Masters Project

January 2026

Introduction

In the study of partial differential equations, the advection-diffusion reaction equation is used to describe how physical quantities, such as liquids, change over time due to drift (or advection), random spreading (or diffusion), and a reaction or reactions such as generation or consumption. In this paper, we begin with the steady state advection-diffusion reaction equation:

$$-D\Delta u + \boldsymbol{V} \cdot \nabla u + \alpha u = f \quad \text{in } \Omega \tag{1}$$

with Dirichlet boundary conditions, $u = 0$ on $\partial\Omega$. Here,

- $D > 0$ is the constant of diffusivity
- $\boldsymbol{V} \in \mathbb{R}^d, d \in \{1, 2, 3\}$ is a constant vector field
- $\alpha \geq 0$ generates a linear reaction
- We assume $f \in L^2(\Omega)$

We aim to develop a functional analytic framework the coupled version of this problem. We then aim to construct a finite element approximation to the solution.

The Coupled Problem

Let $u(x,t)$ be the population density of the prey. By conservation of mass,

$$u_t + \nabla \cdot J_u = R_u$$

where J_u is the flux and R_u is the local birth/death interactions. From Fick's Law, the diffusion term of J_u is given as $-D_u \nabla u$ (we assume $D_u > 0$ is constant). If individuals drift with velocity field $\boldsymbol{V}_u(x)$ then the advection term of J_u is given by $\boldsymbol{V}_u u$. Thus,

$$\begin{aligned} J_u &= \boldsymbol{V}_u u - D_u \nabla u. \\ \Longrightarrow\ r(u) &= u_t + \nabla \cdot (\boldsymbol{V}_u u - D_u \nabla u) \\ &= u_t + \boldsymbol{V}_u \cdot \nabla u + u \nabla \cdot \boldsymbol{V}_u - D_u \Delta u \end{aligned}$$

we assume $\nabla \cdot \boldsymbol{V}_u = 0$ so we arrive at

$$u_t - D_u u + \boldsymbol{V}_u \cdot \nabla u = r(u).$$

We will assume logistic population growth that is $u_t = r(u)u$ where

$$r(u) = r_{max} \frac{K - u}{K}$$

where K is the carrying capacity. We may rescale K to 1 so that

$$u_t = u(1-u).$$

The predation term is given by $-\alpha uv$ where v is the population density of the predator species. Then

$$R_u = u(1-u) - \alpha uv$$

Similarly,

$$R_v = \beta uv - \gamma v.$$

So our equations are

$$u_t - D_u \Delta u + \boldsymbol{V}_u \cdot \nabla u = u(1-u) - \alpha uv \tag{2}$$

$$v_t - D_v \Delta v + \boldsymbol{V}_v \cdot \nabla v = \beta uv - \gamma v \tag{3}$$

We consider the system on a bounded domain $\Omega \subset \mathbb{R}^d$ with homogeneous Dirichlet boundary conditions. We will consider solutions in $H_0^1(\Omega)$.

Theory

Preliminaries

Def. Weak Derivative

Let Ω be an open subset of $\mathbb{R}^d$. Let $u, v \in L^1_{loc}(\Omega)$. We say that v is the weak partial derivative of u in the direction i if

$$\int_\Omega u \partial_i \varphi \, dx = -\int_\Omega v \varphi \, dx, \qquad \forall \varphi \in C_0^\infty(\Omega).$$

Def. Sobolev Space

Let Ω be an open subset of $\mathbb{R}^d$, and let p, s be a positive integers. Then the

Sobolev Space $W^{s,p}(\Omega)$ is defined by

$$W^{s,p}(\Omega) = \{f \in L^p(\Omega) : \forall |t| < s, \partial_x^t f \in L^p(\Omega)\}.$$

When $p = 2$, our space is called a Hilbert Space and is denoted $H^s(\Omega)$.

