

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

REDUCCIÓN DE UN MODELO DE DISPERSIÓN POBLACIONAL
UTILIZANDO EL MÉTODO POD

PROYECTO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
MATEMÁTICO

SOFÍA GABRIELA JIJÓN ALBÁN

sofia.jijon@gmail.com

DIRECTOR: PEDRO MERINO, PH.D.

pedro.merino@epn.edu.ec

Quito, Enero 2013

DECLARACIÓN

Yo, SOFÍA GABRIELA JIJÓN ALBÁN, declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Sofía Gabriela Jijón Albán

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por SOFÍA GABRIELA JIJÓN ALBÁN, bajo mi supervisión

Pedro Merino, Ph.D.
DIRECTOR

AGRADECIMIENTOS

Agradezco a quienes significaron para mí un soporte, a lo largo de la carrera y durante el proceso de escribir este proyecto:

A Pedro Merino, mi director, por haber dedicado tanto tiempo a este documento, por sus minuciosas correcciones y por haber estado disponible siempre que tuve dudas.

A mis papás y mi hermana, por su apoyo desde el instante en el que decidí estudiar Matemáticas y por ser para mí los brazos abiertos que me recibieron exhausta tantas veces, incondicionalmente.

A mis amigos: Alan, Juan, Andrés y el resto de los Green, quienes me han tenido paciencia, me han hecho barras y me han acompañado siempre que los he necesitado; a Erick y Sebastián por ser los físicos que me han acompañado y apoyado durante toda la carrera.

DEDICATORIA

A mi familia: mi ñaña, mi mami y mi papi.

Con mucho cariño.

Índice de Contenido

| | |
|---|------------|
| Lista de Figuras | ii |
| Lista de Tablas | iii |
| Resumen | iv |
| Abstract | v |
| Introducción | 1 |
| 1 Modelo de Fisher | 3 |
| 1.1 Modelos de Dinámicas de Población con Ecuaciones de Reacción-Difusión | 3 |
| 1.1.1 Dinámica de la especie | 3 |
| 1.1.2 Difusión de la especie | 4 |
| 1.2 Descripción del Modelo de Fisher | 5 |
| 2 Problema de Valor Inicial | 8 |
| 2.1 Propiedades A-Priori de la Solución | 9 |
| 2.2 Existencia y Unicidad | 12 |
| 2.3 Regularidad | 19 |
| 3 Resolución Numérica de la Ecuación de Fisher | 21 |
| 3.1 Aproximación por el Método de Galerkin y de los Elementos Finitos . . | 21 |
| 3.1.1 Linealización de la Ecuación de Fisher | 21 |
| 3.1.2 Esquema de Newton | 27 |
| 3.1.3 Esquema de Euler implícito (EI) | 29 |
| 3.1.4 Esquema de Crank-Nicholson (CN) | 30 |
| 3.2 Implementación Numérica | 32 |
| 3.2.1 Resolución por el método de Euler Implícito | 33 |
| 3.2.2 Resolución por el método de Crank-Nicholson estabilizado . . . | 37 |

| | |
|--|-----------|
| 4 Reducción del Modelo de Fisher Mediante el Método POD | 40 |
| 4.1 Introducción al Método POD | 40 |
| 4.1.1 Descripción del método | 41 |
| 4.1.2 Modelo de Fisher reducido | 46 |
| 4.1.3 Construcción numérica del esquema en la base POD | 48 |
| 4.2 Experimentos Numéricos para el Método POD | 51 |
| 4.3 Experimentación Numérica y Simulación del Modelo de Fisher | 59 |
| | |
| Conclusiones y Recomendaciones | 66 |
| | |
| Referencias | 68 |
| | |
| Anexos | 69 |
| | |
| A Algunas Definiciones | 70 |
| | |
| B Algunos Resultados y Propiedades | 71 |

Lista de Figuras

| | | |
|------|---|----|
| 3.1 | Malla con $h = 0,5$ | 34 |
| 3.2 | Malla con $h \approx 0,0625$ | 35 |
| 3.3 | Ejemplo 3.1: solución exacta en $t = 1$ | 36 |
| 3.4 | Ejemplo 3.1: aproximación por método EI, en $t = 1$ con $h = 0,5$ | 37 |
| 3.5 | Ejemplo 3.1: aproximación por EI, en $t = 1$ $h = 0,0625$ | 37 |
| 3.6 | Ejemplo 3.2: aproximación por método CN, con $h = 0,5$ | 39 |
| 3.7 | Ejemplo 3.2: aproximación por CN, $h = 0,0625$ | 39 |
| 4.1 | Ejemplo 3.2: valores propios de la matriz de correlación | 52 |
| 4.2 | Ejemplo 3.2: 1er elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.3 | Ejemplo 3.2: 2do elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.4 | Ejemplo 3.2: 3er elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.5 | Ejemplo 3.2: 4to elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.6 | Ejemplo 3.2: 5to elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.7 | Ejemplo 3.2: 6to elemento de la Base POD, $h = 0,016$ y $N_t = 800$ | 53 |
| 4.8 | Errores POD+EI y POD+CN en función de Δt | 54 |
| 4.9 | Ejemplo 3.2: EI vs. CN. | 54 |
| 4.10 | Ejemplo 3.2: errores con $d = 2, 4, 6 + EI$ | 55 |
| 4.11 | Ejemplo 3.2: errores con $d = 2, 4, 6 + CN$ | 55 |
| 4.12 | Ejemplo 3.2: errores con $N_t = 300, 500, 800 + EI$ | 55 |
| 4.13 | Ejemplo 3.2: errores con $N_t = 300, 500, 800 + CN$ | 55 |
| 4.14 | Ejemplo 4.1: Solución Aproximada en $t = 0$ | 56 |
| 4.15 | Ejemplo 4.1: Solución Aproximada en $t = 1$ | 56 |
| 4.16 | Ejemplo 4.1: valores propios de la matriz de correlación | 56 |
| 4.17 | Ejemplo 4.1: 1er elemento de la Base POD | 57 |
| 4.18 | Ejemplo 4.1: 2do elemento de la Base POD | 57 |
| 4.19 | Ejemplo 4.1: 3er elemento de la Base POD | 57 |
| 4.20 | Ejemplo 4.1: 5to elemento de la Base POD | 57 |
| 4.21 | Errores POD+EI y POD+CN en función de Δt | 58 |
| 4.22 | Ejemplo 4.1: CN vs EI. | 58 |

| | | |
|------|--|----|
| 4.23 | Ejemplo 4.1: errores con $d = 3, 5, 10 + \text{EI}$ | 59 |
| 4.24 | Ejemplo 4.1: errores con $d = 3, 5, 10 + \text{CN}$ | 59 |
| 4.25 | Ejemplo 4.1: errores con $N_t = 300, 500, 800 + \text{EI}$ | 59 |
| 4.26 | Ejemplo 4.1: errores con $N_t = 300, 500, 800 + \text{CN}$ | 59 |
| 4.27 | Simulación: mapa del Ecuador | 60 |
| 4.28 | Simulación: estado inicial y | 60 |
| 4.29 | Simulación: función $u(\cdot, t)$ | 61 |
| 4.30 | Simulación: función u | 62 |
| 4.31 | Simulación: errores de aproximación CN estabilizado vs. POD. | 62 |
| 4.32 | Simulación: valores propios de la matriz de Correlación. | 63 |
| 4.33 | Simulación: 1er elemento de la Base POD | 63 |
| 4.34 | Simulación: 2do elemento de la Base POD | 63 |
| 4.35 | Simulación: 3er elemento de la Base POD | 64 |
| 4.36 | Simulación: 4to elemento de la Base POD | 64 |
| 4.37 | Simulación: error de aproximación del problema perturbado. | 64 |

Lista de Tablas

| | | |
|-----|--|----|
| 3.1 | Ejemplo 3.1: error en norma $L^2(\Omega)$, método EI. | 35 |
| 3.2 | Ejemplo 3.1: error en norma $H^1(\Omega)$, método EI. | 36 |
| 3.3 | Ejemplo 3.2: error en norma $L^2(\Omega)$, método CN. | 38 |
| 3.4 | Ejemplo 3.2: error en norma $H^1(\Omega)$, método CN. | 39 |
| 4.1 | Ejemplo 3.2: errores de aproximación del método POD. | 54 |
| 4.2 | Ejemplo 4.1: errores de aproximación del método POD. | 57 |

Resumen

Los modelos de dinámica de población que no sólo consideran el crecimiento poblacional de una especie a lo largo del tiempo, sino que también consideran su dispersión en el espacio, pueden ser modelados matemáticamente mediante ecuaciones de reacción-difusión. Desde el punto de vista práctico, es de interés saber de qué manera es posible predecir y controlar dichos procesos. Por ejemplo: limitar la expansión de plagas, planificar el método de pesca en una piscina, la recolección de productos de una plantación, limitar el crecimiento de una población de bacterias, etc. En este proyecto de titulación estudiaremos el problema dado por un sencillo modelo particular de crecimiento poblacional, descrito por la ecuación de Fisher.

Para su aplicación práctica, estos problemas requieren ser resueltos numéricamente en un computador. En el caso de la resolución de ecuaciones de reacción-difusión en dos dimensiones, luego de un proceso de discretización del modelo, nos encontramos rápidamente con la dificultad de resolver problemas a gran escala; los cuales difícilmente pueden ser trabajados con alta precisión en computadores personales debido al gran costo computacional que requiere su resolución numérica. Por esta razón, y como tema central de este proyecto, nos proponemos aplicar a la ecuación de Fisher la técnica de reducción de modelo conocida como POD por sus siglas en inglés: “Proper Orthogonal Decomposition”. Diversos ensayos numéricos son presentados a lo largo de este trabajo para ilustrar distintos aspectos del modelo estudiado y las técnicas numéricas aplicadas para resolverlo.

PALABRAS Y CONCEPTOS CLAVE: Proper Orthogonal Decomposition (POD), Modelos de Reacción-Difusión, Métodos de Semidiscretización para Ecuaciones Diferenciales Parciales Parabólicas, Método de los Elementos Finitos, Dinámicas de Población.

Abstract

Population dynamics models which not only consider population growth of a species over time, but also consider their dispersion in space, can be modeled mathematically using reaction-diffusion equations. From a practical point of view, it is interesting to know how these processes can be predicted and controlled. For example: limiting the spread of pests, planing fishing method in a pool, collecting a plantation products, limiting the growth of a bacteria population, etc. In this project, we study a problem given by a particular simple model of population growth described by the Fisher's equation.

For practical application, these problems need to be solved numerically on a computer. For solving reaction-diffusion equations in two dimensions, after a process of discretization of the model, we quickly find ourselves with the difficulty of solving large-scale problems, which can hardly be solved with high precision in a personal computer due to the large amount of computational effort required for its numerical resolution. For this reason, we propose the application of a model reduction technique known as POD (Proper Orthogonal Decomposition) to Fisher's equation, which is the focus of this project. Several numerical examples are presented throughout this document to illustrate different aspects of the model and the numerical techniques applied to solve it.

KEY WORDS AND CONCEPTS: Proper Orthogonal Decomposition (POD), Reaction-Diffusion Models, Semi Discretization Methods for Parabolic Partial Differential Equations, Finite Element Method, Population Dynamics

Introducción

“Una vez que aceptemos el hecho de que un perfecto entendimiento de un sistema muy complejo está fuera del alcance, y que la noción de la explicación última no es más que un sueño, somos libres de hacer un progreso científico.”

SHONKWILER, R. Y HEROLD, J., [14].

El alto nivel de complejidad en los procesos biológicos ha obligado a que el estudio de los mismos sea extendido a otras disciplinas. La Biología busca comprender los fenómenos biológicos a profundidad, y desde un punto de vista práctico, busca predecir comportamientos. Con este fin, las Matemáticas han aportado a la Biología con herramientas para describir distintos fenómenos.

En [14] se introduce una relación sinérgica que se produce cuando la Biología hace conocer la existencia de problemas que pueden ser comprendidos a través de la modelización Matemática. Posteriormente, la Biología se nutre de estos modelos matemáticos y verifica su validez en la práctica.

De una manera muy resumida, la idea es la siguiente: se observa un fenómeno biológico y se intenta modelarlo matemáticamente. ¿A qué nos referimos con esto? Se busca explicar matemáticamente el fenómeno; por ejemplo mediante funciones que describan el proceso y sistemas de ecuaciones que hagan de restricciones, eligiendo aquellas que son determinantes (y deben necesariamente ser tomados en cuenta en el modelo) y aquellas que no lo son. A este conjunto de ecuaciones se le llama *modelo*. El objetivo de un modelo no es sólo el comprender mejor el fenómeno desde un punto de vista simplificado, sino también el dar el salto a la predicción: cuando un modelo está formulado, de ser posible resolverlo, se puede predecir los resultados del fenómeno. La Biología se beneficia de la solución obtenida matemáticamente, puesto que puede aplicarla a un problema real, incluso siendo capaces de modificar las condiciones del fenómeno para conseguir un resultado esperado.

En vista de que el modelo no es más que una aproximación al fenómeno, no siempre se obtienen resultados suficientemente favorables en la práctica. En ese caso, se reconsidera el problema matemático mejorando el modelo: se introducen nuevas restricciones, se eliminan otras, se consideran también nuevos parámetros, se ajustan los métodos numéricos; en fin, se busca el mejor acercamiento posible al problema para encontrar una solución cercana a la solución real, comprendiendo que no estamos todavía en la capacidad de modelar completamente un proceso aceptando las limitaciones de los modelos y sus métodos de resolución.

En este proyecto de titulación, trataremos matemáticamente un problema particular de crecimiento y dispersión poblacional: el Modelo de Fisher, considerando una única población en una región de dos dimensiones. Lo que nos interesa es conocer la densidad de individuos de la población en un tiempo determinado y en un punto específico de la región considerada. Esta información viene a ser llamada el *estado* de la población en un punto y tiempo determinados.

La ecuación que describe el crecimiento de la población a través del paso del tiempo y su extensión en el espacio considerado, hacen de éste un problema de Reacción-Difusión, ya que está conformado por una ecuación parabólica semilineal, cuyo término no lineal viene dado por una función de crecimiento logístico; cf. [11, 2]. Ésta contempla además dos parámetros inherentes al fenómeno: la tasa de crecimiento de la especie y la capacidad máxima de individuos que puede ser sostenida por el entorno.

Adicionalmente, se considera un término que de alguna manera controla el crecimiento poblacional y que puede interpretarse como la tasa de mortalidad propia de la especie o como la tasa de extracción o aniquilación de individuos.

Una vez formulado el modelo, se puede proceder a la resolución del mismo. La teoría desarrollada en el campo de las ecuaciones diferenciales parciales permite conocer de la existencia y unicidad de la solución del problema, además de conocer ciertas propiedades de la misma que son importantes para la aplicación de los resultados.

Por otro lado, la resolución numérica de este tipo de problemas puede llegar a ser muy costosa computacionalmente. En particular, la resolución de problemas de Reacción-Difusión por el método de Elementos Finitos y métodos de integración numérica, genera problemas que involucran matrices de dimensiones muy grandes. Por esta razón, es necesario aplicar un método alternativo de resolución que permita obtener buenos resultados con un costo computacional mucho menor.

Uno de los métodos desarrollados que reduce significativamente las dimensiones de este tipo de problemas, conservando la información más importante del mismo, es el método POD (Proper Orthogonal Decomposition). Este método se basa en la información de una solución aproximada del modelo, y considera la representación de dicho modelo mediante funciones ricas en información inherente al problema, de manera que necesita un número significativamente menor de variables y por tanto, los problemas matriciales que deben resolverse tienen una dimensión significativamente menor.

En este proyecto de titulación, aplicaremos el método POD para reducir el problema de valor inicial dado por la ecuación de Fisher, y constataremos la efectividad de esta técnica en la reducción de modelos de dispersión poblacional.

Capítulo 1

Modelo de Fisher para la Dinámica Poblacional Espacial

1.1 Modelos de Dinámicas de Población con Ecuaciones de Reacción-Difusión

En esta sección, vamos a estudiar los fenómenos de dinámica y difusión poblacional de una manera detallada para comprender y justificar la presencia de cada uno de los elementos de la ecuación de Reacción-Difusión.

Comenzamos discutiendo sobre la conocida *Ecuación Logística*, que modela el crecimiento poblacional dominado por dos parámetros: la tasa de crecimiento natural de la especie y la capacidad de carga del entorno. Luego, justificaremos la inclusión del término de difusión, a partir del concepto del *Paseo Aleatorio*.

1.1.1 Dinámica de la especie: una breve discusión sobre la ecuación logística

Consideremos una población de cierta especie no diferenciada¹ cuya *densidad poblacional* (es decir, la cantidad de individuos vivos de la población) en un cierto instante de tiempo t está dada por una función $y(t)$. Observar el comportamiento *dinámico* de la especie significa observar el comportamiento de esta función que depende del tiempo.

Supongamos que r es la tasa de natalidad de la población, y que κ es la cantidad máxima de habitantes naturalmente sostenibles por el hábitat o *capacidad de carga*. En el modelo estudiado en este proyecto de titulación no se toma en cuenta al término correspondiente a la mortalidad natural de la especie, pues vamos a considerar un término general que puede referirse tanto a la mortalidad natural como a un agente externo de control, que puede ser a su vez la aniquilación o recolección de los individuos durante la evolución del fenómeno.

¹Sin tomar en cuenta ningún tipo de categorización intraespecífica como la edad de los individuos, o el sexo, o interacciones intraespecíficas.

Si queremos conocer la razón de cambio de la cantidad de individuos con respecto al tiempo, podemos modelar el comportamiento dinámico de la población tomando en cuenta cómo se incrementa la cantidad de individuos de la especie y cómo disminuye por la limitación de la capacidad de carga del entorno, mediante la ecuación:

$$\frac{dy}{dt} = ry(t) - \frac{r}{\kappa}y(t).$$

El término $-r/\kappa$ representa una tasa de mortandad referente a la sobrepoblación. Así, el modelo que considera un crecimiento poblacional autoregulatorio dependiendo únicamente de una variable temporal, queda dado por la llamada *ecuación logística* (cf. [11]) que tiene la siguiente forma:

$$\frac{dy}{dt} = ry(t) \left(1 - \frac{y(t)}{\kappa}\right). \quad (1.1)$$

1.1.2 Difusión de la especie

Ahora, analicemos el esparcimiento o *difusión* de la población en un territorio, es decir, el fenómeno de movimiento de la población que viene dado por muchos pequeños movimientos de cada uno de los individuos de la población en direcciones al azar. El *estado* de la población representa la cantidad de individuos en un punto x del espacio y en un instante t ; es decir, es una función $y = y(x, t)$.

Para justificar el término de difusión vamos a seguir el modelo que se encuentra en [5], en el que el movimiento de los individuos se considera dado por el proceso estocástico del Paseo Aleatorio (o Random Walk, en inglés). Para simplificar el planteamiento del problema, en [5] se considera el paseo aleatorio de un solo individuo y en una sola dimensión.

Supongamos que nuestro individuo se encuentra en la posición x en el instante t , y que se mueve sobre la línea recta dando pasos de longitud δx o bien hacia la derecha con probabilidad $1/2$, o bien hacia la izquierda con la misma probabilidad, en cada intervalo de tiempo $\delta t > 0$ transcurrido; de manera que en el instante $t - \delta t$, el individuo pudo haber dado un paso a la izquierda o a la derecha, y luego se movió hacia la derecha o hacia la izquierda antes de llegar a la posición x . Si $p(x, t)$ es la probabilidad de que el individuo se encuentre en la posición x en el instante t , entonces

$$p(x, t) = \frac{1}{2}p(x + \delta x, t - \delta t) + \frac{1}{2}p(x - \delta x, t - \delta t).$$

Restamos $p(x, t - \delta t)$ de ambos lados de la ecuación anterior, y dividimos para δt :

$$\frac{p(x, t) - p(x, t - \delta t)}{\delta t} = \frac{1}{2\delta t}[p(x + \delta x, t - \delta t) - 2p(x, t - \delta t) + p(x - \delta x, t - \delta t)]. \quad (1.2)$$

Ahora, ponemos:

$$\frac{(\delta x)^2}{\delta t} \equiv 2\gamma,$$

para cierto coeficiente $\gamma > 0$, al cual lo llamaremos *coeficiente de difusión*. Esta razón representa el crecimiento difusivo (diffusive scaling) y tiene sentido en el modelo que vamos a estudiar puesto que en él se considera una *velocidad* de difusión constante, cf. [4, 5].

Esta interpretación del coeficiente de difusión γ es válido para cualquier número finito de dimensiones, en particular para dos dimensiones, que es lo que requiere nuestro problema, cf. [2, 5]. Con estas

consideraciones, (1.2) queda de la forma:

$$\frac{p(x, t) - p(x, t - \delta t)}{\delta t} = \frac{\gamma}{(\delta x)^2} [p(x + \delta x, t - \delta t) - 2p(x, t - \delta t) + p(x - \delta x, t - \delta t)]. \quad (1.3)$$

Tomando el límite cuando $\delta x, \delta t \rightarrow 0$ en (1.3), resulta que el lado izquierdo de la ecuación es un cociente de diferencias con respecto a t , mientras que el lado derecho es una diferencia de segundo orden con respecto a x . En efecto,

$$\frac{\partial p}{\partial t} = \lim_{\delta t, \delta x \rightarrow 0} \frac{p(x, t) - p(x, t - \delta t)}{\delta t},$$

y

$$\begin{aligned} & \lim_{\delta t, \delta x \rightarrow 0} \frac{\gamma}{(\delta x)^2} [p(x + \delta x, t - \delta t) - 2p(x, t - \delta t) + p(x - \delta x, t - \delta t)] \\ &= \gamma \lim_{\delta x \rightarrow 0} \frac{p(x + \delta x, t) - 2p(x, t) + p(x - \delta x, t)}{(\delta x)^2} \\ &= \gamma \lim_{\delta x \rightarrow 0} \frac{1}{\delta x} \left[\frac{p(x + \delta x, t) - p(x, t)}{\delta x} - \frac{p(x, t) - p(x - \delta x, t)}{\delta x} \right] \\ &= \gamma \lim_{\delta x \rightarrow 0} \frac{1}{\delta x} \left[\frac{\partial}{\partial x} p(x + \delta x, t) - \frac{\partial}{\partial x} p(x, t) \right] \\ &= \gamma \frac{\partial^2 p}{\partial x^2}. \end{aligned}$$

Con lo que, tomando el límite cuando $\delta x, \delta t \rightarrow 0$ en la ecuación(1.3), tenemos

$$\frac{\partial p}{\partial t} = \gamma \frac{\partial^2 p}{\partial x^2}. \quad (1.4)$$

Esta ecuación tiene la forma de la ecuación de calor, de la cual se obtiene la noción de difusión. De manera que, si sólo consideramos el movimiento de los individuos, sin tomar en cuenta las dinámicas poblacionales referentes a la reproducción o la mortalidad, podemos relacionar el estado de la población (referente al movimiento) con la probabilidad de que los individuos se encuentren en cierta posición x y en cierto instante t . Así, obtenemos la siguiente expresión:

$$\frac{\partial y}{\partial t} = \gamma \frac{\partial^2 y}{\partial x^2},$$

donde $y = y(x, t)$, $\gamma > 0$.

Análogamente, considerando un espacio de dos dimensiones, la parte derecha de la ecuación anterior corresponde al Laplaciano de la función y . Entonces, el fenómeno difusivo de la especie queda representado como sigue:

$$\frac{\partial y}{\partial t} = \gamma \Delta y, \quad (1.5)$$

con $y = y(x, t)$, $x \in \mathbb{R}^2$ y $\gamma > 0$.

1.2 Descripción del Modelo de Fisher

Para modelar el fenómeno que describe cómo una población crece de acuerdo a la ecuación logística conforme pasa el tiempo y, a su vez, se dispersa sobre un espacio plano, combinamos las expresiones

obtenidas en las ecuaciones (1.1) y (1.5). Así, tenemos una primera forma de describir nuestro modelo:

$$\frac{\partial}{\partial t}y(x, t) - \gamma\Delta y(x, t) = ry(x, t) \left(1 - \frac{y(x, t)}{\kappa}\right), \quad (1.6)$$

con $x \in \mathbb{R}^2$ y los parámetros r, κ, γ mayores que 0.

Al término correspondiente al crecimiento poblacional se lo conoce como el término de *reacción*, por lo que la ecuación (1.6) es conocida como la *ecuación de reacción-difusión*.

El modelo de Fisher considera el crecimiento poblacional dado por una ecuación de reacción-difusión, pero también considera un control adicional del crecimiento de la población que puede ser la tasa de mortalidad natural de los individuos, la acción de recolección de los individuos al alcanzar cierta madurez o la tasa de muerte provocada de individuos.

Este modelo toma en cuenta un punto de partida; es decir, una cantidad inicial de individuos de la población en un tiempo inicial. Además, limita el problema considerando un espacio acotado y aislado sobre el cual se dispersa la población y un tiempo final, que representa el instante de tiempo en el que queremos conocer la densidad poblacional de la especie sobre los distintos puntos del espacio.

Consideramos un hábitat aislado $\Omega \subset \mathbb{R}^2$ acotado y con frontera $\partial\Omega$, un intervalo de tiempo $[0, T]$ y una densidad inicial y_0 de la población. A la ecuación (1.6) se le adiciona un término correspondiente a un control adicional del crecimiento de la población, representado por la función $u(x, t)$. Así, el crecimiento poblacional queda dado por:

$$\frac{\partial}{\partial t}y(x, t) - \gamma\Delta y(x, t) + u(x, t)y(x, t) = ry(x, t) \left(1 - \frac{y(x, t)}{\kappa}\right), \quad \text{para } (x, t) \in Q_T,$$

donde $Q_T = \Omega \times (0, T)$.

Por otro lado, la condición por la cual la población se encuentra aislada en el hábitat Ω se representa en el modelo mediante las condiciones en la frontera $\partial\Omega$, considerando que el flujo de la población sobre la misma es nula. Esta condición está dada por las condiciones de frontera tipo Neumann homogéneas:

$$\partial_\nu y = 0, \quad (x, t) \in \Sigma_T, \quad t \geq 0,$$

donde $\Sigma_T = \partial\Omega \times (0, T)$, y $\partial_\nu y = \frac{\partial y}{\partial \nu}$ es la derivada con respecto a la normal exterior en la frontera. Las condiciones de Neumann homogéneas garantizan que el flujo de la población es nulo en la frontera; es decir, que ningún individuo entra o sale del espacio geográfico determinado que está representado por Ω .

Finalmente, el modelo de Fisher queda dado por el siguiente problema de valor inicial:

$$\boxed{\begin{cases} \frac{\partial y}{\partial t} - \gamma\Delta y + u(x, t)y = ry \left(1 - \frac{y}{\kappa}\right), & \text{para } (x, t) \in Q_T, \\ \partial_\nu y = 0, & \text{para } (x, t) \in \Sigma_T, \\ y(x, 0) = y_0(x), & \text{para } x \in \Omega. \end{cases}} \quad (1.7)$$

Donde:

- $\Omega \subset \mathbb{R}^2$ es el espacio acotado sobre el cual la población crece; $x \in \Omega$ es la variable espacial.

- $\partial\Omega$ es la frontera de Ω .
- $[0, T]$ es el intervalo de tiempo considerado, $T > 0$; $t \in [0, T]$ es la variable temporal.
- $Q_T = \Omega \times (0, T)$ es el dominio del modelo, $\Sigma_T = \partial\Omega \times (0, T)$ es la frontera del mismo.
- $y = y(x, t)$; $\frac{\partial y}{\partial t}$ es la derivada de y con respecto al tiempo.
- y_0 es el estado inicial del sistema.
- r es la tasa natural de crecimiento de la especie, κ es la capacidad máxima de sostenibilidad de individuos en el espacio considerado, y ambas son consideradas constantes positivas.
- γ es la velocidad de dispersión de la especie sobre el territorio.

Más adelante, definiremos los espacios funcionales a los cuales deben pertenecer las funciones involucradas y ciertas condiciones sobre las funciones y la frontera.

El problema planteado consiste en una ecuación parabólica no lineal y lo que nos interesa es resolverlo numéricamente. Para investigar sobre el mismo, se necesita tomar en cuenta sus características particulares: para comenzar, nos ubicamos dentro del área de las ecuaciones diferenciales parciales parabólicas, por la forma de la ecuación de reacción-difusión. Luego, es necesario considerar sus condiciones de frontera de Neumann. Además, hay que subrayar que este problema tiene una semi-linealidad en el término correspondiente a la función logística, lo cual representa algunas dificultades en el desarrollo del trabajo y requiere tratar el problema con herramientas específicas de la teoría de las ecuaciones diferenciales.

