 0]^T \) is an eigenvector for \( H_0 \), corresponding to the eigenvalues \( \lambda_p, p = 1, \ldots, P \) of \( L_h \). In turn, the remaining eigenvectors of \( H_0 \) are in the form of \( [0, Y]^T \) where \( Y \in \mathbb{R}^P \) is arbitrary, with eigenvalue \( \lambda = 1 \) of multiplicity \( P \). The set of eigenvalues for \( H_0 \) is \( \{1 - \lambda_p\} \). Since \( H_1 \) is self-adjoint, conditional stability of (30) follows by checking that the eigenvalues of \( H_0 \) are not exceeding 1.

Corollary 3  The scheme (29) for \( n^* = n - 1 \) is conditionally strongly stable in the sense of (27) if \( \frac{2d\tau}{h^2} \leq 1 \).

Next, we briefly mention that handling periodic and Neumann problems with FD is done differently than by constructing the simple analogue of (29). Also, \( L_h \) is typically not symmetric. We do not discuss these cases here.

4.3 FE discretization for \( L = D \) and general boundary conditions

Next we consider \( L = D \) with Dirichlet, or Neumann, or periodic boundary conditions covered by Rem. 1 so that \( L \) is m-accretive on \( D(L) \subset L^2(\Omega) \). Note that \( L \) can have variable coefficients and possibly correspond to some other boundary conditions, as long as \( L \) is m-accretive.

Next considered piecewise linear finite elements forming the approximating subspace \( V_h \subset H^1(\Omega) \), where \( V_h \) accounts properly for the essential boundary conditions. The nodal degrees of freedom \( u^n_j, v^n_j, j = 1, \ldots, N_h \) are in the space \( H_h = \mathbb{R}^{N_h} \). It is well known [27, 28] that the matrix \( L_h \) inherits the properties of the operator \( L \) and, in particular, it is symmetric and nonnegative definite, thus m-accretive. For this setting we have the result as follows.

Corollary 4  Finite element discretization (24) for \( n^* = n \) is strongly stable in the sense of (27).

4.4 Stability for advection and advection–diffusion via extended von-Neumann analysis for systems

The von-Neumann framework for stability analysis of finite difference schemes on uniform spatial grids for scalar linear equations with constant coefficients
Table 1: Amplification factors and stability conditions for scalar diffusion and upwind advection schemes. Here $s_D(\xi) = 2D_D^2(1 - \cos(\xi h))$, $\lambda = q\frac{\tau}{h}$, and $s_\lambda = \lambda(1 - e^{-i\xi h})$.

is well known and is covered in various textbooks see, e.g., [29], (Chapters 9 and 10). The classical monograph [30] deals also with nonlinearity, non-constant coefficients and coupled systems with non-normal amplification matrix; we adopt their notation.

**Notation** First we establish the notation and recall the usual steps. We consider the true solution $s(x, t) \in \mathbb{R}$, $x \in \mathbb{R}$, $t > 0$, to a scalar differential equation, which is approximated by a finite difference equation with uniform spatial and temporal grid parameters $h$ and $\tau$ defining $s_j^n \approx s(x_j, t_n)$, with $x_j = jh$, $j = 0, \pm 1, \pm 2, \ldots$ and $t_n = n\tau, n = 0, 1, \ldots$. The vector $S_h^n = (s_j^n)_{j=-\infty}^{\infty}$, and its grid 2-norm $\|S_h^n\| = \sqrt{h \sum_j (s_j^n)^2}$ is equivalent to and indistinguished from the norm in $L^2(\mathbb{R})$. The discrete Fourier transform applied to $S_h^n = (s_j^n)_j$ gives $\hat{S}^n = (\hat{s}^n(\xi))_{\xi}$, with $-\pi/h \leq \xi \leq \pi/h$. By Parseval’s relation, the study of the evolution of $\|S_h^n\|$ is equivalent to the study of $\|\hat{S}^n\|$ defined through $L^2(-\frac{\pi}{h}, \frac{\pi}{h})$. For one step scheme from $t_{n-1} \rightarrow t_n$ we derive a formula for

$$\hat{S}^n(\xi) = g(\xi)\hat{S}^{n-1}(\xi).$$

The amplification factor $g(\xi)$ for $L = D$ and $L = A$ is well known; see Tab. 1 for the concise summary of conditions required to establish a bound $|g(\xi)| \leq 1$, from which the strong stability $\|S_h^n\| \leq \|S_h^{n-1}\|$ follows, the scheme is strongly stable, and the error propagates linearly.