Weak Formulation

Consider the prey population equation on a bounded domain $\Omega \subset \mathbb{R}^d$.

$$u_t - D_u \Delta u + \boldsymbol{V}_u \cdot \nabla u = u(1-u) - \alpha uv \quad \text{in } \Omega.$$

Let $\varphi \in H_0^1(\Omega)$. We multiply (2) by φ, here φ is known as a test function. We then integrate over Ω:

$$\int_\Omega u_t \varphi \, dx - D_u \int_\Omega \Delta u \varphi \, dx + \int_\Omega (\boldsymbol{V}_u \cdot \nabla u) \, dx = \int_\Omega (u(1-u) - \alpha uv) \varphi \, dx.$$

We then convert this to the weak form by integrating the Laplacian term by parts.

$$-\int_\Omega \Delta u \varphi \, dx = \int_\Omega \nabla u \cdot \nabla \varphi \, dx - \int_{\partial\Omega} u_n \varphi \, dx$$

Since $\varphi \in H_0^1(\Omega)$ by assumption, $\varphi = 0$ on $\partial\Omega$, so the latter integral is 0. So the weak form of (2) is

$$\int_\Omega u_t \varphi \, dx + D_u \int_\Omega \nabla u \cdot \nabla \varphi \, dx + \int_\Omega (\boldsymbol{V}_u \cdot \nabla u) \, dx = \int_\Omega (u(1-u) - \alpha uv) \, dx \tag{4}$$

Similarly, we find the weak formulation of (3):

$$\int_\Omega v_t \varphi \, dx - D_v \int_\Omega \Delta v \varphi \, dx + \int_\Omega (\boldsymbol{V}_v \cdot \nabla v) \, dx = \int_\Omega \beta uv\varphi - \gamma v \varphi \, dx.$$

The Maximal Linear Problem

The "maximal" linear versions (2) and (3) look like

$$u_t = D_u u - \boldsymbol{V}_u \cdot \nabla u + au + bv \tag{5}$$

$$v_t = D_v v - \boldsymbol{V}_v \nabla v + cu + dv. \tag{6}$$

We tackle these problems through the use of *semigroups*. For equations of the form $du/dt = f(u)$, where $u(0) = u_0$, if there is a unique solution then we can define a solution operator $S(t)$ by $u(t) = S(t)u_0$. These soltion operators form a semigroup.

Def. Continuous Semigroup of Operators

For the linear system $dx/dt = Ax$ with initial condition $x(0) = x_0$, $x \in X$, the solution operator $S(t) : X \to X$ is e^{At} with

$$\begin{cases} S(t+s) = S(t)S(s) & \forall t, s \geq 0 \\ S(0) = I & \\ S(t)x = S(t_0)x & \forall x \in X \text{ whenever } t \to t_0. \end{cases}$$

We view $S(t)$ as defining a flow on X. Now, apply this to the maximal linear versions, (5) and (6), of our problem. Take $a = b = c = d = 0$ for the time being. Define

$$w := \begin{pmatrix} u \\ v \end{pmatrix}$$

Then

$$
Aw = \begin{pmatrix} D_u - \boldsymbol{V}_u + au + bv \\ D_v - \boldsymbol{V}_v + cu + dv \end{pmatrix}
$$

where $X = L^2(\Omega) \times L^2(\Omega)$ and the domain of A, $D(A) = (H^2(\Omega) \cap H_0^1(\Omega) \times (H^2(\Omega) \cap H_0^1(\Omega))$.

Theorem Lumer-Phillips

Let A be a linear operator defined on a linear subspace $D(A)$ of a Banach space X. Then A generates a semigroup if and only if

1. $D(A)$ dense in X
2. A is disspative
3. $\lambda I - A$ is surjective for some $\lambda > 0$

Since X is reflexive, $D(A)$ being dense follows from the other conditions, so we need not verify 1.

Def. Dissipative

A linear operator A is dissipative if $(Ay, y) \leq 0$ for all $y \in D(A)$. In other words, energy of the system does not increase.

So we want to show that $(Aw, w) \leq 0$. Recall that $(u, v)_X = (u_1, v_1)_{L^2} + (u_2, v_2)_{L^2}$, so

$$
(Aw, w)_X = (Au, u)_{L^2} + (Av.v)_{L^2}
$$

.