Capítulo 2

Problema de Valor Inicial con Ecuaciones Diferenciales Parciales Parabólicas Semilineales

En este capítulo, discutiremos sobre la existencia, unicidad y ciertas propiedades cualitativas del problema dado por la solución de la ecuación de Fisher, sujeta a ciertas condiciones iniciales. El modelo de Fisher consiste en el siguiente problema de valor inicial:

$$\begin{cases} \frac{\partial}{\partial t}y(x, t) - \gamma\Delta y(x, t) + u(x, t)y(x, t) = ry(x, t) \left(1 - \frac{y(x, t)}{\kappa}\right), & \text{en } Q_T, \\ \partial_\nu y(x, t) = 0, & \text{sobre } \Sigma_T, \\ y(x, 0) = y_0, & \text{en } \Omega, \end{cases} \quad (2.1)$$

donde los parámetros, el estado inicial del sistema y condiciones de frontera están definidos como en el capítulo anterior. Para realizar el análisis del modelo de Fisher, consideramos la hipótesis que se presenta a continuación.

Hipótesis 1. *Hacemos las siguientes suposiciones:*

- i) $\Omega \subset \mathbb{R}^2$ es un abierto cuya frontera $\partial\Omega$ es Lipschitz.
- ii) La condición inicial y_0 es elemento de $L^\infty(\Omega)$ y verifica

$$y_0(x) \geq 0 \quad \text{c.t.p. } x \in \Omega, \quad \text{y } y_0 \not\equiv 0.$$

- iii) La función u es elemento de $C(\overline{Q_T})$ y verifica

$$u(x, t) \geq 0, \quad \forall (x, t) \in \Omega \times [0, T].$$

- iv) γ, κ y r son constantes positivas.

2.1 Propiedades A-Priori de la Solución

Las siguientes propiedades de la solución son esenciales en el estudio de existencia de soluciones del modelo. Éstas son conocidas como *estimaciones a-priori* de la solución, pues vamos a suponer la existencia de soluciones del problema (2.1) y probar la positividad y acotación de las mismas, sin haber demostrado aún que tal solución efectivamente existe. Este procedimiento es usual para demostrar la existencia de soluciones de problemas de EDP con valores iniciales utilizando el teorema del punto fijo de Banach¹; por ejemplo, se puede observar la demostración de existencia de una única solución de problemas de valor inicial que presenta Smoeller en [15], pág. 72-75.

Las demostraciones se realizarán con los métodos de comparación propuestos en [5], los cuales son consecuencia del principio del máximo. Comencemos enunciando el siguiente resultado, conocido como *principio de comparación*.

Teorema 1. *Sea $\Omega \subset \mathbb{R}^N$ un conjunto abierto, acotado y con frontera de clase C^1 . Sea L un operador elíptico definido como sigue*

$$L = \sum_{i,j=1}^N a_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x,t) \frac{\partial}{\partial x_i} + c(x,t),$$

el cual verifica la condición de elipticidad; es decir, para $\gamma_0 > 0$ se tiene

$$\sum_{i,j=1}^N a_{ij}(x,t) \xi_i \xi_j \geq \gamma_0 |\xi|^2 \quad \forall \xi \in \mathbb{R}^N.$$

Supongamos que $a_i(x,t)$, $b_i(x,t)$ son uniformemente acotados sobre $\Omega \times (0, T]$, $c(x,t) \equiv 0$, $f(x,t,y)$ y $\partial f(x,t,y)/\partial y \in C(\overline{\Omega} \times [0, T] \times \mathbb{R})$. Si $\overline{y}, \underline{y} \in C^{1,2}(\overline{\Omega} \times (0, T]) \cap C(\overline{\Omega} \times (0, T])$ con:

$$\begin{aligned} \frac{\partial}{\partial t} \overline{y}(x,t) - L\overline{y}(x,t) &\geq f(x,t, \overline{y}(x,t)) \text{ en } \Omega \times (0, T], \\ \frac{\partial}{\partial t} \underline{y}(x,t) - L\underline{y}(x,t) &\leq f(x,t, \underline{y}(x,t)) \text{ en } \Omega \times (0, T], \end{aligned} \quad (2.2)$$

$\overline{y}(x,0) \geq \underline{y}(x,0)$ sobre Ω , y se satisface una de las siguientes condiciones:

$$\begin{aligned} \overline{y}(x,t) &\geq \underline{y}(x,t), \\ \gamma(x)\overline{y}(x,t) + \beta(x)\partial_\nu \overline{y}(x,t) &\geq \gamma(x)\underline{y}(x,t) + \beta(x)\partial_\nu \underline{y}(x,t), \end{aligned} \quad (2.3)$$

sobre $\Omega \times (0, T]$. Entonces, o bien se tiene que $\overline{y} \equiv \underline{y}$, o bien $\overline{y} > \underline{y}$ sobre $\Omega \times (0, T]$.

Para comenzar, vamos a demostrar que las soluciones a este problema son no negativas, lo cual es coherente con el modelo de dispersión poblacional, pues no tendría sentido obtener un número de individuos negativo. Luego, demostraremos que las soluciones están acotadas; lo cual también es coherente, pues no es de esperarse que la población crezca indefinidamente, debido a las limitaciones del medio ambiente.

Lema 1 (Estimaciones a-priori). *Supongamos que existe una solución $y \in C^{2,1}(Q_T)$ del problema*

¹Ver Teorema 7, Anexo B.

(2.1). Entonces existe una constante $M = \|y_0\|_\infty e^{rT}$ tal que

$$0 \leq y(x, t) \leq M.$$

Demostración. Vamos a dividir esta demostración en dos, comenzando por la positividad de la solución y siguiendo con la acotación.

1. Consideremos el operador elíptico L como en la Definición 9 en el Anexo A; en referencia a nuestro problema, los coeficientes de L están dados por:

$$a_{ij}(x) = \begin{cases} \gamma & \text{si } i = j \\ 0 & \text{si no} \end{cases}, \quad b_i(x) = 0, \quad c(x) \equiv 0,$$

para $i, j = 1, \dots, n$. También definimos la función $f : \Omega \times (0, T) \times \mathbb{R} \rightarrow \mathbb{R}$ como sigue²

$$f(x, t, \xi) = r\xi - \frac{r}{\kappa}\xi^2 - u(x, t)\xi,$$

donde $u = u(x, t)$ es una función fija que verifica el literal iii) de la Hipótesis 1. También por la Hipótesis 1, tenemos que $f(x, t, y)$ y $\partial f(x, t, y)/\partial y \in C(\overline{\Omega} \times [0, T] \times \mathbb{R})$.

Es fácil verificar que 0 es solución del problema. Ponemos $\underline{y} = 0$, y suponemos que existe una solución y de (2.1) que es distinta de 0.

Tomando $\overline{y} = y$, se verifica (2.1):

$$\frac{\partial}{\partial t}\overline{y}(x, t) - L\overline{y}(x, t) \geq f(x, t, \overline{y}(x, t)),$$

y

$$\frac{\partial}{\partial t}\underline{y}(x, t) - L\underline{y}(x, t) \leq f(x, t, \underline{y}(x, t)),$$

pues \overline{y} y \underline{y} son soluciones. Además, se verifica la condición (2.3):

$$0 = \partial_\nu \overline{y}(x, t) \geq \frac{\partial}{\partial \nu} \underline{y}(x, t) = 0,$$

puesto que ambas son soluciones del sistema y verifican las condiciones de frontera.

Con todas estas condiciones satisfechas, y tomando en cuenta que habíamos supuesto que $y \not\equiv 0$, aplicamos el Teorema 1 y concluimos que necesariamente $y > 0$. Es decir, de existir una solución al problema (2.1) distinta de la trivial, ésta sería positiva.

2. Ahora definimos \overline{y} como:

$$\overline{y}(x, t) = \|y_0\|_\infty e^{rt}, \quad \text{para } (x, t) \in [0, T] \times \Omega,$$

Una vez más, verificamos la condición (2.1): por un lado,

$$\begin{aligned} \frac{\partial}{\partial t}\overline{y}(x, t) - L\overline{y}(x, t) &= \frac{\partial}{\partial t}(\|y_0\|_\infty e^{rt}) - \gamma \Delta \overline{y}(x, t) \\ &= r\|y_0\|_\infty e^{rt} - 0 \\ &= r\|y_0\|_\infty e^{rt}. \end{aligned}$$

²Para nuestro caso, por simplicidad hemos considerado el término lineal uy incluido en el término no lineal f . Esto no es estrictamente necesario, ver Corolario 1.18 pág. 79, en [5].

Por otro lado,

$$\begin{aligned} f(x, t, \bar{y}(x, t)) &= r\bar{y}(x, t) - \frac{r}{\kappa}\bar{y}^2(x, t) - u(x, t)\bar{y}(x, t). \\ &= r\|y_0\|_\infty e^{rt} - \frac{r}{\kappa}\|y_0\|_\infty^2 e^{2rt} - u(x, t)\|y_0\|_\infty e^{rt}. \end{aligned}$$

Con lo que, gracias a no negatividad de $u(x, t)$, κ y r , tenemos que $\bar{y}(x, t)$ verifica:

$$\begin{aligned} \frac{\partial}{\partial t}\bar{y}(x, t) - L\bar{y}(x, t) &= r\|y_0\|_\infty e^{rt} \\ &\geq r\|y_0\|_\infty e^{rt} - \frac{r}{\kappa}\|y_0\|_\infty^2 e^{2rt} - u(x, t)\|y_0\|_\infty e^{rt} \\ &= f(x, t, \bar{y}(x, t)). \end{aligned} \tag{2.4}$$

Nuevamente, supongamos que $\underline{y} = y$ es una solución distinta de cero del problema (2.1). Entonces $y \geq 0$ y verifica:

$$\frac{\partial}{\partial t}\underline{y}(x, t) - L\underline{y}(x, t) \leq f(x, t, \underline{y}(x, t)).$$

Además,

$$\begin{aligned} \bar{y}(x, 0) &= \|y_0\|_\infty \\ &\geq y(x, 0) = \underline{y}(x, 0). \end{aligned} \tag{2.5}$$

Con respecto a las condiciones de frontera, tenemos:

$$\partial_\nu \bar{y} = \partial_\nu \underline{y} = 0, \tag{2.6}$$

pues $\underline{y} = y$ es solución y $\partial_\nu \bar{y} = 0$ pues es constante con respecto a x .

Entonces podemos aplicar el Teorema 1 a (2.4), (2.5) y (2.6) y deducir que

$$\|y_0\|_\infty e^{rt} = \bar{y} \geq \underline{y} = y.$$

Finalmente, usando el resultado del primer literal y tomando $M = \|y_0\|_\infty e^{rT}$ concluimos que:

$$0 \leq y \leq M.$$

□

Observación.- Estas estimaciones justifican la incorporación de la Hipótesis 2, establecida más adelante, que nos servirá para demostrar que en efecto existe una solución de la Ecuación de Fisher.

Los métodos de comparación de distintas soluciones al problema de reacción-difusión permiten realizar estimaciones importantes de la solución. El siguiente lema (presentado en [4]) es una versión del principio del máximo para ecuaciones elípticas y nos servirá para realizar una nueva acotación de la solución y en los puntos de equilibrio.

Definición 1 (Punto de equilibrio). *Sea $y = y(x, t)$ una solución del problema (2.1). Se llama punto*

de equilibrio al valor de y cuando se verifica

$$\frac{\partial y}{\partial t} = 0.$$

A continuación, vamos a exponer algunas ideas que están detrás del Lema 1.16 en [4], para determinar propiedades de los puntos de equilibrio. Supongamos que y es un punto de equilibrio de (2.1); es decir, $\frac{\partial y}{\partial t} = 0$. Entonces, y verifica el siguiente problema elíptico:

$$\begin{cases} -\gamma\Delta y + u(x, t)y - ry\left(1 - \frac{y}{\kappa}\right) = 0, & \text{en } Q_T, \\ \partial_\nu y = 0, & \text{sobre } \Sigma_T, \\ y(x, 0) = y_0, & \text{en } \Omega. \end{cases} \quad (2.7)$$

Si y es al menos dos veces diferenciable sobre $\bar{\Omega}$, entonces debe alcanzar su máximo en algún punto de $\bar{\Omega}$. Llamemos x^* al punto donde se alcanza cierto valor máximo M^* .

Supongamos que $x^* \in \Omega$. Entonces para $x = x^*$, se tiene

$$\frac{\partial y}{\partial x_i} = 0 \quad \text{y} \quad \frac{\partial^2 y}{\partial x_i^2} \leq 0, \quad i = 1, 2. \quad (2.8)$$

Ahora, supongamos que $M^* > \kappa$. Recordando la positividad de las constantes r , γ , κ y de la función $u(x, t)$, y utilizando (2.7) y (2.8) tenemos

$$0 \leq -\gamma\Delta M^* = rM^* \left(1 - \frac{M^*}{\kappa}\right) - u(x, t)M^* < 0;$$

lo que es una contradicción. Con lo que se concluye que, si y es un punto de equilibrio de (2.1), entonces $y(x, t) \leq \kappa$.

Por otro lado, supongamos que $x^* \in \partial\Omega$. Entonces, si además $y \in \mathcal{C}(\bar{\Omega})$, podemos aplicar el Principio del Máximo (Teorema 11, Anexo B) y concluir que $y(x, t) \equiv M^*$. De nuevo, supongamos que $M^* > \kappa$. Dado que y es constante, se satisface (2.8) y se llega de nuevo a una contradicción, por lo que se concluye que necesariamente

$$y(x, t) \leq \kappa.$$

Observación.- El hecho de que la densidad poblacional no cambie con respecto al tiempo se representa por el equilibrio del sistema. Entonces, la importancia de este resultado radica en la coherencia con el fenómeno biológico, puesto que se demuestra que al alcanzar el equilibrio del sistema, la densidad poblacional $y(x, t)$ no puede superar el valor κ , correspondiente a la capacidad máxima que el entorno puede soportar.

2.2 Existencia y Unicidad de la Solución

La existencia de soluciones de (2.1) puede ser abordada utilizando varias técnicas, por ejemplo las técnicas de monotonía (cf. [16]) o las de semigrupos (cf. [10]). En particular, las técnicas de monotonía nos llevan a formular hipótesis poco realistas sobre los coeficientes en el contexto del Modelo de Fisher, por lo que no son adecuadas para tratar nuestro modelo, aunque matemáticamente sean aplicables.

Por esta razón, fue necesario utilizar una técnica alternativa: la aplicación del teorema del punto fijo de Banach a una formulación adecuada de la ecuación (ver [6], cap. 9) utilizando las estimaciones a-priori obtenidas.

Para comenzar, vamos a analizar ciertas características de la no linealidad del problema, dada por el término logístico. Comencemos demostrando que la función definida por

$$x \mapsto rz(x) \left(1 - \frac{z(x)}{\kappa}\right) - v(x)z(x),$$

para $z \in H^1(\Omega)$ y v fija continua sobre $\overline{\Omega}$. En efecto, podemos poner $C_v = \max_{x \in \overline{\Omega}} |v(x)|$ pues v es continua sobre $\overline{\Omega}$, y tener

$$\begin{aligned} \int_{\Omega} \left| rz(x) \left(1 - \frac{z(x)}{\kappa}\right) - v(x)z(x) \right|^2 dx &\leq r \int_{\Omega} \left| z(x) \left(1 - \frac{z(x)}{\kappa}\right) \right|^2 dx + 2rC_v \int_{\Omega} \left| (z(x))^2 \left(1 - \frac{z(x)}{\kappa}\right) \right| dx \\ &\quad + C_v^2 \int_{\Omega} |z(x)|^2 dx. \end{aligned}$$

Gracias al Teorema de Rellich-Kondrachov (ver Teorema 6, Anexo B), obtenemos que $V \hookrightarrow L^4(\Omega)$ y $V \hookrightarrow L^3(\Omega)$; es decir, $z^3 \in L^1(\Omega)$ y $z^4 \in L^1(\Omega)$. Además, sabemos que $z \in L^2(\Omega)$ y $z \geq 0$, con lo cual tenemos

$$\begin{aligned} \int_{\Omega} \left| rz(x) \left(1 - \frac{z(x)}{\kappa}\right) - v(x)z(x) \right|^2 dx &\leq r \int_{\Omega} \left| z(x) - \frac{(z(x))^2}{\kappa} \right|^2 dx + 2rC_v \int_{\Omega} \left| (z(x))^2 - \frac{(z(x))^4}{\kappa} \right| dx \\ &\quad + C_v^2 \int_{\Omega} |z(x)|^2 dx. \\ &= r \int_{\Omega} \left(|(z(x))^2| - \frac{2|(z(x))^3|}{\kappa} + \frac{|(z(x))^4|}{\kappa} \right) dx \\ &\quad + 2rC_v \int_{\Omega} \left(|(z(x))^2| + \left| \frac{(z(x))^4}{\kappa} \right| \right) dx + C_v^2 \int_{\Omega} |z(x)|^2 dx \\ &= (r + 2rC_v + C_v^2) \int_{\Omega} |z(x)|^2 dx - \frac{2r}{\kappa} \int_{\Omega} |(z(x))^3| dx \\ &\quad + \frac{r(1 + 2C_v)}{\kappa} \int_{\Omega} |(z(x))^4| dx \\ &< \infty. \end{aligned}$$

Gracias a este argumento, podemos establecer el siguiente lema.

Lema 2. *Considérese el operador de Nemytskii $f : H^1(\Omega) \rightarrow L^2(\Omega)$ definido por*

$$z(\cdot) \mapsto f(z(\cdot)) = rz(\cdot) \left(1 - \frac{z(\cdot)}{\kappa}\right) - v(\cdot)z(\cdot),$$

donde v es una función continua en Ω dada. Entonces se tiene la siguiente condición tipo Lipschitz: para dos elementos $z = z(\cdot, t), \tilde{z} = \tilde{z}(\cdot, t)$ de $H^1(\Omega)$ que verifican $0 \leq z, \tilde{z} \leq M$, con M constante

positiva, existe una constante $C > 0$ tal que

$$\|f(z) - f(\tilde{z})\|_{L^2(\Omega)}^2 \leq C \|z - \tilde{z}\|_{L^2(\Omega)}^2 \quad \text{c.t.p. } 0 \leq t \leq T. \quad (2.9)$$

Demostración. Por simplicidad de la notación, omitimos la dependencia de z y \tilde{z} con respecto al tiempo. Un cálculo directo nos da q como resultado que

$$\begin{aligned} \|f(z) - f(\tilde{z})\|_{L^2(\Omega)}^2 &= \int_{\Omega} (f(z(x)) - f(\tilde{z}(x)))^2 dx \\ &= \int_{\Omega} \left(rz(x) - \frac{r}{\kappa} (z(x))^2 - v(x)z(x) - rz(\tilde{x}) + \frac{r}{\kappa} \tilde{z}^2(x) + v(x)\tilde{z}(x) \right)^2 dx \\ &= \int_{\Omega} \left((r - v(x))(z(x) - \tilde{z}(x)) - \frac{r}{\kappa} ((z(x))^2 - \tilde{z}(x)^2) \right)^2 dx \\ &= \int_{\Omega} \left((r - v(x)) - \frac{r}{\kappa} (z(x) + \tilde{z}(x)) \right)^2 (z(x) - \tilde{z}(x))^2 dx. \end{aligned} \quad (2.10)$$

Pero

$$\begin{aligned} \left((r - v(x)) - \frac{r}{\kappa} (z(x) + \tilde{z}(x)) \right)^2 &\leq \left| (r - v(x)) - \frac{r}{\kappa} (z(x) + \tilde{z}(x)) \right|^2 \\ &\leq \left| (r + v(x)) + \frac{r}{\kappa} (z(x) + \tilde{z}(x)) \right|^2 \\ &= |r + v(x)|^2 + \frac{2r}{\kappa} |r + v(x)| |z(x) + \tilde{z}(x)| + \frac{r^2}{\kappa^2} |z(x) + \tilde{z}(x)|^2 \\ &\leq |r + v(x)|^2 + \frac{4Mr}{\kappa} |r + v(x)| + \frac{4M^2r^2}{\kappa^2} \\ &\leq \left(1 + \frac{4Mr}{\kappa} \right) |r + v(x)|^2 + \frac{4M^2r^2}{\kappa^2}. \end{aligned} \quad (2.11)$$

Reemplazando (2.11) en (2.10), resulta que

$$\begin{aligned} \|f(z) - f(\tilde{z})\|_{L^2(\Omega)}^2 &\leq \int_{\Omega} \left(\left(1 + \frac{4Mr}{\kappa} \right) |r + v(x)|^2 + \frac{4M^2r^2}{\kappa^2} \right) (z(x) - \tilde{z}(x))^2 dx \\ &\leq \int_{\Omega} \left(\left(1 + \frac{4Mr}{\kappa} \right) |r + v(x)|^2 + \frac{4M^2r^2}{\kappa^2} \right) |z(x) - \tilde{z}(x)|^2 dx. \end{aligned}$$

Tomando el máximo de $|r + v(x)|^2$ sobre Q_T , tenemos finalmente que

$$\begin{aligned} \|f(z) - f(\tilde{z})\|_{L^2(\Omega)}^2 &\leq \int_{\Omega} \left(\left(1 + \frac{4Mr}{\kappa} \right) \max_{(x,t) \in Q_T} |r + v(x)|^2 + \frac{4M^2r^2}{\kappa^2} \right) |z(x) - \tilde{z}(x)|^2 dx \\ &\leq C \int_{\Omega} |z(x) - \tilde{z}(x)|^2 dx \\ &= C \|z - \tilde{z}\|_{L^2(\Omega)}^2, \end{aligned}$$

para cierta $C > 0$. Con lo que se verifica (2.9) y queda demostrado el lema. \square

A continuación, presentamos algunas definiciones y resultados que serán necesarios para la demostración de la existencia de una única solución débil de (2.1).

Notación 1. Para facilitar la lectura de algunos de los resultados a lo largo del documento, usaremos paralelamente la siguiente notación:

$$y_t = \frac{\partial y}{\partial t},$$

para la derivada parcial con respecto al tiempo de la función y .

Definición 2. En lo que sigue del documento, vamos a notar por V al espacio $H^1(\Omega)$, con la norma usual

$$\|y\|_V = \left(\int_{\Omega} |y(x)|^2 + |\nabla y(x)|^2 dx \right)^{1/2}.$$

El espacio V^* corresponde al espacio dual de V y está dotado de la norma:

$$\|v\|_{V^*} = \sup_{y \in V, y \neq 0} \frac{|\langle v, y \rangle_{V^*, V}|}{\|y\|_V}.$$

Para resolver nuestro problema es necesario introducir el concepto de los espacios $L^2(a, b; X)$, los cuales nos permiten tratar con funciones que asignan a cada elemento del intervalo de tiempo $[a, b]$ un elemento del espacio de Banach X . Así, el espacio $L^p(a, b; X)$ con $1 \leq p < \infty$ está definido como el espacio de todas las funciones vectoriales $y : [a, b] \rightarrow X$, tales que

$$\|y\|_{L^p(a, b; X)} = \left(\int_a^b \|y\|_X dt \right)^{1/p} < \infty.$$

En particular, definimos el siguiente espacio considerando $X = V$, $a = 0$ y $b = T$.

Definición 3 (Espacio $W(0, T)$). Se nota por $W(0, T)$ al espacio de las funciones $y \in L^2(0, T; V)$ cuyas derivadas y_t pertenecen al espacio $L^2(0, T; V^*)$. El espacio $W(0, T)$ está dotado de la norma:

$$\|y\|_{W(0, T)} = \left(\int_0^T \|y(t)\|_V^2 + \|y_t(t)\|_{V^*}^2 dt \right)^{1/2};$$

mientras que el espacio $L^2(0, T; V^*)$ está dotado de la norma:

$$\|v\|_{L^2(0, T; V^*)} = \left(\int_0^T \|v\|_{V^*}^2 dt \right)^{1/2}.$$

Definición 4 (Solución Débil). Se dice que una función $y \in W(0, T)$ es una solución débil de (2.1) si verifica

$$\langle y_t, \phi \rangle + B[y, \phi] = (f(y), \phi) \quad \forall \phi \in V,$$

c.t.p. $0 \leq t \leq T$, con $y(0) = y_0$, donde:

$\langle \cdot, \cdot \rangle$ denota el producto en dualidad entre V^* y V ;

(\cdot, \cdot) denota el producto interno en $L^2(\Omega)$, y

$B[\cdot, \cdot]$ es la forma bilineal de $V \times V \rightarrow \mathbb{R}$ asociada con el Laplaciano $-\Delta$ en V definida por

$$B[u, v] = \gamma(\nabla u, \nabla v), \quad \text{para todo } (u, v) \in V \times V.$$

Hipótesis 2. Considerando el problema de valor inicial (2.1), suponemos que

$$0 \leq y \leq M,$$

para una constante $M > 0$.

A pesar de ser un problema ampliamente conocido, no encontramos en la literatura correspondiente una referencia precisa de la demostración de existencia y unicidad para nuestro caso. A continuación presentamos una demostración de la existencia y unicidad de una solución débil para el problema de Fisher, siguiendo las líneas de la demostración de [6] (Teorema 9.2.2, Pág. 500) para la existencia de soluciones débiles de problemas de valor inicial parabólicos no lineales con condiciones de Dirichlet, adaptando la demostración considerando condiciones de frontera de Neumann y una no linealidad localmente Lipschitz.

Teorema 2 (Existencia y unicidad). *Bajo las Hipótesis 1 y 2, existe una única solución débil para el problema de valor inicial dado por (2.1).*

Demostración. Esta demostración se desarrolla en cuatro partes: en la primera, consideramos un problema auxiliar cuya solución nos permite asociar un operador a nuestro problema. En la segunda, demostramos que este operador es una contracción. En la tercera, aplicamos el teorema del punto fijo de Banach (Teorema 7 del Anexo B) para demostrar la existencia de soluciones débiles. Finalmente, en la cuarta parte, demostramos la unicidad de la solución utilizando la desigualdad de Gronwall (Teorema 8 del Anexo B).

1. Definimos el espacio

$$X = C([0, T]; L^2(\Omega)),$$

dotado con la norma

$$\|y\|_X = \max_{0 \leq t \leq T} \|y(t)\|_{L^2(\Omega)};$$

donde $y(t) = y(\cdot, t)$. Recordemos que $W(0, T) \hookrightarrow X$, ver [16], Teorema 3.10, Pág. 148. Fijamos una función $y \in X$ y definimos

$$h_y(t) := f(y(\cdot, t)), \quad \text{para } t \in [0, T].$$

Gracias a que f verifica (2.9) y que, de la Hipótesis 2, sabemos que y está acotada, resulta que $h_y \in L^2(0, T; L^2(\Omega))$.

Planteamos el siguiente problema auxiliar:

$$\begin{cases} \frac{\partial}{\partial t} z(x, t) - \gamma \Delta z(x, t) = h_y(x, t), & \text{en } Q_T, \\ \partial_\nu z(x, t) = 0, & \text{sobre } \Sigma_T, \\ z(x, 0) = y_0, & \text{en } \Omega. \end{cases} \quad (2.12)$$

Nótese que (2.12) es un problema parabólico lineal que tiene una única solución débil $z \in W(0, T)$; ver [16], Teorema 3.9, pág. 140.

Así, z verifica:

$$\langle z_t(t), \phi \rangle + B[z(t), \phi] = (h_y(t), \phi), \quad \forall \phi \in V$$

c.t.p. $0 \leq t \leq T$.

Podemos definir entonces un operador $A : X \rightarrow X$ que asigna a cada y fijo la solución z de (2.12) correspondiente:

$$A[y] = z.$$

2. Ahora vamos a demostrar que si T es suficientemente pequeño, entonces A es una contracción. Fijamos $y, \tilde{y} \in L^2(\Omega)$ y definimos $z = A[y]$, $\tilde{z} = A[\tilde{y}]$, con $h_{\tilde{y}} = f(\tilde{y})$. De esta manera, z y \tilde{z}

pertencen a $W(0, T)$, son soluciones de (2.12) y verifican:

$$\langle z_t(t), \phi \rangle + B[z(t), \phi] = (h_y(t), \phi), \quad \forall \phi \in V,$$

y

$$\langle \tilde{z}_t(t), \phi \rangle + B[\tilde{z}(t), \phi] = (h_{\tilde{y}}(t), \phi), \quad \forall \phi \in V,$$

respectivamente, c.t.p $0 < t \leq T$ y verifican las condiciones de frontera. Restando ambas ecuaciones, tenemos:

$$\langle z_t(t) - \tilde{z}_t(t), \phi \rangle + B[z(t) - \tilde{z}(t), \phi] = (h_y(t) - h_{\tilde{y}}(t), \phi), \quad (2.13)$$

c.t.p. $0 < t \leq T$ y para cada $\phi \in V$.

