

**UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA**

TESIS PRESENTADA COMO PARTE DE LOS REQUISITOS DE LA UNIVERSIDAD
NACIONAL DEL LITORAL PARA LA OBTENCIÓN DEL GRADO ACADÉMICO DE

Maestría en Matemática

EN EL CAMPO DE: **Análisis Numérico**

TÍTULO DE LA TESIS:

**ESTIMACIÓN A POSTERIORI PARA UN
PROBLEMA PARABÓLICO CON COEFICIENTE
DE DIFUSIÓN DISCONTINUO:**

**Simulación adaptativa de liberación de droga
en dispositivos poliméricos**

Departamento de Matemática, FIQ (UNL)

TESISTA:

Egle E. Haye

DIRECTORES:

Dr. Pedro Morin – Mg. Marta B. Bergallo

PRESENTACIÓN: Mayo, 2009

AGRADECIMIENTOS

Agradezco a Dios por haberme permitido llegar a esta meta.

No hubiera podido lograrlo sin las enseñanzas, la generosidad y la valiosa guía de mis directores, Pedro y Marta, quienes han sido mis maestros y cuya paciencia y comprensión agradezco enormemente.

Gracias a mis amigos de la F.I.C.H., Mariel, Mario, Fabiana, Laura, Carlos y Silvia, por el apoyo, la compañía y el aliento permanente.

Por último, va mi gratitud a mi familia, pero muy especialmente a mis padres, por su amor incondicional y porque en todo momento me han sostenido y animado para transformar las dificultades en una posibilidad de crecimiento.

RESUMEN

En la industria agroveterinaria, existen dispositivos de liberación de drogas por difusión, que son utilizados como un medio para el control de insectos en el ganado.

El objetivo final que motiva este trabajo es el diseño de la forma óptima de un dispositivo capaz de liberar una cantidad *casi constante* de droga por unidad de tiempo. Este es un problema específico dentro de la clase de problemas de optimización de formas.

Un primer paso para concretar nuestro propósito, es el desarrollo de un modelo que simule tales procesos y el diseño de un método numérico eficiente. Ya que estamos interesados en la difusión de drogas desde un dispositivo hacia el medio que lo rodea y específicamente, en el efecto de la *forma del dispositivo* en el comportamiento del sistema, proponemos como modelo, una ecuación de difusión que toma como dominio la unión de dos partes, una correspondiente al dispositivo y otra al medio exterior. En este contexto, la ecuación que modela esta aplicación, es una ecuación parabólica con coeficiente de difusión discontinuo, que toma en el dispositivo un valor distinto que en el medio que lo rodea. Más aún, esos coeficientes son en general de muy diferentes órdenes de magnitud, lo que causa una pérdida de regularidad en la solución y crea dificultades para resolver las ecuaciones numéricamente.

Dada la pérdida de regularidad mencionada, proponemos un método de elementos finitos adaptativo para resolver esta ecuación, en el que utilizamos estimadores a posteriori del error robustos cuyas cotas del error no presentan el cociente entre los valores máximo y mínimo del coeficiente de difusión, lo que disminuye considerablemente el costo computacional para una tolerancia dada. Con este método desarrollamos una herramienta de cálculo eficiente para simular, dada una forma específica, la difusión en un dispositivo y en el medio circundante simultáneamente. Esto resulta un primer módulo necesario y fundamental para poder resolver el problema de optimización de forma que motivó la tesis. Finalmente, para mostrar sus propiedades, y hacer un estudio experimental del efecto de la forma en el perfil de liberación, utilizamos la herramienta desarrollada sobre algunas formas particulares, lo que nos permite obtener algunas conclusiones preliminares.

Índice general

1. Introducción	9
2. Análisis matemático de la ecuación	13
2.1. Planteo del problema	13
2.2. Formulación débil del problema	14
2.2.1. Consideraciones generales y notación	14
2.2.2. Solución débil	16
2.2.3. Equivalencia entre solución clásica y solución débil	17
2.3. Existencia y unicidad de solución débil	19
2.3.1. Existencia de solución débil	19
2.3.2. Unicidad de solución débil	25
2.4. Estabilidad de la solución de la ecuación de difusión	26
3. Discretización y estimación a priori del error	29
3.1. Semidiscretización espacial	29
3.2. Estimación a priori del error en la semidiscretización espacial	30
3.3. Semidiscretización temporal	35
3.4. Estimación a priori del error en la semidiscretización temporal	36
3.5. Discretización completa	41
3.6. Estimación a priori del error en la discretización completa	42
4. Estimaciones a posteriori del error	45
4.1. Discretización numérica	46
4.2. Operadores de interpolación	46
4.3. La condición de casi monotonía	49
4.4. Un operador de interpolación robusto	50
4.5. Estimadores a posteriori del error	54
4.6. Cota superior global	56
4.7. Un ejemplo numérico	59
5. Algoritmo adaptativo	65
5.1. Introducción	65
5.2. Aspectos generales de la implementación	66
5.3. Diseño de un algoritmo adaptativo	67
5.4. Implementación	70

6. Simulación de liberación de droga	73
6.1. Ejemplo 1: Cuadrado	75
6.2. Ejemplo 2: Cuadrado con concavidad	76
6.3. Ejemplo 3: Cuadrado con parte del borde aislada	80
6.4. Ejemplo 4: Cuadrado con concavidad y parte del borde aislada	81
A. Apéndice	89
A.1. Los espacios L^p	89
A.2. Distribuciones y derivada débil	89
A.3. Los espacios de Sobolev H^s	90
A.4. Sobre Elementos Finitos	94
A.4.1. Triangulaciones del dominio	94
A.4.2. Un mapeo afín y algunas propiedades	95
A.5. El problema elíptico con coeficiente discontinuo	96
A.5.1. Forma continua y débil	96
A.5.2. Lema de Céa	97
A.5.3. Regularidad y Estabilidad	97
A.5.4. Estimación del error	99

Capítulo 1

Introducción

Uno de los mecanismos más importantes para la liberación de drogas a través de membranas o matrices es la difusión. Por su simplicidad de fabricación y la existencia de menos riesgos, la difusión a través de matrices poliméricas representa uno de los sistemas más utilizados. Un ejemplo concreto se da en la industria agroveterinaria, donde se fabrican dispositivos matriciales de liberación de droga con el objeto de atacar a la *mosca del cuerno* de la vaca. Estos dispositivos que se cuelgan de la oreja del animal, liberan la droga al ambiente exterior (aire) a una velocidad y durante un tiempo preestablecidos y así eliminan las moscas que se acercan a la cabeza del ganado.

El usuario de este tipo de dispositivos pretende alcanzar una liberación de droga casi constante por unidad de tiempo, de manera de evitar dos problemas fundamentales:

- que la cantidad de droga sea insuficiente para controlar al insecto;
- que la cantidad de droga en el aire sea excesiva y resulte tóxica para el animal.

Otra cuestión a tener en cuenta al diseñar estos dispositivos, es la cantidad de droga que queda almacenada en los mismos al terminar el período. Es de interés que esta cantidad sea la menor posible.

El logro de los objetivos mencionados puede depender de muchas variables, a saber: la concentración y distribución inicial de la droga en la matriz; el tamaño, forma o material del dispositivo o las condiciones del medio exterior en el que se libera la droga. Muchas investigaciones pusieron su atención particularmente en analizar el grado de influencia de la forma geométrica del dispositivo en el perfil de liberación de droga.

El modelo matemático del fenómeno que ocurre en el dispositivo está dado por la ecuación de difusión. Los primeros estudios del efecto de la forma del dispositivo sobre la liberación de la droga fueron realizados utilizando el método de separación de variables. Este método restringe seriamente las formas que se pueden estudiar, a aquellas que puedan escribirse como productos cartesianos de las variables: rectángulos, círculos, bolas, paralelepípedos, etc. Más recientemente se realizaron algunas simulaciones con el método de elementos finitos sobre las formas para las que ya se poseía la solución por el método de separación de variables [15].

El objetivo principal que motivó el desarrollo de esta tesis es el modelado del problema de difusión de droga a través de una matriz. Hemos puesto especial énfasis en la construcción de una herramienta computacional robusta y eficiente para simular el fenómeno

en cuestión. Esta herramienta permite experimentar con diferentes formas de dispositivos para buscar la óptima, en el sentido que, la cantidad de droga liberada por unidad de tiempo sea lo más constante posible y que a la vez el residuo dentro del dispositivo sea mínimo al finalizar el período de tiempo de interés. De esta manera, la herramienta computacional desarrollada podría ser usada como un módulo dentro de un programa de optimización que requiera su utilización repetidas veces con diferentes parámetros, formas y condiciones de contorno; es por eso que la misma debe ser robusta y eficiente.

Hasta el momento se ha estudiado la ecuación de difusión en el dispositivo, utilizando condiciones de borde de tipo Dirichlet, Neumann o Newton en diferentes partes de la frontera del mismo, considerando todo el tiempo que la concentración de droga es nula en el exterior (ver [15] y las referencias allí citadas). Es decir, se considera velocidad de difusión infinita en el exterior y sólo se ha estudiado o simulado el fenómeno de difusión *dentro* del dispositivo.

En general, el parámetro que interesa a los diseñadores es el flujo por unidad de tiempo hacia fuera del dispositivo. Existen para esta variable un límite superior sobre el cual el nivel de toxicidad es inaceptable e insalubre para el ganado, y también un límite inferior debajo del cual el nivel de toxicidad es ineficiente para controlar al insecto. Si bien esta es la variable que los diseñadores pretenden controlar, dado que queremos comprender el efecto de la forma del dispositivo, y de eventuales concavidades que pueda tener, entendemos que es igualmente (o más) importante estudiar el flujo de droga en una región un poco más grande que el dispositivo. Es decir, la región por donde circularán los insectos y/o donde se encuentra la parte de los animales susceptible a la droga (ojos, nariz, boca, etc.). Por este motivo, y con el objetivo de comprender mejor el efecto de la forma del dispositivo, estudiaremos la difusión simultáneamente en su interior y en parte del medio exterior.

Para ello, consideramos en esta tesis un dominio poligonal $\Omega \subset \mathbb{R}^2$ particionado en dos subdominios también poligonales Ω_1 y Ω_2 que representan al dispositivo y a parte del ambiente exterior que lo rodea, respectivamente (Ver Figura 1.1). Más precisamente, $\overline{\Omega_1} \cup \overline{\Omega_2} = \overline{\Omega}$ y $\Omega_1 \cap \Omega_2 = \emptyset$.

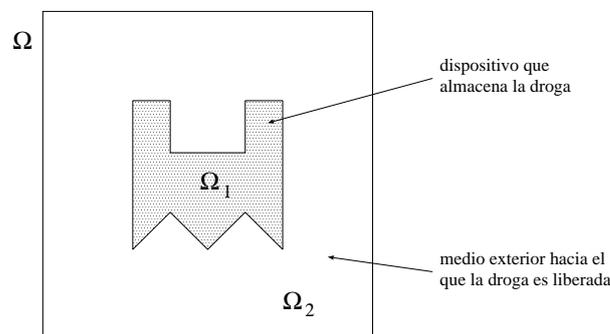


Figura 1.1: Esquema del dominio Ω y subdominios Ω_1 y Ω_2 .

El modelo matemático que gobierna este fenómeno es un problema parabólico con

coeficiente de difusión discontinuo de la forma:

$$\left\{ \begin{array}{ll} \frac{\partial u}{\partial t} - \operatorname{div}(k_i \nabla u) = f & \text{en } \Omega_i \times (0, T), \quad i = 1, 2 \\ u(x, t) = 0 & \text{en } \partial\Omega \times (0, T), \\ u(x, 0) = u^0(x) & \text{en } \Omega, \\ u_1 = u_2 & \text{en } \Gamma \times (0, T), \\ -k_1 \nabla u_1 \cdot n_1 = k_2 \nabla u_2 \cdot n_2 & \text{en } \Gamma \times (0, T). \end{array} \right.$$

donde u representa la concentración de droga; k_1 y k_2 son las constantes (positivas) del coeficiente de difusión de la droga en la matriz (Ω_1) y en el medio exterior (Ω_2), respectivamente; $(0, T)$ es un intervalo de tiempo dado, u^0 es la distribución inicial de droga y f es un eventual término fuente. La curva poligonal que llamamos interfaz se define como $\Gamma := \overline{\Omega_1} \cap \overline{\Omega_2}$ y denotamos con n_i al vector unitario normal a Γ y exterior a Ω_i . La condición de borde Dirichlet homogénea en todo $\partial\Omega$ indica que, como la difusión en el aire es más rápida, la droga se diluye inmediatamente en el medio exterior.