**Von-Neumann analysis for systems** When approximating

$$w(x, t) = [u(x, t), v(x, t)]^T \in \mathbb{R}^2,$$

the Fourier analysis is applied to each component of $w(x, t)$. For the evolution system considered in this paper, instead of (31), we derive the system

$$H_1\hat{w}^n(\xi) = H_0\hat{w}^{n-1}(\xi),$$

where $H_1, H_0 \in \mathbb{C}^{2 \times 2}$ are matrices dependent on $h, \tau$, the Fourier variable $\xi$, and the coefficients of the PDE. The form directly resembling (31) is

$$\hat{w}^n = G(\xi)\hat{w}^{n-1},$$

14
with the amplification matrix \( G = G(h, \tau; \xi) = (H_1)^{-1}H_0 \).

Our stability analysis establishes the conditions upon which the n'th power \( G(\cdot; \xi)^n \) of \( G \) is uniformly bounded for all \( 0 \leq n \tau \leq T \). For normal matrices, it suffices to study the spectral radius \( \rho(G) \), since, when \( G^*G = GG^* \), all three members in the inequality \( \rho(G)^n \leq \|G^n\| \leq \|G\|^n \), \( n \geq 1 \), involving the 2-matrix norm \( \|G\| \), are equal. For non-normal matrices, one must study the spectral radius \( G^*G \) matrix, i.e., the first singular value of \( G \), and the analysis of \( \|G\| \) gets quickly quite complicated. The symmetrization and change of variables help in the calculations which otherwise are not easily manageable.

**Proposition 2** Let \( k = 1 \) and \( q \geq 0, d \geq 0 \). The fully discrete finite difference schemes for (21) are strongly stable in \( \sqrt{cu,v} \) variables under the same conditions that apply to the scalar diffusion, advection shown in Table 1. In particular, the implicit schemes are unconditionally strongly stable for \( L = D \) and \( L = A \), and the explicit-implicit schemes are conditionally strongly stable for each \( L = D \) and \( L = A \). The scheme for \( L = D + A \) in which diffusion is implicit, and advection is explicit is strongly stable under the same conditions as that for explicit advection.

The proof is established in the individual subsections below.

### 4.5 Stability of finite difference scheme for diffusion

We provide details for \( L = D \) for the sake of exposition, since for a bounded domain the case was already handled in Cor. 2 and 3 via MOL.

The row of a system (23) with \( L_h \) as in (29c) is equivalent to

\[
\begin{align*}
\frac{u^n_j - u^{n-1}_j}{\tau} + \frac{v^n_j - v^{n-1}_j}{\tau} + d \frac{2u^n_j - u^{n-1}_j - u^{n+1}_j}{h^2} &= 0 \quad (35a) \\
\frac{v^n_j - v^{n-1}_j}{\tau} + \alpha(u^n_j - cu^n_j) &= 0 \quad (35b)
\end{align*}
\]

As suggested in Sec. 3, we first symmetrize (35) by substituting (35b) in (35a), and rescale (35a) by the factor \( \sqrt{c} \). Then we follow the usual non-Neumann analysis steps applied to both components of \( w^n_j = [\sqrt{c}u^n_j, v^n_j]^T \) and its Fourier transforms \( [\sqrt{c}u^n_j(\xi), v^n_j(\xi)]^T \), which we denote by \( \tilde{w}^n \). (According to the convention adopted in Sec. 3 we should use \( \tilde{\tilde{w}}^n \) but we will skip the tilde.)

#### 4.5.1 Implicit scheme for diffusion

If \( n* = n \), we rewrite (35) in the form (33) with

\[
H_1 = \begin{bmatrix}
1 + bc + sD(\xi) & -b\sqrt{c} \\
-b\sqrt{c} & 1 + b
\end{bmatrix}, \quad H_0 = I,
\]

and where \( b = \alpha\tau \).
Since $s_D$ is real, thus $H_1$ is real and symmetric, and its eigenvalues $\lambda_1, \lambda_2$ are real. Let $\lambda_1 \leq \lambda_2$ while $\lambda_1 + \lambda_2 = \text{Trace}(H_1) = 2 + b(c + 1) + s_D$. Since $s_D \geq 0$, both the trace and the determinant of $H_1$ are positive, with the latter given by

$$det(H_1) = 1 + b + bc + b^2c - b^2c + s_D(1 + b).$$

We have thus

$$\lambda_1 \leq \frac{\lambda_1 + \lambda_2}{2} \leq \lambda_2 = \rho(H_1) = \|H_1\|,$$

and we see that $\|H_1\| \geq 1 + \frac{b(c + 1)}{2} + \frac{s_D(\xi)}{2} \geq 1$ for all $\xi$. Since $s_D \geq 0$, we get

$$\|G\| = \|H_1^{-1}\|_2 \leq \frac{2}{2 + b(c + 1)} < 1,$$

which completes this case.