En particular, tomamos $\phi = z(t) - \tilde{z}(t)$; con lo que (2.13) queda como sigue:

$$\langle z_t(t) - \tilde{z}_t(t), z(t) - \tilde{z}(t) \rangle + B[z(t) - \tilde{z}(t), z(t) - \tilde{z}(t)] = (h_y(t) - h_{\tilde{y}}(t), z(t) - \tilde{z}(t)), \quad (2.14)$$

c.t.p. $0 < t \leq T$. Por un lado, tenemos que

$$\begin{aligned} \langle z_t(t) - \tilde{z}_t(t), z - \tilde{z}(t) \rangle &= \int_{\Omega} \left[\frac{\partial}{\partial t} (z(t) - \tilde{z}(t)) \right] (z(t) - \tilde{z}(t)) dx \\ &= \int_{\Omega} \frac{\partial}{\partial t} \left[\frac{(z(t) - \tilde{z}(t))^2}{2} \right] dx \\ &= \frac{1}{2} \frac{\partial}{\partial t} \int_{\Omega} (z(t) - \tilde{z}(t))^2 dx \\ &= \frac{1}{2} \frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2. \end{aligned} \quad (2.15)$$

Por otro lado,

$$\begin{aligned} B[z(t) - \tilde{z}(t), z(t) - \tilde{z}(t)] &= \gamma (\nabla(z(t) - \tilde{z}(t)), \nabla(z(t) - \tilde{z}(t))) \\ &= \gamma \|\nabla(z(t) - \tilde{z}(t))\|_{L^2(\Omega)}^2. \end{aligned} \quad (2.16)$$

Reemplazando (2.15) y (2.16) en (2.14) nos queda:

$$\frac{1}{2} \frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2 + \gamma \|\nabla(z(t) - \tilde{z}(t))\|_{L^2(\Omega)}^2 = (h_y(t) - h_{\tilde{y}}(t), z(t) - \tilde{z}(t)) \quad \text{c.t.p. } 0 < t \leq T;$$

lo que es equivalente a

$$\frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2 + 2\gamma \|\nabla(z(t) - \tilde{z}(t))\|_{L^2(\Omega)}^2 = 2(h_y(t) - h_{\tilde{y}}(t), z(t) - \tilde{z}(t)) \quad \text{c.t.p. } 0 < t \leq T.$$

Utilizando la positividad del término $2\gamma \|\nabla(z(t) - \tilde{z}(t))\|_{L^2(\Omega)}^2$ y la desigualdad de Cauchy-Schwarz, se tiene que

$$\begin{aligned} \frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2 &\leq 2(h_y(t) - h_{\tilde{y}}(t), z(t) - \tilde{z}(t)) \\ &\leq 2 \|h_y(t) - h_{\tilde{y}}(t)\|_{L^2(\Omega)} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}. \end{aligned}$$

La desigualdad de Young (ver anexo B, teorema 10) implica la siguiente desigualdad:

$$2 \|h_y(t) - h_{\tilde{y}}(t)\|_{L^2(\Omega)} \|z(t) - \tilde{y}(t)\|_{L^2(\Omega)} \leq 2\frac{1}{\varepsilon} \|h_y(t) - h_{\tilde{y}}(t)\|_{L^2(\Omega)}^2 + 2\varepsilon \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2,$$

para un $\varepsilon > 0$ arbitrario.

Si tomamos ε suficientemente pequeño, se puede garantizar que

$$\frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2 \leq C \|h_y(t) - h_{\tilde{y}}(t)\|_{L^2(\Omega)}^2,$$

para cierta constante $C > 0$. En lo que sigue de la demostración, C representará una constante genérica independiente de y y de T , que representa a todas las constantes que se obtienen a lo largo de los cálculos realizados.

Tomando en cuenta las estimaciones a-priori demostradas en el Lema 2, concluimos que

$$\frac{\partial}{\partial t} \|z(t) - \tilde{z}(t)\|_{L^2(\Omega)}^2 \leq C \|y - \tilde{y}\|_{L^2(\Omega)}^2, \quad \text{c.t.p. } 0 \leq t \leq T.$$

Ahora, integrando con respecto a t ambas partes de la ecuación, obtenemos para todo $s \in [0, T]$:

$$\|z(s) - \tilde{z}(t)(s)\|_{L^2(\Omega)}^2 \leq C \int_0^s \|y(t) - \tilde{y}(t)\|_{L^2(\Omega)}^2 dt. \quad (2.17)$$

Maximizando el lado derecho con respecto a t , tenemos:

$$\begin{aligned} \|z(s) - \tilde{z}(s)\|_{L^2(\Omega)}^2 &\leq C \int_0^s \max_{0 \leq t \leq T} \|y(t) - \tilde{y}(t)\|_{L^2(\Omega)}^2 dt \\ &= C \|y - \tilde{y}\|_X^2 \int_0^s dt \\ &= Cs \|y - \tilde{y}\|_X^2. \end{aligned} \quad (2.18)$$

Tomando el máximo con respecto a s en (2.18), obtenemos

$$\|z - \tilde{z}\|_X^2 \leq CT \|y - \tilde{y}\|_X^2.$$

Es decir,

$$\|A[y] - A[\tilde{y}]\|_X \leq CT^{\frac{1}{2}} \|y - \tilde{y}\|_X. \quad (2.19)$$

Por lo tanto, A es una contracción siempre y cuando $T > 0$ sea lo suficientemente pequeño para que $CT^{\frac{1}{2}} < 1$.

3. Vamos a demostrar la existencia de la solución aplicando el teorema del punto fijo a (2.19), en el espacio X . Dado que C es independiente de T , tomamos un $T > 0$ cualquiera y elijamos un T_1 tal que verifique.

$$\sqrt{CT_1} < 1.$$

Entonces podemos aplicar el teorema del punto fijo de Banach (Teorema 7 del Anexo B) al operador A y afirmar que el operador A tiene un único punto fijo. Es decir, existe un y tal que

$$A[y] = y;$$

y que pertenece a $W(0, T)$, gracias a la definición de A .

Hemos demostrado que existe una solución débil del problema (2.1) en el intervalo $[0, T_1]$. La solución débil $y(t)$ pertenece a V en casi todo punto $0 \leq t \leq T_1$, de manera que, de ser necesario, podemos redefinir T_1 para poder asumir que $y(T_1) \in V$ en el intervalo $[0, T_1]$.

Repitiendo este proceso, se puede extender nuestra solución al intervalo $[T_1, 2T_1]$ y así, sucesivamente, para un cierto número finito de intervalos disjuntos que cubren $[0, T]$. De esta forma, podemos finalmente obtener una solución débil sobre todo el intervalo $[0, T]$, i.e. la solución pertenece a X .

4. Finalmente, vamos a demostrar la unicidad de la solución. Para ello, vamos a suponer que tanto y como \tilde{y} son dos soluciones débiles distintas del problema. Entonces resulta que la ecuación (2.17) se verifica para $z = y$ y $\tilde{z} = \tilde{y}$, sobre cierto subintervalo de $[0, T]$. Si consideramos todos los subintervalos obtenidos en el paso 3, podemos afirmar que

$$\|y(s) - \tilde{y}(s)\|_{L^2(\Omega)}^2 \leq C \int_0^s \|y(t) - \tilde{y}(t)\|_{L^2(\Omega)}^2 dt$$

para $0 \leq s \leq T$.

Aplicando la desigualdad de Gronwall a la expresión anterior, obtenemos que

$$\|y(s) - \tilde{y}(s)\|_{L^2(\Omega)}^2 = 0,$$

de donde concluimos finalmente que

$$y = \tilde{y}.$$

□

2.3 Regularidad de la Solución

Finalizaremos este capítulo con una breve discusión sobre la regularidad extra de la solución débil de (2.1). Para la aplicación de la solución del problema, es de interés conocer de la existencia de soluciones clásicas, por lo que ahora vamos a enunciar un resultado sobre la regularidad de la solución correspondiente al problema de valor inicial de naturaleza parabólica y lineal que se presentan en [16].

Teorema 3 (Regularidad de la Solución). *Sea $\Omega \subset \mathbb{R}^N$ un dominio de Lipschitz acotado y sean $y_0 \in C(\overline{\Omega})$ y $h \in L^s(\Omega)$ dos funciones dadas; donde $s > N/2 + 1$. Entonces, la solución débil de un problema parabólico lineal con valor inicial y condiciones de frontera tipo Neumann dado por*

$$\begin{cases} \frac{\partial}{\partial t} y(x, t) - \gamma \Delta y(x, t) = h(t), & \text{en } Q_T, \\ \partial_\nu y(x, t) = 0, & \text{sobre } \Sigma_T, \\ y(0) = y_0, & \text{en } \Omega, \end{cases} \quad (2.20)$$

pertenece a $W(0, T) \cap C(\overline{Q_T})$. Además, existe una constante $C_s > 0$ que depende de s pero es independiente de y_0 y h , tal que

$$\|z\|_{W(0, T)} + \|z\|_{C(\overline{Q_T})} \leq C_s \left(\|h\|_{L^s(Q_T)} + \|y_0\|_{C(\overline{\Omega})} \right). \quad (2.21)$$

Consideramos $h = h_y$ como en la demostración del Teorema 2; de manera análoga a esa demostración, fijamos y solución del problema (2.1) y tomamos $z = A[y]$. Por otro lado, gracias al Teorema 9 el Anexo B, se tiene que $h_y \in L^\infty(\Omega)$.

Entonces podemos aplicar el Teorema 3 y concluir que $z \in W(0, T) \cap C(\overline{Q_T})$ y que existe una constante $C > 0$ que verifica

$$\|z\|_{W(0, T)} + \|z\|_{C(\overline{Q_T})} \leq C \left(\|h_y\|_{L^2(Q_T)} + \|y_0\|_{C(\overline{\Omega})} \right).$$

Esto se verifica en un subintervalo $[0, T_i]$ suficientemente pequeño tal que A sea una contracción. Es necesario extender esta propiedad en todo el intervalo aplicando el teorema sobre un número finito de subintervalos de la forma $[T_i, T_{i+1}]$ y construir así una solución y del problema (2.1) que verifica (2.21) sobre todo el intervalo $[0, T]$.

Capítulo 3

Resolución Numérica de la Ecuación de Fisher

Este capítulo está dedicado a la resolución numérica de la ecuación de Fisher. Para diseñar un esquema de resolución numérica, empezamos por establecer un esquema iterativo linealizado mediante la aplicación del método de Newton; luego, utilizaremos el método conocido de semidiscretización sobre la formulación variacional de la ecuación linealizada de Fisher aplicando el método de Galerkin para las variables espaciales, de donde se obtiene un sistema de ecuaciones diferenciales ordinarias (EDO). Para la construcción del subespacio discreto del método de Galerkin, utilizaremos el método de los Elementos Finitos. Mientras que, para la resolución del sistema de EDO resultante, emplearemos dos esquemas de resolución: el método de Euler Implícito (EI) y un esquema de Crank-Nicholson estabilizado (CN). En particular, el método de Euler Implícito servirá para la estabilización del método de Crank-Nicholson. Desarrollaremos ambos esquemas en detalle y formularemos los sistemas lineales respectivos. Al final del capítulo discutiremos sobre los resultados numéricos obtenidos en diversos experimentos.

3.1 Aproximación por el Método de Galerkin y de los Elementos Finitos

La ecuación de Fisher es un problema de Ecuaciones Diferenciales Parciales Parabólicas. En esta sección, vamos a aproximar la solución de este problema en base al método de los Elementos Finitos y la aproximación de Galerkin, atacando la semilinealidad de la ecuación aplicando el método de Newton al problema.

3.1.1 Linealización de la Ecuación de Fisher

Recordemos el problema de valor inicial:

$$\begin{cases} y_t - \gamma \Delta y + uy = ry \left(1 - \frac{y}{\kappa}\right), & \text{en } Q_T, \\ \partial_\nu y = 0, & \text{sobre } \Sigma_T, \\ y(x, 0) = y_0, & \text{en } \Omega, \end{cases} \quad (3.1)$$

donde $y = y(x, t)$, $u = u(x, t)$ y los parámetros γ , r y κ como se especificó en los capítulos previos. Recordemos también la notación $V = H^1(\Omega)$ y V^* para su espacio dual (cf. sección 2.2). Reorganizando los términos de (3.1), tenemos la siguiente ecuación:

$$\frac{\partial}{\partial t}y(x, t) - \gamma\Delta y(x, t) + u(x, t)y(x, t) - ry(x, t) + \frac{1}{\kappa}(y(x, t))^2 = 0 \quad (3.2)$$

Asociamos un funcional F al lado izquierdo de (3.2) para poder encontrar una expresión que linealice el problema (3.1). Los cálculos que se realizan a continuación son cálculos formales, sin detenernos ante detalles teóricos puesto que buscamos solamente exponer la idea de formular variacionalmente el problema, y luego aplicar el método de Newton para linealizarlo. Procedemos asignando a cada función $y \in W(0, T)$ un valor $F(y)(\cdot)$ de la siguiente manera:

$$y \mapsto F(y)(t) = \frac{\partial}{\partial t}y(\cdot, t) - \gamma\Delta y(\cdot, t) + u(\cdot, t)y(\cdot, t) - ry(\cdot, t) + \frac{y(\cdot, t)^2}{\kappa}.$$

Por tanto, definimos $F : W(0, T) \rightarrow L^2(0, T; V^*)$ como sigue: para $\phi \in V$,

$$\langle F(y), \phi \rangle_{V^*, V} = \langle y_t, \phi \rangle_{V^*, V} - \gamma \langle \Delta y, \phi \rangle_{V^*, V} + (uy, \phi)_{L^2(\Omega)} - r(y, \phi)_{L^2(\Omega)} + \frac{r}{\kappa} (y^2, \phi)_{L^2(\Omega)}.$$

Por simplicidad en la notación, hemos omitido la dependencia con respecto a t de la expresión anterior y en lo que queda del capítulo, entenderemos (\cdot, \cdot) como el producto interno en el espacio $L^2(\Omega)$.

Aplicando integración por partes al término $-\gamma \langle \Delta y, \phi \rangle_{V^*, V}$ y utilizando las condiciones de frontera, resulta que

$$\langle F(y), \phi \rangle_{V^*, V} = \langle y_t, \phi \rangle_{V^*, V} + \gamma(\nabla y, \nabla \phi) + (uy, \phi) - r(y, \phi) + \frac{r}{\kappa} (y^2, \phi), \quad (3.3)$$

para todo $\phi \in V$; c.t.p. $t \in (0, T)$. La ecuación

$$\langle F(y), \phi \rangle_{V^*, V} = 0$$

corresponde a la formulación variacional del problema (3.1), el cual tiene una solución débil en $W(0, T)$, como demostramos en la Sección 2.2.

Éste resulta ser un problema no lineal, al cual aplicaremos el método de Newton para establecer un esquema linealizado del mismo, para poder aproximar la solución del problema original. Dado que el método de Newton requiere el cálculo de la derivada de F , planteamos el siguiente lema.

Lema 3. *Sea $F : W(0, T) \rightarrow L^2(0, T; V^*)$ definida como en (3.3). Entonces la derivada de F en y , $F'(y) : W(0, T) \rightarrow L^2(0, T; V^*)$, está dada por:*

$$\langle F'(y)z, \phi \rangle_{V^*, V} = \langle z_t, \phi \rangle_{V^*, V} + \gamma(\nabla z, \nabla \phi) + (uz, \phi) - r(z, \phi) + \frac{2r}{\kappa}(yz, \phi), \quad (3.4)$$

para todo $\phi \in V$, $z \in W(0, T)$ y c.t.p. $t \in (0, T)$.

Demostración. Tomemos como candidata de la derivada de Fréchet de F a su derivada direccional. Sea z una función en $W(0, T)$, entonces debemos calcular

$$\lim_{\tau \rightarrow 0} \frac{\langle F(y + \tau z) - F(y), \phi \rangle_{V^*, V}}{\tau},$$

Dado que los primeros términos son operadores lineales y continuos, nos concentraremos en los cálculos correspondientes al término cuadrático del lado derecho de (3.3). A lo largo del siguiente cálculo, τ representa una constante que nos sirve para calcular la derivada direccional y que no está relacionada a la variable temporal. Tenemos entonces que:

$$\begin{aligned} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left[\frac{r}{\kappa} ((y + \tau z)^2, \phi) - \frac{r}{\kappa} (y^2, \phi) \right] &= \frac{r}{\kappa} \lim_{\tau \rightarrow 0} \frac{1}{\tau} [(y^2 + 2\tau yz + \tau^2 z^2, \phi) - (y^2, \phi)] \\ &= \frac{r}{\kappa} \lim_{\tau \rightarrow 0} \frac{1}{\tau} [(y^2, \phi) + (2\tau yz, \phi) + (\tau^2 z^2, \phi) - (y^2, \phi)] \\ &= \frac{r}{\kappa} \lim_{\tau \rightarrow 0} [(2yz, \phi) + (\tau z^2, \phi)] \\ &= \frac{2r}{\kappa} (yz, \phi). \end{aligned}$$

Luego, gracias a la linealidad con respecto a y de los primeros 4 términos de (3.3), tenemos como candidata a derivada de Fréchet, la derivada direccional de F :

$$\langle F'(y)z, \phi \rangle_{V^*, V^*} = \langle z_t, \phi \rangle_{V^*, V^*} + \gamma(\nabla z, \nabla \phi) + (uz, \phi) - r(z, \phi) + \frac{2r}{\kappa} (yz, \phi).$$

Ahora, es necesario demostrar que $F'(y) : W(0, T) \rightarrow L^2(0, T; V^*)$ es lineal y continua. En efecto, sean $z_1, z_2 \in W(0, T)$, y $\alpha \in \mathbb{R}$, se sigue que

$$\begin{aligned} \langle F'(y)(z_1 + \alpha z_2), \phi \rangle_{V^*, V} &= \left\langle \frac{\partial}{\partial t} (z_1 + \alpha z_2), \phi \right\rangle_{V^*, V} + \gamma(\nabla(z_1 + \alpha z_2), \nabla \phi) + (u(z_1 + \alpha z_2), \phi) - r(z_1 + \alpha z_2, \phi) \\ &\quad + \frac{2r}{\kappa} (y(z_1 + \alpha z_2), \phi) \\ &= \langle F'(y)z_1, \phi \rangle_{V^*, V} + \alpha \langle F'(y)z_2, \phi \rangle_{V^*, V}, \end{aligned} \quad (3.5)$$

con lo que $F'(y)$ es lineal. Para demostrar la continuidad, es necesario demostrar que $F'(y)$ es acotado, es decir:

$$\|F'(y)z\|_{L^2(0, T; V^*)} \leq C \|z\|_{W(0, T)}, \quad \forall z \in W(0, T).$$

En efecto, tomemos $z \in W(0, t)$. Entonces tenemos

$$\|F'(y)z\|_{L^2(0, T; V^*)} = \left(\int_0^T \left(\sup_{\substack{\phi \in V \\ \phi \neq 0}} \frac{|\langle F'(y)z, \phi \rangle_{V^*, V}|}{\|\phi\|_V} \right)^2 dt \right)^{1/2}. \quad (3.6)$$

Si consideramos el integrando de (3.6), tenemos que

$$\begin{aligned} \left| \langle F'(y)z, \phi \rangle_{V^*, V} \right| &= \left| \langle z_t, \phi \rangle_{V^*, V} + \gamma(\nabla z, \nabla \phi) + (uz, \phi) - r(z, \phi) + \frac{2r}{\kappa} (yz, \phi) \right| \\ &\leq |\langle z_t, \phi \rangle_{V^*, V}| + \gamma |(\nabla z, \nabla \phi)| + |(uz, \phi)| + r |(z, \phi)| + \frac{2r}{\kappa} |(yz, \phi)|. \end{aligned}$$

Por la definición de la norma en el espacio V^* en la Definición 2, tenemos que

$$\left| \langle z_t, \phi \rangle_{V^*, V} \right| \leq \|z_t\|_{V^*} \|\phi\|_V, \quad \forall \phi \in V.$$

Nótese que los términos que incluyen a z_t , z , y y u en la expresión anterior dependen de t , por lo que podemos aplicar la desigualdad de Cauchy-Shwarz (cf. Anexo B) y obtener:

$$\begin{aligned} \left| \langle F'(y)z, \phi \rangle_{V^*, V} \right| &\leq \|z_t\|_{V^*} \|\phi\|_V + \gamma \|\nabla z\|_{L^2(\Omega)} \|\nabla \phi\|_{L^2(\Omega)} \\ &\quad + \|uz\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} + r \|z\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)} + \frac{2r}{\kappa} \|yz\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}. \end{aligned} \quad (3.7)$$

Reemplazamos (3.7) en (3.6) y, gracias a que V está continuamente inyectado en $L^2(\Omega)$ (ver Anexo B), sabemos que $\|\phi\|_{L^2(\Omega)} \leq \|\phi\|_V$. Además, $\|\nabla v\|_{L^2(\Omega)} \leq \|v\|_V, \forall v \in V$ por la definición de la norma de V ; con lo que tenemos:

$$\begin{aligned} \|F'(y)z\|_{L^2(0,T;V^*)}^2 &\leq \int_0^T \left(\|z_t\|_{V^*} + \gamma \|\nabla z\|_{L^2(\Omega)} + \|uz\|_{L^2(\Omega)} + r \|z\|_{L^2(\Omega)} + \frac{2r}{\kappa} \|yz\|_{L^2(\Omega)} \right)^2 dt \\ &\leq \int_0^T \left(\|z_t\|_{V^*} + \gamma \|z\|_V + \|uz\|_{L^2(\Omega)} + r \|z\|_{L^2(\Omega)} + \frac{2r}{\kappa} \|yz\|_{L^2(\Omega)} \right)^2 dt. \end{aligned} \quad (3.8)$$

El teorema de Hölder (Teorema 5, Anexo B) nos permite obtener las siguientes desigualdades:

$$\|uz\|_{L^2(\Omega)} = \int_{\Omega} |uz| dx \leq \|u\|_{L^2(\Omega)} \|z\|_{L^2(\Omega)},$$

y

$$\|yz\|_{L^2(\Omega)} = \int_{\Omega} |yz| dx \leq \|y\|_{L^2(\Omega)} \|z\|_{L^2(\Omega)}.$$

Con ellas, (3.8) queda como sigue:

$$\|F'(y)z\|_{L^2(0,T;V^*)}^2 \leq \int_0^T \left(\|z_t\|_{V^*} + \gamma \|z\|_V + \|u\|_{L^2(\Omega)} \|z\|_{L^2(\Omega)} + r \|z\|_{L^2(\Omega)} + \frac{2r}{\kappa} \|y\|_{L^2(\Omega)} \|z\|_{L^2(\Omega)} \right)^2 dt.$$

Usando de nuevo la inyección de V en $L^2(\Omega)$, resulta que

$$\begin{aligned} \|F'(y)z\|_{L^2(0,T;V^*)}^2 &\leq \int_0^T \left(\|z_t\|_{V^*} + \gamma \|z\|_V + \|u\|_{L^2(\Omega)} \|z\|_V + r \|z\|_V + \frac{2r}{\kappa} \|y\|_{L^2(\Omega)} \|z\|_V \right)^2 dt \\ &= \int_0^T \left(\|z_t\|_{V^*} + \left(\gamma + \|u\|_{L^2(\Omega)} + r + \frac{2r}{\kappa} \|y\|_{L^2(\Omega)} \right) \|z\|_V \right)^2 dt. \end{aligned} \quad (3.9)$$

Dado que $u \in \mathcal{C}(\overline{Q_T})$ y $y \in W(0, T) \hookrightarrow \mathcal{C}(0, T; L^2(\Omega))$, podemos definir

$$C_y = \max \left\{ 1; \left(\gamma + \max_t \|u\|_{L^2(\Omega)} + r + \frac{2r}{\kappa} \max_t \|y\|_{L^2(\Omega)} \right) \right\}.$$

Luego, de (3.9) se tiene:

$$\begin{aligned} \|F'(y)z\|_{L^2(0,T;V^*)}^2 &\leq C_y \int_0^T (\|z_t\|_{V^*} + \|z\|_V)^2 dt, \\ &\leq 2C_y \int_0^T \|z_t\|_{V^*}^2 + \|z\|_V^2 dt, \\ &\leq C \|z\|_{W(0,T)}^2. \end{aligned} \quad (3.10)$$

Con (3.5) y (3.10) se ha demostrado que $F'(y)$ es lineal y continua.

Finalmente, para demostrar que (3.4) es en efecto la derivada de Fréchet de (3.3), debemos demostrar que

$$\lim_{z \rightarrow 0} \frac{\|F(y+z) - F(y) - F'(y)z\|_{L^2(0,T;V^*)}}{\|z\|_{W(0,T)}} = 0,$$

donde $z \in W(0, T)$.

Pongamos $G = F(y+z) - F(y) - F'(y)z$. Entonces $G \in L^2(0, T; V^*)$ y, por la Definición 2 de la página 15, sabemos que

$$\|G\|_{L^2(0, T; V^*)} = \left(\int_0^T \|G(t)\|_{V^*}^2 dt \right)^{1/2} = \left(\int_0^T \left(\sup_{\substack{\phi \in V \\ \phi \neq 0}} \frac{|\langle G(t), \phi \rangle|}{\|\phi\|_V} \right)^2 dt \right)^{1/2}, \quad (3.11)$$

donde $\langle \cdot, \cdot \rangle$ denota el producto en dualidad de V^* y V . Por tanto, debemos demostrar que

$$\lim_{z \rightarrow 0} \frac{\|G\|_{L^2(0, T; V^*)}}{\|z\|_{W(0, T)}} = 0$$

Para ello, calculamos $\langle G, \phi \rangle_{V^*, V}$ formalmente. Para comenzar,

$$\begin{aligned} F(y+z) - F(y) &= \frac{\partial}{\partial t}(y+z) - \gamma \Delta(y+z) + u(y+z) - r(y+z) + \frac{r}{\kappa}(y+z)^2 \\ &\quad - \frac{\partial y}{\partial t} + \gamma \Delta y - uy + ry - \frac{r}{\kappa}y^2 \\ &= \frac{\partial}{\partial t}(y+z) - \gamma \Delta(y+z) + u(y+z) - r(y+z) + \frac{r}{\kappa}(y^2 + 2yz + z^2) \\ &\quad - \frac{\partial y}{\partial t} + \gamma \Delta y - uy + ry - \frac{r}{\kappa}y^2 \\ &= \frac{\partial z}{\partial t} - \gamma \Delta z + uz - rz + \frac{r}{\kappa}z^2 + \frac{2r}{\kappa}zy. \end{aligned}$$

Con lo que, para $\phi \in V$ y cierto $z \in W(0, T)$, obtenemos

$$\begin{aligned} \langle G, \phi \rangle_{V^*, V} &= \langle F(y+z) - F(y) - F'(y)z, \phi \rangle \\ &= \langle F(y+z) - F(y), \phi \rangle - \langle F'(y)z, \phi \rangle \\ &= \langle z_t, \phi \rangle_{V^*, V} + \gamma(\nabla z, \nabla \phi) + (uz, \phi) - r(z, \phi) + \frac{r}{\kappa}(z^2, \phi) + \frac{r}{\kappa}(2yz, \phi) \\ &\quad - \langle z_t, \phi \rangle_{V^*, V} - \gamma(\nabla z, \nabla \phi) - (uz, \phi) + r(z, \phi) - \frac{2r}{\kappa}(yz, \phi) \\ &= \frac{r}{\kappa}(z^2, \phi), \end{aligned} \quad (3.12)$$

c.t.p. $t \in (0, T)$. Nuevamente, por simplicidad en la notación, hemos omitido la dependencia de la variable t . Reemplazando (3.12) en (3.11), tenemos

$$\|G\|_{L^2(0, T; V^*)} = \frac{r}{\kappa} \left(\int_0^T \left(\sup_{\substack{\phi \in V \\ \phi \neq 0}} \frac{|(z^2, \phi)|}{\|\phi\|_V} \right)^2 dt \right)^{1/2}, \quad (3.13)$$

donde $z = z(\cdot, t)$. Pero, gracias a la desigualdad de Cauchy-Schwarz, sabemos que

$$|(z^2, \phi)| \leq \|z^2\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}.$$

Usando este resultado en (3.13), tenemos

$$\|G\|_{L^2(0,T;V^*)} \leq \frac{r}{\kappa} \left(\int_0^T \left(\sup_{\substack{\phi \in V \\ \phi \neq 0}} \frac{\|z^2\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)}}{\|\phi\|_V} \right)^2 dt \right)^{1/2}.$$

Gracias a la Propiedad 1 del Anexo B, sabemos que $\|\phi\|_{L^2(\Omega)} \leq C\|\phi\|_V$, para cierta constante $C > 0$. A partir de ahora y a lo largo de esta sección, C notará una constante genérica que representa las constantes que se obtienen en los cálculos; además, tendremos en mente la dependencia con respecto al tiempo, es decir $z = z(\cdot, t)$. Con esto, obtenemos la estimación:

$$\|G\|_{L^2(0,T;V^*)} \leq C \left(\int_0^T \|z^2\|_{L^2(\Omega)}^2 dt \right)^{1/2}. \quad (3.14)$$