En el capítulo 2 de esta tesis formulamos el problema en forma débil y definimos su solución (débil) sobre espacios mixtos que involucran a las variables espacio y tiempo, y demostramos que la solución clásica y la débil son equivalentes bajo ciertas condiciones. Se demuestra además, la existencia y unicidad de solución débil del problema y la estabilidad de la solución clásica, para el caso $f \equiv 0$.

El capítulo 3 está dividido en tres partes donde básicamente discretizamos el problema débil y obtenemos cotas del error de aproximación de dicha discretización. Cada parte del capítulo se distingue de acuerdo a la variable que se discretiza: primero sólo la espacial, segundo sólo la temporal y por último ambas variables simultáneamente. En los tres casos, las acotaciones obtenidas se llaman *estimaciones a priori del error*.

Estas estimaciones a priori del error, dependen de la solución exacta del problema débil como también del tamaño h de una partición con refinamiento uniforme del dominio y de los tamaños Δt^n de una partición del intervalo de tiempo. Lo que ocurre en general es que, no tenemos conocimiento de la solución u o bien, que una inadecuada elección de h y Δt^n puede originar errores muy grandes que no son representativos del error verdadero en todo el dominio.

Para reducir el error, proponemos en el capítulo 4 refinamientos con adaptatividad en las dos variables, espacio y tiempo, sobre la base de *estimadores a posteriori del error*. El cálculo de los estimadores se realiza a partir de los datos del problema y de la solución discreta calculada. Para derivar estos estimadores del error, proponemos un operador de interpolación robusto en cuyas cotas del error no figure el cociente entre los valores máximo y mínimo del coeficiente de difusión. Para ello, se necesitará que el coeficiente k cumpla con la condición de casi monotonía (esta condición se cumple siempre para el tipo de dispositivos que estamos considerando). Las estimaciones del error de interpolación con este operador nos permitirán demostrar que los estimadores a posteriori resultan ser una cota superior local (en el tiempo) del error y global. Por último, presentamos un ejemplo cuya solución exacta es conocida con el fin de comparar el error exacto y el error aproximado utilizando los estimadores a posteriori y las acotaciones para el error demostradas anteriormente.

En el capítulo 5, en base a los estimadores a posteriori ya definidos y a la cota global del error obtenida, se desarrolla un algoritmo adaptativo robusto y eficiente capaz de resolver numéricamente el problema y se comentan algunas consideraciones generales sobre su implementación.

Finalmente, en el capítulo 6, presentamos los resultados numéricos y conclusiones de cuatro experimentos que muestran el comportamiento del algoritmo adaptativo. En particular, en el dominio de cada ejemplo se representa una forma diferente de dispositivo. Utilizando el algoritmo, se realiza la simulación de liberación de drogas desde el dispositivo hacia el exterior y se analiza del efecto de las distintas formas con las curvas de flujo por unidad de tiempo.

Capítulo 2

Análisis matemático de la ecuación

2.1. Planteo del problema

En este capítulo estudiaremos un problema de difusión definido sobre un dominio abierto $\Omega \subset \mathbb{R}^2$ Lipschitz, poligonal y acotado con borde $\partial\Omega$ y $[0, T]$ un intervalo de tiempo dado. Supongamos además que Ω está particionado en dos subdominios Ω_1 y Ω_2 , Lipschitz poligonales abiertos, de manera que $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$ y $\Omega_1 \cap \Omega_2 = \emptyset$. El parámetro de difusión $k : \Omega \rightarrow \mathbb{R}$ es una función positiva, con valor constante k_i sobre cada $\Omega_i \subset \Omega$, con $k_1 \neq k_2$. Llamaremos *interfaz* entre los subdominios a la curva poligonal Γ que separa Ω_1 de Ω_2 , definida por $\Gamma = \bar{\Omega}_1 \cap \bar{\Omega}_2$ (ver Figura 2.1).

Concretamente, nuestro propósito es analizar el siguiente problema parabólico de valores iniciales y de borde:

$$\left\{ \begin{array}{ll} \frac{\partial u}{\partial t} - \operatorname{div}(k(x)\nabla u) = f(x, t) & \text{en } \Omega_1 \cup \Omega_2 \times (0, T), \\ u(x, t) = 0 & \text{en } \partial\Omega \times (0, T), \\ u(x, 0) = u^0(x) & \text{en } \Omega, \\ u_1 = u_2 & \text{en } \Gamma \times (0, T), \\ -k_1 \nabla u_1 \cdot n_1 = k_2 \nabla u_2 \cdot n_2 & \text{en } \Gamma \times (0, T), \end{array} \right. \quad (2.1)$$

cuya solución es una función $u : \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}$. Aquí, f y u^0 son funciones dadas. Como la función k es discontinua, la ecuación diferencial principal se cumple en ambos dominios Ω_1 y Ω_2 , pero no se cumple *a través* de la interfaz. En cambio, tanto u como $k(x)\nabla u \cdot n$ son funciones continuas a través de Γ (donde n denota un vector unitario normal a Γ). Estas últimas condiciones están representadas por las dos últimas ecuaciones escritas en (2.1), donde para $i = 1, 2$, u_i representa $u|_{\Omega_i}$ y n_i representa el vector unitario normal a Γ y exterior a Ω_i .

Observación 2.1. *El término de divergencia en la ecuación diferencial de (2.1), también puede escribirse en la forma $-\operatorname{div}(A\nabla u)$, donde $A = (a_{ij}(x))$ es la matriz múltiplo escalar (discontinuo) de la matriz identidad de orden 2, es decir, $A = k(x)I_2$. Esta matriz es*

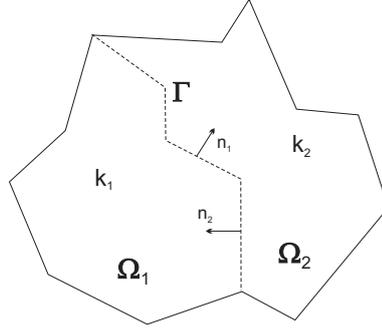


Figura 2.1: Dominio del problema (2.1)

simétrica, definida positiva y verifica para cada $t \in [0, T]$:

$$a_{ij} \in L^\infty(\Omega), \quad \forall i, j = 1, 2,$$

$$\exists \theta = \min_{i=1,2} k_i > 0 \text{ tal que } \sum_{i,j=1}^2 a_{ij}(x) \xi_i \xi_j \geq \theta |\xi|^2, \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^2. \quad (2.2)$$

La definición que veremos a continuación nos dará las condiciones mínimas que debe satisfacer una función $u(x, t)$ para ser solución del problema presentado. A esta función la llamaremos *solución clásica* del problema.

Definición 2.2 (Solución clásica). Supongamos que $f(\cdot, t) \in L^2(\Omega)$, $\forall t \in [0, T]$, y $u^0 \in L^2(\Omega)$. Llamaremos *solución clásica del problema (2.1)* a una función $u : \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}$ que satisfice:

- en la variable espacial $u(\cdot, t) \in C^2(\Omega_i) \cap C^1(\bar{\Omega}_i) \cap C(\bar{\Omega})$, $\forall t \in [0, T]$ fijo, $i = 1, 2$;
- en la variable temporal $u(x, \cdot) \in C^1((0, T)) \cap C([0, T])$, $\forall x \in \bar{\Omega}$ fijo;
- y además satisfice (2.1).

2.2. Formulación débil del problema

2.2.1. Consideraciones generales y notación

Establecemos aquí algunas de las notaciones que usaremos de ahora en adelante.

El espacio $L^2(\Omega)$ (Sección A.1) es un espacio de Hilbert con el producto escalar

$$(v, w) := \left(\int_{\Omega} vw \, dx \right)^{1/2},$$

y a su correspondiente norma la denotaremos por:

$$\|v\| := \|v\|_{L^2(\Omega)} = (v, v)^{1/2}.$$

Si $S \subset \Omega$ y S medible, simbolizaremos con $\|\cdot\|_S$ a la norma $L^2(S)$.

El espacio de Sobolev $H^1(\Omega)$ (sección A.3) es un espacio de Hilbert con el producto escalar definido por:

$$(v, w)_1 := \left(\int_{\Omega} vw \, dx + \int_{\Omega} \nabla v \cdot \nabla w \, dx \right)^{1/2},$$

y la seminorma y norma correspondientes las denotaremos, respectivamente, por

$$|v|_1 := |v|_{1,\Omega} = \|\nabla v\|, \quad \|v\|_1 := \|v\|_{1,\Omega} = (\|v\|^2 + |v|_1^2)^{1/2}.$$

Del mismo modo que en el primer espacio, denotaremos con $|v|_{1,S}$ o $\|v\|_{1,S}$ a la seminorma y norma $H^1(S)$ respectivamente, definidas en $S \subset \Omega$ medible.

En $H_0^1(\Omega) = \{v \in H^1(\Omega) : \gamma v = 0\}$, donde $\gamma : H^1(\Omega) \rightarrow L^2(\partial\Omega)$ es el operador traza [2, Sec. 5.5], la norma y seminorma son las mismas que las definidas en $H^1(\Omega)$.

Ya que Ω es acotado, gracias a la *desigualdad de Poincaré*:

$$\text{existe } c(\Omega) = \text{diam}(\Omega) > 0 \text{ tal que } \|v\| \leq c(\Omega)\|\nabla v\|, \quad \forall v \in H_0^1(\Omega), \quad (2.3)$$

la seminorma $|\cdot|_1$ resulta ser una norma sobre el espacio $H_0^1(\Omega)$ equivalente a $\|\cdot\|_1$, debido a que

$$|v|_1 \leq \|v\|_1 \leq (c(\Omega)^2 + 1)^{1/2}|v|_1.$$

Para simplificar la notación, consideraremos a las funciones u y f como mapeos entre el intervalo de tiempo $[0, T]$ y los espacios $H_0^1(\Omega)$ y $L^2(\Omega)$ respectivamente, escribiendo de ahora en más, $u(t) = u(\cdot, t)$, $f(t) = f(\cdot, t)$ y $u' = \frac{\partial u}{\partial t}$.

Denotaremos con $H^{-1}(\Omega)$ al dual de $H_0^1(\Omega)$, es decir, el espacio de todos los funcionales lineales y acotados sobre $H_0^1(\Omega)$. Para un funcional $g \in H^{-1}(\Omega)$, escribiremos como $\langle g, v \rangle$ a su aplicación $g(v)$ sobre una función $v \in H_0^1(\Omega)$.

Consideraremos a $H^{-1}(\Omega)$ equipado con la norma

$$\|g\|_{-1} = \sup_{v \in H_0^1(\Omega), \|v\|_1=1} |\langle g, v \rangle| = \sup_{v \in H_0^1(\Omega), \|v\|_1 \neq 0} \frac{|\langle g, v \rangle|}{\|v\|_1}.$$

Antes de introducir el concepto de solución débil para el problema planteado (2.1), será necesario definir los siguientes espacios de funciones que mapean el intervalo de tiempo $[0, T]$ sobre un espacio de Banach ([2, Sec. 5.9.2]).

Definición 2.3 (Espacios $L^p(0, T, X)$ y $C(0, T, X)$). Sea X un espacio Banach con norma $\|\cdot\|_X$.

- El espacio $L^p(0, T, X)$ es el conjunto de todas las funciones medibles $u : [0, T] \rightarrow X$ con

$$\|u\|_{L^p(0, T, X)} := \left(\int_0^T \|u(t)\|_X^p \, dt \right)^{1/p} < \infty,$$

si $1 \leq p < \infty$ y con

$$\|u\|_{L^\infty(0, T, X)} := \sup_{0 \leq t \leq T} \|u(t)\|_X < \infty,$$

cuando $p = \infty$.