### 4.5.2 Explicit diffusion

In this case $n^* = n - 1$, and we rewrite (35) in the form (33) with

$$H_1 = \begin{bmatrix} 1 + bc & -b\sqrt{c} \\ -b\sqrt{c} & 1 + b \end{bmatrix}, \quad H_0 = \begin{bmatrix} 1 - s_D(\xi) & 0 \\ 0 & 1 \end{bmatrix}$$

and both $H_1, H_0$ are real symmetric matrices.

First we want to find a lower bound for $\|H_1\| = \rho(H_1)$. Denoting the eigenvalues of $H_1$ by $\lambda_1, \lambda_2$, we calculate that $det(H_1) = 1 + b(1 + c) = \lambda_1\lambda_2$, and $\text{Trace}(H_1) = 2 + b(1 + c) = \lambda_1 + \lambda_2$. From this we see $\lambda_2(\lambda_1 - 1) = \lambda_1 - 1$, thus either $\lambda_1$ or $\lambda_2$ must equal 1. Assuming, wlog, $\lambda_1 = 1$, we conclude, from $1 + b(1 + c) = \lambda_2$ that $\lambda_2 > 1$, and thus $\|H_1^{-1}\|_2 = 1$.

On the other hand, the eigenvalues for $H_0$ are on its diagonal, thus the spectral radius is $\rho(H_0) = \max\{1, |1 - s_D|\}$. In order to guarantee $\|G\| = \|H_1^{-1}H_0\| \leq \|H_1^{-1}\||H_0\| \leq \|H_0\| = \rho(H_0) \leq 1$, we must therefore have that $s_D \leq 2$ which requires

$$\frac{d\tau}{h^2} \leq \frac{1}{2}.$$

In summary, the scheme is strongly stable if (39) holds. This is the same result as that in Cor. 3 obtained by MOL.

### 4.6 Scheme for advection

We begin by writing the upwind advection scheme for $L = A$

$$\frac{u_j^n - u_j^{n-1}}{\tau} + \frac{v_j^n - v_j^{n-1}}{\tau} + q\frac{u_j^{n*} - u_j^{n*}_{j-1}}{h} = 0 \quad (40a)$$

$$\frac{v_j^n - v_j^{n-1}}{\tau} + \alpha(v_j^n - cu_j^n) = 0 \quad (40b)$$
We proceed as in Section 4.5, with symmetrization and rescaling, to determine the matrices $H_1$ and $H_0$ in (33).

### 4.6.1 Explicit advection

We find that when $n^* = n - 1$

$$H_1 = \begin{bmatrix} 1 + bc & -b\sqrt{c} \\ -b\sqrt{c} & 1 + b \end{bmatrix}, \quad H_0 = \begin{bmatrix} 1 - s\lambda & 0 \\ 0 & 1 \end{bmatrix}. \quad (41)$$

Our analysis here is similar to that in Section 4.5.2 from which we have $\rho(H_1^{-1}) \leq 1.\) We find that the system is strongly stable provided $\rho(H_0) = \max(1, |1 - s\lambda|) \leq 1$, which holds provided $0 \leq \lambda \leq 1$, and requires $q \geq 0$ and $\tau \leq \frac{h}{q}$, i.e., the usual CFL condition.

### 4.6.2 Implicit advection

Intuitively, we expect to find unconditional stability for $n^* = n$. Proceeding as in Section 4.5 we find that

$$H_1 = \begin{bmatrix} 1 + bc + s\lambda(\xi) & -b\sqrt{c} \\ -b\sqrt{c} & 1 + b \end{bmatrix}, \quad H_0 = I_2. \quad (42)$$

where $s\lambda(\xi)$ is given as in Tab. 1, and has a positive real part $Re(s\lambda)$.

Now $H_1$ is complex symmetric, but not normal, and this requires extra work as compared to the cases before. To show $\|H_1\| \geq 1$ which will demonstrate unconditional stability, we need to estimate the spectral radius of $K = H_1 H_1^*$. Since $\sqrt{|\det(K)|} \leq \rho(K)$, if we prove that $\det(K) \geq 1$, we are done.

We first calculate $K$, simplifying some notation in $H_1 = \begin{bmatrix} X + iY & -\beta \\ -\beta & \gamma \end{bmatrix}$ where we substituted $X = 1 + bc + Re(s\lambda)$, $Y = Im(s\lambda)$, and $\beta = b\sqrt{c}$, and $\gamma = 1 + b$. We get

$$K = \begin{bmatrix} X^2 + Y^2 + \beta^2 & -\beta(X + iY) - \beta\gamma \\ -\beta(X - iY) - \beta\gamma & \beta^2 + \gamma^2 \end{bmatrix},$$