Por otro lado, gracias al Teorema de Rellich-Kondrachov (Teorema 6, Anexo B), sabemos que $V \hookrightarrow L^4(\Omega)$, con lo que podemos estimar (3.14) como sigue:

$$\begin{aligned} \|G\|_{L^2(0,T;V^*)} &\leq C \left(\int_0^T \|z^2\|_{L^2(\Omega)}^2 dt \right)^{1/2} \\ &= C \left(\int_0^T \|z\|_{L^4(\Omega)}^4 dt \right)^{1/2} \\ &\leq C \left(\int_0^T \|z\|_V^4 dt \right)^{1/2}. \end{aligned} \quad (3.15)$$

De (3.11), y usando la desigualdad (3.15), tenemos que

$$\begin{aligned} \lim_{z \rightarrow 0} \frac{\|G\|_{L^2(0,T;V^*)}}{\|z\|_{W(0,T)}} &\leq \lim_{z \rightarrow 0} \frac{C \left(\int_0^T \|z\|_V^4 dt \right)^{1/2}}{\|z\|_{W(0,T)}} \\ &\leq C \lim_{z \rightarrow 0} \frac{\left(\int_0^T \left(\|z\|_V^2 + \|z_t\|_{V^*}^2 \right)^2 dt \right)^{1/2}}{\|z\|_{W(0,T)}}. \end{aligned} \quad (3.16)$$

Pero, de la Definición 3 sabemos que la aplicación $t \mapsto \|z\|_V^2 + \|z_t\|_{V^*}^2$ es elemento de $L^2([0, T])$, por lo que podemos aplicar la desigualdad de Hölder (ver Teorema 5, Anexos B) y así obtener

$$\begin{aligned} \int_0^T \left(\|z\|_V^2 + \|z_t\|_{V^*}^2 \right)^2 dt &= \int_0^T \left| \|z\|_V^2 + \|z_t\|_{V^*}^2 \right| \left| \|z\|_V^2 + \|z_t\|_{V^*}^2 \right| dt \\ &\leq \left(\int_0^T \|z\|_V^2 + \|z_t\|_{V^*}^2 dt \right) \left(\int_0^T \|z\|_V^2 + \|z_t\|_{V^*}^2 dt \right) \\ &= \|z\|_{W(0,T)}^4. \end{aligned} \quad (3.17)$$

Reemplazando (3.17) en (3.16), finalmente deducimos que

$$\lim_{z \rightarrow 0} \frac{\|G\|_{L^2(0,T;V^*)}}{\|z\|_{W(0,T)}} \leq C \lim_{z \rightarrow 0} \frac{\left(\|z\|_{W(0,T)}^4 \right)^{1/2}}{\|z\|_{W(0,T)}}$$

$$\begin{aligned}
&= C \lim_{z \rightarrow 0} \|z\|_{W(0,T)} \\
&= 0,
\end{aligned}$$

con lo cual queda demostrado el lema. \square

3.1.2 Esquema de Newton

Recordemos que (\cdot, \cdot) denota el producto interno en $L^2(\Omega)$ y $\langle \cdot, \cdot \rangle$ el producto en dualidad entre V^* y V . Supongamos que nos encontramos en cierto paso k del método de Newton caracterizado por la ecuación lineal:

$$\langle F'(y_k)(y_{k+1} - y_k), \phi \rangle = - \langle F(y_k), \phi \rangle, \quad (3.18)$$

para todo $\phi \in V$, c.t.p. $t \in (0, T)$. Por un lado, tomamos la dirección $z = (y_{k+1} - y_k)$ en (3.4), y obtenemos:

$$\begin{aligned}
\langle F'(y_k)(y_{k+1} - y_k), \phi \rangle &= \left\langle \frac{\partial}{\partial t}(y_{k+1} - y_k), \phi \right\rangle + \gamma(\nabla(y_{k+1} - y_k), \nabla\phi_j) + (u(y_{k+1} - y_k), \phi) \\
&\quad - r((y_{k+1} - y_k), \phi) + \frac{2r}{\kappa}(y_k(y_{k+1} - y_k), \phi) \\
&= \left\langle \frac{\partial}{\partial t}y_{k+1}, \phi \right\rangle - \left\langle \frac{\partial}{\partial t}y_k, \phi \right\rangle + \gamma(\nabla y_{k+1}, \nabla\phi) - \gamma(\nabla y_k, \nabla\phi) \\
&\quad + (uy_{k+1}, \phi) - (uy_k, \phi) - r(y_{k+1}, \phi) + r(y_k, \phi) + \frac{2r}{\kappa}(y_k y_{k+1}, \phi) \\
&\quad - \frac{2r}{\kappa}(y_k y_k, \phi).
\end{aligned}$$

Por otro lado, gracias a la definición de F en (3.3), sabemos que

$$\langle F(y_k), \phi \rangle = \left\langle \frac{\partial}{\partial t}y_k, \phi \right\rangle + \gamma(\nabla y_k, \nabla\phi) + (uy_k, \phi) - r(y_k, \phi) + \frac{r}{\kappa}(y_k^2, \phi).$$

Reemplazando estas dos cantidades en (3.18), tenemos:

$$\begin{aligned}
&\left\langle \frac{\partial}{\partial t}y_{k+1}, \phi \right\rangle - \left\langle \frac{\partial}{\partial t}y_k, \phi \right\rangle + \gamma(\nabla y_{k+1}, \nabla\phi) - \gamma(\nabla y_k, \nabla\phi) + (uy_{k+1}, \phi) \\
&\quad - (uy_k, \phi) - r(y_{k+1}, \phi) + r(y_k, \phi) + \frac{2r}{\kappa}(y_k y_{k+1}, \phi) - \frac{2r}{\kappa}(y_k y_k, \phi) \\
&\quad = - \left\langle \frac{\partial}{\partial t}y_k, \phi \right\rangle - \gamma(\nabla y_k, \nabla\phi) - (uy_k, \phi) + r(y_k, \phi) - \frac{r}{\kappa}(y_k^2, \phi).
\end{aligned}$$

Eliminando los términos semejantes, obtenemos la siguiente ecuación

$$\begin{aligned}
&\left\langle \frac{\partial}{\partial t}y_{k+1}, \phi \right\rangle + \gamma(\nabla y_{k+1}, \nabla\phi) + (uy_{k+1}, \phi) \\
&\quad - r(y_{k+1}, \phi) + \frac{2r}{\kappa}(y_k y_{k+1}, \phi) = \frac{r}{\kappa}(y_k^2, \phi), \quad \forall \phi \in V.
\end{aligned}$$

Nótese que $\left\langle \frac{\partial}{\partial t}y_{k+1}, \phi \right\rangle = \frac{\partial}{\partial t}(y_{k+1}, \phi_j)$, por lo que finalmente obtenemos la siguiente expresión:

$$\begin{aligned} \frac{\partial}{\partial t}(y_{k+1}, \phi_j) + \gamma(\nabla y_{k+1}, \nabla \phi) + (uy_{k+1}, \phi) \\ - r(y_{k+1}, \phi) + \frac{2r}{\kappa}(y_k y_{k+1}, \phi) = \frac{r}{\kappa}(y_k^2, \phi), \quad \forall \phi \in V, \end{aligned} \quad (3.19)$$

que constituye el ingrediente principal del esquema de Newton para nuestro problema.

Fijando $t \in (0, T]$ en (3.19), obtenemos un problema variacional lineal definido en V . A continuación, vamos a realizar la semidiscretización de dicho problema aplicando el método de Galerkin en la variable espacial. Para el espacio discreto, consideramos elementos finitos triangulares lineales definidos por un mallado del espacio plano. Definimos nuestro espacio de elementos finitos lineales $V_h \subset V$ como:

$$V_h := \{\phi \in C(\Omega) : \phi \text{ es un polinomio de grado 1 definido sobre cada triángulo del mallado}\}.$$

Ver [7] para los detalles de la construcción de V_h . Fijamos $N_m \in \mathbb{N}$ y recordamos que el conjunto $\{\phi_j\}_{j=1}^{N_m} \subset V_h$ constituye una base ortonormal en V_h en $L^2(\Omega)$. El problema semidiscreto que aproxima (3.19) es el siguiente: buscamos $y^h = y^h(x, t)$ de la forma

$$y^h(x, t) = \sum_{i=1}^{N_m} w_i(t) \phi_i(x), \quad (3.20)$$

para ciertas funciones $w_i(t)$, que aproxime a y en el siguiente sentido:

$$\begin{aligned} \frac{\partial}{\partial t}(y_{k+1}, \phi_j) + \gamma(\nabla y_{k+1}, \nabla \phi_j) + (uy_{k+1}, \phi_j) \\ - r(y_{k+1}, \phi_j) + \frac{2r}{\kappa}(y_k y_{k+1}, \phi_j) = \frac{r}{\kappa}(y_k^2, \phi_j), \quad \forall j = 1, \dots, N_m, \end{aligned} \quad (3.21)$$

para casi todo $t \in [0, T]$. Finalmente, reemplazando la aproximación (3.20) en (3.21), obtenemos el siguiente sistema de ecuaciones diferenciales ordinarias lineales que constituye el paso principal del algoritmo de Newton:

$$\begin{aligned} Mw'_{k+1}(t) + \gamma K w_{k+1}(t) + M_u(t) w_{k+1}(t) \\ - r M w_{k+1}(t) + \frac{2r}{\kappa} M_{y_k}(t) w_{k+1}(t) = \frac{r}{\kappa} M \tilde{w}_k(t), \end{aligned} \quad (3.22)$$

para $k = 0, 1, \dots$, c.t.p. $t \in [0, T]$; donde:

- M , es la *Matriz de Masa*, con entradas $M(j, l) = (\phi_j, \phi_l)_{L^2(\Omega)}$,
- K , es la *Matriz de Rigidez*, con entradas $K(j, l) = (\nabla \phi_j, \nabla \phi_l)_{L^2(\Omega)}$,
- $M_u = M_u(t)$, es la matriz con $M_u(j, l) = (u(\cdot, t) \phi_j, \phi_l)_{L^2(\Omega)}$,
- $M_{y_k} = M_{y_k}(t)$, es la matriz con $M_{y_k}(j, l) = (y_k(\cdot, t) \phi_j, \phi_l)_{L^2(\Omega)}$, y
- $\tilde{w}_k = \tilde{w}_k(t)$, es el vector con coordenadas $\tilde{w}_k(j) = (w_k(j))^2$;

para $j, l = 1, \dots, N_m$.

El sistema de EDO obtenido en (3.22) será resuelto utilizando dos esquemas: el método de Euler Implícito y el de Crank-Nicholson. Este último requiere ser estabilizado debido al tipo de condiciones iniciales, lo cual se detallará más adelante. En las secciones siguientes se detalla la aplicación de ambos métodos.

3.1.3 Esquema de Euler implícito (EI)

En esta sección, vamos a resolver el sistema de EDO (3.22) utilizando el método de Euler Implícito. Primero recordemos el método para una EDO con condición inicial de la forma :

$$\begin{cases} y'(t) = f(t, y(t)), \\ y(0) = y_0. \end{cases}$$

Sea $\{[t_{i-1}, t_i]\}_{i=1}^{N_t}$ una partición uniforme del intervalo $[0, T]$, de manera que

$$0 = t_0 < \dots < t_i < t_{i+1} < \dots < t_{N_t} = T, \quad (3.23)$$

con $i = 0, 1, \dots, N_t$. El paso de discretización Δt , está definido por $\Delta t = \frac{T}{N_t}$; donde N_t es el número de subintervalos en el que dividimos el intervalo de tiempo considerado. El método de EI se plantea como sigue, cf. [7]:

$$y(t_{i+1}) = y(t_i) + \Delta t f(t_{i+1}, y(t_{i+1})), \quad i = 0, 1, 2, \dots, N_t.$$

Si tomamos $f(t_{i+1}, y(t_{i+1})) = w'_{k+1}(t_{i+1})$ en la ecuación (3.22), entonces

$$w'_{k+1}(t_{i+1}) = \frac{1}{\Delta t} (w_{k+1}(t_{i+1}) - w_{k+1}(t_i)),$$

donde Δt es suficientemente pequeño y $w_{k+1}(t_{i+1}) = w_{k+1}(t_i + \Delta t)$. Notemos que el índice k hace referencia al método de Newton y el índice i al método de Euler Implícito.

La aplicación del método de EI a la ecuación (3.22) da lugar a la siguiente expresión:

$$\begin{aligned} \frac{1}{\Delta t} M (w_{k+1}(t_{i+1}) - w_{k+1}(t_i)) + \gamma K w_{k+1}(t_{i+1}) + M_u(t_{i+1}) w_{k+1}(t_{i+1}) - \\ r M w_{k+1}(t_{i+1}) + \frac{2r}{\kappa} M_{y_k}(t_{i+1}) w_{k+1}(t_{i+1}) = \frac{r}{\kappa} M \tilde{w}_k(t_{i+1}). \end{aligned}$$

Reorganizando los términos, obtenemos finalmente el esquema de EI para (3.22):

$$\begin{aligned} \left(M + \gamma \Delta t K + \Delta t M_u(t_{i+1}) - r \Delta t M + \frac{2r}{\kappa} \Delta t M_{y_k}(t_{i+1}) \right) w_{k+1}(t_{i+1}) = \\ M w_{k+1}(t_i) + \Delta t \frac{r}{\kappa} M \tilde{w}_k(t_{i+1}). \end{aligned} \quad (3.24)$$

Proposición 1 (Newton + EI). *Si al sistema (3.22) se le aplica el método de Euler Implícito, se obtiene un sistema lineal de la forma:*

$$A_k w_{k+1}(t_{i+1}) = b_k, \quad (3.25)$$

donde:

$$A_k = M + \Delta t \left(\gamma K + M_u(t_{i+1}) - rM + \frac{2r}{\kappa} M_{y_k}(t_{i+1}) \right),$$

$$b_k = Mw_{k+1}(t_i) + \Delta t \frac{r}{\kappa} M \tilde{w}_k(t_{i+1}).$$

Demostración. La demostración de esta Proposición queda expuesta en el desarrollo de la Sección 3.1.3 que concluye con la ecuación (3.24). \square

Usando la Proposición 1, formulamos el siguiente algoritmo:

Input: N_n (Número de pasos para el método de Newton), N_t (Número de subintervalos del intervalo de tiempo), valor inicial y_0 , valor de u .

Output: Solución aproximada de (3.1) notada por W_{new}

- 1: Inicializar W_{old} resolviendo el problema de valor inicial lineal.
- 2: **while** ($k < N_n$ y criterio de parada) **do**
- 3: **for** $i = 1 : N_t$ **do**
- 4: Calcular $W_{new}(:, t_i)$ resolviendo (3.25).
- 5: **end i**
- 6: Asignar $W_{old} = W_{new}$.
- 7: **end k**
- 8: **return** W_{new}

Algoritmo 1: Euler Implícito

3.1.4 Esquema de Crank-Nicholson (CN)

En esta sección, aplicaremos el método de Crank-Nicholson a (3.22). Recordemos de igual manera que, para la EDO con condición inicial de la forma:

$$\begin{cases} y'(t) = f(t, y(t)), \\ y(0) = y_0, \end{cases}$$

el esquema de Crank-Nicholson utiliza la siguiente aproximación:

$$y(t_{i+1}) = y(t_i) + \Delta t \left(\frac{f(t_{i+1}, y(t_{i+1})) + f(t_i, y(t_i))}{2} \right).$$

Si en el contexto de la ecuación (3.22) tomamos $f(t_{i+1}, y(t_{i+1})) = w'_{k+1}(t_{i+1})$, entonces tenemos:

$$\frac{w_{k+1}(t_{i+1}) - w_{k+1}(t_i)}{\Delta t} = \frac{1}{2} (w'_{k+1}(t_{i+1}) + w'_{k+1}(t_i)),$$

donde Δt es el tamaño de paso en la variable temporal, como en la sección anterior.

Para facilitar la lectura y comprensión de la aplicación del método CN, comencemos reorganizando los términos de la ecuación (3.22) como sigue:

$$Mw'_{k+1}(t) = -\gamma Kw_{k+1}(t) - M_u(t)w_{k+1}(t) + rMw_{k+1}(t) - \frac{2r}{\kappa} M_{y_k}(t)w_{k+1}(t) + \frac{r}{\kappa} M \tilde{w}_k(t).$$

Aplicamos el método CN a la ecuación (3.22) y obtenemos la siguiente expresión:

$$\begin{aligned} \frac{1}{\Delta t} M (w_{k+1}(t_{i+1}) - w_{k+1}(t_i)) &= \frac{1}{2} \left[\frac{r}{\kappa} M \tilde{w}_k(t_{i+1}) - \gamma K w_{k+1}(t_{i+1}) \right. \\ &\quad - M_u(t_{i+1}) w_{k+1}(t_{i+1}) + r M w_{k+1}(t_{i+1}) - \frac{2r}{\kappa} M_{y_k}(t_{i+1}) w_{k+1}(t_{i+1}) + \frac{r}{\kappa} M \tilde{w}_k(t_i) \\ &\quad \left. - \gamma K w_{k+1}(t_i) - M_u(t_i) w_{k+1}(t_i) + r M w_{k+1}(t_i) - \frac{2r}{\kappa} M_{y_k}(t_i) w_{k+1}(t_i) \right]. \end{aligned}$$

Finalmente, reorganizamos los términos de la ecuación previa y obtenemos la expresión:

$$\begin{aligned} \left[M + \frac{\Delta t}{2} \left(\gamma K + M_u(t_{i+1}) - r M + \frac{2r}{\kappa} M_{y_k}(t_{i+1}) \right) \right] w_{k+1}(t_{i+1}) &= \\ \left[M - \frac{\Delta t}{2} \left(\gamma K + M_u(t_i) - r M + \frac{2r}{\kappa} M_{y_k}(t_i) \right) \right] w_{k+1}(t_i) &+ \\ + \frac{r \Delta t}{2 \kappa} M (\tilde{w}_k(t_{i+1}) + \tilde{w}_k(t_i)), & \quad (3.26) \end{aligned}$$

a partir de la cual se plantea la siguiente proposición.

Proposición 2 (Newton + CN). *Al aplicar al sistema (3.22) el método de Crank-Nicholson, se obtiene un sistema lineal de la forma:*

$$A_k w_{k+1}(t_{i+1}) = b_k, \quad (3.27)$$

donde

$$\begin{aligned} A_k &= M + \frac{\Delta t}{2} \left(\gamma K + M_u(t_{i+1}) - r M + \frac{2r}{\kappa} M_{y_k}(t_{i+1}) \right), \\ b_k &= \left[M - \frac{\Delta t}{2} \left(\gamma K + M_u(t_i) - r M + \frac{2r}{\kappa} M_{y_k}(t_i) \right) \right] w_{k+1}(t_i) \\ &\quad + \frac{r \Delta t}{2 \kappa} M (\tilde{w}_k(t_{i+1}) + \tilde{w}_k(t_i)). \end{aligned}$$

A partir de la Proposición 2, formulamos el siguiente algoritmo:

Input: N_n (Número de pasos para el método de Newton), N_t (Número de subintervalos del intervalo de tiempo), valor inicial y_0 , valor de u .

Output: Solución aproximada de (3.1) notada por W_{new}

- 1: Inicializar W_{old} resolviendo el problema de valor inicial lineal.
- 2: **while** ($k < N_n$ y criterio de parada) **do**
- 3: **for** $i = 1 : N_t$ **do**
- 4: Calcular $W_{new}(:, t_i)$ resolviendo (3.27)
- 5: **end i**
- 6: Asignar $W_{old} = W_{new}$.
- 7: **end k**
- 8: **return** W_{new}

Algoritmo 2: Crank-Nicholson.

3.2 Implementación Numérica

Para la resolución de los sistema de ecuaciones dados en (3.25) y (3.27) vamos a utilizar la plataforma *MATLAB*[®] y su complemento para EDP *pdetoolbox*. Para comenzar, establecemos una triangulación del dominio Ω , que consiste en una estructura compuesta por los nodos, los triángulos, lados del dominio y su geometría. A esta estructura la llamamos *malla* y llamamos N_m al número de nodos de la misma.

Similarmente, vamos a discretizar el intervalo del tiempo $[0, T]$ como en (3.23), pág. 29, donde $N_t + 1$ es el número de nodos del intervalo de tiempo discretizado.

Las funciones dadas en el problema corresponderán a matrices cuyas filas representan los valores de las funciones en los nodos espaciales y cuyas columnas representan los valores en los nodos temporales.

Llamaremos *aproximación de la solución* a la matriz $W \in \mathbb{R}^{N_m \times N_t}$, cuyas columnas están determinadas por los vectores w_{k+1} obtenidos al resolver los sistemas matriciales (3.25) y (3.27). Cuando realicemos la implementación numérica, esta matriz es calculada en cada iteración del método de Newton.

Así, la i -ésima columna de la matriz W , notada por $W(:, i)$, está dada por el vector w_{k+1}^i , es decir:

$$W(:, i) = w_{k+1}(\cdot, t_i);$$

de lo que resulta que la matriz W nos da como información la evolución del estado del sistema conforme pasa el tiempo.

Utilizando los resultados de la sección anterior, vamos a resolver los esquemas matriciales (3.25) y (3.27) en dos niveles iterativos: el primero, con las iteraciones del método de Newton (asociadas al índice k), para lo cual es necesario considerar un número N_n de iteraciones del método de Newton. El segundo, el ciclo de las iteraciones de los métodos de Euler o de CN estabilizado, respectivamente, y en el cual corre el índice temporal i .

Las matrices M y K no dependen del tiempo y pueden construirse utilizando la herramienta *assem* del complemento *pdetoolbox*.

Por el contrario, las matrices M_u y M_{y_k} sí dependen del tiempo, de manera que deberán ser calculadas en cada punto de la partición del intervalo $[0, T]$. La matriz M_u se construye usando la herramienta *assem*, la cual ensambla matrices cuyas entradas equivalen a las de la matriz de masa M multiplicadas por cierto coeficiente. Para ello, es necesario primero evaluar la función u sobre los centros de masa de los triángulos de la malla (usando el comando *pdeintrp*) y usar estos valores como coeficientes de peso de la matriz M .

La matriz M_{y_k} también se obtiene utilizando la herramienta *assem*, pero con el peso dado por el coeficiente $\frac{2r}{\kappa} y_k$; ver la Ecuación (3.25). Nótese que esta matriz no sólo depende del tiempo, sino que también depende del valor de la aproximación de y calculado en el paso anterior del método de Newton, es decir, hace falta guardar la información de esta matriz en el paso k para realizar los cálculos correspondientes al paso $k + 1$.

Una vez ensambladas todas las matrices, podemos escribir el problema $A_k w_{k+1}(t_{i+1}) = b_k$ como en (3.25) y (3.27), el cual será resuelto utilizando el comando `\` (el cual utiliza métodos directos para

resolver el sistema lineal) en la siguiente forma:

$$w_{k+1}(t_{i+1}) = A_k \setminus b_k.$$

En cada iteración del método de Newton, se obtiene una matriz a la cual llamamos W_{new} , correspondiente a la aproximación de la solución en el paso k , que depende de la matriz obtenida en el paso anterior ($k - 1$), a la cual llamamos W_{old} . Así, la matriz W_{old} es la matriz cuya i -ésima columna corresponde al vector $w_k(t_i)$ obtenido al resolver el sistema matricial $A_{k-1}w_k(t_i) = b_{k-1}$. Análogamente, la matriz W_{new} tiene como i -ésima columna la aproximación de y en el tiempo t_i : $y(\cdot, t_i)$. Es decir, el vector $W_{new}(:, t_i)$ corresponde a la aproximación de la solución dada por $w_{k+1}(t_i)$.

Es necesario entonces dar un valor inicial a la matriz W_{old} . Para ello, asignamos a la primera columna de la matriz el vector dado por los valores de la condición inicial y_0 , y el resto de las columnas se calculan resolviendo el sistema lineal dado por el problema considerando sólo los términos lineales:

$$\begin{cases} W_{old}(:, 1) = y_0, \\ W_{old}(:, i) = (M + \Delta t(\gamma K + M)) \setminus (M W_{old}(:, i - 1)), \quad i = 2, \dots, N_t. \end{cases} \quad (3.28)$$

3.2.1 Resolución numérica con el método de Euler Implícito

A continuación, probaremos el Algoritmo 1 en un ejemplo sencillo dado por una función que verifica el problema (3.1).

Ejemplo 3.1. *El siguiente ejemplo fue determinado a partir de su solución exacta; dada por*

$$y(x_1, x_2; t) = \frac{1}{t+1} \left[\frac{1}{3} \left(x_1 - \frac{1}{2} \right)^3 - \frac{1}{4} x_1 + 1 \right], \quad (3.29)$$

donde $x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2$ es la variable espacial del problema y $t \in [0, 1]$. Nótese que y satisface las condiciones de frontera de Neuman homogéneas. En efecto,

$$\nabla y(x_1, x_2, t) = \begin{pmatrix} \frac{1}{t+1} \left[\left(x_1 - \frac{1}{2} \right)^2 - \frac{1}{4} \right] \\ 0 \end{pmatrix}.$$

De donde $\nabla y(x_1, x_2, t)|_{x_1=0} = 0$, $\nabla y(x_1, x_2, t)|_{x_1=1} = 0$. Además,

$$\nabla y(x_1, x_2, t)|_{x_2=0} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 0 \quad y \quad \nabla y(x_1, x_2, t)|_{x_2=0} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0.$$

Con lo que $\nabla y(x_1, x_2, t)|_{\Gamma} \cdot \nu = 0$. Es decir, $\partial_\nu y = 0$.

Recordemos el problema de Fisher:

$$y_t - \gamma \Delta y + u(x, t)y(x, t) = ry(x, t) \left(1 - \frac{y(x, t)}{\kappa} \right) \quad (x, t) \in Q_T,$$

sujeto a condiciones de frontera de Neumann homogéneas. Para construir el ejemplo, vamos a tomar $\Omega = (0, 1) \times (0, 1)$ y $T = 1$. El problema de Fisher se satisface tomando las siguientes cantidades:

$$\begin{aligned}
\cdot y_t &= -\frac{1}{(t+1)^2} \left[\frac{1}{3} \left(x_1 - \frac{1}{2} \right)^3 - \frac{1}{4} x_1 + 1 \right], \\
\cdot \Delta y &= \frac{1}{t+1} (2x_1 - 1), \\
\cdot u &= r - \frac{r}{\kappa(t+1)} \left[\frac{1}{3} \left(x_1 - \frac{1}{2} \right)^3 - \frac{1}{4} x_1 + 1 \right] + \frac{(2x_1 - 1)}{\left[\frac{1}{3} \left(x_1 - \frac{1}{2} \right)^3 - \frac{1}{4} x_1 + 1 \right]} + \frac{1}{t+1}, \\
\cdot y_0 &= \begin{cases} x_1 + x_2 & \text{si } |[x_1, x_2] - [0,5, 0,5]| \leq 0,1 \\ 0 & \text{si no.} \end{cases},
\end{aligned}$$

donde $|\cdot|$ denota la norma euclídeana. Los parámetros r y κ tendrán el valor de 1.

El Algoritmo 1 fue implementado sobre una malla generada en *MATLAB*[®], triangulizando regularmente nuestro conjunto Ω como sigue. Para efectos de comparación, llamaremos h al mayor de los lados de los triángulos de la malla; como la partición del intervalo temporal es uniforme, mantenemos la notación Δt .

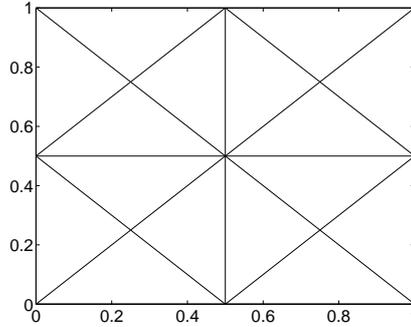


Figura 3.1: Ejemplo 3.1, malla inicial, $h = 0,5$.

Dado que el Ejemplo 3.1 se construyó a partir del conocimiento de la solución exacta (3.29), es posible para nosotros hacer un cálculo exacto del error de aproximación de la solución, comparando la solución aproximada con los valores de la función exacta y en los nodos de la malla.

Así, vamos a utilizar dos parámetros de cálculo del error: la discretización espacial y la temporal. De manera que vamos a realizar algunos refinamientos de la malla y del intervalo de tiempo $[0, T]$, para observar cómo funciona el algoritmo a medida que se generan más intervalos temporales o más triángulos en el espacio Ω . Estas pruebas nos indicarán si estamos resolviendo el problema numérico.

Por otro lado, se analizan dos errores: el error en la norma inducida por el espacio $L^2(\Omega)$ y el error en la norma inducida por el espacio V , con el fin de verificar el error en los distintos espacios. En particular, la norma en $V = H^1(\Omega)$ nos indica el error de la variación de la solución.