Breaking this down,

$$\begin{aligned}(Au,u) &= -D_u\int_\Omega \nabla u\cdot\nabla u\;dx - \int_\Omega(\boldsymbol{V}_u\cdot\nabla u)u\;dx\\ &= -D_u||\nabla u||_2^2 - \frac{1}{2}\int_\Omega \boldsymbol{V}_u\cdot\nabla(u^2)\;dx\\ &= -D_u||\nabla u||_2^2 + \frac{1}{2}\int_\Omega(\nabla\cdot\boldsymbol{V}_u)u^2.\end{aligned}$$

We assume divergence free flow, so this becomes

$$(Au,u) = -D_u||\nabla u||_2^2.$$

Similarly,

$$\begin{aligned}&(Av,v) = -D_v||\nabla v||_2^2.\\ \implies &(Aw,w) = -D_u||\nabla u||_2^2 - D_v||\nabla v||_2^2 \le 0.\end{aligned}$$

The second condition of Lumer-Phillips is a direct result of coercivity and the Lax-Milgrim Theorem, covered in the next section. So by Lumer-Phillips, A generates a C^0-semigroup on X, which implies that the linear problem has a unique solution, $w(t) = e^{tA}w_0$. What if don't assume $a=b=c=d=0$? Then our new operator is

$$\begin{aligned}Aw &= \begin{pmatrix} D_u\Delta u - \boldsymbol{V}_u\nabla u\\ D_v\Delta v - \boldsymbol{V}_v\nabla v\end{pmatrix} + \begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix} u\\ v\end{pmatrix}\\ &=: (A'+B)w.\end{aligned}$$

We already have shown that $(A'w,w)\le 0$, now let's consider (Rw,w).

$$(Rw,w) = \int_\Omega au^2 + (b+c)uv + dv^2\;dx.$$

The diffusion-advection operator generates a C^0-semigroup. The linear reaction terms cause a perturbation so a C^0-semigroup is still generated. However, the dissipativity depends on the symmetric part of the reaction matrix. Define

$$\hat{R} := \frac{1}{2}(R + R^T) = \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{pmatrix}.$$

Then,

$$(Rw, w) = \int_\Omega (\hat{R}w, w)_{\mathbb{R}^2}.$$

Since we want $(Aw, w) \leq 0$, we require $(\hat{R}w, w) \leq 0$ which occurs exactly when $\hat{R}$ is negative semi-definite. Which gives us conditional dissipativity based on the conditions:

1. $a \leq 0$
2. $d \leq 0$
3. $ad - \frac{(b+c)^2}{4} \geq 0.$

For what assumptions on $f, \boldsymbol{V}$, and D does a weak solution exist?

Def. Coercive Functional

A bilinear functional ϕ on a normed space E is called coercive if there exists a positive constant K such that $\phi(x, x) \geq K||x||^2$ for all $x \in E$.

Theorem Lax-Milgram

Let ϕ be a bounded coercive bilinear form on a Hilbert space H. Then, for every bounded linear functional f on H, there exists a unique $x_f \in H$ such that

$f(x) = \phi(x, x_f)$ for all $x \in H$.

Let $\Phi = H_0^1(\Omega)$, that is Φ is the space of functions with first-order weak derivatives that disappear in the boundary. on Ω. To use Lax-Milgram to guarantee existence of a weak solution, we must find $u \in \Phi$ such that

$$a(u, \phi) = l(\phi) \ \ \forall \phi \in \Phi,$$

where $|a(u, \phi)| \leq C||u||_\Phi ||\phi||_\Phi$ and $a(\phi, \phi) \geq K||\phi||_\Phi^2$. Consider the case when $\boldsymbol{V} = 0$ so our equation becomes:

$$-D\Delta u + \alpha u = f.$$

Then

$$a(u, \phi) = \int_\Omega D\nabla u \cdot \nabla \phi + \int_\Omega \alpha u \phi.$$

Hence, for any $\phi \in \Phi$,

$$a(\phi, \phi) = D||\nabla \phi||^2 + \alpha ||\phi||^2.$$

Recall that the norm on H^1 is given by

$$||\phi||_{H^1} = \left(||\phi||_{L^2(\Omega)}^2 + ||\nabla \phi||_{L^2(\Omega)}^2 \right)^{1/2}.$$