After some lengthy calculations and simplifications, we find that $\det(K) = Y^2\gamma^2 + (X\gamma - \beta^2)^2$ which has a lower bound of $(X\gamma - \beta^2)^2$. We can estimate this term from below, reverting to the original constants in $H_1$ and using $Re(s\lambda) = \lambda(1 - \cos(\xi h)) \geq 0$, to see that

$$X\gamma - \beta^2 = 1 + bc + b + (1 + b)\lambda(1 - \cos(\xi h)) \geq 1 + bc + b \geq 1,$$

and we’re done.
4.7 IMEX scheme for explicit advection and implicit diffusion

The discrete system is as follows

\[
\begin{align*}
\frac{u_j^n - u_j^{n-1}}{\tau} - \alpha (v_j^n - cu_j^n) + q \frac{u_j^{n-1} - u_j^{n-1}}{h} \\
+ \frac{d}{h^2} (-u_j^{n+1} + 2u_j^n - u_j^{n-1}) &= 0 \\
v_j^n - v_j^{n-1} + \alpha (v_j^n - cu_j^n) &= 0
\end{align*}
\]

(43a)

(43b)

We quickly see that the matrix \(H_1\) is the same as in Section 4.5.1 and the matrix \(H_0\) is the same as in Section 4.6.1. The analysis in these sections therefore gives the strong stability of the scheme provided the CFL condition holds.

Summary With the last case we conclude the proof of Proposition 2.

5 Numerical examples

In this Section we illustrate the theory we developed in Sec. 4 with three examples. First, we show the lack of stability of the discrete system in the natural product norm; our simple example motivates the use of weighted spaces and symmetrization. Second, we consider an advection example for which we test the convergence of the numerical scheme, and compare the solutions for different \(\alpha\) to the equilibrium case. Third, we provide an example and convergence rates for diffusion.

5.1 Instability in Euclidean norm on \(\mathbb{R}\)

Here we let \(H = \mathbb{R}\), with \(L = 0.1\), and \(f = 0\), and we consider a fully implicit time discretization (21) of (8). The initial condition \(w^0 = [1, 1]^T\) is given.

The discrete system is solved for the approximations \(w^n = [u^n, v^n]^T\) with a fully implicit scheme

\[w^n = (I + \tau B)^{-1} w^{n-1},\]

(44)

We use \(c = 5\), \(\alpha = 0.1\), and \(\tau = 0.2\).

In Fig. 1 we illustrate the evolution of \(w^n\); these are close to those obtained to MATLAB’s ode45 close to \(w^n\). It is clear that the solutions quickly tend to an asymptote and then start decaying towards the origin. What is interesting is that, the magnitude \(w^n\) grows and the trajectory is “above” the circle \(\|w\| = \|w^0\| = \sqrt{2}\), before it heads towards the origin along the asymptotic.

To explain, we examine \(I + \tau B\) which is not normal when \(c \neq 1\). In fact, even though its eigenvalues can be proven to be greater than 1, its singular values are
Figure 1: Illustration of the lack of strong stability discussed in Sec. 5.1. Left: the phase plot \((u, v)\) shows that the norm \(\|w\|_{H \times H}\) does not necessarily decrease. Right: the plot of the weighted norm \(\|w\|_{W_c}(t)\) decreases while \(\|w\|_{H \times H}(t)\) does not.

not both greater than 1. For example, \(\|(I + \tau B)^{-1}\| \approx 1.00741\), even though its largest eigenvalue is \(\approx 0.9971\).

For independent interest, we study the asymptotics. To determine the asymptotics, we solve for \(v^n\) in terms of \(u^n\), and substitute back to (21). Taking limits of both sides proves that the limit, if it exists, must be 0. For the continuous problem \(w' + Bw = 0\) we clearly expect that close to \([0, 0]^T\) we will have \(v\) follow close to \(v = cu\). However, we find that \(w^n\) actually follows rather the asymptotics for the discrete system, \(v \approx \gamma u\). We can calculate the slope \(\gamma\) directly from \([u, \gamma u]^T = (I + \tau B)^{-1}[u, u]^T\).

In Figure 1, we illustrate both lines \(v = cu\) and \(v = \gamma u\).

On the other hand, after symmetrization, the matrix \((I + \tau \overline{B})\) is symmetric positive definite. We can calculate directly the eigenvalues of \(\overline{B} = \overline{C} + A\), or simply show that for this symmetric \(2 \times 2\) matrix, \(\det(\overline{B}) > 0\) thus both of its eigenvalues \(\lambda_{1,2} \geq 0\). From this we conclude that the eigenvalues of \(I + \tau \overline{B}\) are given by \(1 + \tau \lambda_{1,2} \geq 1\) and thus \(\|(I + \tau \overline{B})^{-1}\| < 1\). (For the numerical example as above, we find \(\|(I + \tau \overline{B})^{-1}\| \approx 0.99732002\).

For illustration, we show that \(\|w^n\|_c\) is a decreasing sequence but \(\|w^n\|\) is not. This is illustrated also in Figure 1.