Las pruebas se realizaron refinando la malla 6 veces usando el comando `refinemesh` del `pdetoolbox`, tomando como dato de referencia al mayor de los lados de todos los triángulos de la malla, h . Refinamos también la discretización del intervalo temporal con $\Delta t = \frac{1}{N_t}$ con $N_t = 100, 300, 500$ y 1000 . La malla refinada queda como sigue:

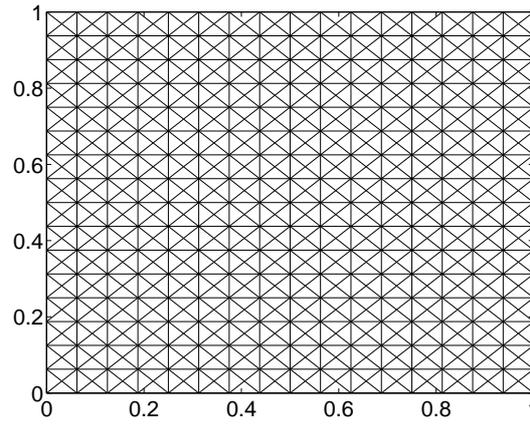


Figura 3.2: Malla resultante después de refinar 6 veces la malla inicial, $h \approx 0,0625$.

Los cálculos del error en norma $L^2(\Omega)$ obtenidos al correr el Algoritmo 1, refinando regularmente la malla y realizando una partición más fina del intervalo $[0, T]$ se pueden observar en la siguiente tabla. Los datos están aproximados a la sexta cifra decimal.

| | | $t = \Delta t$ | $t = 0,25$ | $t = 0,5$ | $t = 1$ |
|-----------------------------|--------------|----------------|------------|-----------|----------|
| $\Delta t = \frac{1}{100}$ | $h = 0,5$ | 0.000926 | 0.003941 | 0.003989 | 0.004400 |
| | $h = 0,125$ | 0.000186 | 0.001013 | 0.001221 | 0.001026 |
| | $h = 0,0884$ | 0.000092 | 0.001500 | 0.001169 | 0.001435 |
| | $h = 0,0625$ | 0.000086 | 0.001543 | 0.001543 | 0.001494 |
| $\Delta t = \frac{1}{300}$ | $h = 0,5$ | 0.000506 | 0.004283 | 0.004651 | 0.005292 |
| | $h = 0,125$ | 0.000157 | 0.000311 | 0.000271 | 0.000188 |
| | $h = 0,0884$ | 0.000044 | 0.000440 | 0.000360 | 0.000388 |
| | $h = 0,0625$ | 0.000024 | 0.000482 | 0.000482 | 0.000447 |
| $\Delta t = \frac{1}{500}$ | $h = 0,5$ | 0.000356 | 0.004371 | 0.004805 | 0.005479 |
| | $h = 0,125$ | 0.000145 | 0.000267 | 0.000253 | 0.000311 |
| | $h = 0,0884$ | 0.000042 | 0.000229 | 0.000199 | 0.000181 |
| | $h = 0,0625$ | 0.000022 | 0.000270 | 0.000270 | 0.000238 |
| $\Delta t = \frac{1}{1000}$ | $h = 0,5$ | 0.000206 | 0.004441 | 0.004924 | 0.005620 |
| | $h = 0,125$ | 0.000122 | 0.000298 | 0.000343 | 0.000449 |
| | $h = 0,0884$ | 0.000040 | 0.000084 | 0.000090 | 0.000048 |
| | $h = 0,0625$ | 0.000021 | 0.000117 | 0.000117 | 0.000088 |

Tabla 3.1: Ejemplo 3.1, cálculo del error en norma $L^2(\Omega)$ con distintos valores de Δt y h , al aplicar el método de Euler Implícito.

Como se puede observar en la Tabla 3.1, el error disminuye rápida y significativamente. Sin embargo, a partir de un cierto refinamiento de la malla inicial (en este caso, el sexto) el error en el espacio $L^2(\Omega)$ vuelve a crecer conforme el tiempo crece. Para subsanar este problema, fue necesario aplicar el algoritmo sobre una malla espacial y una partición del intervalo temporal sumamente finos, lo cual es computacionalmente muy costoso. Este fenómeno evidenció la necesidad de utilizar un mejor método de integración numérica, por lo que se eligió utilizar el método de Crank-Nicholson estabilizado (los detalles se discutirán en la siguiente sección).

La siguiente tabla contiene los datos del error en la norma $H^1(\Omega)$, aproximados a la sexta cifra decimal.

| | | $t = \Delta t$ | $t = 0,25$ | $t = 0,5$ | $t = 1$ |
|-----------------------------|--------------|----------------|------------|-----------|----------|
| $\Delta t = \frac{1}{100}$ | $h = 0,5$ | 0.007872 | 0.006734 | 0.005603 | 0.012359 |
| | $h = 0,125$ | 0.000186 | 0.001014 | 0.001221 | 0.001026 |
| | $h = 0,0884$ | 0.000092 | 0.001500 | 0.001169 | 0.001435 |
| | $h = 0,0625$ | 0.000085 | 0.001543 | 0.001542 | 0.001494 |
| $\Delta t = \frac{1}{300}$ | $h = 0,5$ | 0.006740 | 0.018460 | 0.016105 | 0.012500 |
| | $h = 0,125$ | 0.007286 | 0.006713 | 0.005598 | 0.004194 |
| | $h = 0,0884$ | 0.004079 | 0.002827 | 0.003394 | 0.002115 |
| | $h = 0,0625$ | 0.002935 | 0.002405 | 0.002002 | 0.001498 |
| $\Delta t = \frac{1}{500}$ | $h = 0,5$ | 0.004830 | 0.018483 | 0.016121 | 0.012528 |
| | $h = 0,125$ | 0.006754 | 0.006709 | 0.005597 | 0.004199 |
| | $h = 0,0884$ | 0.004000 | 0.002826 | 0.003392 | 0.002116 |
| | $h = 0,0625$ | 0.002907 | 0.002404 | 0.002006 | 0.001500 |
| $\Delta t = \frac{1}{1000}$ | $h = 0,5$ | 0.002827 | 0.018501 | 0.016134 | 0.012549 |
| | $h = 0,125$ | 0.005704 | 0.0067068 | 0.005598 | 0.004204 |
| | $h = 0,0884$ | 0.003805 | 0.002825 | 0.003390 | 0.002118 |
| | $h = 0,0625$ | 0.002834 | 0.002404 | 0.002003 | 0.001502 |

Tabla 3.2: Ejemplo 3.1, cálculo del error en norma $H^1(\Omega)$ al aplicar el método de Euler Implícito.

Al tomar $N_t = 1000$, $h \approx 0,16$, hacen falta 4 iteraciones del método de Newton, para alcanzar una aproximación con error de orden menor a 10^{-8} (valor que fue considerado como criterio de parada). A continuación, algunos gráficos obtenidos durante la resolución:

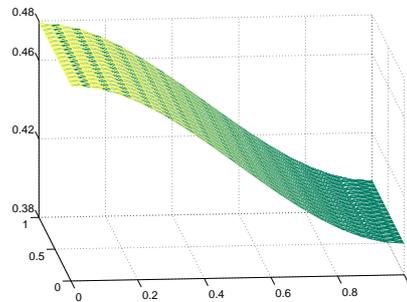


Figura 3.3: Ejemplo 3.1, Solución exacta en el tiempo $t = 1$.

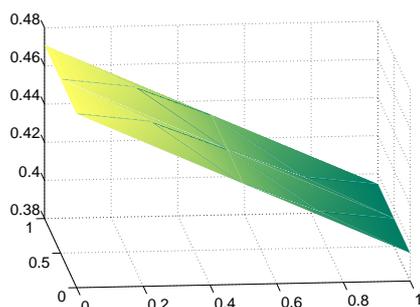


Figura 3.4: Ejemplo 3.1, primera aproximación de la solución calculada al tiempo $t = 1$ por el método de Euler Implícito, con $h = 0,5$.

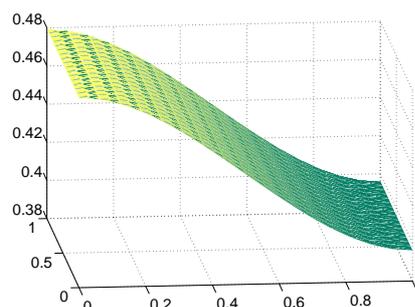


Figura 3.5: Ejemplo 3.1, aproximación de la solución calculada al tiempo $t = 1$ por el método de Euler Implícito, con $h = 0,0625$.

3.2.2 Resolución numérica con el método de Crank-Nicholson estabilizado

La resolución numérica utilizando el método de Crank-Nicholson (CN) es análoga a la implementación con el método de Euler, salvo que las matrices M_u y M_{y_k} no sólo dependen del tiempo, sino que también es necesario considerar tanto el valor en el tiempo t_i como en el tiempo t_{i-1} . Por esta razón, definimos las matrices

$$M_{u_{i-1}} := M_u(t_{i-1}) \quad \text{y} \quad M_{u_i} := M_u(t_i)$$

que corresponden al cálculo (también utilizando el comando `assem`) de la matriz M_u en el instante de tiempo anterior t_{i-1} y en el actual t_i , respectivamente. De manera análoga, introducimos las matrices

$$M_{y_{k_{i-1}}} := M_{y_k}(t_{i-1}) \quad \text{y} \quad M_{y_{k_i}} := M_{y_k}(t_i).$$

Establecemos el siguiente algoritmo, referente al método de Crank-Nicholson Estabilizado (CN).

Input: N_n (Número de pasos para el método de Newton), N_t (Número de subintervalos del intervalo de tiempo), valor inicial y_0 , valor de u .

Output: Solución aproximada de (3.1) notada por W_{new} .

- 1: Inicializar W_{old} resolviendo el problema de valor inicial lineal.
- 2: **while** ($k < N_n$ y criterio de parada) **do**
- 3: **for** $i = 1, 2$ **do**
- 4: Calcular $W_{new}(:, t_i)$ resolviendo (3.25)
- 5: **end i**
- 6: **for** $i = 3 : N_t$ **do**
- 7: Calcular $W_{new}(:, t_i)$ resolviendo (3.27)
- 8: Asignar $M_{u_{i-1}} = M_{u_i}$
- 9: Asignar $M_{y_{k_{i-1}}} = M_{y_{k_i}}$
- 10: **end i**
- 11: Asignar $W_{old} = W_{new}$.
- 12: **end k**
- 13: **return** W_{new}

Algoritmo 3: Crank Nicholson estabilizado

En [13] se puede encontrar una discusión sobre cómo los problemas de difusión que involucran datos iniciales que no son continuos sino que, por ejemplo, son elementos de $L^\infty(\Omega)$ pueden generar irregularidades también en la aproximación de la solución; es decir, se observan problemas de estabilidad numérica, los cuales involucran oscilaciones en la aproximación a la solución (las cuales también se observaron en los primeros experimentos numéricos del desarrollo de este proyecto de titulación). Por esta razón, el autor de [13] recomienda subsanar este problema utilizando algún método de estabilización del método. Seguiremos las recomendaciones de la conclusión de dicho artículo: iniciamos la resolución del problema matricial utilizando el método de EI las primeras dos iteraciones del método de Newton antes de utilizar el método de Crank-Nicholson clásico, procedimiento que se puede ver en el Algoritmo 3.

Ejemplo 3.2. *Con las mismas cantidades definidas en el ejemplo 3.1, utilizamos el Algoritmo 3 para resolver el problema de Fisher.*

De manera análoga al Ejemplo 1, partimos de una malla gruesa como en la figura 3.1 y aplicamos el algoritmo 3 a siete refinamientos de dicha malla y en las discretizaciones temporales utilizadas en el ejemplo anterior, obteniendo la siguiente tabla con datos del error en la norma $L^2(\Omega)$:

| | | $t = \Delta t$ | $t = 0,25$ | $t = 0,5$ | $t = 1$ |
|---|--------------|----------------|------------|-----------|----------|
| $\frac{1}{100}$ $\Delta t = \frac{1}{100}$ | $h = 0,5$ | 0.000926 | 0.004422 | 0.004973 | 0.005725 |
| | $h = 0,125$ | 0.000186 | 0.000304 | 0.000408 | 0.000558 |
| | $h = 0,0884$ | 0.000092 | 0.000065 | 0.000080 | 0.000111 |
| | $h = 0,0625$ | 0.000086 | 0.000059 | 0.000059 | 0.000060 |
| $\frac{1}{300}$ $\Delta t = \frac{1}{300}$ | $h = 0,5$ | 0.000506 | 0.004496 | 0.005034 | 0.005758 |
| | $h = 0,125$ | 0.000157 | 0.000363 | 0.000466 | 0.000592 |
| | $h = 0,0884$ | 0.000044 | 0.000111 | 0.000084 | 0.000144 |
| | $h = 0,0625$ | 0.000024 | 0.000076 | 0.000076 | 0.000088 |
| $\frac{1}{500}$ $\Delta t = \frac{1}{500}$ | $h = 0,5$ | 0.000356 | 0.004505 | 0.005040 | 0.005761 |
| | $h = 0,125$ | 0.000145 | 0.000369 | 0.000471 | 0.000595 |
| | $h = 0,0884$ | 0.000042 | 0.000115 | 0.000089 | 0.000147 |
| | $h = 0,0625$ | 0.000022 | 0.000079 | 0.000079 | 0.000091 |
| $\frac{1}{1000}$ $\Delta t = \frac{1}{1000}$ | $h = 0,5$ | 0.000206 | 0.004510 | 0.005043 | 0.005762 |
| | $h = 0,125$ | 0.000122 | 0.000372 | 0.000473 | 0.000596 |
| | $h = 0,0884$ | 0.000040 | 0.000117 | 0.000092 | 0.000148 |
| | $h = 0,0625$ | 0.000021 | 0.000081 | 0.000081 | 0.000092 |

Tabla 3.3: Ejemplo 3.2, cálculo del error en norma $L^2(\Omega)$ al aplicar el método de Crank-Nicholson estabilizado.

Ahora, en la Tabla 3.3 observamos que el error sólo disminuye, que lo hace rápidamente y que no vuelve a aumentar. Esto se debe a que hemos aplicado un método más estable, que garantiza una mejor aproximación a la solución.

Por otro lado, la Tabla 3.4 contiene los datos del error en la norma $H^1(\Omega)$, y las Figuras 3.6 y 3.7 muestran los gráficos obtenidos durante la resolución.

| | | $t = \Delta t$ | $t = 0,25$ | $t = 0,5$ | $t = 1$ |
|-----------------------------|--------------|----------------|------------|-----------|----------|
| $\Delta t = \frac{1}{100}$ | $h = 0,5$ | 0.011166 | 0.018574 | 0.016230 | 0.012565 |
| | $h = 0,125$ | 0.007872 | 0.006755 | 0.005633 | 0.004207 |
| | $h = 0,0884$ | 0.004140 | 0.002843 | 0.003414 | 0.002119 |
| | $h = 0,0625$ | 0.002947 | 0.002421 | 0.002017 | 0.001504 |
| $\Delta t = \frac{1}{300}$ | $h = 0,5$ | 0.006740 | 0.018542 | 0.016177 | 0.012570 |
| | $h = 0,125$ | 0.007286 | 0.006722 | 0.005610 | 0.004208 |
| | $h = 0,0884$ | 0.004079 | 0.002831 | 0.003397 | 0.002120 |
| | $h = 0,0625$ | 0.002935 | 0.002410 | 0.002009 | 0.001505 |
| $\Delta t = \frac{1}{500}$ | $h = 0,5$ | 0.004830 | 0.018533 | 0.016165 | 0.012570 |
| | $h = 0,125$ | 0.006754 | 0.006715 | 0.005605 | 0.004208 |
| | $h = 0,0884$ | 0.003999 | 0.002829 | 0.003394 | 0.002120 |
| | $h = 0,0625$ | 0.002907 | 0.002407 | 0.002007 | 0.001505 |
| $\Delta t = \frac{1}{1000}$ | $h = 0,5$ | 0.002827 | 0.018526 | 0.016156 | 0.012571 |
| | $h = 0,125$ | 0.005704 | 0.006710 | 0.005602 | 0.004208 |
| | $h = 0,0884$ | 0.003805 | 0.002827 | 0.003391 | 0.002120 |
| | $h = 0,0625$ | 0.002834 | 0.002406 | 0.002006 | 0.001505 |

Tabla 3.4: Ejemplo 3.2, cálculo del error en norma $H^1(\Omega)$ al aplicar el método de Crank-Nicholson estabilizado.

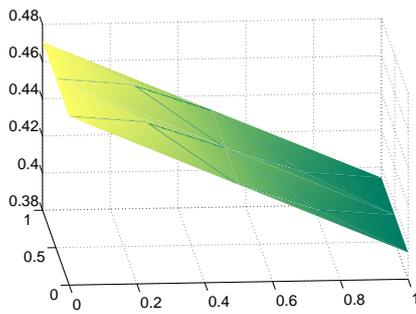


Figura 3.6: Primera aproximación de la solución calculada por el método de Crank-Nicholson estabilizado; $h = 0,5$.

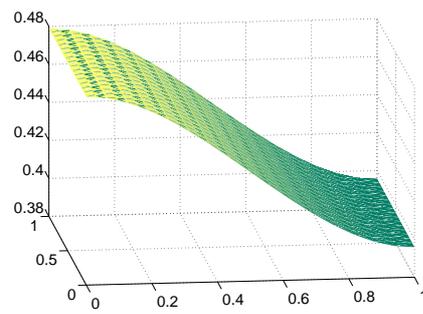


Figura 3.7: Aproximación de la solución calculada por el método de Crank-Nicholson estabilizado, con $h = 0,0625$.

Con estos datos, podemos confirmar numéricamente que el Algoritmo de CN devuelve una aproximación aceptable de la solución exacta de (3.1).

Capítulo 4

Reducción del Modelo de Fisher Mediante el Método POD

4.1 Introducción al Método POD

Resolver numéricamente los problemas (3.18) y (3.15) resulta computacionalmente costoso, más aún cuando el refinamiento de la malla (que es lo que nos permitirá obtener una mejor aproximación a la solución) produce matrices asociadas a la discretización del modelo que crecen rápidamente. Por ello, es de interés disponer de un modelo alternativo que involucre un número considerablemente menor de variables.

La idea del método POD¹ consiste en construir una base ortonormal *conveniente* para el espacio de las soluciones, en el sentido en que logra recuperar una gran cantidad de información del modelo, con un número significativamente menor de elementos que reconstruyen la aproximación de la solución; de manera que el objetivo que alcanza este método es el de “encontrar un subespacio que aproxime un cierto conjunto de datos dado, en el sentido de lo mínimos cuadrados”, ver [12].

El método POD permite aproximar las matrices y vectores que componen el modelo discreto con matrices y vectores correspondientes de tamaño mucho menor. Con este nuevo modelo, es posible calcular una solución numérica aproximada para los problemas de sistemas dinámicos apoyándose en el método de proyección de Galerkin: utilizando una nueva base que genera un subespacio del espacio original, en el cual podemos proyectar el sistema dinámico.

Como ya se mencionó, la construcción de la base ortonormal se realiza utilizando datos inherentes al sistema, que contengan las características especiales del mismo. Para ello, primero necesitamos obtener un número finito de *snapshots* o *imágenes* (cf. [9]), es decir, tener una primera observación de la solución del problema evolutivo en ciertos instantes de tiempo, una vez discretizado el intervalo $[0, T]$. Más adelante introducimos una definición para los snapshots.

La solución numérica de los problemas con ecuaciones diferenciales parciales y, en particular los problemas difusivos en varias dimensiones, pueden dar lugar a considerar un espacio discreto con una

¹Proper Orthogonal Decomposition, también conocido como descomposición Karhunen-Loève.

cantidad muy grande de variables, que a su vez provocan la resolución de problemas matriciales muy costosos. El método de los snapshots resulta ser una técnica natural para obtener información de un sistema dinámico (cf. [12] y las referencias ahí recomendadas).

Otra ventaja de utilizar el método POD es que sólo se requieren operaciones matriciales estándares para obtener resultados, a pesar de ser aplicado a un problema no lineal, como podrá observarse más adelante. Su eficacia radica en la capacidad para guardar información de la evolución del sistema en la base ortonormal que genera.

4.1.1 Descripción del método

A lo largo de este capítulo, vamos a seguir el esquema propuesto por K. Kunish y S. Volkwein en [9], junto con ciertas definiciones presentes en [4] por R. Cantrell y C. Cosner, justificaciones presentadas por R. Pinnau en [12] y el trabajo realizado por E. Kamman y R. Tröltzsch en [8] para describir y aplicar el método POD a nuestro problema de valor inicial.

Consideremos el espacio de Hilbert $L^2(\Omega)$ dotado del producto interior $(\cdot, \cdot)_{L^2(\Omega)}$ usual y de la norma inducida $\|\cdot\|_{L^2(\Omega)}$, como en el Capítulo 2. Introducimos la siguiente

Definición 5 (Snapshots o Imágenes). *Sea $y = y(x, t)$ una solución aproximada de (2.1) en $W(0, T) \hookrightarrow C([0, T; L^2(\Omega)])$ y sea $0 = t_1 < \dots < t_j < \dots < t_m = T$ una partición del intervalo $[0, T]$. Para $j = 1, \dots, m$, $y_j = y(x, t_j)$ es un snapshot o imagen de y al tiempo t_j .*

Sea $\mathcal{I} = \{y_j\}_{j=1}^m \subset L^2(\Omega)$ un conjunto de snapshots del sistema, con al menos uno de ellos distinto de cero. Estos snapshots corresponden a aproximaciones de la solución que se obtienen de una primera observación a la resolución del problema.

Sea $\mathcal{V}_m \subset L^2(\Omega)$ el espacio generado por los snapshots; i.e. $\mathcal{V}_m = \text{span}\{y_1, y_2, \dots, y_m\}$. Sea $\{\psi_k\}_{k=1}^d$ una base ortonormal de \mathcal{V}_m con $d = \dim(\mathcal{V}_m)$, $d \leq m$. Entonces, cada elemento del conjunto de snapshots puede ser escrito como sigue:

$$y_j = \sum_{k=1}^d (y_j, \psi_k)_{L^2(\Omega)} \psi_k, \quad \forall j \in \{1, 2, \dots, m\}. \quad (4.1)$$

El método POD consiste en escoger una base ortonormal en V tal que para todo $\ell \in \{1, 2, \dots, d\}$, se minimice el error cuadrático promedio entre las funciones y_j y la ℓ -ésima suma parcial de (4.1); es decir, se busca una familia $\{\psi_k\}_{k=1}^\ell$ que resuelva:

$$\left\{ \begin{array}{l} \min_{\{\psi_k\}_{k=1}^\ell} \sum_{j=1}^m \left\| y_j - \sum_{k=1}^\ell (y_j, \psi_k)_{L^2(\Omega)} \psi_k \right\|_{L^2(\Omega)}^2 \\ \text{sujeto a } (\psi_i, \psi_j)_{L^2(\Omega)} = \delta_{ij} \quad \text{para } 1 \leq j, k \leq \ell. \end{array} \right. \quad (4.2)$$

A la familia $\{\psi_k\}_{k=1}^\ell$ que verifica (4.2) se le llama *base POD de rango ℓ* .

A continuación, vamos a discutir sobre la elección de los snapshots y la construcción de la base POD, la cual será utilizada para realizar una aproximación de Galerkin de (2.1) y así establecer un esquema reducido.

Definición 6 (Elección de los Snapshots). *Supongamos que tomamos m snapshots equidistantes. Para ello, introducimos el término $\Delta t = \frac{T}{m-1}$, correspondiente al tamaño de paso del tiempo entre ellos, con lo que los instantes de tiempo quedan dados por $t_j = j\Delta t$, con $j = 1, 2, \dots, m$; cf. [9]. Los snapshots $Y_{snap,j}$ quedan definidos por:*

$$y_j = y(t_j), \quad j = 1, \dots, m,$$

y

$$y_j = \bar{\partial}y_{j-m}, \quad j = m+2, \dots, 2m,$$

donde

$$\bar{\partial}y_j = \frac{1}{\Delta t} [y(t_j) - y(t_{j-1})].$$

De manera que tomamos también como snapshots las combinaciones lineales de los snapshots que pueden interpretarse como las aproximaciones de las derivadas de y en los tiempos t . Así, por construcción, todos los snapshots pertenecen a V .

Para construir la base POD, es necesario introducir la siguiente

Definición 7 (Matriz de Correlación). *La Matriz de Correlación se construye a partir de los datos o snapshots, y está dada por: $C \in \mathbb{R}^{m \times m}$, cuyas entradas están dadas por*

$$C(i, j) = (y_j, y_i)_{L^2(\Omega)}, \quad i, j = 1, \dots, m. \quad (4.3)$$

Notemos que por su definición, la matriz C es simétrica semidefinida positiva; y en consecuencia, sus valores propios son reales y positivos.

Buscamos construir un subespacio de $L^2(\Omega)$ que contenga el mayor porcentaje posible de información sobre el problema; es decir, buscamos $\{\psi_k\}_{k=1}^d$ que sea solución de (4.2). Para fijar ideas sobre la elección de la base POD consideramos el problema en dimensión finita:

$$\begin{cases} \min_{\{\psi_k\}_{k=1}^{\ell}} \sum_{j=1}^m \left\| y_j - \sum_{k=1}^{\ell} (y_j, \psi_k)_{\mathbb{R}^{N_m}} \psi_k \right\|_{\mathbb{R}^{N_m}}^2 \\ \text{sujeto a } (\psi_i, \psi_j)_{\mathbb{R}^{N_m}} = \delta_{ij} \quad \text{para } 1 \leq j, k \leq \ell \leq d, \end{cases} \quad (4.4)$$

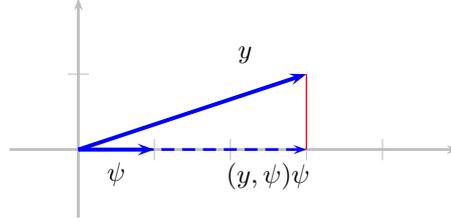
donde N_m es el número de nodos de la discretización espacial, $\|\cdot\|_{\mathbb{R}^{N_m}}$ y $(\cdot, \cdot)_{\mathbb{R}^{N_m}}$ representan la norma y el producto escalar usuales en \mathbb{R}^{N_m} , respectivamente, y

$$\delta_{kj} = \begin{cases} 1 & \text{si } k = j, \\ 0 & \text{caso contrario.} \end{cases}$$

Observación.- Nótese que la solución de (4.4) a grosso modo busca minimizar la diferencia entre las funciones y_j y su proyección sobre el subespacio generado por la base POD $\{\psi_k\}_{k=1}^{\ell}$, con $1 \leq \ell \leq d$.

Pensemos en una sola función y y un sólo elemento de la base POD ψ . La diferencia entre y su proyección sobre ψ está dada por $\|y - (y, \psi)_{\mathbb{R}^{N_m}} \psi\|_{\mathbb{R}^{N_m}}^2$. Buscamos minimizar la diferencia $\|y - (y, \psi)_{\mathbb{R}^{N_m}} \psi\|_{\mathbb{R}^{N_m}}$, lo cual intuitivamente sugiere ser equivalente a maximizar el coeficiente de proyección $(y, \psi)_{\mathbb{R}^{N_m}}$.

La siguiente figura esclarece esta idea.