Then

$$a(\phi, \phi) \geq \min\{D, \alpha\} ||\phi||_{H^1}^2.$$

This gives the coercivity condition for Lax-Milgram. For boundedness, recall the Poincare inequality:

$$||u - u_\Omega||_{L^2(\Omega)} \leq A||\nabla u||_{L^2(\Omega)} \text{ for some constant } A.$$

Define $C = D + \alpha A^2$ From above,

$$\begin{aligned} a(u,\phi) &= \int_\Omega D\nabla u \cdot \nabla\phi + \int_\Omega \alpha u\phi \\ &\leq D||\nabla u||\ ||\nabla\phi|| + \alpha||u||\ ||\phi|| \\ &\leq D||\nabla u||\ ||\nabla\phi|| + \alpha A^2||\nabla u||\ ||\nabla\phi|| \\ &\leq C||\nabla u||\ ||\nabla\phi|| \\ &\leq C||u||_{H^1}||\phi||_{H^1}. \end{aligned}$$

Thus, we have satisfied boundedness for Lax-Milgram. Hence, a weak solution exists.

Regularity

Consider first, the uncoupled case

$$-D\Delta u + \boldsymbol{V} \cdot \nabla u + cu = u(1-u) \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

The weak formulation is to find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega D\nabla u \cdot \nabla\varphi + \boldsymbol{V} \cdot \nabla u\varphi + cu\varphi - (u(1-u))\varphi dx = 0 \quad \forall\varphi \in H_0^1(\Omega).$$

We want to know if the nonlinear term

$$\int_\Omega u(1-u)\varphi dx$$

can be understood for $u, \varphi \in H_0^1(\Omega)$.

Def. Sobolev Space

The Sobolev space $W^{s,p}$ defined on some subset $\Omega \subseteq \mathbb{R}^d$ is defined by

$$W^{s,p} = \{f \in L^p(\Omega) : \forall |\alpha| < s, \partial_x^\alpha f \in L^p(\Omega)\}.$$

The special case when $p = 2$ is called a Hilbert Space, denoted $W^{s,2} = H^s$.

Theorem Sobolev Embedding

Let Ω be a bounded connected subset of $\mathbb{R}^n$. Then the sobolev space

Numerical Analysis

The Finite Element Method

Consider the stationary problem:

$$\begin{aligned} -u''(x) &= f(x), \text{ for } 0 < x < 1 \\ u(0) &= u(1) = 0 \end{aligned} \tag{7}$$

where f is given. This is a 1D Poisson equation with Dirichelet boundary conditions. The purpose of the finite element method is to approximate the solution $u(x)$ using a finite-dimensional subspace $V_h \subset V$, where $V = \{v : v \text{ is a continuous function on } [0,1], \frac{dv}{dx} \text{ is piecewise continuous and bounded on } [0,1] \text{ and } v(0) = v(1) = 0\}$. We then define $V_h = \{v : v \text{ continuous on } [0,1], v \text{ linear on each subinterval } I_i, v(0) = v(1) = 0\}$. Finally, we define our *basis functions* $\varphi_i \in V_H, i = 1, 2, ..., M$ by:

$$\varphi_i(x_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

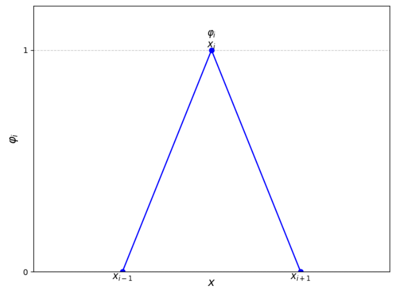

Figure 1: A basis function in one dimension

Variational Formulation

The finite element method starts by rewriting (1) in an equivalent variational formulation. Define the inner product of two vectors v and w by:

$$\int_0^1 v(x)w(x)dx.$$

Multiplying both sides of (1) by a test function v we obtain

$$-u''(x)v(x) = f(x)v(x)$$

integrating,

$$(-u'', v) = (f, v).$$

Integrating the lefthand side by parts, see that

$$(-u'', v) = u'(x)v(x)|_0^1 - (u', v')$$

And since $v(1) = v(0) = 0$,

$$(-u'', v) = -(u', v')$$

Thus, we arrive at the *Galerkin variational form* of (1):

$$(u', v') = (f, v) \tag{8}$$

Then the *Galerkin finite element method* is the problem:

$$\text{Find } u_h \in V_h \text{ such that } (u', v') = (f, v) \text{ for all } v \in V_h. \tag{9}$$