5.2 Convergence of the schemes for advection and for diffusion

With the stability results developed above, we expect the error for the case \(L = A\) to be of first order, as long as the true solution is smooth enough. While the study of the regularity of the solutions is outside the scope of this paper,
we see that the case \( L = A \) with Riemann data develops enough smoothness to warrant first order error in all \( L_p \) spaces for \( 1 \leq p < \infty \) and even for \( p = \infty \), similarly to what was observed in [5]. In turn, for \( L = D \), with optimal smoothness, we expect second order convergence, which is confirmed.

To test convergence, we use fine grid solution \( u_{h,fine} \) instead of manufacturing solutions which would require nonhomogeneous right-hand side in (7b). To simplify matters, we only report on convergence rate at a fixed stopping time \( T \).

We define the error quantities (classical, and new quantity of interest)

\[
E_{CQ} = \sqrt{\| u - u_h \|^2_{L_2} + \| v - v_h \|^2_{L_2}},
\]

\[
E_{QoI} = \sqrt{\| u - u_h \|^2_{L_2} + \| v - v_h \|^2_{L_2}},
\]

where the \( L_p \) grid norm for \( 1 \leq p < \infty \) is defined as usual

\[
\| u - u_h \|_{L_p} = \left( \sum_i h |u(x_i, T) - u_h(x_i, T)|^p \right)^{1/p}.
\]

In tables below, we report on the errors in different quantities of interest \( E_r \) as well in different norms \( \| \cdot \|_p \), and calculate the respective orders of the error \( \alpha_r, \alpha_p \).

### 5.2.1 Advection case

We consider the problem

\[
\begin{aligned}
  u_t + v_t + u_x &= 0, \quad x \in \mathbb{R} \\
  v_t + \alpha(v - cu) &= 0.
\end{aligned}
\]

and its approximation by the upwind scheme (40). To satisfy the CFL condition, we use \( \lambda = 0.99 \), and we vary \( \tau \) with \( h \) in convergence test. We choose initial data

\[
u(x, 0) = \text{“box”}(x) = \begin{cases} 
1 & \text{if } x \in [-1, 0] \\
0 & \text{otherwise}
\end{cases}.
\]

and \( c = 0.1, \alpha = 2 \). We also set

\[
v(x, 0) = cu(x, 0).
\]

which coresponds to (6). This helps to relate our convergence rates to those obtained in [5].

Since the true solution is not known, we use \( M_{fine} = 5050 \) and \( T = 4.8 \). In Table 2 we show that the error in every quantity of interest is of first order.
\[ M \parallel u - u_h \parallel_{L_2} \alpha_{L_2,u} \parallel v - v_h \parallel_{L_2} \alpha_{L_2,v} \parallel u - u_h \parallel_{L_1} \alpha_{L_1,u} \\
20 \quad 0.03682 \quad - \quad 0.004641 \quad - \quad 0.06217 \quad - \\
50 \quad 0.01655 \quad 0.8728 \quad 0.002557 \quad 0.6504 \quad 0.02607 \quad 0.9483 \\
100 \quad 0.007575 \quad 1.127 \quad 0.0009912 \quad 1.367 \quad 0.01281 \quad 1.026 \\
200 \quad 0.003687 \quad 1.039 \quad 0.0004855 \quad 1.03 \quad 0.006244 \quad 1.036 \\
500 \quad 0.001329 \quad 1.113 \quad 0.0001771 \quad 1.101 \quad 0.002254 \quad 1.112 \\
\]

Table 2: Errors for advection case, with parameters \( c = 0.1, \alpha = 2, \) and “box” as the initial condition. Here \( M_{\text{fine}} = 5050 \) and \( T = 4.8 \)

\[
M \parallel u - u_h \parallel_\infty \alpha_{\text{inf}} E_{\text{CQ}} \alpha_{\text{CQ}} E_{\text{QoI}} \alpha_{\text{QoI}} \\
20 \quad 0.03396 \quad - \quad 0.03711 \quad - \quad 0.1221 \quad - \\
50 \quad 0.02598 \quad 0.2922 \quad 0.01674 \quad 0.8685 \quad 0.05488 \quad 0.8728 \\
100 \quad 0.007129 \quad 1.866 \quad 0.00764 \quad 1.132 \quad 0.02512 \quad 1.127 \\
200 \quad 0.003529 \quad 1.014 \quad 0.003719 \quad 1.038 \quad 0.01223 \quad 1.036 \\
500 \quad 0.001331 \quad 1.064 \quad 0.001341 \quad 1.113 \quad 0.004409 \quad 1.113 \\
\]

Table 3: Errors for diffusion case, with parameters \( c = 5, \alpha = 1.2, \) and the “bell” as the initial condition. Here \( M_{\text{fine}} = 2000; \) \( T = 3.2 \)