- El espacio $C(0, T, X)$ es el conjunto de todas las funciones continuas $u : [0, T] \rightarrow X$ con

$$\|u\|_{C(0, T, X)} := \max_{0 \leq t \leq T} \|u(t)\|_X < \infty.$$

2.2.2. Solución débil

Definición 2.4 (Solución débil). Supongamos que $f \in L^2(0, T, L^2(\Omega))$ y $u^0 \in L^2(\Omega)$. Decimos que una función $u \in L^2(0, T, H_0^1(\Omega))$ con $u' \in L^2(0, T, H^{-1}(\Omega))$ es solución débil del problema parabólico (2.1) si cumple:

$$\langle u'(t), v \rangle + a(u(t), v) = \langle f(t), v \rangle, \quad \forall v \in H_0^1(\Omega) \text{ y c.t. } t \in [0, T] \quad (2.4)$$

$$u(0) = u^0(x), \quad \forall x \in \Omega \quad (2.5)$$

donde (\cdot, \cdot) es el producto interno en $L^2(\Omega)$, $\langle \cdot, \cdot \rangle$ representa la aplicación de un funcional de $H^{-1}(\Omega)$ sobre un elemento de $H_0^1(\Omega)$ y $a : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ es la forma bilineal definida por:

$$a(v, w) := \int_{\Omega} k(x) \nabla v \cdot \nabla w \, dx. \quad (2.6)$$

La ecuación (2.4) junto con la condición (2.5) es lo que llamamos *forma débil* del problema (2.1).

Debido a las hipótesis sobre u y u' en la definición, se tiene que $u \in C(0, T, L^2(\Omega))$ (ésto se prueba en [2, Sec. 5.9.2]), lo que es importante para que tenga sentido (2.5).

El hecho de pedir que las funciones $f \in L^2(0, T, L^2(\Omega))$, $u \in L^2(0, T, H_0^1(\Omega))$ y $u' \in L^2(0, T, H^{-1}(\Omega))$, es consecuencia del resultado de estabilidad (Teorema 2.7) y de la demostración de existencia de solución débil (Teorema 2.10) que exhibiremos más adelante.

La forma bilineal $a(\cdot, \cdot)$ tiene las siguientes propiedades:

- **es simétrica:** $a(v, w) = a(w, v)$, $\forall v, w \in H_0^1(\Omega)$.

- **es continua:**

$$\exists \beta = \max_{x \in \Omega} k(x) > 0 : |a(v, w)| \leq \beta \|v\|_1 \|w\|_1, \quad \forall v, w \in H_0^1(\Omega). \quad (2.7)$$

En efecto, por la desigualdad de Hölder se tiene que

$$\begin{aligned} |a(v, w)| &= \left| \int_{\Omega} k(x) \nabla v \cdot \nabla w \, dx \right| \\ &\leq \max_{x \in \Omega} k(x) \|\nabla v\| \|\nabla w\| \leq \max_{x \in \Omega} k(x) \|v\|_1 \|w\|_1. \end{aligned}$$

- **es coerciva sobre $H_0^1(\Omega)$:**

$$\exists \alpha = \frac{\min_{i=1,2} k_i}{c(\Omega)^2 + 1} > 0 : a(v, v) \geq \alpha \|v\|_1^2, \quad \forall v \in H_0^1(\Omega), \quad (2.8)$$

pues debido a la desigualdad (2.2) y a la desigualdad de Poincaré (2.3), resulta que

$$a(v, v) = \int_{\Omega} k(x) \nabla v \cdot \nabla v \, dx \geq \theta |v|_1^2 \geq \frac{\theta}{c(\Omega)^2 + 1} \|v\|_1^2.$$

- **induce una norma:** Como a es una forma bilineal simétrica y definida positiva, resulta ser un producto interno en $H_0^1(\Omega)$ que induce una norma en dicho espacio, llamada *norma energía*, y se define como

$$\|v\|_a = a(v, v)^{1/2} = \|\sqrt{k} \nabla v\|, \quad \forall v \in H_0^1(\Omega). \quad (2.9)$$

La norma energía es equivalente a la norma definida en el espacio $H_0^1(\Omega)$ ya que por la continuidad y coercividad de la forma bilineal a se tiene :

$$\sqrt{\alpha}\|v\|_1 \leq \|v\|_a \leq \sqrt{\beta}\|v\|_1, \quad \forall v \in H_0^1(\Omega).$$

Con $\|\cdot\|_{a,S}$ denotaremos la norma energía definida sobre un subconjunto medible $S \subset \Omega$.

El uso de $\langle \cdot, \cdot \rangle$ en la forma débil está motivado por el siguiente razonamiento. Debido a la hipótesis sobre u en la definición de solución débil, es claro que, para cada t fijo, $u(t) \in H_0^1(\Omega)$ y $f(t) \in L^2(\Omega)$. Entonces, si se define $\langle u'(t), v \rangle = \int_{\Omega} u'(t)v \, dx, \forall v \in H_0^1(\Omega)$, podemos decir que $u'(t) \in H^{-1}(\Omega)$. En efecto, $u'(t)$ es un funcional lineal y además acotado ya que de (2.4), la desigualdad de Cauchy - Schwartz y (2.7), se deduce que

$$|\langle u'(t), v \rangle| \leq \|f(t)\| \|v\| + \beta \|u(t)\|_1 \|v\|_1 \leq (\|f(t)\| + \beta \|u(t)\|_1) \|v\|_1, \quad \forall v \in H_0^1(\Omega).$$

2.2.3. Equivalencia entre solución clásica y solución débil

Es sencillo probar mediante integración por partes y argumentos de densidad que la solución clásica del problema (2.1) es también solución débil, pero el recíproco no siempre es cierto. El siguiente teorema mostrará que bajo ciertas condiciones de suavidad en la solución, las definiciones de solución clásica y de solución débil del problema (2.1), coinciden.

Teorema 2.5. *Sea $u : \bar{\Omega} \times [0, T] \rightarrow \mathbb{R}$ tal que $u(\cdot, t) \in C^2(\Omega_i) \cap C^1(\bar{\Omega}_i) \cap C(\bar{\Omega}), \forall t \in [0, T], i = 1, 2$ y $u(x, \cdot) \in C^1((0, T)) \cap C([0, T]), \forall x \in \bar{\Omega}$. Supongamos además $f \in L^2(0, T, L^2(\Omega))$ y $u^0 \in L^2(\Omega)$. Entonces: u es solución clásica de (2.1) si y sólo si u es solución débil de (2.1).*

Demostración. Supongamos que u es solución clásica de (2.1). Entonces para cada $t \in (0, T)$ vale

$$u'(t) - \operatorname{div}(k(x)\nabla u(t)) = f(t) \quad \text{en } \Omega_1 \cup \Omega_2.$$

Si multiplicamos esta ecuación por una función $v \in C_0^\infty(\Omega)$ e integramos sobre Ω , tenemos para todo $t \in (0, T)$:

$$\int_{\Omega} u'(t)v \, dx - \sum_{i=1}^2 \int_{\Omega_i} \operatorname{div}(k_i \nabla u_i(t)) v \, dx = \int_{\Omega} f v \, dx.$$

con $u_i = u|_{\Omega_i}$.

Integrando por partes sobre Ω_i (ver Teorema A.18) y utilizando que $v|_{\partial\Omega} = 0$, llegamos a que

$$\int_{\Omega} u'(t)v \, dx + \int_{\Omega} k(x)\nabla u(t) \cdot \nabla v \, dx - \int_{\Gamma} (k_1\nabla u_1(t) \cdot n_1 + k_2\nabla u_2(t) \cdot n_2) v \, d\sigma = \int_{\Omega} f(t)v \, dx,$$

donde n_i es el vector unitario normal a la interfaz y exterior a Ω_i .

Pero de acuerdo a la condición sobre Γ que cumple la solución clásica u en (2.1), la igualdad anterior resulta

$$\int_{\Omega} u'(t)v \, dx + \int_{\Omega} k(x)\nabla u(t)\nabla v \, dx = \int_{\Omega} f(t)v \, dx, \quad \forall v \in C_0^{\infty}(\Omega).$$

De aquí, ya que las funciones de $C_0^{\infty}(\Omega)$ son densas en $H_0^1(\Omega)$, se tiene:

$$\int_{\Omega} u'(t)v \, dx + \int_{\Omega} k(x)\nabla u(t)\nabla v \, dx = \int_{\Omega} f(t)v \, dx, \quad \forall v \in H_0^1(\Omega) \text{ y c.t. } t \in [0, T],$$

y como además u satisface la condición inicial, hemos probado que u es solución de la forma débil del problema (2.1).

Supongamos ahora que u es solución débil de (2.1) y satisface las hipótesis del teorema. Para demostrar que u es solución clásica, es suficiente probar que u satisface la ecuación diferencial en cada Ω_i y la condición de continuidad de flujo en la interfaz.

En efecto, como la ecuación (2.4) vale $\forall v \in H_0^1(\Omega)$ también será válida para $\forall v \in H_0^1(\Omega_i)$:

$$\int_{\Omega_i} u'(t)v \, dx + \int_{\Omega_i} k_i\nabla u(t) \cdot \nabla v \, dx = \int_{\Omega_i} f(t)v \, dx, \quad \forall v \in H_0^1(\Omega_i) \text{ y c.t. } t \in [0, T].$$

Integrando por partes la segunda integral del lado izquierdo y agrupando, nos queda

$$\int_{\Omega_i} (u'(t) - \operatorname{div}(k_i\nabla u(t)) - f(t))v \, dx = 0, \quad \forall v \in H_0^1(\Omega_i), \quad (2.10)$$

ya que v se anula en $\partial\Omega_i$. Luego, debido al resultado dado en el Lema A.3 del Apéndice, se cumple que

$$u'(t) - \operatorname{div}(k_i\nabla u(t)) = f(t), \quad \forall (x, t) \in \Omega_i \times (0, T), \quad i = 1, 2.$$

Resta probar que $-k_1\nabla u_1 \cdot n_1 = k_2\nabla u_2 \cdot n_2$, $\forall (x, t) \in \Gamma \times (0, T)$. Como u es solución débil, vale la ecuación (2.4) que podemos escribirla como:

$$\sum_{i=1}^2 \left(\int_{\Omega_i} u'v \, dx + \int_{\Omega_i} k_i\nabla u \cdot \nabla v \, dx \right) = \sum_{i=1}^2 \int_{\Omega_i} fv \, dx, \quad \forall v \in H_0^1(\Omega_i).$$

Análogamente a pasos anteriores, integrando por partes los términos que contienen los gradientes, llegamos a

$$\begin{aligned} \sum_{i=1}^2 \left(\int_{\Omega_i} (u'v - \operatorname{div}(k_i\nabla u)v) \, dx + \int_{\partial\Omega_i \cap \Gamma} k_i\nabla u \cdot n_i v \, d\sigma + \underbrace{\int_{\partial\Omega_i \cap \partial\Omega} k_i\nabla u \cdot n_i v \, d\sigma}_{=0} \right) \\ - \sum_{i=1}^2 \int_{\Omega_i} fv \, dx = 0 \end{aligned}$$

De aquí que, por (2.10) se desprende:

$$\int_{\Gamma} (k_1 \nabla u_1 \cdot n_1 + k_2 \nabla u_2 \cdot n_2) v \, d\sigma = 0 \quad \forall v \in H_0^1(\Omega).$$

En consecuencia, la afirmación $-k_1 \nabla u_1 \cdot n_1 = k_2 \nabla u_2 \cdot n_2$ en $\Gamma \times (0, T)$ queda demostrada. \square

2.3. Existencia y unicidad de solución débil

En la sección anterior, hemos definido el problema (2.1) en su forma débil. Nuestro interés ahora es demostrar que existe solución de este problema débil y además es única. Los dos apartados siguientes están dedicados a la prueba de existencia y unicidad, separadamente.