Any $v \in V_h$ has a unique representation

$$v(x) = \sum_{i=1}^{M} v_i \varphi_i(x), 0 \leq x \leq 1,$$

where $v_i = v(x_i)$. For every $j \in \{1, 2, ..., M\}$, substitute v for φ_j in (3) to get the discretized version of (3):

$$(u', \varphi_j') = (f, \varphi_j) \tag{10}$$

which then gives us,

$$\sum_{i=1}^{M} (\varphi_i', \varphi_j') u_i = (f, \varphi_j) \tag{11}$$

Equivalent Minimization Problem

Consider the linear functional $F : V \to \mathbb{R}$ defined by

$$F(v) = \frac{1}{2}(v', v') - (f, v).$$

We claim that the minimization problem

$$\text{Find } u \in V \text{ such that } F(u) \leq F(v) \text{ for all } v \in V$$

is equivalent to the variational problem given above. The quantity, $\frac{1}{2}(v', v')$ represents the internal elastic energy of a system, (f, v) represents the load potential, and $F(v)$ is the total potential energy. The this minimization problem corresponds to the *fundamental principal of minimum potential energy.* Suppose u is a solution to (1). Then from (2) we see that $(u', v') = (f, v)$ so that u is also a solution to the variational problem. Let $w = v - u$ so that $v = u + w$ and $w \in V$. Then

$$\begin{aligned} F(v) = F(u+w) &= \frac{1}{2}(u' + w', u' + w') - (f, u + w) \\ &= \frac{1}{2}(u', u') - (f, u) + (u', w') + \frac{1}{2}(w', w') - (f, w) \end{aligned}$$

Since u is a solution to (2), we have that $(u', w') - (f, w) = 0$ and $w \in V$ implies that $(w', w') \geq 0$. Thus,

$$F(u) = \frac{1}{2}(u', u') - (f, u) \leq \frac{1}{2}(u', u') - (f, u) + \frac{1}{2}(w', w')$$

which implies u is a solution to the minimization problem.

Conversely, let u be a minimizer of F. Then for any $v \in V$ and $\epsilon > 0$, we have that $F(u) \leq F(u + \epsilon v)$. Define $G(\epsilon) = F(u + \epsilon v)$ which must be nonnegative so it has a minimizer at $\epsilon = 0$. Therefore, $G'(0) = 0$ so that $(u', v') - (f, v) = G'(0) = 0$ implying that u is a solution to (2). We conclude that the variational problem and the minimization problem have the same solution.

Appendix

Miscellaneous Definitions

Def. Reflexive Space

Let U be a normed vector space and U' its topological dual. Let $u \in U$ and $f \in U'$. Then U is called reflexive if the evaluation of f at u is a surjective map.

Def. Weak Topology

Let U be a normed space and U' its dual. For each continuous linear functional $f \in U$, define $p_f(u) := |f(u)| = |(f, u)|$. The p_f seminorms can be used to construct a locally convex topology on U, called the weak topology.

Proofs of Theorems

Theorem Weak Sequential Compactness

Let U be a separable reflexive Banach space and $\{u_n\}$ be any sequence of U that is bounded in the norm of U by M. Then there exists a subsequence $\{U_{n_k}\}$ that converges weakly to an element $u \in U$ such that $||u|| \leq M$.

Proof:

Since U is separable, U' must also be separable. So let $\{f_i\}$ be a countable dense subset of U'. We have

$$|f_i(u_n)| \leq ||f_i||_{U'}||u_n||_U \leq ||f_i||_{U'}M$$

which implies that $f_i(u_n)$ is bounded. Hence, $\{f_i(u_n)\}$ contains a convergent subsequence $\{f_i(u_{n_k})\}$. Then $\{u_{n_k}\}$ is weakly convergent which implies $u_n \rightharpoonup u$ for some $u \in U$ with

$$||u|| \liminf_{n \longrightarrow \infty} ||u_n|| \leq M$$

.