Al considerar todas las aproximaciones y_j , con $j = 1, \dots, m$ y los elementos de la base POD ψ_k con $k = 1, \dots, \ell$, para $\ell \leq d$, el problema de minimización (4.2) equivale a resolver el siguiente problema de maximización:

$$\begin{cases} \text{máx}_{\{\psi_k\}_{k=1}^{\ell}} \sum_{k=1}^{\ell} \sum_{j=1}^m |(y_j, \psi_k)_{\mathbb{R}^{N_m}}|^2 \\ \text{sujeto a } (\psi_k, \psi_j)_{\mathbb{R}^{N_m}} = \delta_{kj} \quad \text{para } 1 \leq k, j \leq \ell. \end{cases} \quad (4.5)$$

Para resolver el problema de optimización dado en (4.5), seguimos [17] e introducimos el Lagrangiano $\mathcal{L} : \mathbb{R}^{N_m} \times \dots \times \mathbb{R}^{N_m} \times \mathbb{R}^{\ell \times \ell}$ como sigue

$$\mathcal{L}(\psi_1, \dots, \psi_{\ell}, \Lambda) = \sum_{k=1}^{\ell} \sum_{j=1}^m |(y_j, \psi_k)_{\mathbb{R}^{N_m}}|^2 + \sum_{j,k=1}^{\ell} \lambda_{jk} [\delta_{kj} - (\psi_k, \psi_j)_{\mathbb{R}^{N_m}}],$$

para $\psi_1, \dots, \psi_{\ell} \in \mathbb{R}^{N_m}$ y la matriz diagonal $\Lambda \in \mathbb{R}^{\ell \times \ell}$ cuyas entradas son λ_{kj} , los valores propios de la matriz de correlación. Las condiciones de optimalidad de primer orden están dadas por:

$$\frac{\partial \mathcal{L}}{\partial \psi_k}(\psi_1, \dots, \psi_{\ell}, \Lambda) \delta \psi_k = 0, \quad \text{para toda } \delta \psi_k \in \mathbb{R}^{N_m} \text{ y } k \in \{1, \dots, \ell\}. \quad (4.6)$$

Pero

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi_{\mu}}(\psi_1, \dots, \psi_{\ell}, \Lambda) \delta \psi_{\mu} &= 2 \sum_{k=1}^{\ell} \sum_{j=1}^m (y_j, \psi_k)_{\mathbb{R}^{N_m}} (y_j, \delta \psi_{\mu})_{\mathbb{R}^{N_m}} \delta_{k\mu} \\ &\quad - \sum_{j,k=1}^{\ell} \lambda_{jk} (\psi_k, \delta \psi_{\mu})_{\mathbb{R}^{N_m}} \delta_{j\mu} - \sum_{j,k=1}^{\ell} \lambda_{kj} (\delta \psi_{\mu}, \psi_j)_{\mathbb{R}^{N_m}} \delta_{\mu k} \\ &= 2 \sum_{j=1}^m (y_j, \psi_{\mu})_{\mathbb{R}^{N_m}} (y_j, \delta \psi_{\mu})_{\mathbb{R}^{N_m}} - \sum_{k=1}^{\ell} (\lambda_{k\mu} + \lambda_{\mu k}) (\psi_k, \delta \psi_{\mu})_{\mathbb{R}^{N_m}} \\ &= \left(2 \sum_{j=1}^m (y_j, \psi_{\mu})_{\mathbb{R}^{N_m}} y_j - \sum_{k=1}^{\ell} (\lambda_{k\mu} + \lambda_{\mu k}) \psi_k, \delta \psi_{\mu} \right)_{\mathbb{R}^{N_m}}, \end{aligned} \quad (4.7)$$

para todo $\mu \in \{1, \dots, \ell\}$. Con (4.6) y (4.7), tenemos que

$$\sum_{j=1}^m (y_j, \psi_{\mu})_{\mathbb{R}^{N_m}} y_j = \frac{1}{2} \sum_{k=1}^{\ell} (\lambda_{k\mu} + \lambda_{\mu k}) \psi_k, \quad (4.8)$$

en \mathbb{R}^{N_m} y para todo $\mu \in \{1, \dots, \ell\}$. Nótese que

$$C \psi = \sum_{j=1}^m (y_j, \psi)_{\mathbb{R}^{N_m}} y_j, \quad \text{para } \psi \in \mathbb{R}^{N_m}.$$

Así, la condición (4.8) puede expresarse como

$$C \psi_\mu = \frac{1}{2} \sum_{k=1}^{\ell} (\lambda_{k\mu} + \lambda_{\mu k}) \psi_k, \quad (4.9)$$

en \mathbb{R}^{N_m} y para todo $\mu \in \{1, \dots, \ell\}$. Ponemos $\lambda_i = \lambda_{ii}$ y procedemos por inducción. Para $\ell = 1$, tenemos $\mu = 1$ y gracias a (4.9), obtenemos

$$C \psi_1 = \lambda_1 \psi_1,$$

para \mathbb{R}^{N_m} . Luego, suponemos que para $\ell \geq 1$ se verifica

$$C \psi_\mu = \lambda_\mu \psi_\mu, \quad \text{en } \mathbb{R}^{N_m} \text{ y para } \mu \in \{1, \dots, \ell\}. \quad (4.10)$$

Para demostrar que las condiciones de optimalidad de primer orden necesarias para una base POD $\{\psi_\mu\}_{\mu=1}^{\ell+1}$ de rango $\ell + 1$, están dadas por

$$C \psi_\mu = \lambda_\mu \psi_\mu, \quad \text{en } \mathbb{R}^{N_m} \text{ y para } \mu \in \{1, \dots, \ell + 1\}$$

consideramos la hipótesis de inducción (4.10); por lo que sólo hace falta probar

$$C \psi_{\ell+1} = \lambda_{\ell+1} \psi_{\ell+1}, \quad \text{en } \mathbb{R}^{N_m}.$$

Gracias a (4.9), tenemos

$$C \psi_{\ell+1} = \frac{1}{2} \sum_{k=1}^{\ell+1} (\lambda_{k,\ell+1} + \lambda_{\ell+1,k}) \psi_k, \quad \text{en } \mathbb{R}^{N_m}. \quad (4.11)$$

Dado que $\{\psi_k\}_{k=1}^{\ell+1}$ es una base POD, tenemos

$$(\psi_{\ell+1}, \psi_j)_{\mathbb{R}^{N_m}} = 0, \quad \text{para } 1 \leq j \leq \ell.$$

Usando (4.10) y la simetría de la matriz de correlación C , tenemos que para todo $j \in \{1, \dots, \ell\}$, se verifica

$$\begin{aligned} 0 &= \lambda_j (\psi_{\ell+1}, \psi_j)_{\mathbb{R}^{N_m}} \\ &= (\psi_{\ell+1}, C \psi_j)_{\mathbb{R}^{N_m}} \\ &= (C \psi_{\ell+1}, \psi_j)_{\mathbb{R}^{N_m}} \\ &= \frac{1}{2} \sum_{k=1}^{\ell+1} (\lambda_{k,\ell+1} + \lambda_{\ell+1,k}) (\psi_k, \psi_j)_{\mathbb{R}^{N_m}} \\ &= (\lambda_{j,\ell+1} + \lambda_{\ell+1,j}). \end{aligned}$$

De donde

$$\lambda_{j,\ell+1} = -\lambda_{\ell+1,j}, \quad \forall k \in \{1, \dots, \ell\}. \quad (4.12)$$

Reemplazando (4.12) en (4.11), se tiene

$$\begin{aligned} C \psi_{\ell+1} &= \frac{1}{2} \sum_{k=1}^{\ell} (\lambda_{k,\ell+1} + \lambda_{\ell+1,k}) \psi_k + \lambda_{\ell+1} \psi_{\ell+1} \\ &= \frac{1}{2} \sum_{k=1}^{\ell} (\lambda_{k,\ell+1} + \lambda_{k,\ell+1}) \psi_k + \lambda_{\ell+1} \psi_{\ell+1} \\ &= \lambda_{\ell+1} \psi_{\ell+1}, \end{aligned}$$

con lo que queda demostrado (4.1.1) y por tanto, $\{\psi_k\}_{k=1}^{\ell}$ con $\ell \in \{1, \dots, d\}$ es solución de (4.4).

Para el caso funcional (4.2) se tiene un resultado análogo al anterior y su demostración es similar (cf. [17]) lo cual permite establecer la siguiente proposición (cf. [9]).

Proposición 3. Sean $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ los primeros d valores propios más grandes de la matriz de correlación C y $v_1, \dots, v_d \in \mathbb{R}^n$ los vectores propios asociados. Entonces, el conjunto $\{\psi_k\}_{k=1}^d$ definido por:

$$\psi_k(x) = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^m v_k^j y_j(x, \cdot), \quad (4.13)$$

donde $k = 1, \dots, d \leq m$ y v_k^j es la j -ésima componente del vector propio v_k , constituye un conjunto ortonormal de funciones en $L^2(\Omega)$ que resuelve el problema de minimización dado en (4.2). En lo que sigue del documento, llamaremos a este conjunto con el nombre de base POD.

Una de las características más importantes de esta base es que sus elementos están definidos sobre todo el dominio y ya no están definidos localmente como en el caso de las funciones base del método de los elementos finitos, las cuales son funciones a soporte compacto. Además, los elementos fueron originados por los snapshots y los vectores y valores propios de la matriz de correlación, lo cual nos permite guardar una gran cantidad de información del fenómeno en la base POD.

Ahora, la pregunta que surge inmediatamente es: ¿cómo elegir los valores m y d ? Con respecto a m , el número de snapshots a elegir, se recomienda aprovechar la primera observación del comportamiento del sistema dinámico para tomar más snapshots de las zonas donde se observen mayores cambios. Además, de [12] se puede extraer la propuesta de elegir al menos un número $m \leq \dim(C)$ para evitar elegir snapshots linealmente dependientes.

Por otro lado, con respecto al valor d , el número de valores propios reales positivos que tomemos, en [12] se recomienda tomar tantos cuanto porcentaje de información sobre el problema se quiera guardar, tomando en cuenta que los valores propios grandes corresponden a la principal información del sistema dinámico.

Para ello, se define el *contenido relativo de información* como sigue:

$$I_d = \frac{\sum_{j=1}^d \lambda_j}{\sum_{j=1}^m \lambda_j},$$

de manera que I_d es más cercano a 1 conforme d se acerca a m ; cf. [12]. Con esto, el subespacio \mathcal{V}_d contiene un porcentaje p de la información del sistema. Así, el autor de [12] sugiere elegir un valor d tal que

$$d = \min_{\delta} \left\{ I_{\delta} : I_{\delta} \geq \frac{p}{100} \right\} \quad (4.14)$$

De la definición de I_d se puede concluir que, mientras más rápido decaigan los valores propios, más información está contenida en pocos valores propios y, por lo tanto, el problema se puede reducir significativamente. Aunque I_d es una medida de la cantidad de información que contiene la base POD, un análisis de error es necesario para determinar cotas del error.

El conjunto $\{\psi_j\}_{j=1}^d$ es una base ortonormal que genera un subespacio de $L^2(\Omega)$. De manera que podemos utilizar esta base en la proyección de Galerkin, y resolver el sistema dado por la formulación variacional. Es decir, con esta nueva base ortonormal, podemos formular un esquema análogo a (3.8):

$$\frac{\partial}{\partial t}(y_{k+1}, \psi_j) + \gamma(\nabla y_{k+1}, \nabla \psi_j) + (u y_{k+1}, \psi_j) - r(y_{k+1}, \psi_j) + \frac{2r}{\kappa}(y_k y_{k+1}, \psi_j) = \frac{r}{\kappa}(y_k^2, \psi_j), \quad (4.15)$$

donde (\cdot, \cdot) corresponde al producto interior en el espacio $L^2(\Omega)$.

En los documentos [9] y [12], podemos encontrar resultados que muestran cómo los valores propios de la matriz de correlación decaen rápidamente, y por tanto el número de elementos de la base requerida para la solución del problema se reduce significativamente.

Por otro lado, es de interés conocer las estimaciones de los errores de aproximación del método POD. Sin embargo, éste sigue siendo un tema en estudio actualmente, como se pueden ver en las referencias recomendadas en [12]. En la siguiente proposición, presentamos las estimaciones a-priori obtenidas en [9].

Proposición 4 (Estimación del Error de Aproximación del Método POD). *Para todo $\ell \leq d$ tenemos la siguiente fórmula de estimación del error de aproximación:*

$$\sum_{j=1}^m \left\| y(t_j) - \sum_{k=1}^{\ell} (y(t_j), \psi_k)_{L^2(\Omega)} \psi_k \right\|_{L^2(\Omega)}^2 + \sum_{j=2}^m \left\| \bar{\partial} y(t_j) - \sum_{k=1}^{\ell} (\bar{\partial} y(t_j), \psi_k)_{L^2(\Omega)} \psi_k \right\|_{L^2(\Omega)}^2 = \sum_{k=\ell+1}^d \lambda_k. \quad (4.16)$$

4.1.2 Modelo de Fisher reducido

Ahora, vamos a discutir algunos detalles del nuevo esquema dado por (4.15), el cual va a ser resuelto numéricamente utilizando el Algoritmo 3, pero tomando como funciones test a las funciones ψ_k , con $k = 1, \dots, d$ definidas como en (4.13). Es decir, se considera la siguiente aproximación de la solución:

$$y(x, t) \approx \sum_{\mu=1}^d w_{\mu}(t) \psi_{\mu}(x).$$

Para establecer el esquema reducido introducimos las siguientes matrices M^{ψ} , K^{ψ} , M_u^{ψ} , $M_{y_k}^{\psi}$ análogas a las definidas en (3.22), pero dadas por los elementos $\{\psi_j\}_{j=1}^d$ de la base POD construida. Por un lado, la *Matriz de Masa en la base POD*, notada por M^{ψ} tiene por componentes las cantidades $M^{\psi}(j, l) = (\psi_j, \psi_l)_{L^2(\Omega)}$, con $j, l = 1, \dots, d$. Pero $\{\psi_j\}_{j=1}^d$ es una base ortonormal en $L^2(\Omega)$, de manera análoga al método de los Elementos Finitos. Entonces

$$M^{\psi} = I \in \mathbb{R}^{d \times d}, \quad (4.17)$$

es la matriz identidad de dimensión d .

Por otro lado, la *Matriz de Rigidez en la base POD*, notada por K^ψ , tiene por componentes:

$$K^\psi(j, l) = (\nabla\psi_j, \nabla\psi_l)_{L^2(\Omega)}, \quad j, l = 1, \dots, d. \quad (4.18)$$

Finalmente, las matrices M_u^ψ y $M_{y_k}^\psi$ tienen por componentes los valores dados respectivamente por

$$M_u^\psi(j, l) = (u\psi_j, \psi_l)_{L^2(\Omega)} = \int_{\Omega} u(x, t)\psi_j(x)\psi_l(x)dx, \quad (4.19)$$

y

$$M_{y_k}^\psi(j, l) = (y_k\psi_j, \psi_l)_{L^2(\Omega)} = \int_{\Omega} y_k(x, t)\psi_j(x)\psi_l(x)dx, \quad (4.20)$$

c.t.p. $t \in (0, T]$ y para $j, l = 1, \dots, d$.

De manera análoga al sistema completo (ver sección 3.1.2, página 28), una vez hechos los cálculos de (4.18), (4.19), (4.20) y del vector $\tilde{w}_k = w_k^2(t)$, podemos establecer el siguiente sistema de EDO en la base POD:

$$M^\psi w'_{k+1}(t) + \gamma K^\psi w_{k+1}(t) + M_u^\psi(t)w_{k+1}(t) - rM^\psi w_{k+1}(t) + \frac{2r}{\kappa}M_{y_k}^\psi(t)w_{k+1}(t) = \frac{r}{\kappa}M^\psi \tilde{w}_k(t), \quad (4.21)$$

con $t \in [0, T]$. Resolveremos este problema en la siguiente sección utilizando el método de CN estabillizado utilizando los esquemas de Euler Implícito y de Cranck-Nicholson análogos a los obtenidos en la sección anterior:

Proposición 5 (Newton + EI, en la base POD). *Si al sistema (4.21) se le aplica el método de Euler Implícito, se obtiene un sistema lineal de la forma:*

$$A_k^\psi w_{k+1}(t_{i+1}) = b_k^\psi, \quad (4.22)$$

donde:

$$A_k^\psi = M^\psi + \Delta t \left(\gamma K^\psi + M_u^\psi(t_{i+1}) - rM^\psi + \frac{2r}{\kappa}M_{y_k}^\psi(t_{i+1}) \right),$$

$$b_k^\psi = M^\psi w_{k+1}(t_i) + \Delta t \frac{r}{\kappa}M^\psi \tilde{w}_k(t_{i+1}).$$

Proposición 6 (Newton + CN, en la base POD). *Al aplicar al sistema (4.21) el método de Crank-Nicholson, se obtiene un sistema lineal de la forma:*

$$A_k^\psi w_{k+1}(t_{i+1}) = b_k^\psi, \quad (4.23)$$

donde

$$A_k^\psi = M^\psi + \frac{\Delta t}{2} \left(\gamma K^\psi + M_u^\psi(t_{i+1}) - rM^\psi + \frac{2r}{\kappa}M_{y_k}^\psi(t_{i+1}) \right),$$

$$b_k^\psi = \left[M^\psi - \frac{\Delta t}{2} \left(\gamma K^\psi + M_u^\psi(t_i) - rM^\psi + \frac{2r}{\kappa} M_{y_k}^\psi(t_i) \right) \right] w_{k+1}(t_i) + \frac{r\Delta t}{2\kappa} M^\psi (\tilde{w}_k(t_{i+1}) + \tilde{w}_k(t_i)).$$

De manera análoga a la Sección 3.2.2 en la pág. 37, introducimos las matrices $M_{u_{i-1}}^\psi$ y $M_{u_i}^\psi$ cuyas entradas corresponden a los valores de la función u en los tiempos t_{i-1} y t_i respectivamente, y en la base POD. De la misma manera, introducimos también las matrices $M_{y_{k_{i-1}}}^\psi$ y $M_{y_{k_i}}^\psi$.

Con las Proposiciones 5 y 6 podemos establecer el siguiente algoritmo, análogo al Algoritmo 3, considerando los esquemas en la base POD:

Input: N_n (Número de pasos para el método de Newton), N_t (Número de subintervalos del intervalo de tiempo), valor inicial y_0 , valor de u .

Output: Solución aproximada de (3.1) notada por W_{new}^ψ

- 1: Inicializar W_{old}^ψ resolviendo el problema de valor inicial lineal.
- 2: **while** ($k < N_n$ y criterio de parada) **do**
- 3: **for** $i = 1, 2$ **do**
- 4: Calcular $W_{new}^\psi(:, t_i)$ resolviendo (4.22)
- 5: **end i**
- 6: **for** $i = 3 : N_t$ **do**
- 7: Calcular $W_{new}^\psi(:, t_i)$ resolviendo (4.23)
- 8: Asignar $M_{u_{i-1}}^\psi = M_{u_i}^\psi$
- 9: Asignar $M_{y_{k_{i-1}}}^\psi = M_{y_{k_i}}^\psi$
- 10: **end i**
- 11: Asignar $W_{old}^\psi = W_{new}^\psi$.
- 12: **end k**
- 13: **return** W_{new}^ψ

Algoritmo 4: Crank Nicholson estabilizado, aplicado al esquema con la base POD

4.1.3 Construcción numérica del esquema en la base POD

En esta sección, vamos a explicar la aplicación del método POD a nuestro problema (2.1) y vamos a resolverlo numéricamente. El esquema de resolución se verá *reducido* en el sentido en que seremos capaces de resolver el sistema matricial del problema con matrices de tamaño significativamente menor y obtener una *buena*² aproximación de la solución.

Comenzamos resolviendo numéricamente el esquema (3.20) de la Sección 3.1.4. Según la triangularización de la malla que hayamos tomado y según la partición del intervalo de tiempo que hayamos hecho, la matriz solución W puede tener dimensiones muy grandes: tiene tantas filas como nodos en la malla y tantas columnas como instantes de tiempo considerados.

A continuación, vamos a elegir un número m de snapshots, que en nuestro caso, corresponden a m columnas de la matriz W . Tomamos m snapshots equidistantes, en los tiempos $t_k = k\Delta t$, con

²Ver resultados en las tablas de la Sección 4.2.

$k = 1, 2, \dots, m$, donde Δt el tamaño de paso correspondiente.

A partir de estos m vectores, vamos a construir la matriz Y_{snap} , de los snapshots. Siguiendo el esquema de [9], esta matriz se define como sigue.

Definición 8 (Matriz de Snapshots). *Sea $\Delta t = \frac{T}{m-1}$ el tamaño de paso en el tiempo entre los snapshots. Entonces, la Matriz de Snapshots tiene por elementos:*

$$Y_{snap}(:, j) = W(:, j), \quad j = 1, \dots, m,$$

y

$$Y_{snap}(:, j) = \bar{\partial}W(:, j - m), \quad j = m + 2, \dots, 2m,$$

donde

$$\bar{\partial}W(:, j) = \frac{1}{\Delta t} [W(:, j) - W(:, j - 1)].$$

Por construcción, todos los snapshots pertenecen a V .

Ahora, construimos la Matriz de Correlación C definida en (4.3). Dado que hemos obtenido la matriz Y_{snap} numéricamente, no contamos con las funciones y_j sino con aproximaciones numéricas de ellas, por lo que no podemos calcular las entradas exactas de la matriz C . Entonces, calculamos una aproximación de la matriz C como se propone en [12]:

$$C \approx Y_{snap}^T M Y_{snap}, \quad (4.24)$$

donde Y_{snap}^T es la matriz transpuesta de la matriz Y_{snap} y M es la matriz de masa en la base de elementos finitos.

Para obtener los valores y vectores propios de la matriz C , utilizamos el comando `eig`, el cual nos devuelve una matriz diagonal con los valores propios λ_k de C y una matriz cuyas columnas corresponden a los vectores propios v_k correspondientes.

Utilizando (4.14), calculamos el valor de d según el porcentaje de *información* del sistema que deseemos guardar. Una vez determinado el número d de valores y vectores propios asociados que vamos a utilizar, construimos la matriz Ψ , cuya k -ésima columna está dada por el vector correspondiente a la aproximación de la función test ψ_k definida en (4.13), de la siguiente forma:

$$\Psi(:, k) = \frac{1}{\sqrt{\lambda_k}} Y_{snap} v_k. \quad (4.25)$$

Ahora, vamos a ensamblar las matrices dadas por (4.17), (4.18), (4.19) y (4.20). Como ya se indicó, la Matriz de Masa en la base POD está dada por la matriz de Identidad de tamaño d :

$$M^\psi = I \in \mathbb{R}^{d \times d}.$$

Mientras que la Matriz de Rigidez se aproxima de la siguiente manera (ver [9]):

$$K^\psi \approx \Psi^T K \Psi,$$

donde K es la matriz de rigidez en la base de los elementos finitos.

Por otro lado, calcular las coordenadas de las matrices M_u^ψ y $M_{y_k}^\psi$ requiere de un trabajo numérico extra, puesto que hemos obtenido las funciones ψ mediante una aproximación numérica; es decir, no contamos con funciones continuas sino con vectores correspondientes a la aproximación de ellas en los determinados nodos de la malla y en cierto instante de tiempo.

En este caso, ya no podemos usar el comando `assema` de Matlab, puesto que funciona con elementos lineales y a soporte compacto; pero en el sistema dinámico que estamos estudiando, no tratamos con elementos lineales necesariamente, y en general, tampoco son a soporte compacto, sino que están definidos sobre todo el dominio. Esto es beneficioso puesto que ya no tratamos con información local, sino que esta nueva base ortonormal guarda más información del sistema dinámico. Sin embargo, nos resulta una dificultad puesto que no podemos calcular analíticamente los valores de las integrales

$$\int_{\Omega} u(x, t) \psi_j(x) \psi_l(x) dx \quad \text{y} \quad \int_{\Omega} y_k(x, t) \psi_j(x) \psi_l(x) dx.$$

Para subsanar este problema, vamos a aproximar el valor del producto en $L^2(\Omega)$ aplicando la *fórmula prismoidal* utilizada en [8], que se basa en la interpolación lineal de los nodos de las funciones base del POD con los baricentros de los triángulos de los elementos finitos.

Supongamos que tenemos un número N_ρ de triángulos que aproximan Ω . Siguiendo la notación de [8], llamamos x_{ρ_ν} al centro de masa y A_{ρ_ν} al área del triángulo correspondiente ρ_ν , con $\nu = 1, \dots, N_\rho$. Entonces, podemos encontrar un valor aproximado de la siguiente integral:

$$\int_{\Omega} \psi_j(x) \psi_l(x) \psi_\mu(x) dx \approx \sum_{\nu=1}^{N_\rho} A_{\rho_\nu} y_k(x_{\rho_\nu}) \psi_j(x_{\rho_\nu}) \psi_l(x_{\rho_\nu}) \psi_\mu(x_{\rho_\nu}). \quad (4.26)$$

Consideramos la aproximación de las funciones y_k y u con respecto a la base POD:

$$y_k(x, t) \approx \sum_{\mu=1}^d y_k^\mu(t) \psi_\mu(x), \quad u(x, t) \approx \sum_{\mu=1}^d u^\mu(t) \psi_\mu(x). \quad (4.27)$$

para ciertas funciones $y_k^\mu(t)$, $u^\mu(t)$ con $\mu = 1, \dots, d$. Con (4.26) y (4.27), se puede aproximar el valor de las integrales dadas en (4.19) y (4.20) como sigue:

$$\begin{aligned} \int_{\Omega} y_k(x, t) \psi_j(x) \psi_l(x) dx &\approx \int_{\Omega} \left(\sum_{\mu=1}^d y_k^\mu(t) \psi_\mu(x) \right) \psi_j(x) \psi_l(x) dx \\ &= \sum_{\mu=1}^d y_k^\mu(t) \int_{\Omega} \psi_\mu(x) \psi_j(x) \psi_l(x) dx \\ &\approx \sum_{\mu=1}^d y_k^\mu(t) \sum_{\nu=1}^{N_\rho} A_{\rho_\nu} \psi_\mu(x_{\rho_\nu}) \psi_j(x_{\rho_\nu}) \psi_l(x_{\rho_\nu}) \end{aligned}$$

y

$$\int_{\Omega} u(x, t) \psi_j(x) \psi_l(x) dx \approx \int_{\Omega} \left(\sum_{\mu=1}^d u^\mu(t) \psi_\mu(x) \right) \psi_j(x) \psi_l(x) dx$$

$$\begin{aligned}
&= \sum_{\mu=1}^d u^\mu(t) \int_{\Omega} \psi_\mu(x) \psi_j(x) \psi_l(x) dx \\
&\approx \sum_{\mu=1}^d u^\mu(t) \sum_{\nu=1}^{N_p} A_{\rho_\nu} \psi_\mu(x_{\rho_\nu}) \psi_j(x_{\rho_\nu}) \psi_l(x_{\rho_\nu})
\end{aligned}$$

con $t \in [0, T]$. Por otro lado, las componentes del vector \tilde{w}_k están dados por (y_k^2, ψ_j) . Si consideramos la aproximación de y_k

$$y_k(x, t) \approx \sum_{\mu=1}^d y_k^\mu(t) \psi_\mu(x),$$

entonces las componentes de \tilde{w}_k quedan dadas como sigue

$$\begin{aligned}
(y_k^2, \psi_l) &\approx \left(\left(\sum_{\mu=1}^d y_k^\mu(t) \psi_\mu(x) \right)^2, \psi_l(x) \right) \\
&= \left(\sum_{\mu, j=1}^d y_k^\mu(t) \psi_\mu(x) y_k^j(t) \psi_j(x), \psi_l(x) \right) \\
&= \int_{\Omega} \sum_{\mu, j=1}^d y_k^\mu(t) y_k^j(t) \psi_\mu(x) \psi_j(x) \psi_l(x) dx \\
&= \sum_{\mu, j=1}^d y_k^\mu(t) y_k^j(t) \int_{\Omega} \psi_\mu(x) \psi_j(x) \psi_l(x) dx,
\end{aligned}$$

para $l = 1, \dots, d$. Luego, la integral $\int_{\Omega} \psi_\mu(x) \psi_j(x) \psi_l(x) dx$ se calcula como en (4.26), utilizando la fórmula prismoidal.

Nótese que para calcular el valor de estas integrales, es necesario además aproximar los valores de las funciones y_k y u con respecto a la base POD, para $t \in [0, T]$. Para ello, es necesario realizar un *cambio de base* de estas funciones en cada tiempo t , utilizando la proyección de Galerkin. Llamemos y_k^h y u^h , respectivamente, a las aproximaciones de dichas funciones en la base de elementos finitos. Entonces, recordando (4.18), las aproximaciones de éstas funciones en la base POD está dada por:

$$y_k(x, t) \approx (y_k^h)^T M \Psi, \quad u(x, t) \approx (u^h)^T M \Psi,$$

donde T denota la transpuesta del vector.

4.2 Experimentos Numéricos para el Método POD

La reducción del Modelo de Fisher se realiza sobre una primera resolución numérica del sistema, utilizando el algoritmo 3 sobre la malla dada por la figura 3.2; con lo que obtenemos la matriz W , que corresponde a la aproximación de la solución del problema (2.1).

Para la experimentación numérica utilizamos el toolbox de ecuaciones diferenciales parciales de MATLAB para ensamblar las matrices de masa y de rigidez del método de los elementos finitos asociados al problema.

Para efectos experimentales, los ejemplos fueron resueltos en un servidor HP con 2 procesadores XEON X5670, con sistema Linux CentOS 64bit con 96GB de memoria RAM.

Aplicamos el método POD al Ejemplo 3.2 considerando una discretización mucho más fina, lo cual genera un problema matricial de tamaño mucho mayor, para observar cómo funciona. Para ello, refinamos la malla dada por la figura 3.2 hasta obtener 16384 triángulos ($h \approx 0,016$).

De la aproximación numérica obtenida W , tomamos una cantidad de snapshots igual al número de columnas de la matriz W . En este caso, elegimos los snapshots con un intervalo de tiempo Δt igual al tamaño de paso dado por la partición del intervalo $[0, T]$, como en la Definición 8; de manera que tomamos todas las observaciones obtenidas tiempo a tiempo.