2.3.1. Existencia de solución débil

La demostración que haremos de la existencia de una solución débil del problema (2.1) es una adaptación de la presentada en [2, Sec. 7.1.2] y la desarrollaremos en tres etapas. La idea general es dar forma a una solución débil construyendo primero soluciones de ciertas aproximaciones en dimensión finita y luego pasando al límite. Este procedimiento es conocido como método de Galerkin.

En la primera etapa obtendremos una sucesión de soluciones aproximadas de (2.4)–(2.5). Para ello, en primer lugar, analizaremos la forma débil en un subespacio de $H_0^1(\Omega)$ de dimensión finita, para determinar los coeficientes de la correspondiente solución definida en este espacio finito-dimensional. En segundo lugar, utilizando propiedades de los espacios de Sobolev, definiremos una sucesión de tales subespacios y con ellos una sucesión de soluciones, que serán las aproximaciones de Galerkin correspondientes a estos espacios.

En la segunda etapa se probará un resultado de estabilidad que vale uniformemente para las soluciones obtenidas en el paso anterior. Este resultado intermedio será de gran utilidad para la etapa final donde, con la utilización de límite, se probará que existe una solución débil en el espacio $H_0^1(\Omega)$.

Etapa 1

Teorema 2.6. *Sea V un subespacio de dimensión finita de $H^1(\Omega)$. Sean $f \in L^1(0, T, L^2(\Omega))$, $u^0 \in L^2(\Omega)$ y $a : V \times V \rightarrow \mathbb{R}$ la forma bilineal definida en (2.6). Entonces existe una única $u : [0, T] \rightarrow V$ absolutamente continua tal que*

$$\langle u'(t), v \rangle + a(u(t), v) = (f(t), v), \quad \forall v \in V \text{ y c.t. } t \in [0, T] \quad (2.11)$$

$$u(0) = P_V(u^0), \quad (2.12)$$

donde $P_V(u^0)$ es la proyección $L^2(\Omega)$ de u^0 sobre el espacio de dimensión finita V , esto es, $\forall v \in V : (u^0, v) = (P_V(u^0), v)$.

Demostración. Sea $\{\phi_1, \phi_2, \dots, \phi_m\}$ una base para V . Entonces, para cada $t \in [0, T]$, si $u(t) \in V$ existirán m escalares únicos $\xi_i(t)$ tal que $u(t)$ puede escribirse como

$$u(t) = \sum_{i=1}^m \xi_i(t) \phi_i. \quad (2.13)$$

La ecuación (2.11) y la condición (2.12) se cumplirán si y sólo si

$$\begin{aligned} \langle u'(t), \phi_j \rangle + a(u(t), \phi_j) &= (f(t), \phi_j), \quad \forall j = 1, 2, \dots, m \\ (u(0), \phi_j) &= (P_V(u^0), \phi_j), \quad \forall j = 1, 2, \dots, m. \end{aligned}$$

Reemplazando la expresión (2.13) de $u(t)$ en estas ecuaciones, obtenemos un sistema de m ecuaciones diferenciales con condición inicial, de la forma

$$\begin{cases} \sum_{i=1}^m (\xi_i)'(t) (\phi_i, \phi_j) + \sum_{i=1}^m \xi_i(t) a(\phi_i, \phi_j) = (f, \phi_j), & j = 1, \dots, m, \quad t \in [0, T] \\ \sum_{i=1}^m \xi_i(0) (\phi_i, \phi_j) = (u^0, \phi_j), & j = 1 \dots m, \end{cases}$$

donde las m incógnitas a determinar son las funciones $\xi_i(t), i = 1, \dots, m$.

Este sistema lineal de ecuaciones diferenciales, puede ser expresado de manera equivalente como

$$\begin{cases} M \xi'(t) + K \xi(t) = F(t) \\ M \xi(0) = U^0 \end{cases} \quad (2.14)$$

donde ξ, F y U^0 son vectores columnas de \mathbb{R}^m y M, K son matrices simétricas de orden m , todos ellos definidos por

$$\begin{aligned} (\xi(t))_j &= \xi_j(t), & (F(t))_j &= (f(t), \phi_j), & (U^0)_j &= (u^0, \phi_j), \quad j = 1, 2, \dots, m \\ M_{ij} &= (\phi_i, \phi_j), & K_{ij} &= a(\phi_i, \phi_j), \quad i, j = 1, 2, \dots, m \end{aligned}$$

La matriz M es conocida como *matriz de masa* y la matriz K como *matriz de rigidez*. En particular, notamos que M es definida positiva ya que

$$\xi^t \cdot M \xi = \sum_{i=1}^m \sum_{j=1}^m \xi_i(\phi_i, \phi_j) \xi_j = \left\| \sum_{i=1}^m \xi_i \phi_i \right\|^2 > 0$$

para todo $\xi \neq 0$ en \mathbb{R}^m y así M es invertible. Esto implica que el problema (2.14) es equivalente a encontrar un único vector $\xi(t)$ solución de

$$\begin{cases} \xi'(t) + M^{-1} K \xi(t) = M^{-1} F(t) \\ \xi(0) = M^{-1} U^0. \end{cases}$$

Si definimos $h : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ por $h(t, x) = -M^{-1} K x + M^{-1} F(t)$, entonces nuestro problema se reduce a demostrar que existe un único vector $\xi(t)$ tal que

$$\begin{cases} \xi'(t) = h(t, \xi(t)) \\ \xi(0) = M^{-1} U^0. \end{cases} \quad (2.15)$$

Es fácil ver que $h(t, x)$ es Lipschitz continua en la variable x y que existe una función $m(t)$ integrable que la acota superiormente para todo $|\xi| \leq r$, como se muestra a continuación:

$$\begin{aligned} |h(t, \xi)| &\leq \|M^{-1}K\| |\xi| + \|M^{-1}\| |F(t)| \\ &\leq \|M^{-1}K\| r + \|M^{-1}\| \sqrt{\sum_i \|\phi_i\|^2} \|f(t)\| := m(t), \quad \forall |\xi| \leq r \end{aligned}$$

Luego, de acuerdo a la teoría de ecuaciones diferenciales ordinarias (Teorema A.19), existe una única solución $\xi(t)$ absolutamente continua que satisface el sistema (2.15) en casi todo $t \in [0, T]$. \square

Presentamos a continuación, la construcción de los subespacios de dimensión finita y de las aproximaciones en tales subespacios, de la solución del problema débil.

Construcción de los subespacios anidados V_m :

Los espacios $H^k(\Omega)$ con $k \in \mathbb{N}_0$, son separables (ver [1]). En particular, $H_0^1(\Omega)$ es un subespacio separable de $H^1(\Omega)$. En consecuencia, existe $\{\phi_i\}_{i=1}^\infty \subset H_0^1(\Omega)$ denso en $H_0^1(\Omega)$ con la norma de $H^1(\Omega)$.

A partir de ello, definimos para cada $m = 1, 2, \dots$ el espacio de dimensión finita

$$V_m = \text{gen}\{\phi_1, \phi_2, \dots, \phi_m\}. \quad (2.16)$$

Dichos espacios tienen las siguientes propiedades:

- $V_m \subset V_{m+1} \subset H_0^1(\Omega)$, $m = 1, 2, \dots$
- $\bigcup_{m \in \mathbb{N}} V_m$ es denso en $H_0^1(\Omega)$ y también en $L^2(\Omega)$, puesto que $H_0^1(\Omega)$ es denso en $L^2(\Omega)$.

La utilización de estos subespacios anidados de dimensión finita V_m en la hipótesis del Teorema 2.6, dará lugar a una sucesión de soluciones aproximadas del problema (2.4)–(2.5), que servirán más adelante para dar forma a una solución débil.

Aclaremos que, en la construcción que sigue a continuación, usaremos la inclusión

$$L^2(0, T, L^2(\Omega)) \subset L^1(0, T, L^2(\Omega)).$$

Construcción de soluciones aproximadas:

Consideremos los espacios V_m dados por (2.16). Sean $f \in L^2(0, T, L^2(\Omega))$, $u^0 \in L^2(\Omega)$ y $a : V_m \times V_m \rightarrow \mathbb{R}$ la forma bilineal definida en (2.6). Ya que para cada $m \in \mathbb{N}$, $V_m \subset H_0^1(\Omega)$ es un subespacio de dimensión finita de $H^1(\Omega)$, por Teorema 2.6 se sigue que existe una única $u_m : [0, T] \rightarrow V_m$ absolutamente continua que satisface

$$\langle u'_m(t), v \rangle + a(u_m(t), v) = (f(t), v), \quad \forall v \in V_m \text{ y c.t. } t \in [0, T] \quad (2.17)$$

$$u_m(0) = P_{V_m}(u^0) \quad (2.18)$$

donde P_{V_m} es el operador proyección $L^2(\Omega)$ sobre V_m .

Etapa 2

La estimación que a continuación ofrece el Teorema 2.7, es válida para cada solución u_m de los problemas (2.17)–(2.18) y será de gran importancia para la prueba de existencia en la que estamos encaminados.

Teorema 2.7 (Estabilidad). *Sean $f \in L^2(0, T, L^2(\Omega))$, $u^0 \in L^2(\Omega)$ y $\{u_m\}_{m \in \mathbb{N}}$ las soluciones obtenidas de los problemas discretos (2.17)–(2.18). Entonces existe una constante C que depende de Ω , T y del coeficiente de difusión $k(x)$, pero no de m , tal que*

$$\max_{0 \leq t \leq T} \|u_m(t)\| + \|u_m\|_{L^2(0, T, H_0^1(\Omega))} \leq C (\|f\|_{L^2(0, T, L^2(\Omega))} + \|u^0\|), \quad (2.19)$$

para $m = 1, 2, \dots$

Demostración. Consideremos $v = u_m(t) \in V_m$ en (2.17):

$$\langle u'_m(t), u_m(t) \rangle + a(u_m(t), u_m(t)) = (f(t), u_m(t)), \quad \text{c.t. } t \in [0, T]. \quad (2.20)$$

Si analizamos el primer término en particular, vemos que

$$\langle u'_m(t), u_m(t) \rangle = \int_{\Omega} u'_m(t) u_m(t) dx = \frac{d}{dt} \left(\frac{1}{2} \int_{\Omega} u_m^2(t) dx \right) = \frac{1}{2} \frac{d}{dt} (\|u_m(t)\|^2). \quad (2.21)$$

Entonces, de las expresiones (2.20) y (2.21) junto con la *desigualdad de Young*:

$$ab \leq \epsilon a^2 + \frac{b^2}{4\epsilon}, \quad a, b > 0, \quad \epsilon > 0, \quad (2.22)$$

con $\epsilon = \alpha/2$ y α la constante de coercividad (2.8), obtenemos

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} (\|u_m(t)\|^2) + a(u_m(t), u_m(t)) &\leq |(f(t), u_m(t))| \\ &\leq \|f(t)\|_{-1} \|u_m(t)\|_1 \leq \frac{\|f(t)\|_{-1}^2}{2\alpha} + \alpha \frac{\|u_m(t)\|_1^2}{2}, \end{aligned} \quad (2.23)$$

interpretando a $f(t) \in L^2(\Omega)$ como el funcional lineal acotado de $H_0^1(\Omega)$ cuya aplicación a un elemento $v \in H_0^1(\Omega)$ se define por $\int_{\Omega} f(t)v dx$.