Full Firedrake Code

MMS Verification

```
from firedrake import *

import numpy as np

import matplotlib.pyplot as plt


# final time

t_final = 0.1


discrete = False


# mesh refinement in space *and time*

mesh_sizes = [8, 16, 32, 64, 128]

h_values = [1.0 / size for size in mesh_sizes]

dt_values = [t_final / size for size in mesh_sizes]


errors_u = []

errors_v = []


for size, h, dt in zip(mesh_sizes, h_values, dt_values):

    print(f"solving with {size}x{size} mesh (h={h:.4f}) and dt={dt:.4f} ...")


    mesh = UnitSquareMesh(size,size)

    V = FunctionSpace(mesh,"CG",1)

    W = V*V
```

```
U = Function(W)
U_n = Function(W)

(u,v) = split(U)
(u_n,v_n) = split(U_n)
(psi,phi) = TestFunctions(W)

D_u = Constant(1.0) # diffusivity
D_v = Constant(1.0)
alpha = Constant(1.0)
beta = Constant(1.0)
gamma = Constant(1.0)
V_u = as_vector((1.0, 0.5)) # advective terms
V_v = as_vector((-0.5, 1.0))

x, y = SpatialCoordinate(mesh)
t = Constant(0.0)

u_exact = sin(pi*x)*sin(pi*y)*exp(-t)
v_exact = sin(2*pi*x)*sin(2*pi*y)*exp(-2*t)

if discrete:
    u_exact_n = sin(pi*x)*sin(pi*y)*exp(-(t - dt))
    v_exact_n = sin(2*pi*x)*sin(2*pi*y)*exp(-2*(t - dt))

    f_u = (u_exact - u_exact_n)/dt - D_u*div(grad(u_exact)) + dot(V_u, grad(u_exact)) -
```

```
    f_v = (v_exact - v_exact_n)/dt - D_v*div(grad(v_exact)) + dot(V_v, grad(v_exact)) -
else:
    t = variable(t)
    f_u = diff(u_exact, t) - D_u*div(grad(u_exact)) + dot(V_u, grad(u_exact)) - (u_exact
    f_v = diff(v_exact, t) - D_v*div(grad(v_exact)) + dot(V_v, grad(v_exact)) - (beta*u_
    t = Constant(t)

# weak forms
F_u = ((u - u_n)/dt * psi * dx + D_u*dot(grad(u), grad(psi)) * dx + dot(V_u, grad(u)) *
F_v = ((v - v_n)/dt * phi * dx + D_v*dot(grad(v), grad(phi)) * dx + dot(V_v, grad(v)) *
F = F_u + F_v

bcs = [DirichletBC(W.sub(0), 0.0, "on_boundary"),DirichletBC(W.sub(1), 0.0, "on_boundary

# initial conditions
U_n.sub(0).interpolate(u_exact)
U_n.sub(1).interpolate(v_exact)
U.assign(U_n)

t_curr = 0.0
while t_curr < t_final - 1e-12:
    t.assign(t_curr + dt)
    solve(F == 0,U,bcs=bcs,solver_parameters={
            "snes_type": "newtonls",
            "snes_converged_reason":None,
            "ksp_type": "gmres",
            "pc_type": "ilu",
```

```
            }, options_prefix=f"t={t_curr + dt:.4f}"
        )

        U_n.assign(U)
        t_curr += dt

    u_h, v_h = U.subfunctions
    u_e = Function(V).interpolate(u_exact)
    v_e = Function(V).interpolate(v_exact)

    error_u = errornorm(u_e, u_h, norm_type="L2")
    error_v = errornorm(v_e, v_h, norm_type="L2")

    errors_u.append(error_u)
    errors_v.append(error_v)

plt.figure()
plt.loglog(h_values, errors_u, "o-", label="u error")
plt.loglog(h_values, errors_v, "s-", label="v error")
plt.xlabel("h")
plt.ylabel("Error")
plt.legend()
plt.grid(True, which="both")
plt.show()
```