### 5.2.2 Convergence for diffusion

We consider the problem

\[
\begin{align*}
u_t + v_t - u_{xx} &= 0, \quad x \in (0,1) \\
v_t + \alpha(v - cu) &= 0.
\end{align*}
\]

with the homogenous Dirichlet boundary conditions, and initial conditions

\[
\begin{align*}
u(x,0) &= \text{“bell”} = \exp\left(-\frac{(x - 0.5)^2}{0.3}\right), \\
v(x,0) &= cu(x,0).
\end{align*}
\]

In all experiments for this case we use \( d = 2, \) stopping time \( T = 3.2 \) and \( M_{\text{fine}} = 2000. \) We vary \( \tau = O(h^2) \) and expect optimal second order convergence. Indeed, Table 3 shows that error is \( O(h^2) \) in every quantity of interest.
5.3 Illustration of equilibrium vs non-equilibrium models

Now we are ready to show simulation results which illustrate the kinetic effects in contrast to the equilibrium case. They are most interesting for \( L = A \), similar to (10). We use \( \Omega = (-1, 3) \) and periodic boundary conditions for \( u \). We set \( \alpha = 2 \) and \( c = 0.5 \).

Fig. 2 shows the evolution of \([u(x,t), v(x,t)]^T\) at three time steps as shown. In addition, we show the evolution of an equilibrium solution in which \( \alpha \to \infty \).

6 Extensions

Above we have shown a unified framework for the analysis of explicit-implicit schemes for the kinetic problems with a linear non-equilibrium relationship. Further work is underway to generalize these results, see below for nonlinear systems and systems with multiple immobile sites. Error analysis and stability for time-discrete schemes is underway.

6.1 Nonlinear equilibrium

Here we consider the nonlinear extension of (2) in which \( cU \) in (2b) is replaced by \( g(u) \), with a monotone increasing function \( g: \mathbb{R} \to \mathbb{R} \). We only consider a finite dimensional case and \( H = \mathbb{R}^P \) since the proper setup with nonlinearity in, e.g., \( L^2(\Omega) \) is extensive and outside the present scope. Here \( g(U) = (g(u_j))_j, u \in \mathbb{R}^P \) is understood pointwise. The problem is

\[
\begin{align*}
U' - \alpha(V - g(U)) + LU &= 0 \quad (51a) \\
V' + \alpha(V - g(U)) &= 0, \quad (51b)
\end{align*}
\]

and we will show stability of a particular new quantity of interest.

**Lemma 2** Let \( L \) and \( g \) satisfy

\[
\langle LU, g(U) \rangle \geq 0 \quad (52)
\]

Then it holds that

\[
\frac{d}{dt} \left[ G(U) + \frac{1}{2} \|V\|^2 \right] \leq 0. \quad (53)
\]

where \( G \) is the primitive of \( g(\cdot) \).

**Proof:** To show the stability, we take the inner product of (51a) with \( g(U) \) and of (51b) with \( V \)

\[
\begin{align*}
\langle U', g(U) \rangle - \alpha \langle V, g(U) \rangle + \langle g(U), g(U) \rangle + \langle AU, g(U) \rangle &= 0 \\
\langle V', V \rangle + \alpha \langle V, V \rangle - \alpha \langle g(U), V \rangle &= 0.
\end{align*}
\]
Figure 2: Evolution for $L = A$ described in Sec. 5.3.
Adding the two equations up we have
\[ \langle U', g(U) \rangle + \langle V', V \rangle + \alpha \langle V, V \rangle - 2\alpha \langle V, g(U) \rangle + \alpha \langle g(U), g(U) \rangle + \langle AU, g(U) \rangle = 0. \]

Rewriting we obtain
\[ \langle U', g(U) \rangle + \langle V', V \rangle + \alpha \langle V - g(U), V - g(U) \rangle + \langle LU, g(U) \rangle = 0. \]

Next step is to define a primitive \( G : \mathbb{R} \rightarrow \mathbb{R} \) of \( g(\cdot) \) so that \( G'(r) = g(r) \) and \( \frac{d}{dr} G(U_j) = g(u_j) \frac{d}{dr} u_j \). We can thus write \( \langle U', g(U) \rangle = \frac{1}{m} \sum_j G(u_j) \).

Thus, if (52) holds, we obtain stability since and we have proven Lemma 2.

Next we provide sufficient conditions for (52) to hold. By a difference matrix \( D \) we mean \( D \) is a difference matrix, and \( K \) is a diagonal matrix with positive entries. Then (52) holds.

**Proposition 3** Let \( L = D^T KD \) where \( D \) is a difference matrix, and \( K \) is a diagonal matrix with positive entries. Then (52) holds.