Por otro lado, como $u_m(t) \in H_0^1(\Omega)$ y a es coerciva

$$\alpha \|u_m(t)\|_1^2 \leq a(u_m(t), u_m(t)), \quad \forall t \in [0, T].$$

Teniendo en cuenta esta desigualdad en (2.23) y el hecho de que $\|f(t)\|_{-1} \leq \|f(t)\|$, resulta:

$$\frac{d}{dt} (\|u_m(t)\|^2) + \alpha \|u_m(t)\|_1^2 \leq \frac{\|f(t)\|_{-1}^2}{\alpha} \leq \frac{\|f(t)\|^2}{\alpha} \quad (2.24)$$

e integrando entre 0 y t

$$\|u_m(t)\|^2 + \alpha \int_0^t \|u_m(s)\|_1^2 ds \leq \|u_m(0)\|^2 + \frac{1}{\alpha} \int_0^t \|f(s)\|^2 ds. \quad (2.25)$$

Por la definición de proyección y la desigualdad de Cauchy-Schwartz,

$$\|u_m(0)\|^2 = (u_m(0), u_m(0)) = (u^0, u_m(0)) \leq \|u_m(0)\| \|u^0\|,$$

de aquí que

$$\|u_m(0)\| \leq \|u^0\|. \quad (2.26)$$

Ahora, retomando (2.25) y usando (2.26) tenemos que para cada $t \in [0, T]$

$$\|u_m(t)\|^2 + \alpha \int_0^t \|u_m(s)\|_1^2 ds \leq \|u^0\|^2 + \frac{1}{\alpha} \int_0^t \|f(s)\|^2 ds \leq C_1 (\|u^0\|^2 + \int_0^T \|f(t)\|^2 dt),$$

donde $C_1 = \max\{1, 1/\alpha\}$.

Obviamente esta acotación vale también para el término simple $\|u_m(t)\|^2$; luego, tomando el máximo sobre todo el intervalo de tiempo:

$$\max_{0 \leq t \leq T} \|u_m(t)\|^2 \leq C_1 (\|u^0\|^2 + \|f\|_{L^2(0,T,L^2(\Omega))}^2). \quad (2.27)$$

Análogamente, si integramos (2.24) entre 0 y T obtendríamos

$$\|u_m\|_{L^2(0,T,H_0^1(\Omega))}^2 \leq C_2 (\|u^0\|^2 + \|f\|_{L^2(0,T,L^2(\Omega))}^2), \quad (2.28)$$

con $C_2 = \max\{1/\alpha, 1/\alpha^2\}$.

Finalmente, de las cotas (2.27) y (2.28) deducimos la desigualdad (2.19) que queríamos probar. \square

Etapa 3

Esta última etapa consiste en pasar al límite cuando m tiende a infinito con el propósito de construir una solución débil (Teorema 2.10) para luego demostrar que es única (Teorema 2.11). Antes haremos dos observaciones.

Observación 2.8. Si $v \in L^2(0, T, H_0^1(\Omega))$ entonces $T_v(g) := \int_0^T a(v(t), g(t)) dt$, con $g \in L^2(0, T, H_0^1(\Omega))$, es un funcional lineal y acotado sobre $L^2(0, T, H_0^1(\Omega))$.

En efecto, por (2.7)

$$\begin{aligned} |T_v(g)| &\leq \int_0^T |a(v(t), g(t))| dt \leq \beta \int_0^T \|v(t)\|_1 \|g(t)\|_1 dt \\ &\leq \beta \left(\int_0^T \|v(t)\|_1^2 dt \right)^{1/2} \left(\int_0^T \|g(t)\|_1^2 dt \right)^{1/2} \\ &= \beta \|v\|_{L^2(0,T,H_0^1(\Omega))} \|g\|_{L^2(0,T,H_0^1(\Omega))} \end{aligned}$$

De aquí que T_v pertenece al dual de $L^2(0, T, H_0^1(\Omega))$.

Observación 2.9. Sea $v \in L^2(0, T, H_0^1(\Omega))$. Definimos $T_v : L^2(0, T, H^{-1}(\Omega)) \rightarrow \mathbb{R}$ por $T_v(g) := \int_0^T \langle v(t), g(t) \rangle dt$. Luego T_v es un funcional lineal y acotado sobre $L^2(0, T, H^{-1}(\Omega))$ ya que

$$\begin{aligned} |T_v(g)| &\leq \int_0^T |\langle v(t), g(t) \rangle| dt \leq \int_0^T \|v(t)\|_1 \|g(t)\|_{-1} dt \\ &\leq \left(\int_0^T \|v(t)\|_1^2 dt \right)^{1/2} \left(\int_0^T \|g(t)\|_{-1}^2 dt \right)^{1/2} \\ &= \|v\|_{L^2(0,T,H_0^1(\Omega))} \|g\|_{L^2(0,T,H^{-1}(\Omega))}. \end{aligned}$$

En consecuencia T_v pertenece al dual de $L^2(0, T, H^{-1}(\Omega))$.

Teorema 2.10 (Existencia de solución débil). *Existe una solución débil del problema (2.1).*

Demostración. De la estimación dada por el Teorema 2.7 se desprende que la sucesión $\{u_m\}$ es acotada en $L^2(0, T, H_0^1(\Omega))$. En consecuencia, existen una subsucesión $\{u_{m_\ell}\}_{\ell=1}^\infty$ de $\{u_m\}_{m=1}^\infty$ y una función $u \in L^2(0, T, H_0^1(\Omega))$ tales que

$$u_{m_\ell} \longrightarrow u \quad \text{débilmente en } L^2(0, T, H_0^1(\Omega)).$$

(ver [2, Sec. D.4] y Apéndice Definición A.4).

Queremos mostrar que u es una solución débil. Para ello, probaremos que u satisface las condiciones (2.4)–(2.5) de la definición de solución débil.

Consideremos una función $w \in \bigcup_{m=1}^\infty V_m$. Entonces $w \in V_m$ para algún m y sea $v : [0, T] \longrightarrow V_m$ definida por

$$v(t) = w \xi(t), \quad (2.29)$$

donde $\xi(t)$ es una función con al menos una derivada continua sobre $[0, T]$.

Elijamos ℓ_0 suficientemente grande de modo que $v(t) \in V_m \subset V_{m_{\ell_0}} \subset V_{m_\ell}, \forall \ell \geq \ell_0$.

Entonces se cumple que:

$$\langle u'_{m_\ell}(t), v(t) \rangle + a(u_{m_\ell}(t), v(t)) = (f(t), v(t)) \quad \forall \ell \geq \ell_0, \text{ c.t. } t \in [0, T],$$

y al integrar entre 0 y T nos queda

$$\int_0^T \langle u'_{m_\ell}(t), v(t) \rangle dt = \int_0^T (f(t), v(t)) dt - \int_0^T a(u_{m_\ell}(t), v(t)) dt. \quad (2.30)$$

En particular $v \in L^2(0, T, H_0^1(\Omega))$. De la convergencia débil y la observación 2.8 tenemos que

$$\int_0^T \langle u'_{m_\ell}(t), v(t) \rangle dt \longrightarrow \int_0^T (f(t), v(t)) dt - \int_0^T a(u(t), v(t)) dt \quad \text{cuando } \ell \longrightarrow \infty.$$

Esto último vale también si reemplazamos $v(t)$ por combinaciones lineales finitas de funciones con la forma (2.29) y como éstas son densas en $L^2(0, T, H_0^1(\Omega))$, se sigue que el límite anterior vale para toda $v \in L^2(0, T, H_0^1(\Omega))$.

De aquí que, teniendo en cuenta la observación 2.9, podemos decir que $\{u'_{m_\ell}\}_{\ell \in \mathbb{N}}$ converge débilmente en $L^2(0, T, H^{-1}(\Omega))$. Este límite debe ser u' , debido a la convergencia débil de $\{u_{m_\ell}\}$ a u .

Luego, recordando (2.29) obtenemos

$$\int_0^T (\langle u'(t), w \rangle + a(u(t), w)) \xi(t) dt = \int_0^T (f(t), w) \xi(t) dt \quad (2.31)$$

Por consiguiente,

$$\langle u'(t), w \rangle + a(u(t), w) = (f(t), w) \quad \forall w \in \bigcup_{m=1}^\infty V_m \text{ y c.t. } t \in [0, T],$$

ya que $\xi(t)$ es arbitraria. Y como $\bigcup_{m=1}^{\infty} V_m$ es denso en $H_0^1(\Omega)$, la última igualdad vale $\forall w \in H_0^1(\Omega)$.

Nos falta probar aún que u satisface la condición inicial $u(0) = u^0$. Consideremos una función $v(t)$ con la forma (2.29) que satisfaga la condición $v(T) = 0$. Al integrar por partes en el primer término del lado izquierdo de (2.31) resulta

$$\int_0^T \left(-\langle v'(t), u(t) \rangle + a(u(t), v(t)) \right) dt = \int_0^T (f(t), v(t)) dt + (u(0), v(0)) \quad (2.32)$$

donde $v'(t) = w\xi'(t)$.

Análogamente, de la ecuación (2.30) obtenemos

$$\int_0^T \left(-\langle v'(t), u_{m_\ell}(t) \rangle + a(u_{m_\ell}(t), v(t)) \right) dt = \int_0^T (f(t), v(t)) dt + (u_{m_\ell}(0), v(0)).$$

Haciendo tender ℓ a infinito, de la hipótesis de convergencia débil y puesto que $u_{m_\ell}(0) \rightarrow u^0$ en $L^2(\Omega)$, se tiene que

$$\int_0^T \left(-\langle v'(t), u(t) \rangle + a(u(t), v(t)) \right) dt = \int_0^T (f(t), v(t)) dt + (u^0, v(0)). \quad (2.33)$$

Al comparar (2.32) y (2.33), concluimos que $u(0) = u^0$ pues $v(0)$ es arbitrario en $\bigcup_{m=1}^{\infty} V_m$, que es denso en $L^2(\Omega)$. \square

2.3.2. Unicidad de solución débil

En la demostración del teorema de existencia (Teorema 2.10) se mostró que existe una subsucesión de entre las soluciones u_m , que converge a una solución débil del problema (2.1). Veamos ahora la unicidad de esta solución débil.

Teorema 2.11 (Unicidad de soluciones débiles). *El problema (2.1) tiene única solución débil.*

Demostración. Es suficiente probar que para el caso $f = u^0 = 0$ en el problema (2.1), la única solución débil es $u \equiv 0$. En efecto, llamemos (P) al problema (2.1) cuando $f = u^0 = 0$. Si suponemos que u^1 y u^2 son dos soluciones débiles del problema (2.1), será $u^1 - u^2$ solución débil de (P) , por lo que llegaríamos a que $u^1 \equiv u^2$.

Si $u = u^1 - u^2$ es una solución débil de (P) entonces

$$\langle u'(t), v \rangle + a(u(t), v) = 0, \quad \forall v \in H_0^1(\Omega), \text{ c.t. } t \in [0, T].$$

Tomando $v = u(t)$, por (2.21) se tiene que

$$a(u(t), u(t)) = -\langle u'(t), u(t) \rangle = -\frac{1}{2} \frac{d}{dt} (\|u(t)\|^2).$$

Como $a(u(t), u(t)) \geq 0$, obtenemos de la igualdad anterior, que $\|u(t)\|^2$ es una función de t decreciente. Así, para cualquier tiempo $t \geq 0$: $\|u(t)\|^2 \leq \|u(0)\|^2 = 0$. Entonces solamente puede ocurrir que $u \equiv 0$. \square

2.4. Estabilidad de la solución de la ecuación de difusión

Cuando hablamos de estabilidad nos referimos a la estabilidad respecto a los datos del problema, es decir, a estimaciones que nos permitan inferir que a pequeños cambios o perturbaciones en los datos, se producen pequeños cambios en la solución. Particularmente estudiaremos las estimaciones básicas de estabilidad para la solución exacta de la ecuación de difusión del problema (2.1) cuando $f \equiv 0$, pues es el caso que nos interesa desde el punto de vista de la aplicación. En el siguiente lema, presentamos dichas estimaciones.