A partir de ellas, construimos la matriz de correlación como en (4.24), y obtenemos los valores propios de la misma. En la siguiente figura, se muestra el rápido decrecimiento de los mismos, lo cual demuestra cómo la información del problema se condensa en los primeros valores, permitiéndonos aplicar el método POD con una base bastante pequeña y, por tanto, resolver el mismo sistema con un número significativamente menor de variables.

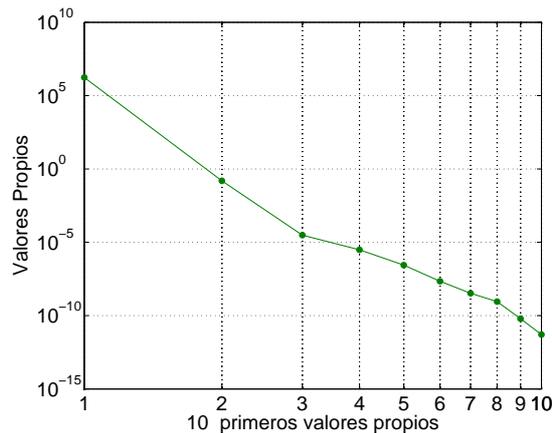


Figura 4.1: Ejemplo 3.2, 10 primeros valores propios de la matriz de correlación. Con $h \approx 0,016$ y $N_t = 800$. (Escala logarítmica).

Para aplicar el método POD tomamos $d = 6$ lo cual, de acuerdo al indicador I_d , corresponde a guardar un 99,99%³ de la información del problema. Luego, construimos la base POD como en (4.25), cuyos elementos construidos quedan como sigue:

³La cantidad de información guardada es muy alta incluso tomando pocos valores propios de la matriz de correlación debido a que estamos resolviendo un problema bastante regular. Más adelante, en el capítulo de la experimentación numérica observaremos un ejemplo en el que es más notoria la diferencia de la información guardada al tomar pocos o muchos valores propios.

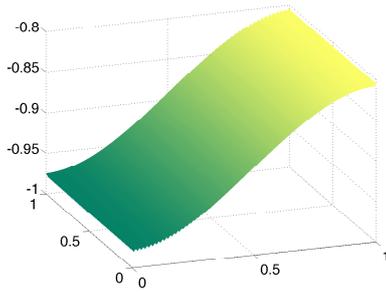


Figura 4.2: Primer elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

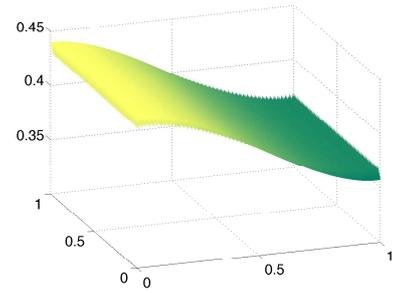


Figura 4.3: Segundo elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

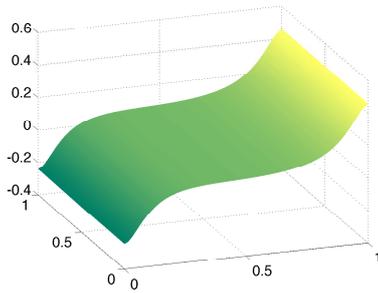


Figura 4.4: Tercer elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

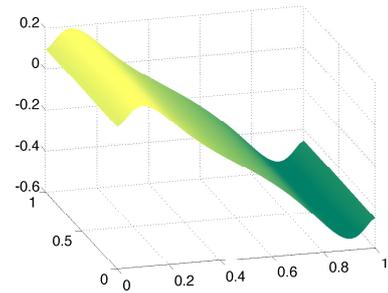


Figura 4.5: Cuarto elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

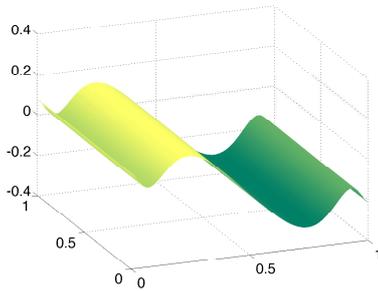


Figura 4.6: Quinto elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

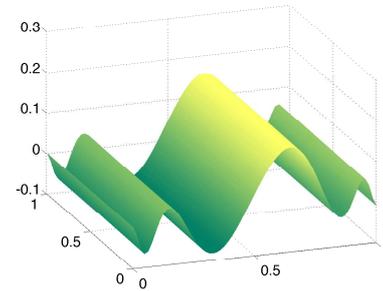


Figura 4.7: Sexto elemento de la base POD, con $h = 0,016$ y $N_t = 800$.

Para efectos de comparación, resolvemos el sistema aplicando el Algoritmo 3 (el método de CN estabilizado) y aplicamos el método POD variando el valor del tamaño de paso temporal: $\Delta t = 1/300, 1/500, 1/800$ y el número de elementos de la base POD: $d = 2, 4, 6$.

La Tabla 4.1 contiene los valores de la suma de los errores de aproximación del método POD en la norma $L^2(\Omega)$; ver (4.2), pág. 41. Estos valores resultan de comparar los resultados obtenidos al aplicar el método POD con los dos métodos de integración numérica: el método de Euler Implícito y el de Cranck Nicholson estabilizado, a los que nos referimos por POD+EI y POD+CN, respectivamente. La Figura 4.8 muestra estos datos en función de Δt , correspondientes a los valores: $h \approx 0,016$ y $d = 6$.

Se puede observar que los errores decrecen conforme se elige un número mayor de d y de N_t , pues esto significa guardar una mayor cantidad de información del modelo y tener una discretización más fina del intervalo temporal, respectivamente.

| | $\Delta t = 1/300$ | | $\Delta t = 1/500$ | | $\Delta t = 1/800$ | |
|---------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | POD + EI | POD + CN | POD + EI | POD + CN | POD + EI | POD + CN |
| $d = 2$ | 0,03891 | 0,04206 | 0,06890 | 0,07215 | 0,11380 | 0,11709 |
| $d = 4$ | $5,9202 \times 10^{-5}$ | $1,0331 \times 10^{-6}$ | $2,7536 \times 10^{-5}$ | $2,4005 \times 10^{-6}$ | $7,3445 \times 10^{-6}$ | $7,7808 \times 10^{-6}$ |
| $d = 6$ | $7,1312 \times 10^{-5}$ | $2,3333 \times 10^{-6}$ | $4,4363 \times 10^{-5}$ | $3,0058 \times 10^{-6}$ | $2,8551 \times 10^{-5}$ | $3,8413 \times 10^{-6}$ |

Tabla 4.1: Ejemplo 3.2, errores de aproximación Esquema Original vs. Esquema reducido (POD+EI, POD+CN).

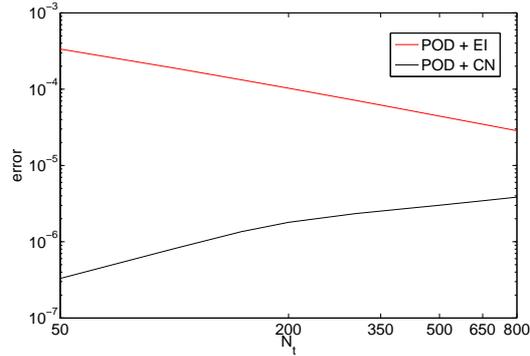


Figura 4.8: Errores de aproximación de los esquemas reducidos (POD+EI y POD+CN) en función de $\Delta t = 1/N_t$, con $d = 6$ y $h \approx 0,016$; en escala logarítmica.

La figura 4.9 corresponde a la comparación de los errores de aproximación en norma $L^2(\Omega)$ al mantener $h \approx 0,016$ y fijar $d = 6$ y $N_t = 800$. En ella se observa cómo el método POD+CN tiene un error de aproximación menor que el método POD+EI, lo cual es esperable dado que el método de CN es de orden 2:

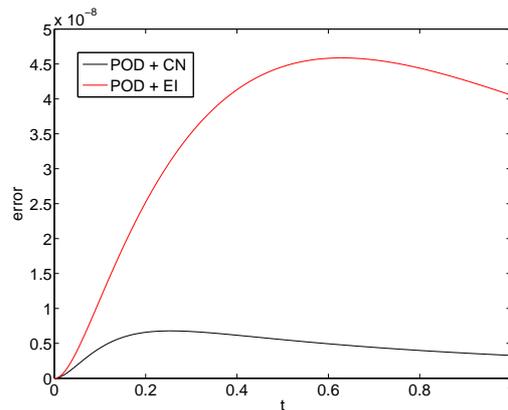


Figura 4.9: Errores de aproximación del esquema reducido + EI vs. esquema reducido + CN.

Al fijar el valor $N_t = 800$, y variando los valores de d , se obtienen los gráficos dados en las siguientes Figuras 4.10 y 4.11, en las que se comparan los errores de aproximación al aplicar el método POD con los distintos métodos de integración numérica: EI y CN.

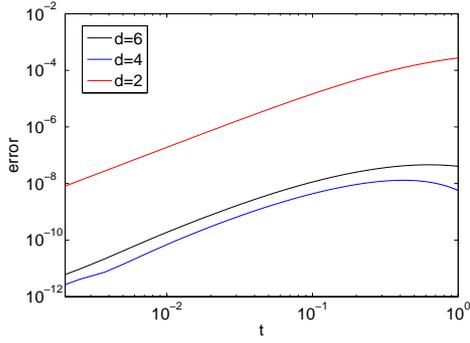


Figura 4.10: Ejemplo 3.2: errores de aproximación del método reducido tomando $d = 2, 4, 6$ + método EI, en escala logarítmica.

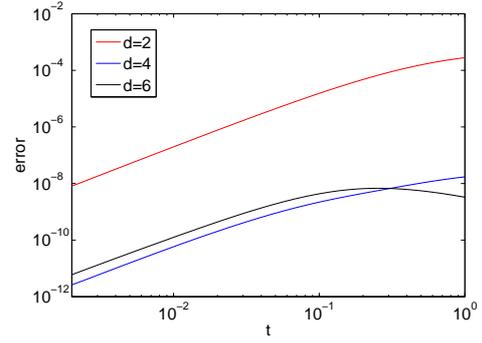


Figura 4.11: Ejemplo 3.2: errores de aproximación del método reducido tomando $d = 2, 4, 6$ + método CN, en escala logarítmica.

Por otro lado, fijando el valor $d = 6$, y variando los valores de N_t , se obtienen los gráficos dados en las Figuras 4.12 y 4.13, que comparan también los errores de aproximación en norma $L^2(\Omega)$ al aplicar el método POD + EI o POD + CN.

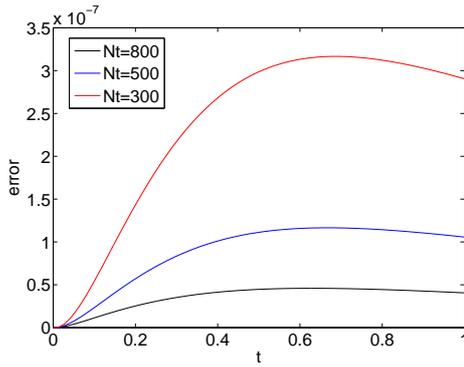


Figura 4.12: Ejemplo 3.2: errores de aproximación del método reducido tomando $N_t = 300, 500, 800$ + método EI.

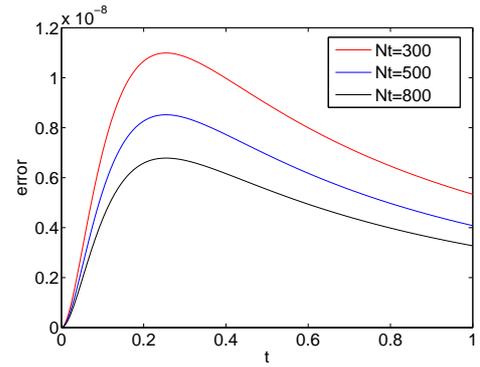


Figura 4.13: Ejemplo 3.2: errores de aproximación del método reducido tomando $N_t = 300, 500, 800$ + método CN.

Ahora, vamos a analizar un ejemplo distinto a los Ejemplos 3.1 y 3.2 para ver el comportamiento del método de resolución. En este ejemplo vamos a considerar una población cuyo estado inicial representa un foco de concentración poblacional.

Ejemplo 4.1. Tomamos $\Omega = (0, 1) \times (0, 1)$, $x = (x_1, x_2) \in \Omega$ y las siguientes cantidades:

$$\cdot y_0(x_1, x_2, 0) = \exp\left(-\frac{(x_1 - 0,5)^2 + (x_2 - 0,5)^2}{0,02}\right) + 5;$$

$$\cdot u(x, t) = r - \frac{r}{\kappa}y(x, t) + \frac{\Delta y}{y} - \frac{y_t}{y}, \quad y \neq 0$$

$$\cdot r = 1, \quad \kappa = 1, \quad T = 1.$$

Además, vamos a seguir trabajando sobre la malla dada por la Figura 3.1, refinada utilizando el comando `refinemesh` hasta obtener 16384 triángulos que dan lugar al valor $h \approx 0,016$.

Las siguientes figuras corresponden a los gráficos de las soluciones aproximadas por el Algoritmo 3, tomando $N_t = 800$. En ellas se puede observar un fenómeno difusivo cuyo estado inicial representa una concentración alta de la población en una zona relativamente pequeña del espacio considerado:

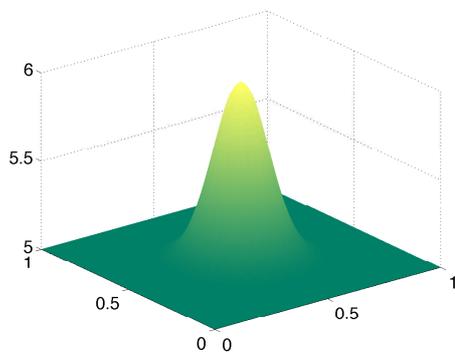


Figura 4.14: Ejemplo 4.1, Solución Aproximada con método CN estabilizado, al tiempo $t = 0$.

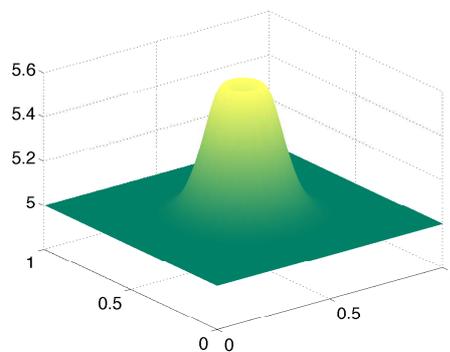


Figura 4.15: Ejemplo 4.1, Solución Aproximada con método CN estabilizado, al tiempo $t = 1$.

Utilizando las aproximaciones obtenidas en cada tiempo, construimos la matriz de correlación y obtenemos sus valores propios:

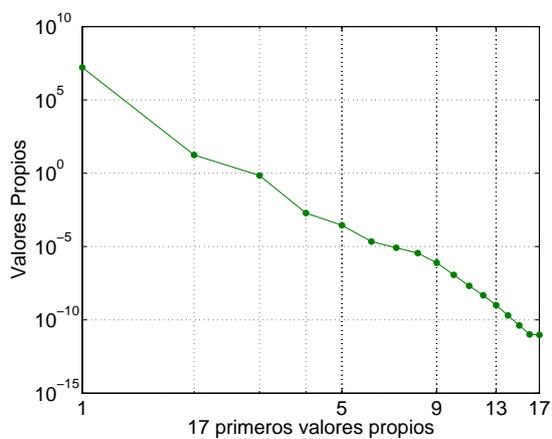


Figura 4.16: Ejemplo 4.1, 17 primeros valores propios de la matriz de correlación. (Escala logarítmica).

Las Figuras 4.17, 4.18, 4.19 y 4.20 corresponden a algunos elementos de la base POD que se construyeron al tener $h \approx 0,016$ y al haber elegido $d = 5$ y $N_t = 800$.

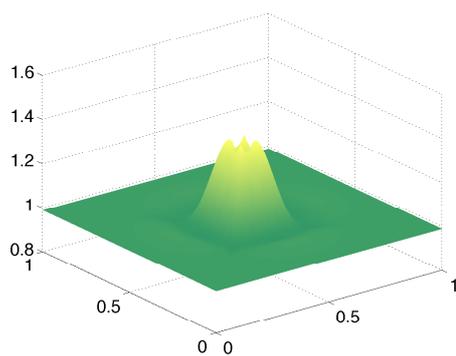


Figura 4.17: Primer elemento de la base POD, para el Ejemplo 4.1.

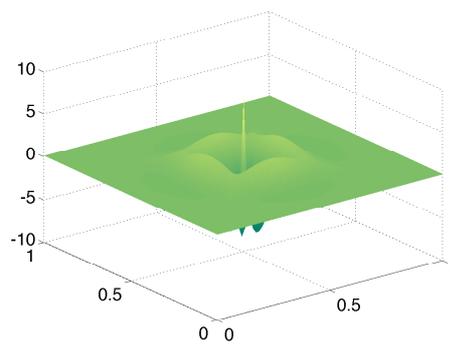


Figura 4.18: Segundo elemento de la base POD, para el Ejemplo 4.1.

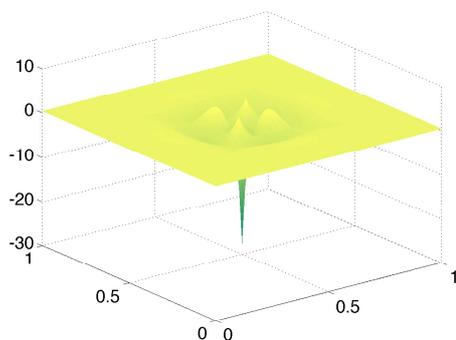


Figura 4.19: Tercer elemento de la base POD, para el Ejemplo 4.1.

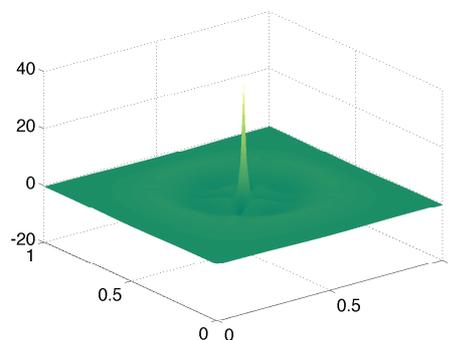


Figura 4.20: Quinto elemento de la base POD, para el Ejemplo 4.1.

Análoga a la Tabla 4.1, la Tabla 4.2 representa los datos tomados al tomar distintos valores de Δt y d . Como este ejemplo es más complejo que el anterior, se obtiene un mayor número de valores propios significativos de la matriz de correlación C . De manera que es necesario tomar un número mayor de elementos de la base POD para construir el esquema reducido que, sin embargo, sigue siendo pequeño en comparación al número de bases del método de elementos finitos.

| | $\Delta t = 1/300$ | | $\Delta t = 1/500$ | | $\Delta t = 1/800$ | |
|----------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | POD + EI | POD + CN | POD + EI | POD + CN | POD + EI | POD + CN |
| $d = 3$ | 0.30361 | 0.30396 | 0.49160 | 0.49220 | 0.74820 | 0.74914 |
| $d = 5$ | 0.15493 | 0.15516 | 0.26123 | 0.26165 | 0.42421 | 0.42491 |
| $d = 10$ | $7,9675 \times 10^{-5}$ | $8,7377 \times 10^{-5}$ | $2,1388 \times 10^{-4}$ | $2,3070 \times 10^{-4}$ | $5,0535 \times 10^{-4}$ | $5,3845 \times 10^{-4}$ |

Tabla 4.2: Ejemplo 4.1, errores de aproximación Esquema Original vs. Esquema reducido (POD+EI, POD+CN).

La figura 4.21 representa los valores del error en función de Δt .

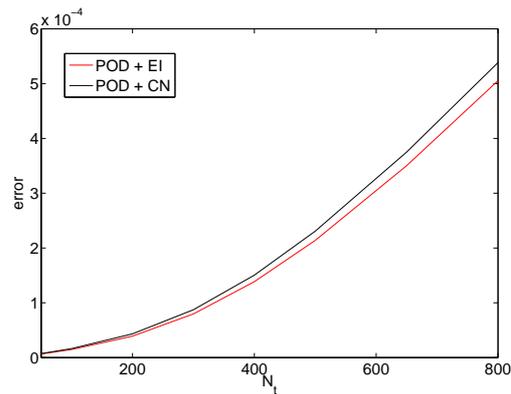


Figura 4.21: Errores de aproximación de los esquemas reducidos (POD+EI y POD+CN) en función de $\Delta t = 1/N_t$, con $d = 10$ y $h \approx 0,016$; en escala logarítmica.

La figura 4.22 corresponde a la comparación de los errores de aproximación en norma $L^2(\Omega)$ al fijar $d = 10$ y $N_t = 800$, en la que se observa que los errores son menores al utilizar el método reducido (POD) aplicando CN, en comparación con el método reducido con EI :

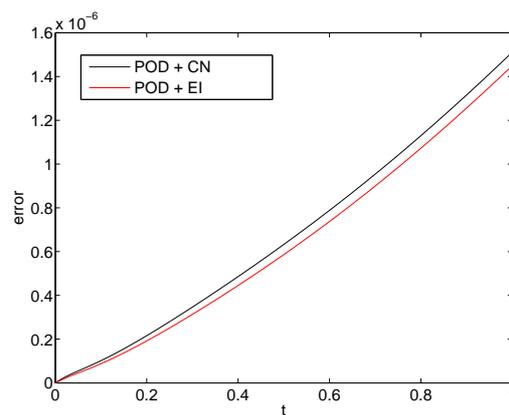


Figura 4.22: Errores de aproximación del esquema reducido + EI vs. esquema reducido + CN.

Los siguientes gráficos corresponden a la comparación de los errores de aproximación en norma $L^2(\Omega)$, variando los valores d y N_t . En ellos se observa cómo se obtienen menores errores de aproximación conforme se toman mayores valores de d y de N_t , a la vez que los errores de aproximación son menores al utilizar el método POD+CN que al utilizar el método POD+EI:

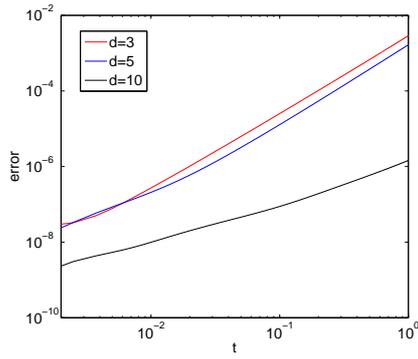


Figura 4.23: Ejemplo 4.1, errores de aproximación del método reducido fijando $N_t = 800$ y tomando $d = 3, 5, 10$ + método EI.

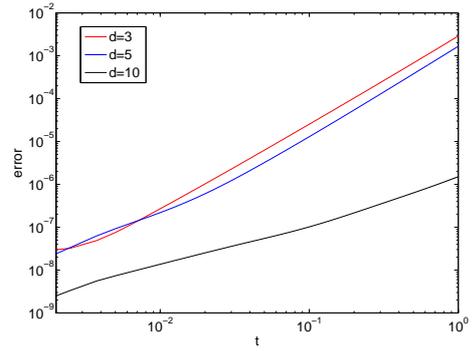


Figura 4.24: Ejemplo 4.1, errores de aproximación del método reducido fijando $N_t = 800$ y tomando $d = 3, 5, 10$ + método CN.

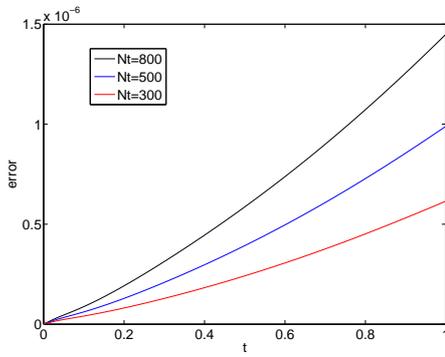


Figura 4.25: Ejemplo 4.1, errores de aproximación del método reducido fijando $d = 10$ y tomando $N_t = 300, 500, 800$ + método EI.

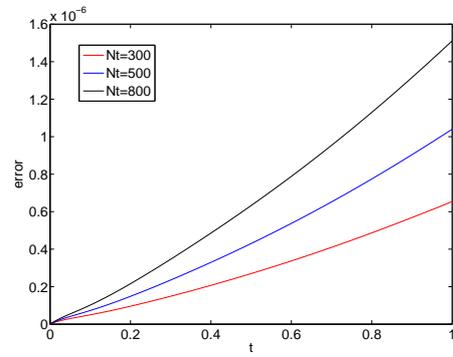


Figura 4.26: Ejemplo 4.1, errores de aproximación del método reducido fijando $d = 10$ y tomando $N_t = 300, 500, 800$ + método CN.

4.3 Experimentación Numérica y Simulación del Modelo de Fisher

Para observar los resultados de la aplicación del método sobre el modelo, hemos construido un ejemplo que, a pesar de no ajustarse por completo a las hipótesis teóricas, es un poco más aplicable ya que permite predecir el comportamiento de una población no diferenciada que crece conforme pasa el tiempo, y que se dispersa sobre el territorio ecuatoriano de manera uniforme. Sin embargo, para realizar estudios más realistas se podría considerar la existencia de un flujo a través de las fronteras, por ejemplo, considerando condiciones de frontera de Robin. Para la simulación, consideramos el dominio $\Omega \subset \mathbb{R}^2$, discretizado dado en la Figura 4.27.

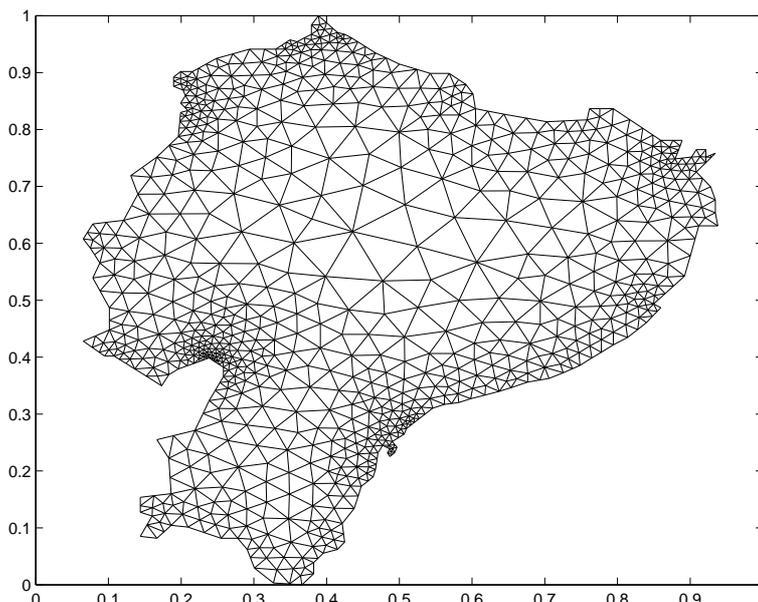


Figura 4.27: Mallado sobre un dominio que representa al mapa del Ecuador; $h \approx 0,096$.

Sean $x = (x_1, x_2) \in \Omega$, $t \in [0, 2]$. La población inicial $y_0 = y(x, 0)$ viene dada por la siguiente función, que representa tres focos de condensación poblacional que pueden verse claramente en la Figura 4.28.

$$y(x_1, x_2, 0) = \begin{cases} x_1 + x_2 & \text{si } \|(x_1, x_2) - (0,3,0,8)\|_{\mathbb{R}^2} \leq 0,1, \\ & \text{o } \|(x_1, x_2) - (0,4,0,3)\|_{\mathbb{R}^2} \leq 0,1, \\ 0 & \text{caso contrario.} \end{cases}$$

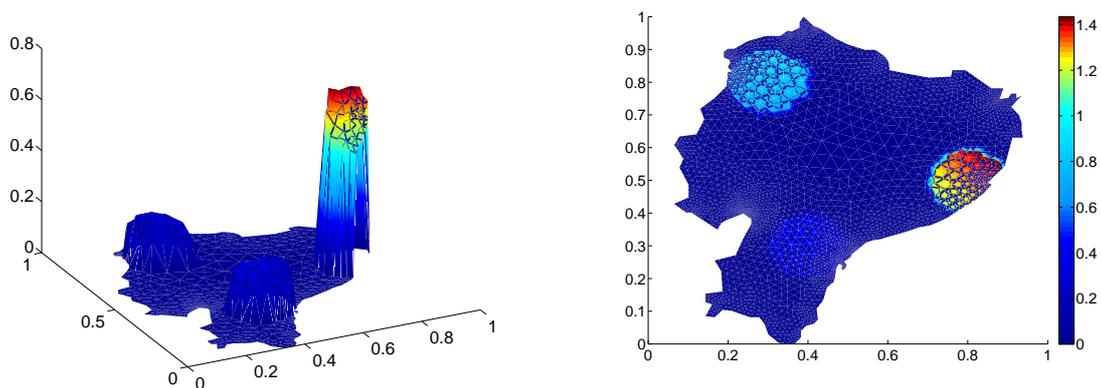


Figura 4.28: Simulación, estado inicial y . Gráficos de la función $y_0 = y(x, 0)$.