**Proof:** It remains to prove that \( L = D^T KD \) satisfies (52). We consider first the case when \( K = I \). The matrix \( D^T D \) is the well known “discrete Laplacian” \( L_h \) in Section 4.2 which is symmetric positive definite, and it is easy to see that \( (D^T Du, u)_{RN} = (Du, Du)_{RN+1} = u_1^2 + \sum_{j=2}^{N} (u_j - u_{j-1})^2 + U_N^2 > 0 \) unless \( U = 0 \). Similarly we obtain \( (D^T Du, g(U))_{RN} = (Du, Dg(U))_{RN+1} = u_1 g(u_1) + \sum_{j=2}^{N} (u_j - u_{j-1}) (g(u) - g(u_{j-1}) + u_N g(u_N) \) which is nonnegative by virtue of \( g(\cdot) \) being a monotone increasing function.

In the more general case when \( K \neq I \) we see that the argument above holds for diagonal matrix \( K \) with the the entries \( k_1, \ldots, k_{N+1} \). Then we obtain \( (D^T KDu, g(u))_{RN} = k_1 u_1 g(u_1) + \sum_{j=2}^{N} (k_j) (u_j - u_{j-1}) (g(u) - g(u_{j-1}) + k_{N+1} u_N g(u_N) \). Since each of these entries is nonnegative, we obtain the desired result.

### 6.2 Stability for a system with two species

Here we consider again the finite dimensional space \( H = \mathbb{R}^n \) and write
\[\begin{align*}
U' + V'_1 + V'_2 + LU &= 0 \quad (54) \\
V'_1 + \alpha_1 (V_1 - c_1 U) &= 0 \quad (55) \\
V'_2 + \alpha_2 (V_2 - c_2 U) &= 0 \quad (56)
\end{align*}\]

Here we take the inner product of the first equation with \( c_1 c_2 U \), the second by \( c_2 V_1 \), the third by \( c_1 V_2 \), (notice the crossmultiplications) and add up to get
\[\begin{align*}
c_2 V_1^T V_1 + c_2 V_2^T V_2 + c_1 c_2 U^T U + c_1 c_2 L U^2 + \alpha_1 c_2 (V_1 - c_1 U)^2 + \alpha_2 c_1 (V_2 - c_2 U)^2 &= 0.
\end{align*}\]
from which the stability follows for the following quantity

\[
\frac{d}{dt} \left( c_1 c_2 \| U \|^2 + c_2 \| V_1 \|^2 + c_1 \| V_2 \|^2 \right) \leq 0.
\] (57)

Further extensions to \( m \) species are possible but will not be discussed.

7 Acknowledgements

The authors would like to thank the anonymous reviewers whose remarks helped to improve the paper.

Research presented in this paper was partially supported by NSF grants DMS-1115827 “Hybrid modeling in porous media”, and DMS-1522734 “Phase transitions in porous media across multiple scales”; second author served as a Principal Investigator on these projects. Majority of research was done when first author F. Patricia Medina was a PhD student and later a faculty at Oregon State University.

References

[1] J. W. Barrett, P. Knabner, Finite element approximation of the transport of reactive solutes in porous media. I. Error estimates for nonequilibrium adsorption processes, SIAM J. Numer. Anal. 34 (1) (1997) 201–227. doi:10.1137/S0036142993249024. URL http://dx.doi.org/10.1137/S0036142993249024

[2] J. W. Barrett, H. Kappmeier, P. Knabner, Lagrange-Galerkin approximation for advection-dominated contaminant transport with nonlinear equilibrium or non-equilibrium adsorption, in: Modeling and computation in environmental sciences (Stuttgart, 1995), Vol. 59 of Notes Numer. Fluid Mech., Vieweg, Braunschweig, 1997, pp. 36–48.

[3] C. N. Dawson, C. J. van Duijn, M. F. Wheeler, Characteristic-Galerkin methods for contaminant transport with nonequilibrium adsorption kinetics, SIAM J. Numer. Anal. 31 (4) (1994) 982–999.

[4] H. J. Schroll, A. Tveito, R. Winther, An \( H^1 \)–error bound for a semi-implicit difference scheme applied to a stiff system of conservation laws, SIAM journal on numerical analysis 34 (3) (1997) 1152–1166.

[5] M. Peszynska, Numerical scheme for a conservation law with memory, Numerical Methods for PDEs 30 (2014) 239–264. doi:10.1002/num.21806\&ArticleID=1159335. URL http://www.math.oregonstate.edu/~mpesz/documents/publications/P13NMPDE.pdf
[6] M. Peszynska, R. Showalter, S.-Y. Yi, Flow and transport when scales are not separated: Numerical analysis and simulations of micro- and macro-models, International Journal Numerical Analysis and Modeling 12 (2015) 476–515. URL http://www.math.ualberta.ca/ijnam/Volume-12-2015/No-3-15/2015-03-04.pdf

[7] M. Peszyńska, Finite element approximation of diffusion equations with convolution terms, Math. Comp. 65 (215) (1996) 1019–1037.