Lema 2.12 (Estabilidad). *Sea u solución del problema (2.1) con $f \equiv 0$. Entonces para $0 < t \leq T$ se cumplen las siguientes estimaciones:*

$$\|u(t)\|^2 + 2 \int_0^t a(u(s), u(s)) ds = \|u(0)\|^2, \quad (2.34)$$

$$2 \int_0^t s \|u'(s)\|^2 ds + t a(u(t), u(t)) \leq \frac{1}{2} \|u(0)\|^2, \quad (2.35)$$

$$\|u'(t)\| \leq \frac{\|u(0)\|}{\sqrt{2t}}. \quad (2.36)$$

Demostración. Si multiplicamos la ecuación diferencial de (2.1) por $u(t)$ e integramos sobre Ω , después de la integración por partes (Teorema A.18) obtenemos $\langle u'(t), u(t) \rangle + a(u(t), u(t)) = 0$, que por (2.21), se puede escribir

$$\frac{1}{2} \frac{d}{dt} (\|u(t)\|^2) + a(u(t), u(t)) = 0.$$

Integrando esta expresión entre 0 y t , se tiene

$$\frac{1}{2} \|u(t)\|^2 + \int_0^t a(u(s), u(s)) ds = \frac{1}{2} \|u(0)\|^2,$$

con lo que llegamos a probar la igualdad (2.34) del lema.

Volvamos a la ecuación diferencial de (2.1) y multipliquémosla por $tu'(t)$. Integrando sobre Ω obtenemos

$$t \langle u'(t), u'(t) \rangle + t a(u(t), u'(t)) = 0. \quad (2.37)$$

Por un lado, sabemos que

$$\partial_t (t a(u(t), u(t))) = a(u(t), u(t)) + t \partial_t a(u(t), u(t)), \quad (2.38)$$

y por otro lado, tenemos

$$\begin{aligned} \partial_t a(u(t), u(t)) &= \int_{\Omega} k(x) \partial_t (|\nabla u(t)|^2) dx = \int_{\Omega} k(x) \partial_t \left(\sum_{i=1}^2 (\partial_i u(t))^2 \right) dx \\ &= \int_{\Omega} k(x) \sum_{i=1}^2 2 \partial_i u(t) \partial_t (\partial_i u(t)) dx \\ &= 2 \int_{\Omega} k(x) \nabla u(t) \cdot \nabla u'(t) dx = 2 a(u(t), u'(t)). \end{aligned}$$

Reemplazando este resultado en (2.38) y utilizando (2.37) se tiene que

$$\partial_t(t a(u(t), u(t))) = a(u(t), u(t)) - 2t\|u'(t)\|^2.$$

Entonces

$$t\|u'(t)\|^2 + \frac{1}{2}\partial_t(t a(u(t), u(t))) = \frac{1}{2}a(u(t), u(t)).$$

Si integramos entre 0 y t , como consecuencia de (2.34) nos queda

$$\int_0^t s\|u'(s)\|^2 ds + \frac{1}{2}t a(u(t), u(t)) = \frac{1}{2} \int_0^t a(u(s), u(s)) ds \leq \frac{1}{4}\|u(0)\|^2,$$

que resulta equivalente a (2.35).

Por último, si derivamos respecto de t la ecuación diferencial en (2.1), el lado izquierdo resulta, sobre cada Ω_i

$$u''(t) - \partial_t \left(\sum_{i=1}^2 \partial_i(k(x) \partial_i u(t)) \right) = u''(t) - \sum_{i=1}^2 \partial_i(k(x) \partial_i u'(t)).$$

Es decir, que nos queda

$$u''(t) - \operatorname{div}(k(x)\nabla u'(t)) = 0, \quad \text{en } \Omega_1 \cup \Omega_2.$$

Si a esta igualdad la multiplicamos por $t^2 u'(t)$, integramos sobre Ω y aplicamos integración por partes, obtenemos

$$t^2(u'(t), u''(t)) + t^2 a(u'(t), u'(t)) = 0. \quad (2.39)$$

Por otro lado

$$\partial_t(t^2\|u'(t)\|^2) = 2t\|u'(t)\|^2 + t^2\partial_t \left(\int_{\Omega} u'^2(t) dx \right) = 2t\|u'(t)\|^2 + 2t^2 \int_{\Omega} u'(t)u''(t) dx. \quad (2.40)$$

Reemplazando la ecuación (2.39) en (2.40) y reordenando los términos, nos queda

$$\frac{1}{2}\partial_t(t^2\|u'(t)\|^2) + t^2 a(u'(t), u'(t)) = t\|u'(t)\|^2. \quad (2.41)$$

Si ahora integramos entre 0 y t obtenemos

$$\frac{1}{2}t^2\|u'(t)\|^2 + \int_0^t s^2 a(u'(s), u'(s)) ds = \int_0^t s\|u'(s)\|^2 ds \leq \frac{1}{4}\|u(0)\|^2,$$

como consecuencia de (2.35).

Este resultado implica que

$$t^2\|u'(t)\|^2 \leq \frac{1}{2}\|u(0)\|^2$$

y así es inmediata la desigualdad (2.36). □

Observación 2.13. *Las estimaciones del Lema 2.12 valen en el espacio $H_0^1(\Omega)$. Por lo tanto también valen en los subespacios discretos V_m definidos en 2.16, reemplazando u por u_m y siguiendo los mismos razonamientos.*

Supongamos por el momento que estamos trabajando en un espacio V_m para un cierto m . Conviene notar que todas las desigualdades de estabilidad obtenidas contendrían a $\|u_m(0)\|$ en el lado derecho, en lugar de $\|u(0)\|$, que no es dato en el problema (2.1). Para subsanar esto, recordemos que $u_m(0)$ es la proyección- $L^2(\Omega)$ de $u(0)$ sobre el espacio de dimensión finita V_m . Entonces

$$\|u(0)\|^2 = \|u_m(0)\|^2 + \|u(0) - u_m(0)\|^2 \geq \|u_m(0)\|^2.$$

Así, si trabajamos en el espacio discreto V_m debemos tener en cuenta en las estimaciones de estabilidad que

$$\|u_m(0)\| \leq \|u(0)\|.$$

Las estimaciones de estabilidad que obtuvimos en el Lema 2.12 son muy importantes ya que nos permiten sacar las siguientes conclusiones:

- De (2.34), $\|u(t)\| \leq \|u(0)\|$, $0 \leq t \leq T$. Más aún, $\|u(t)\|$ es una función decreciente en $[0, T]$ (ver demostración en el Teorema 2.11), lo que indica que a medida que t toma valores más grandes, $\|u(t)\|$ se hace más pequeña.
- Las desigualdades (2.35) y (2.36) nos dicen que para valores de t cercanos a cero, no puede asegurarse que $\|\sqrt{k(x)} \nabla u\|$ y $\|u'(t)\|$ estén acotadas. Pero sí podemos afirmar que para valores cada vez más grandes de t , $\|\sqrt{k(x)} \nabla u\|$ y $\|u'(t)\|$ se acercan a cero y permanecen acotadas.

En general, la solución exacta de un problema parabólico puede presentar una *fase inicial* (que corresponde a instantes de tiempos pequeños) donde algunas derivadas respecto de t o de x son grandes, pero luego la solución se va haciendo más suave a medida que t crece.

Este hecho es muy importante desde el punto de vista de la resolución numérica del problema. Por eso la ventaja de utilizar, en el Método de Elementos Finitos, adaptatividad en las variables espacio y tiempo, donde se puede decidir refinar la malla y/o cambiar el tamaño del paso de tiempo, donde la solución no es tan suave.

Ya que no siempre el problema (2.1) tendrá solución clásica, para continuar en el siguiente capítulo con el análisis del error por Elementos Finitos, a partir de ahora nos referiremos como solución exacta, a la solución exacta de la forma débil (2.4)-(2.5).

Capítulo 3

Discretización y estimación a priori del error

En el capítulo anterior, hemos visto que el problema (2.1) en su forma débil consiste en: encontrar una función $u \in L^2(0, T, H_0^1(\Omega))$ con $u' \in L^2(0, T, H^{-1}(\Omega))$ tal que

$$\langle u'(t), v \rangle + a(u(t), v) = (f(t), v), \quad \forall v \in H_0^1(\Omega), \quad \text{c.t. } t \in [0, T] \quad (3.1)$$

$$u(0) = u^0 \quad (3.2)$$

En este capítulo nos basaremos en el problema débil para plantear su análogo discreto y estudiar la cota para el error obtenido de dicha discretización.

Diremos que un problema es *semidiscreto* si se discretiza sólo una de las variables intervinientes, espacio o tiempo y *completamente discreto* si se discretizan las dos variables. En cada caso estudiaremos las estimaciones básicas para el error por elementos finitos, entre la solución exacta del problema débil (3.1)–(3.2) y la solución aproximada obtenida de la discretización de dicho problema.

Todas las cotas del error en este capítulo se obtendrán en la norma $C(0, T, L^2(\Omega))$ y se expresarán en función de la solución exacta (desconocida). Este tipo de acotaciones son conocidas como estimaciones *a priori* del error.

En el próximo capítulo, se estudiarán las llamadas estimaciones *a posteriori*, en donde las cotas del error se expresan en términos de los datos del problema y de la solución discreta calculada.

3.1. Semidiscretización espacial

Para la discretización del problema (3.1)–(3.2) con respecto solamente a la variable espacial utilizaremos el método de Galerkin. En este caso, dada una triangulación de tamaño h del dominio, buscamos una aproximación a la solución $u(x, t)$, representada por $u_h(x, t)$, en donde para cada t fijo, $u_h(t)$ es una función de x lineal a trozos. Así $u_h(t)$ dependerá de un número finito de parámetros.

Sea \mathcal{T}_h una partición de triángulos T de $\bar{\Omega}$ de tamaño $h = \max_{T \in \mathcal{T}_h} h_T > 0$, en donde llamamos h_T al diámetro del elemento T (ver introducción de la sección A.4.1). La partición consiste en triángulos tales que la intersección de dos distintos es a lo sumo

un vértice común o un lado común. Cuando esto ocurre se llama triangulación *admisibile* (ver al comienzo de A.4.1).

Vamos a suponer que la triangulación \mathcal{T}_h induce dos triangulaciones $\mathcal{T}_{h,i}$ sobre los subdominios Ω_i , $i = 1, 2$ de manera que $\mathcal{T}_h = \bigcup_{i=1}^2 \mathcal{T}_{h,i}$.

Denotamos con V_h al subespacio de dimensión finita de $H_0^1(\Omega)$ asociado a \mathcal{T}_h que consiste en el conjunto de funciones continuas lineales a trozos sobre \mathcal{T}_h que se anulan en $\partial\Omega$, es decir,

$$V_h = \{v \in \mathcal{C}(\bar{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h\},$$

donde $\mathcal{P}_1(T)$ es el espacio de los polinomios de grado a lo sumo uno sobre el elemento T . Llamaremos a V_h *espacio de elementos finitos*.

Entonces, al problema análogo a (3.1)–(3.2) planteado semidiscreto en la variable espacial, lo escribimos de la siguiente manera:

Encontrar una función $u_h \in L^2(0, T, V_h)$ con $u'_h \in L^2(0, T, H^{-1})$ tal que

$$\langle u'_h(t), v_h \rangle + a(u_h(t), v_h) = (f(t), v_h), \forall v_h \in V_h, \text{ c.t. } t \in [0, T], \quad (3.3)$$

$$u_h(0) = P_h(u^0), \quad (3.4)$$

donde $P_h(u^0) : L^2(\Omega) \rightarrow V_h$ es la proyección $L^2(\Omega)$ de u^0 sobre el espacio discreto V_h , es decir que, $\forall v_h \in V_h : (u^0 - P_h(u^0), v_h) = 0$.

El conjunto de todos los nodos interiores de la triangulación \mathcal{T}_h lo denotaremos por \mathcal{N}_h . Para cada t , los parámetros o grados de libertad que describen el espacio de elementos finitos V_h son los valores de $u_h(t)$ en los nodos de la triangulación. Esto significa que cualquier función perteneciente a V_h está unívocamente determinada por sus valores sobre los nodos de \mathcal{T}_h .

El conjunto $\{\phi_i\}_{i=1}^m \subset V_h$ tal que $\phi_i(z_j) = \delta_{ij}$ con $z_j \in \mathcal{N}_h$ (es decir, $\phi_i = 1$ en z_i y $\phi_i = 0$ en cualquier otro nodo), constituye una base de V_h que llamaremos *base nodal*. Para cada t fijo, escribiremos a la función $v_h(t) \in V_h$ como

$$v_h(t) = v_h(\cdot, t) = \sum_{i=1}^m \xi_i(t) \phi_i(\cdot), \quad (3.5)$$

donde cada coeficiente $\xi_i(t) = v_h(z_i, t)$ es el *valor nodal* de $v_h(t)$ correspondiente al nodo z_i .