Sin ánimo de incurrir en una rigurosidad geográfica, dividimos el mapa ecuatoriano en tres regiones que representan a: la Costa (si $0 \leq x_1 < 0,3$), la Sierra (si $0,3 \leq x_1 < 0,7$) y la Amazonía (si $0,7 \leq x_1 \leq 1$). Consideramos una población que crece en función de la humedad presente en el ambiente, la cual a su vez depende del tiempo.

La función $u(x, t)$ representa la capacidad de suprimir individuos según la humedad del ambiente y el paso del tiempo. Dicha función se construye tomando valores constantes 1, 2 y 3 según la región: a menor humedad, mayor capacidad de suprimir individuos; por ejemplo, el valor más alto se encuentra en la Sierra, ya que es el lugar en el que el crecimiento poblacional se podría controlar mejor.

Además, se considera la humedad dependiente del tiempo, asumiendo que la humedad crece linealmente (en la Sierra), cuadráticamente (en la Amazonía) y cúbicamente (en la Costa) a lo largo del año asumiendo que el Ecuador tiene 2 momentos: *verano* e *invierno*.

La Figura 4.29 corresponde a los gráficos del porcentaje de humedad dependiente del tiempo. Aunque estas suposiciones no son del todo realistas, queremos experimentar sobre el modelo con situaciones que podrían presentarse en la práctica.

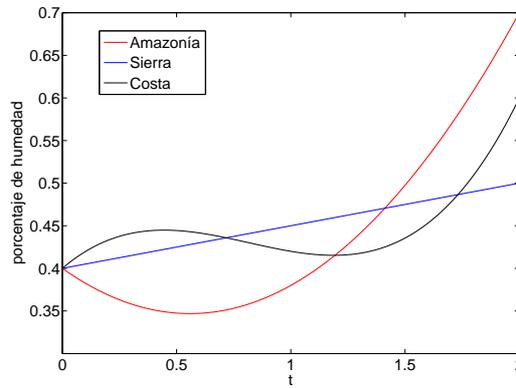


Figura 4.29: Función $u(\cdot, t)$, que representa la humedad presente en las tres regiones del dominio. La función u varía con respecto al tiempo pero es constante a trozos en el espacio (ver Fig. 4.30).

Con todo esto, la función u queda representada en la Figura 4.30 y definida como sigue:

$$u(x_1, x_2, t) = \begin{cases} 1(-0,145 t^3 + 0,3551 t^2 - 0,2301 t + 0,6), & \text{si } 0 \leq x_1 < 0,3 \\ 3(-0,05 t + 0,5), & \text{si } 0,3 \leq x_1 < 0,7 \\ 2(-0,17 t^2 + 0,19 t + 0,7), & \text{si } 0,7 \leq x_1 \leq 1. \end{cases}$$

Los parámetros del modelo de Fisher que tomamos para la experimentación son los siguientes:

$$\gamma = 0,5 \quad r = 1 \quad \kappa = 10 \quad T = 2.$$

Para comenzar, resolvemos el problema utilizando el Algoritmo 3. Construimos la matriz de snapshots Y_{snap} tomando los snapshots de la misma manera que en el ejemplo anterior: tomamos todas las aproximaciones $W(:, i)$ donde cada i corresponde a un nodo de la partición temporal y construyendo a partir de ellas los elementos $\bar{\partial}W$ como en la Definición 8, pág. 49. Luego, construimos la matriz de correlación

$$C = Y_{snap}^T M Y_{snap},$$

y obtenemos sus vectores y valores propios. Con ellos construimos los elementos de la base POD y resolvemos el esquema reducido.

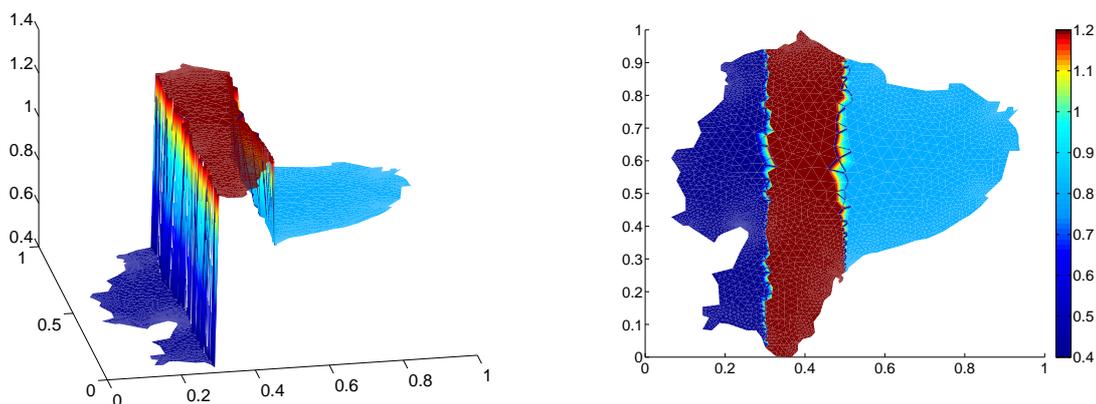


Figura 4.30: Gráfico de la función u evaluada $t = 0$.

La Figura 4.31 muestra los errores de aproximación obtenido al aplicar el método POD a este problema con $d = 10$, sobre el mallado del dominio dado en la Figura 4.27 y $N_t = 400$:

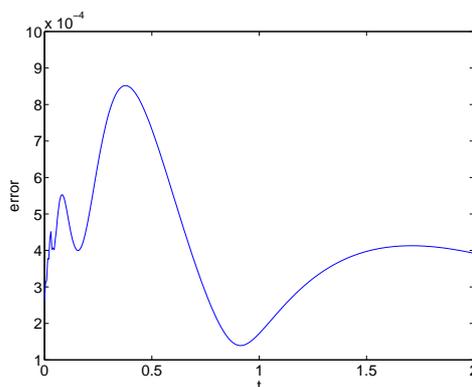


Figura 4.31: Simulación, errores de aproximación del método CN y el método POD + CN, con $d = 10$, $h \approx 0,096$ y $N_t = 400$.

Dado que estamos estudiando un método de reducción del modelo de Fisher, es de interés observar los resultados al considerar un problema muy grande, en términos del costo computacional. Para ello, refinamos la malla que representa al mapa del Ecuador hasta obtener 198 641 nodos que dieron lugar a 393 728 triángulos, con $h \approx 0,006$; y tomamos 1 000 subintervalos del intervalo temporal.

Procedemos de nuevo a resolver el esquema original y obtenemos los snapshots con los cuales construimos la matriz de correlación. El notorio decrecimiento de los valores propios de la misma se puede observar en la Figura 4.32.

Recordemos que los valores propios de la matriz C representan la cantidad de información del problema guardada. En este caso, al ser un problema más complejo, es necesario tomar un mayor número d de valores y vectores propios para construir la base POD de manera que se obtenga una mejor aproximación a la solución.

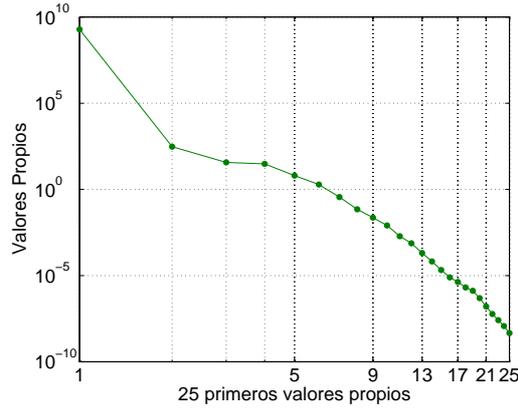


Figura 4.32: Gráfico de los valores propios de la matriz de correlación C , en escala logarítmica.

Al tomar $d = 13$, el porcentaje de información del sistema es casi el 100%:

$$I_d = \frac{\sum_{j=1}^{13} \lambda_j}{\sum_j \lambda_j} = \frac{1,902052875774511 \times 10^9}{1,902052875774615 \times 10^9} = 0,999999999999945.$$

Nótese que tomamos apenas 3 valores propios más y, a pesar de que la discretización es más fina, sigue habiendo un gran porcentaje de información guardado.

Las Figuras 4.33, 4.34, 4.35 y 4.36 muestran los primeros cuatro elementos de la base POD construida.

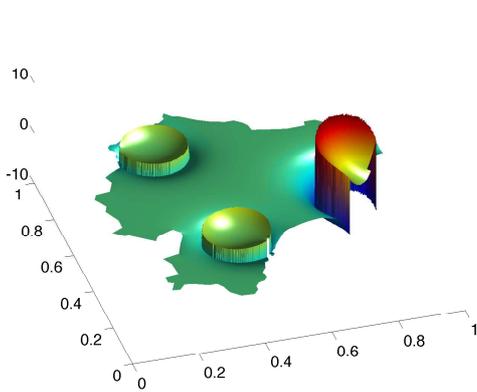


Figura 4.33: Primer elemento de la base POD, asociado al valor propio $\lambda_1 \approx 1,90205249981 \times 10^9$.

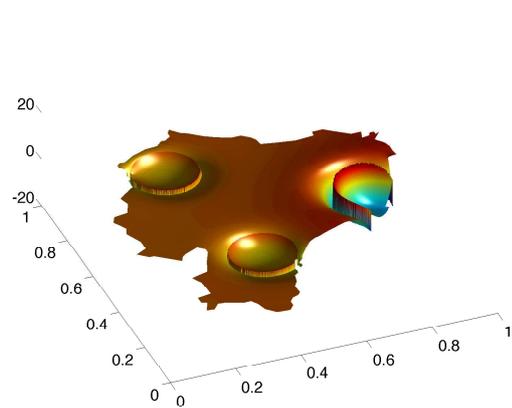


Figura 4.34: Segundo elemento de la base POD, asociado al valor propio $\lambda_2 \approx 3,00170 \times 10^2$.

Al resolver el sistema reducido, se obtuvo una solución aproximada de tamaño 13×2001 , la cual ocupa 208,104 KB y que fue resuelta en aproximadamente 15 minutos; mientras que la solución aproximada obtenida utilizando el esquema de EF es de tamaño 198641×2001 , lo cual ocupa aproximadamente 3,1798 GB y fue resuelta en aproximadamente 10 horas, tomando un tiempo aproximado de una hora para cada iteración del método de Newton. Con este experimento se confirma la reducción del costo computacional que genera el aplicar el método POD.

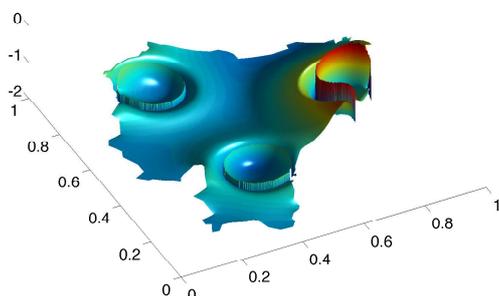


Figura 4.35: Tercer elemento de la base POD, asociado al valor propio $\lambda_3 \approx 36,80345$.

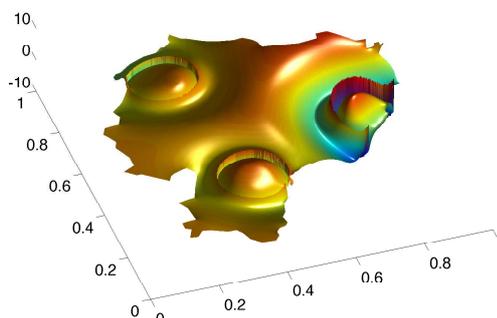


Figura 4.36: Cuarto elemento de la base POD, asociado al valor propio $\lambda_4 \approx 30,28986$.

Este método permite guardar la información del sistema y resolver el problema utilizando un número significativamente menor de variables, pero una duda salta a la vista: ¿cómo nos beneficia esto si para construir la base POD es necesario resolver el esquema utilizando el MEF? Es decir, queremos saber si el método POD es útil para obtener resultados coherentes al perturbar los datos del problema y evitar el tener que primero resolverlo primero mediante un método computacionalmente *costoso*.

Para responder esta pregunta, y pensando en una futura aplicación en el área del Control Óptimo, consideramos una función $\tilde{u}(x, t) = u(x, t) + \varepsilon$, una versión perturbada de la función $u(x, t)$. Tomamos $0 \leq \varepsilon$, y resolvemos el esquema reducido (4.21), con la base POD obtenida al resolver el problema con $u = u(x, t)$.

Consideramos la malla dada en la Figura 4.27 y fijamos $N_t = 400$. Por motivos de comparación, resolvemos también el problema con el esquema dado por el MEF (usando el Algoritmo 3). La Figura 4.37 muestra los errores de aproximación en norma $L^2(\Omega)$ entre el sistema *perturbado* resuelto por el método POD y el sistema perturbado resuelto por MEF, tomando $\varepsilon = 0,001$ y $\varepsilon = 0,1$.

Finalmente, en la Figura 4.38, se pueden observar los errores de aproximación al considerar el problema perturbado tomando un valor de perturbación relativamente alto: $\varepsilon = 0,7$. Se observa que, a pesar de perturbar la función u con un mayor valor, el error de aproximación sigue siendo pequeño.

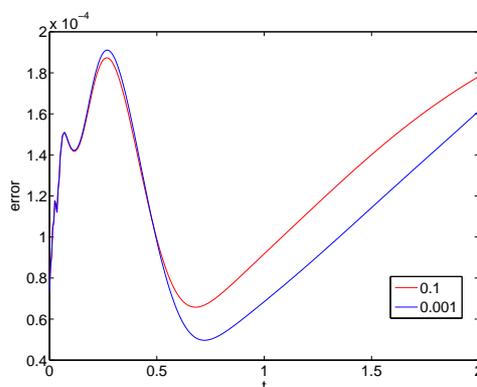


Figura 4.37: Simulación, gráfico de los errores de aproximación obtenidos al aplicar el método POD+CN al problema perturbado, tomando $\varepsilon = 0,1$ y $\varepsilon = 0,001$.

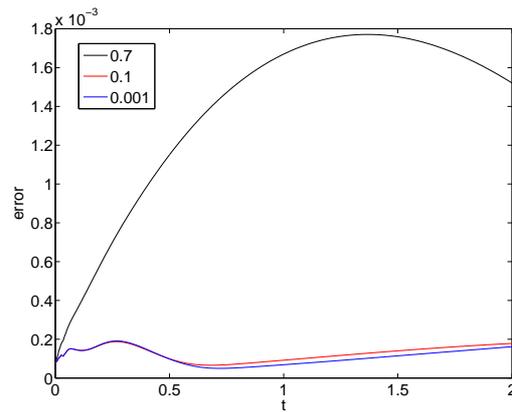


Figura 4.38: Simulación, gráfico de los errores de aproximación obtenidos al aplicar el método POD+CN al problema perturbado, tomando $\varepsilon = 0,1$, $\varepsilon = 0,001$ y $\varepsilon = 0,7$.

Con todo esto, se puede observar que el método POD es útil para obtener resultados coherentes en problemas parecidos al problema original, resolviendo el sistema perturbado utilizando el método POD y aprovechando la información obtenida al resolver el sistema original y la reducción del costo computacional de resolución que ofrece el método POD.

Conclusiones y Recomendaciones

Para escribir este proyecto de titulación fue necesario comprender cómo está establecido el modelo de Fisher para el crecimiento poblacional sobre un espacio de dos dimensiones y a través del tiempo. A pesar de ser sencillo, este modelo sirve de base para el estudio de modelos más complejos. Resolver numéricamente el problema de valor inicial nos permitió encontrar una solución aproximada que puede aplicarse a problemas reales como el de recolección o aniquilación de individuos de una población.

La ecuación de Fisher es de tipo parabólica semilineal, de manera que fue necesario linealizar el problema aplicando el método de Newton para poder aplicar el método de Elementos Finitos. En un primer momento, se resolvió el esquema obtenido utilizando el método de Euler Implícito, pero los resultados obtenidos no fueron satisfactorios: era necesaria una discretización muy fina de la variable temporal para obtener errores de aproximación que decrezcan en cada iteración, y esto provocó un notorio crecimiento del costo computacional asociado a la resolución del sistema. De manera que fue necesario aplicar un método de resolución de ecuaciones diferenciales ordinarias más estable: el método de Crank-Nicholson, con el cual se obtuvieron mejores resultados ya que el error se redujo considerablemente conforme se realizaban refinamientos de la malla y utilizando un número aceptable de nodos temporales.

A partir de la aproximación a la solución obtenida, y como punto central de este proyecto de titulación, aplicamos un método de reducción del modelo: el método POD. Con éste se obtuvieron resultados muy positivos: se logró resolver el modelo de Fisher utilizando un número significativamente menor de variables, ya que se construye el esquema de los Elementos Finitos pero con un conjunto de funciones test (al que llamamos *Base POD*) que contienen mucha información del fenómeno. Esto se debe a que cada elemento de la base fue construido a partir de los valores propios obtenidos de la Matriz de Correlación asociada a los resultados obtenidos en la resolución numérica por el método de Crank-Nicholson. La información guardada se refleja en la base de funciones test que están definidas sobre todo el dominio del problema, el cual no es el caso de la base de funciones test del método de Elementos Finitos clásico, puesto que éstas son lineales y a soporte compacto.

Los experimentos numéricos se realizaron considerando perturbaciones en los datos del problema para comprobar la estabilidad de la aplicación del método POD en problemas perturbados. En particular, consideramos la perturbación de la función u y resolvimos el esquema reducido obteniendo resultados positivos, ya que el error de aproximación en la norma $L^2(\Omega)$ de la solución obtenida fue bastante pequeño.

En conclusión, el método POD es un método que reduce significativamente el tamaño del problema de Fisher, obteniendo muy buenos errores de aproximación. Además, este método es útil para obtener resultados coherentes en problemas *parecidos* al problema original, de manera que es necesario resolverlo utilizando el esquema de elementos finitos una sola vez, y luego se puede proceder a resolver el sistema perturbado utilizando el método POD, lo cual aprovecha la reducción del costo computacional de resolución que ofrece el método POD.

Para aplicar este método a problemas más específicos en el área de Biomatemáticas, recomendamos realizar un estudio previo detallado con respecto a los parámetros del fenómeno, puesto que según las relaciones entre ellos permiten o impiden la aplicación inmediata de resultados obtenidos en la teoría de las Ecuaciones Diferenciales Parciales, por ejemplo, a lo largo de la demostración de existencia y unicidad de la solución, y la aplicación de los métodos numéricos.

Con respecto a la implementación numérica de los esquemas, es posible utilizar otros métodos de resolución de ecuaciones diferenciales ordinarias que podrían resultar más eficientes en términos del error de aproximación, pero considerar el costo computacional que esto implicaría.

Además, se pueden considerar una discretización más diferenciada del dominio, siendo más fina al rededor de los puntos espaciales donde la frontera se muestra más irregular o cerca de los instantes donde las funciones y_0 y u cambian rápidamente su pendiente.

Con respecto a la aplicación del método POD, también se pueden tomar los snapshots de una manera diferenciada: tomando más snapshots en las iteraciones en las que más cambio se nota en los cálculos de la aproximación. Además, se puede considerar el aplicar el método POD en las primeras iteraciones del método de Newton, lo cual posiblemente aceleraría el proceso de Newton.

Como apertura a trabajos más profundizados, este proyecto de titulación puede dar pie al Control Óptimo del problema de Fisher: considerando la función $u(x, t)$ como control y sometiendo al modelo a un proceso de optimización.

También se podrían considerar modelos más complejos que involucren interacciones intraespecíficas como la competencia, interacciones interespecíficas como el fenómeno de predador-presa, o considerar distintos parámetros de supervivencia según una cierta diferenciación entre los individuos de la especie, ya sea por sexo o edad.

Referencias

- [1] Ambrosetti, A. y Prodi, G., (1993). *A Primer of Nonlinear Analysis*, Springer, Cambridge University Press, Inglaterra.
- [2] Anița, S., Arnăutu, V. y Capasso, V., (2010). *An Introduction to Optimal Control Problems in Life Sciences and Economics*, Springer, Estados Unidos.
- [3] Brézis, H., (1984). *Análisis Funcional: Teoría y Aplicaciones*, Alianza Editorial, España.
- [4] Cantrell, R. y Cosner, C., (2003). *Spatial Ecology via Reaction-Diffusion Equations*, John Wiley & Sons, Inglaterra.
- [5] Cosner, C., (2008). *Reaction-Diffusion Equations and Ecological Modeling*, Lecture Notes in Mathematics, Springer, Estados Unidos.
- [6] Evans, L., (1998). *Partial Differential Equations*, AMS, Estados Unidos.
- [7] Grossmann, C., Roos, H. y Stynes, M., (2007). *Numerical Treatment of Partial Differential Equations*, Springer, Alemania.
- [8] Kammann, E., Tröltzsch, F y Volwein, S, (2012). *A Method of A-Posteriori Error Estimation with Application to Proper Orthogonal Decomposition*, aún no publicado.
- [9] Kunisch, K. y Volkwein, S., (2001). *Galerkin Proper Orthogonal Decomposition Methods for Parabolic Problems*, web.
- [10] Lions, J-L. and Dautray, R., (2000). *Mathematical Analysis and Numerical Methods for Science and Technology*, Springer-Verlag, Alemania.
- [11] Murray, J.D., (2002). *Mathematical Biology: I. An Introduction*, Springer, tercera edición, Alemania.
- [12] Pinnau, R., (2008). *Model Reduction via Proper Orthogonal Decomposition*, Fachbereich Mathematik, Alemania.
- [13] Rannacher, R., (1984). *Finite Element Solution of Diffusion Problems with Irregular Data*, Numerische Mathematik, Springer-Verlag.
- [14] Shonkwiler, R. y Herod J., (2009). *Mathematical Biology*, Springer Science, segunda edición, Estados Unidos.
- [15] Smoller, J. (1994). *Shock Waves and Reaction-Diffusion Equations*, Springer Verlag, segunda edición, Estados Unidos.
- [16] Tröltzsch, F. (2010). *Optimal Control of Partial Differential Equations*, AMS, Estados Unidos.
- [17] Volkein, S., (2011). *Model Reduction Using Proper Orthogonal Decomposition*, script.

Anexos

Anexo A

Algunas Definiciones

Definición 9 (Operador Elíptico). Sea Ω un subconjunto abierto y acotado de \mathbb{R}^N , y sea $Q = \Omega \times (0, T]$ para cierto $T > 0$. Se define, para cada tiempo $t \in (0, T]$, el operador elíptico diferencial parcial de segundo orden:

$$L = \sum_{i,j=1}^N a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + c(x), \quad (\text{A.1})$$

para ciertos coeficientes a_{ij}, b_i, c que dependen de $x \in \Omega$.

Definición 10 (Condición de Elipticidad). Sea L un operador elíptico como queda definido en A.1. Supongamos que las funciones coeficientes $a_{ij} \in L^\infty(\Omega)$ y que satisfacen $a_{ij}(x) = a_{ji}(x)$ para todos $i, j \in 1, \dots, N$ y $x \in \Omega$. Se dice que los coeficientes del operador elíptico L verifican la condición de elipticidad si para cierto $\gamma_0 > 0$, se tiene

$$\sum_{i,j=1}^N a_{ij}(x) \xi_i \xi_j \geq \gamma_0 |\xi|^2 \quad \forall \xi \in \mathbb{R}^N$$

para casi todo $x \in \Omega$.

Definición 11 (Operador de Nemytskii). Sea $E \subset \mathbb{R}^N$, $N \in \mathbb{N}$, un conjunto acotado y medible; y sea $\varphi = \varphi(x, y(x)) : E \times \mathbb{R} \rightarrow \mathbb{R}$ una función. La aplicación Φ dada por

$$\Phi(y) = \varphi(\cdot, y(\cdot)),$$

que asigna a cada función $y : E \rightarrow \mathbb{R}$ la función $z : E \rightarrow \mathbb{R}$ con $z(x) = \varphi(x, y(x))$ es llamada un Operador de Nemytskii o un Operador de Superposición.

Anexo B

Algunos Resultados y Propiedades

Teorema 4 (Desigualdad de Cauchy-Schwarz). *Sea H un espacio de Hilbert, entonces*

$$|(u, v)| \leq \|u\| \|v\|, \quad \forall u, v \in H,$$

donde (\cdot, \cdot) denota al producto interno del espacio H y $\|\cdot\|$ denota la norma asociada.

Teorema 5 (Desigualdad de Hölder). *Sea $f \in L^p(\Omega)$ y $g \in L^q(\Omega)$ con $1 \leq p \leq \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$. Entonces, $fg \in L^1$ y*

$$\int |fg| \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}.$$

Propiedad 1 (Inmersión de H^1 en L^2). *Los espacios $H^1(\Omega)$ y $L^2(\Omega)$ son espacios separables, con $H^1(\Omega)$ denso en $L^2(\Omega)$ a través de una inyección continua. Además, $L^2(\Omega)$ se identifica con su espacio dual y es denso en V^* con inyección continua. Por tanto, se tiene*

$$V = H^1(\Omega) \hookrightarrow L^2(\Omega) = (L^2(\Omega))^* \hookrightarrow V^*.$$

Con esto, se tiene que, en particular, existe una constante C que verifica

$$\|y\|_{L^2(\Omega)}^2 \leq C \|y\|_V^2,$$

para todo $y \in V$.

Teorema 6 (Rellich-Kondrachov). *Supongamos que $\Omega \subset \mathbb{R}^N$ es acotado de clase C^1 . Entonces se verifica:*

$$\begin{array}{ll} \text{si } p < N & \text{entonces } W^{1,p}(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, p^*[, \text{ donde } \frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}, \\ \text{si } p = N & \text{entonces } W^{1,p}(\Omega) \hookrightarrow L^q(\Omega), \quad \forall q \in [1, +\infty[, \\ \text{si } p > N & \text{entonces } W^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega}). \end{array}$$

Ver [3].

Teorema 7 (Teorema del Punto Fijo de Banach). *Supongamos que $A : X \rightarrow X$ es una aplicación no lineal, y que*

$$\|A[y] - A[\tilde{y}]\| \leq \gamma \|y - \tilde{y}\|,$$

con $y, \tilde{y} \in X$ y para alguna constante $\gamma < 1$. Entonces A tiene un único punto fijo.

Teorema 8 (Desigualdad de Gronwall). *Sea $\xi(t)$ una función no negativa y medible sobre $[0, T]$ que satisfice para c.t.p. la desigualdad integral*

$$\xi(t) \leq C_1 \int_0^t \xi(s) ds,$$

para cierta constante $C_1 \geq 0$. Entonces

$$\xi(t) = 0 \quad \text{c.t.p.}$$

Teorema 9 (Operador de Nemytskii pertenece a L^∞). *Consideremos la definición 11. Supongamos que la función $\varphi = \varphi(x, y(x)) : E \times \mathbb{R} \rightarrow \mathbb{R}$ es medible con respecto a $x \in E$ para todo $y \in \mathbb{R}$, y supongamos que φ es Lipschitz continua con respecto a y , y satisface $|\varphi(x, 0)| \leq K$ para cierto $K > 0$ y c.t.p. $x \in E$. Entonces el operador de Nemytskii Φ asociado es continuo en $L^\infty(E)$.*

Teorema 10 (Desigualdad de Young (con ε)). *Sean $1 < p, q < \infty$ tales que $\frac{1}{p} + \frac{1}{q} = 1$; y sean para $a, b > 0, \varepsilon > 0$. Entonces*

$$ab \leq \varepsilon a^p + C(\varepsilon)b^q,$$

donde $C(\varepsilon) = (\varepsilon p)^{-a/p} q^{-1}$.

Teorema 11 (Principio del Máximo). *Supongamos que el operador L está definido como en la Definición 9 y que satisface la condición de elipticidad uniforme con $c(x) \leq 0$. Supongamos que $\Omega \subseteq \mathbb{R}^N$ es un dominio acotado y que los coeficientes de L son uniformemente acotados sobre Ω .*

- i) *Supongamos que $y \in C^2(\Omega)$ y $Ly \geq 0$ en Ω . Si y alcanza un máximo $M \geq 0$ en un punto interior de Ω , entonces $y(x) \equiv M$ en Ω .*
- ii) *Supongamos además que $y \in C^2(\Omega) \cap C(\bar{\Omega})$ y que cada punto de $\partial\Omega$ está contenido en la frontera de alguna bola a su vez contenida en Ω . Si $y(x) = M$ en algún punto $x_0 \in \partial\Omega$ para el cual $\partial y / \partial \nu$ existe, siendo ν la normal exterior de la frontera, entonces o bien $\partial y / \partial \nu \geq 0$ en x_0 o bien $y(x) \equiv M$ en Ω .*