[8] R. E. Showalter, Monotone operators in Banach space and nonlinear partial differential equations, Vol. 49 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1997.

[9] J.-Q. Shi, S. Mazumder, K.-H. Wolf, S. Durucan, Competitive methane desorption by supercritical CO2: injection in coal, Transport in Porous Media 75 (2008) 35–54, 10.1007/s11242-008-9209-9. URL http://dx.doi.org/10.1007/s11242-008-9209-9

[10] K. Jessen, W. Lin, A. R. Kovscek, Multicomponent sorption modeling in ECBM displacement calculations, SPE 110258.

[11] K. Jessen, G. Tang, A. R. Kovscek, Laboratory and simulation investigation of enhanced coalbed methane recovery by gas injection, Transport in Porous Media 73 (2008) 141–159.

[12] M. Peszynska, Methane in subsurface: mathematical modeling and computational challenges, in: C. Dawson, M. Gerritsen (Eds.), IMA Volumes in Mathematics and its Applications 156, Computational Challenges in the Geosciences, Springer, 2013.

[13] R. E. Showalter, Hilbert space methods for partial differential equations, Electronic Monographs in Differential Equations, San Marcos, TX, 1994, electronic reprint of the 1977 original.

[14] G. R. King, T. Ertekin, F. C. Schwerer, Numerical simulation of the transient behavior of coal-seam degasification wells, SPE Formation Evaluation 2 (1986) 165–183.

[15] J. Shi, S. Durucan, A bidisperse pore diffusion model for methane displacement desorption in coal by CO2 injection, Fuel 82 (2003) 1219–1229.

[16] M. Peszyńska, R. E. Showalter, A transport model with adsorption hysteresis, Differential Integral Equations 11 (2) (1998) 327–340.

[17] E. DiBenedetto, R. E. Showalter, A pseudoparabolic variational inequality and Stefan problem, Nonlinear Anal. 6 (3) (1982) 279–291.

[18] J. E. Warren, P. J. Root, The behavior of naturally fractured reservoirs, Soc. Petro. Eng. Jour. 3 (1963) 245–255.
[19] G. I. Barenblatt, I. P. Zheltov, I. N. Kochina, Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata), J. Appl. Math. Mech. 24 (1960) 1286–1303.

[20] T. Arbogast, J. Douglas, Jr., U. Hornung, Derivation of the double porosity model of single phase flow via homogenization theory, SIAM J. Math. Anal. 21 (4) (1990) 823–836.

[21] R. E. Showalter, Diffusion in a fissured medium with micro-structure, in: Free boundary problems in fluid flow with applications (Montreal, PQ, 1990), Vol. 282 of Pitman Res. Notes Math. Ser., Longman Sci. Tech., Harlow, 1993, pp. 136–141.

[22] U. Hornung, R. E. Showalter, Diffusion models for fractured media, J. Math. Anal. Appl. 147 (1) (1990) 69–80.

[23] V. Klein, M. Peszynska, Adaptive double-diffusion model and comparison to a highly heterogenous micro-model, Journal of Applied Mathematics 2012 (2012) Article ID 938727, 26 pages. doi:10.1155/2012/938727

[24] M. Peszyńska, R. E. Showalter, Multiscale elliptic-parabolic systems for flow and transport, Electron. J. Diff. Equations 2007 (2007) No. 147, 30 pp. (electronic).

[25] A. Tveito, R. Winther, On the rate of convergence to equilibrium for a system of conservation laws with a relaxation term, SIAM J. Math. Anal. 28 (1) (1997) 136–161. doi:10.1137/S0036141094263755

URL http://dx.doi.org/10.1137/S0036141094263755

[26] M. Böhm, R. E. Showalter, A nonlinear pseudoparabolic diffusion equation, SIAM J. Math. Anal. 16 (5) (1985) 980–999.

[27] V. Thomée, Galerkin finite element methods for parabolic problems, 2nd Edition, Vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2006.

[28] L. Fatone, P. Gervasio, A. Quarteroni, Multimodels for incompressible flows, J. Math. Fluid Mech. 2 (2) (2000) 126–150.

[29] R. J. LeVeque, Finite difference methods for ordinary and partial differential equations, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007, steady-state and time-dependent problems.

[30] R. D. Richtmayer, K. W. Morton, Difference methods for initial-value problems, Second edition. Interscience Tracts in Pure and Applied Mathematics, No. 4. Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[31] G. Strang, Introduction to applied mathematics, Vol. 16, Wellesley-Cambridge Press Wellesley, MA, 1986.