Es importante notar que V_h representa el espacio de dimensión finita V_m dado en (2.16).

Si introducimos la expresión (3.5) de $v_h(t)$ en la forma débil semidiscreta (3.3)–(3.4) obtendremos exactamente el mismo sistema lineal de ecuaciones diferenciales ordinarias (2.14) que nos permitirá determinar de manera única los coeficientes $\xi_i(t)$, es decir, los valores nodales de v_h en el instante t .

3.2. Estimación a priori del error en la semidiscretización espacial

Supondremos que se tiene una familia de triangulaciones $\{\mathcal{T}_h\}_h$ *uniforme* (ver definición A.22). Esto significa que, por un lado, para cada elemento T , h_T y h son del mismo

orden, lo que hace que los triángulos pertenecientes a una misma partición tengan aproximadamente el mismo tamaño, y por otro lado, que el ángulo mínimo de entre todos los triángulos de $\bigcup\{\mathcal{T}_h\}_h$ está acotado inferiormente, no permitiendo así que tienda a cero (*condición del ángulo mínimo*, Definición A.21). Con el uso de triangulaciones uniformes, el error de aproximación se puede medir en términos del tamaño h de la malla.

Para estudiar el error de discretización espacial, será necesario antes definir lo que llamamos *proyección elíptica* y hacer mención a un resultado de regularidad muy significativo para esta tesis que será de utilidad para arribar a la cota del error buscada. Este resultado se deriva en [8] para problemas elípticos con coeficientes discontinuos y con condiciones de borde de tipo Dirichlet homogéneas.

Definición 3.1 (Proyección elíptica). *Llamamos al operador $P_e : H_0^1(\Omega) \longrightarrow V_h$ proyección elíptica si es la proyección ortogonal con respecto al producto interno energía dado por (2.6), esto es,*

$$\text{si } v \in H_0^1(\Omega) : \quad a(v - P_e v, v_h) = 0, \quad \forall v_h \in V_h.$$

Observación 3.2. *Por la propiedad de ortogonalidad de Galerkin en problemas elípticos (igualdad A.7), podemos notar que en la definición de proyección elíptica, $P_e v$ es la aproximación por elementos finitos de la solución exacta v del problema elíptico.*

Las estimaciones a priori del error dependen de la regularidad de la solución exacta del problema. En problemas elípticos, la discontinuidad del coeficiente de difusión $k(x)$, puede provocar la disminución de la regularidad máxima $H^2(\Omega)$, suponiendo que $f \in L^2(\Omega)$. Por ello, hacen falta resultados finos de regularidad, los que suelen involucrar espacios de Sobolev fraccionarios $H^s(\Omega)$ donde s es un número real (ver más detalles en la sección A.3 del Apéndice).

Sintéticamente, si $s \in \mathbb{R}$, se puede descomponer como $s = m + \sigma$, donde $m \in \mathbb{Z}$ y $0 \leq \sigma < 1$. Cuando $s \in \mathbb{Z}$, el espacio de Sobolev H^s coincide con el espacio de Sobolev usual definido para exponentes enteros y si $s > 0$ pero $s \notin \mathbb{N}$, utilizamos la Definición A.8.

El resultado de regularidad dado en el Teorema A.24 que enunciamos en el Apéndice fue demostrado por M. Petzoldt en [8] para problemas elípticos bidimensionales con coeficiente de difusión discontinuo y para el caso general donde no se imponen condiciones en la estructura del coeficiente k . Allí se muestra que $u \in H^{1+r}(\Omega)$ para un cierto $0 < r < 1/2$ que sólo depende de las cotas de k .

Por lo expuesto en dicho teorema no se espera que la solución de nuestro problema parabólico tenga regularidad $H^2(\Omega)$ en cada t . Por lo tanto, en la semidiscretización espacial haremos la estimación del error suponiendo que $u(t) \in H^{1+r}(\Omega)$, $\forall t \in [0, T]$ y para algún $0 < r < 1/2$.

Ahora sí estamos en condiciones de demostrar el Teorema 3.3 que nos da una estimación a priori del error entre la solución exacta $u(t)$ de (3.1)–(3.2) y la solución $u_h(t)$ del problema semidiscreto (3.3)–(3.4). Su demostración se basa en la que se encuentra en [7,

Sec. 8.3], que se refiere al mismo problema parabólico pero con el coeficiente de difusión continuo. En la demostración del Teorema 3.3, se utiliza la técnica de dualidad y se hace uso de una estimación para el error cometido por la proyección elíptica, que no es más que el error de aproximación del correspondiente problema en estado estacionario.

En efecto, para un t fijo, $u(t) \in H_0^1(\Omega)$ y $f(t) \in L^2(\Omega)$ son ambas dependientes de x , y $u(t)$ es solución del siguiente problema elíptico con coeficiente $k(x)$ discontinuo:

$$a(u(t), v) = \langle f(t) - u'(t), v \rangle \quad \forall v \in H_0^1(\Omega).$$

Por lo dicho en la observación de la Definición 3.1, $P_e u(t) \in V_h$ es la solución discreta del problema débil anterior y $u(t) - P_e u(t)$ define el error de aproximación.

En el Apéndice, dedicamos un capítulo especial a este problema elíptico con coeficiente de difusión discontinuo y concluimos con el Teorema A.31 que presenta una estimación para $\|u(t) - P_e u(t)\|$ que utilizaremos al final de la demostración del Teorema 3.3.

Teorema 3.3 (Estimación a priori del error en la semidiscretización espacial). Sean $u \in L^2(0, T, H_0^1(\Omega))$ con $u' \in L^2(0, T, H^{-1}(\Omega))$ solución de (3.1)–(3.2) y $u_h(t) \in V_h$ solución de (3.3)–(3.4). Entonces si $u \in C(0, T, H^{1+r}(\Omega))$ con $0 < r < 1/2$, existe una constante $C_1 > 0$ independiente de T y h tal que :

$$\max_{0 \leq t \leq T} \|u(t) - u_h(t)\| \leq C_1 (1 + |\log \frac{T}{h^2}|) \max_{0 \leq t \leq T} h^{2r} \sum_{i=1}^2 \sqrt{k_i} |u(t)|_{1+r, \Omega_i}.$$

Demostración. Para obtener la representación del error, introducimos para un $t \in (0, T]$ fijo, una función $\varphi_h : [0, t] \rightarrow V_h$ que es solución del siguiente problema (débil) dual:

$$-\langle \varphi_h'(s), v_h \rangle + a(\varphi_h(s), v_h) = 0, \quad \forall v_h \in V_h, s \in (0, t) \quad (3.6)$$

$$\varphi_h(t) = e_h(t) \quad (3.7)$$

donde $e_h(t) := u_h(t) - P_e u(t)$ y P_e es el operador proyección elíptica (definición 3.1).

En particular, tomemos en (3.6) $v_h = e_h(s) \in V_h$:

$$-\langle \varphi_h'(s), e_h(s) \rangle + a(\varphi_h(s), e_h(s)) = 0, \quad s \in (0, t).$$

Si integramos entre 0 y t , aplicamos integración por partes en el tiempo en el primer término y usamos (3.7) nos queda

$$\|e_h(t)\|^2 = (\varphi_h(0), e_h(0)) + \int_0^t \left(\langle \varphi_h(s), e_h'(s) \rangle + a(\varphi_h(s), e_h(s)) \right) ds. \quad (3.8)$$

Ahora analicemos cada término del lado derecho de (3.8) y notemos que:

(i) $\varphi_h(0) \in V_h$ entonces, por definición de $e_h(t)$ y como $u_h(0) = P_h(u^0)$ se tiene:

$$(\varphi_h(0), e_h(0)) = (\varphi_h(0), u_h(0) - P_e u(0) + u^0 - P_h(u^0)) = (\varphi_h(0), u^0 - P_e u(0)) = (\varphi_h(0), \pi(0)),$$

siendo $\pi(s) := u(s) - P_e u(s)$.

3.2. ESTIMACIÓN A PRIORI DEL ERROR EN LA SEMIDISCRETIZACIÓN ESPACIAL 33

(ii) De acuerdo a las ecuaciones (3.1) y (3.3), y ya que $\varphi_h(s) \in V_h \subset H_0^1(\Omega)$, llegamos a que:

$$\langle u'_h(s), \varphi_h(s) \rangle + a(u_h(s), \varphi_h(s)) = \langle u'(s), \varphi_h(s) \rangle + a(u(s), \varphi_h(s)).$$

(iii) De las definiciones de $e_h(s)$ y P_e y de la igualdad obtenida en (ii), el integrando de (3.8) resulta igual a:

$$\begin{aligned} \langle \varphi_h(s), u'_h(s) - P_e u'(s) \rangle + a(\varphi_h, u_h(s) - P_e u(s)) \\ = \langle \varphi_h(s), u'(s) - P_e u'(s) \rangle + \underbrace{a(\varphi_h(s), u(s) - P_e u(s))}_{=0} \\ = \langle \varphi_h(s), \pi'(s) \rangle, \end{aligned}$$

con $\pi(s)$ definida en (i).

Aplicando los resultados de (i) y (iii) en (3.8) obtenemos:

$$\|e_h(t)\|^2 = (\varphi_h(0), \pi(0)) + \int_0^t \langle \varphi_h(s), \pi'(s) \rangle ds,$$

y luego de integrar por partes en el tiempo,

$$\|e_h(t)\|^2 = (\varphi_h(t), \pi(t)) - \int_0^t \langle \varphi'_h(s), \pi(s) \rangle ds.$$

Si aplicamos ahora la desigualdad de Cauchy-Schwartz y (3.7):

$$\begin{aligned} \|e_h(t)\|^2 &\leq |(\varphi_h(t), \pi(t))| + \int_0^t |\langle \varphi'_h(s), \pi(s) \rangle| ds \\ &\leq \|\varphi_h(t)\| \|\pi(t)\| + \int_0^t \|\varphi'_h(s)\| \|\pi(s)\| ds \\ &\leq \max_{s \in [0, t]} \|\pi(s)\| \left(\|e_h(t)\| + \int_0^t \|\varphi'_h(s)\| ds \right). \end{aligned} \quad (3.9)$$

Supongamos por el momento que valen las siguientes afirmaciones que probaremos al final de este teorema:

$$\|\varphi_h(s)\| \leq \|e_h(t)\|, \quad 0 \leq s \leq t \quad (3.10)$$

$$\int_0^t \|\varphi'_h(s)\| ds \leq C^*(1 + |\log \frac{t}{h^2}|) \|e_h(t)\|, \quad (3.11)$$

para una cierta constante $C^* > 0$ independiente de t y $e_h(t)$.

Utilizando la afirmación (3.11), se sigue de (3.9):

$$\|e_h(t)\| \leq (1 + C^*)(1 + |\log \frac{t}{h^2}|) \max_{s \in [0, t]} \|\pi(s)\|.$$

Luego, podemos escribir

$$\begin{aligned} \|u(t) - u_h(t)\| &\leq \|u(t) - P_e u(t)\| + \|P_e u(t) - u_h(t)\| = \|\pi(t)\| + \|e_h(t)\| \\ &\leq \max_{0 \leq s \leq t} \|\pi(s)\| (2 + C^*) (1 + |\log \frac{t}{h^2}|) \\ &\leq (2 + C^*) (1 + |\log \frac{T}{h^2}|) \max_{0 \leq s \leq t} h^{2r} \sum_{i=1}^2 \sqrt{k_i} |u(s)|_{1+r, \Omega_i}, \end{aligned}$$

Aquí hemos utilizado la estimación del error por Elementos Finitos para problemas elípticos dada por el Teorema A.31 que se demuestra en detalle en el Apéndice.

Maximizando sobre $[0, T]$ obtenemos la estimación del error que queríamos probar, con $C_1 = 2 + C^*$.

Nos queda demostrar las afirmaciones (3.10) y (3.11).

Si consideramos $v_h = \varphi_h(s) \in V_h$ en (3.6) y además (2.21) llegamos a que:

$$-\frac{1}{2} \frac{d}{ds} \|\varphi_h(s)\|^2 + a(\varphi_h(s), \varphi_h(s)) = 0.$$

De aquí, la función $\|\varphi_h(s)\|^2$ es creciente en el intervalo $0 \leq s \leq t$ ya que $a(\varphi_h, \varphi_h) \geq 0$. Luego, teniendo en cuenta la condición inicial (3.7), queda probada (3.10).

Para demostrar (3.11), retomemos la ecuación (3.6). Combinando (2.7) y aplicando las acotaciones conocidas como *estimaciones inversas*, en el método de Elementos Finitos (ver [6, Cap. 2, Sec. 6.8]) se tiene que

$$|\langle \varphi'_h(s), v_h \rangle| = |a(\varphi_h(s), v_h)| \leq \beta \|\nabla \varphi_h(s)\| \|\nabla v_h\| \leq \beta \frac{c}{h^2} \|\varphi_h(s)\| \|v_h\|, \quad \forall v_h \in V_h, \quad (3.12)$$

donde c es una constante que depende del ángulo mínimo de la triangulación.

Si además en (3.12) reemplazamos v_h por $\varphi'_h(s)$ y usamos (3.10), tenemos

$$\|\varphi'_h(s)\| \leq \frac{\beta c}{h^2} \|\varphi_h(s)\| \leq \frac{\beta c}{h^2} \|e_h(t)\|. \quad (3.13)$$

Por otro lado, se puede probar (ver observación 3.4) que $\varphi_h(s)$ satisface

$$\|\varphi'_h(s)\| \leq \frac{1}{\sqrt{2}} \frac{\|e_h(t)\|}{t-s}. \quad (3.14)$$

Entonces, si integramos (3.14) entre 0 y $t - h^2$ y (3.13) entre $t - h^2$ y t , al sumar ambas integrales llegamos a:

$$\begin{aligned} \int_0^t \|\varphi'_h(s)\| ds &= \int_0^{t-h^2} \|\varphi'_h(s)\| ds + \int_{t-h^2}^t \|\varphi'_h(s)\| ds \\ &\leq \|e_h(t)\| \left(\frac{1}{\sqrt{2}} \int_0^{t-h^2} \frac{1}{t-s} ds + \frac{\beta c}{h^2} \int_{t-h^2}^t 1 ds \right) \\ &= \|e_h(t)\| \left(\frac{1}{\sqrt{2}} \log \frac{t}{h^2} + \beta c \right) \\ &\leq \|e_h(t)\| C^* (1 + |\log \frac{t}{h^2}|), \end{aligned}$$

donde $C^* = \max\{\frac{1}{\sqrt{2}}, \beta c\}$. Así concluimos la prueba de la afirmación (3.11). \square

Observación 3.4. Para probar (3.14) es suficiente definir la función $\Psi(s) = \varphi_h(t - s)$, para $0 < s < t$ con φ_h solución del problema

$$\begin{aligned} -\langle \varphi_h'(s), v \rangle + a(\varphi_h(s), v) &= 0, \quad \forall v \in V_h \\ \varphi_h(t) &= e_h(t). \end{aligned}$$

Es fácil ver que $\Psi(s)$ es solución del problema

$$\begin{aligned} \langle \Psi'(s), v \rangle + a(\Psi(s), v) &= 0, \quad \forall v \in V_h \\ \Psi(0) &= e_h(t). \end{aligned}$$

En consecuencia, vale la estimación (2.36) del Lema 2.12:

$$\|\Psi'(s)\| \leq \frac{1}{\sqrt{2}} \frac{\|e_h(t)\|}{s}, \quad 0 < s < t.$$

Luego, obtenemos (3.14) por definición de Ψ y reemplazando s por $t - s$.

3.3. Semidiscretización temporal

En las secciones anteriores hemos estudiado el planteo de la discretización del problema (3.1)–(3.2) con respecto sólo a la variable espacial, además de obtener estimaciones a priori del error para dicha discretización.

Nuestro objetivo en esta sección es estudiar el mismo problema pero cuando sólo la variable temporal es discretizada. Para ello, el procedimiento consiste en subdividir el intervalo de tiempo dado $[0, T]$ considerando una sucesión creciente de instantes de tiempos t_i : $0 = t_0 < t_1 < \dots < t_{n-1} < t_n < \dots < t_N = T$.

Denotaremos con $I_n = (t_{n-1}, t_n)$ a cada paso de tiempo y con $\Delta t^n = t_n - t_{n-1}$ al tamaño del paso de tiempo I_n , para $n = 1, \dots, N$.

Sea $U^n = U(t_n)$ la aproximación de la solución $u(\cdot, t_n)$ obtenida de la discretización de (3.1)–(3.2) respecto de la variable tiempo.

Cuando realicemos discretización temporal, consideraremos el *método de Backward-Euler*, en el que se reemplaza $u'(t)$ por una diferencia finita atrasada. Para este caso, la formulación se presenta de la siguiente manera:

Encontrar una sucesión $\{U^n\} \in H_0^1(\Omega)$ para $n = 1, 2, \dots, N$ que sea solución de

$$\left\langle \frac{U^n - U^{n-1}}{\Delta t^n}, v \right\rangle + a(U^n, v) = (\overline{f_n}, v), \quad \forall v \in H_0^1(\Omega), \quad (3.15)$$

$$U^0 = u^0, \quad (3.16)$$

donde $\overline{f_n}$ es una aproximación constante de f sobre el intervalo I_n .

El método de Backward-Euler es implícito en el sentido que para calcular $U(t_n)$ con $U(t_{n-1})$ conocido, necesitamos resolver una ecuación diferencial débil elíptica. Este método tiene la ventaja de poseer la propiedad óptima para problemas parabólicos: el ser

incondicionalmente estable, es decir, estable y no dependiente de ninguna relación entre h y Δt^n . En efecto, tomando $v = U^n$ en (3.15) y ya que $a(U^n, U^n) \geq 0$ tenemos

$$\langle U^n - U^{n-1}, U^n \rangle \leq (\overline{f_n}, U^n) \leq \|\overline{f_n}\| \|U^n\|,$$

o lo que es lo mismo

$$\|U^n\|^2 - (U^{n-1}, U^n) \leq \|\overline{f_n}\| \|U^n\|.$$

Como $(U^{n-1}, U^n) \leq \|U^{n-1}\| \|U^n\|$, la desigualdad anterior implica que

$$\|U^n\| \leq \|U^{n-1}\| + \|\overline{f_n}\|, \quad \text{para } n \geq 1.$$

Si aplicamos esta desigualdad sucesivas veces llegamos a que

$$\|U^n\| \leq \|U^0\| + \sum_{i=1}^n \|\overline{f_i}\|, \quad n = 1, \dots, N.$$

3.4. Estimación a priori del error en la semidiscretización temporal

El Teorema 3.6 que demostraremos al final de la sección, nos permitirá arribar a una estimación a priori del error entre la solución de (3.1)–(3.2) en t_n que llamaremos $u^n = u(t_n)$ y la solución U^n de (3.15)–(3.16), obtenida de la semidiscretización en el tiempo, para $n = 1, \dots, N$.

Requeriremos que la partición del intervalo $(0, T)$ sea tal que la razón $\Delta t^{n-1}/\Delta t^n$ esté acotada superiormente, es decir, que para alguna constante positiva γ , se cumpla:

$$\Delta t^{n-1} \leq \gamma \Delta t^n, \quad \text{para } n \geq 2. \quad (3.17)$$

Para obtener la estimación del error debido a la semidiscretización espacial, tuvimos la necesidad de demostrar las estimaciones (3.10) y (3.11) referidas a la solución de un problema dual. Análogamente, para obtener la estimación del error debido a la semidiscretización temporal, serán necesarias estimaciones similares, pero relacionadas a la solución de un problema discreto dual. Estas estimaciones se obtienen en el Lema 3.5.

Definimos para un entero $1 \leq n \leq N$, el error $e^n = u^n - U^n$ de semidiscretización temporal cometido en el tiempo t_n .

Lema 3.5. *Supongamos se satisface la condición (3.17). Sean $\{Z^n\} \subset H_0^1(\Omega)$, $n = N, N-1, \dots, 0$ soluciones sucesivas obtenidas del problema discreto dual planteado “hacia atrás”*

$$\left\langle \frac{Z^{n-1} - Z^n}{\Delta t^n}, v \right\rangle + a(Z^{n-1}, v) = 0 \quad \forall v \in H_0^1(\Omega), \quad n = N, \dots, 1 \quad (3.18)$$

$$Z^N = e^N, \quad (3.19)$$

3.4. ESTIMACIÓN A PRIORI DEL ERROR EN LA SEMIDISCRETIZACIÓN TEMPORAL 37

con condición inicial dada por el error en el tiempo final. Entonces, existe una constante $C_2 = C_2(\gamma)$ tal que:

$$\|Z^n\| \leq \|e^N\|, \quad n = 0, \dots, N \quad (3.20)$$

$$\left(\sum_{n=1}^N (T - t_{n-1}) \left\| \frac{Z^{n-1} - Z^n}{\Delta t^n} \right\|^2 \Delta t^n \right)^{1/2} \leq C_2 \|e^N\|, \quad (3.21)$$

$$\sum_{n=1}^N \left\| \frac{Z^{n-1} - Z^n}{\Delta t^n} \right\| \Delta t^n \leq C_2 \left(1 + \log \frac{T}{\Delta t^N}\right)^{1/2} \|e^N\|. \quad (3.22)$$

Demostración. Haciendo $v = Z^{n-1}$ en (3.18) y por la desigualdad de Cauchy obtenemos

$$\|Z^{n-1}\|^2 + \Delta t^n a(Z^{n-1}, Z^{n-1}) = (Z^n, Z^{n-1}) \leq \frac{\|Z^n\|^2}{2} + \frac{\|Z^{n-1}\|^2}{2}.$$

Entonces

$$\|Z^{n-1}\| \leq \|Z^n\|, \quad \text{para } n = 1, \dots, N. \quad (3.23)$$

Aplicando esta desigualdad sucesivas veces llegamos a (3.20).

Para probar las dos desigualdades que faltan, haremos un cambio de variables en el problema dado con el fin de facilitar la demostración.

Definimos $s_n = T - t_{N-n}$, con $\Delta s^n = s_n - s_{n-1}$. Por medio del cambio de variables:

$$\begin{aligned} X^n &= Z^{N-n} \quad n = 0, \dots, N \\ \Delta s^n &= \Delta t^{N-n+1}, \quad n = 1, \dots, N \end{aligned}$$

se puede notar que el problema (3.18)–(3.19) se transforma en el siguiente problema discreto planteado “hacia adelante”:

Encontrar $\{X^0, X^1, X^2, \dots, X^N\} \subset H_0^1(\Omega)$ donde $X^n = X(s_n)$ son las soluciones sucesivas de

$$\langle X^n - X^{n-1}, v \rangle + \Delta s^n a(X^n, v) = 0, \quad \forall v \in H_0^1(\Omega), \quad n = 1, \dots, N \quad (3.24)$$

$$X^0 = e^N \quad (3.25)$$

De aquí y dado que vale (3.17), la sucesión de pasos de tiempo del nuevo problema satisface

$$\Delta s^n \leq \gamma \Delta s^{n-1}, \quad n = 2, \dots, N. \quad (3.26)$$

Después de tomar $v = X^n$ en (3.24), aplicar la desigualdad de Cauchy y sumar de 1 a N , obtenemos

$$\sum_{n=1}^N \Delta s^n a(X^n, X^n) \leq \frac{\|X^0\|^2}{2} - \frac{\|X^N\|^2}{2} \leq \frac{\|X^0\|^2}{2}. \quad (3.27)$$

Por otro lado, si elegimos $v = s_n(X^n - X^{n-1})$ en (3.24) y recordamos la definición (2.9) de norma energía, llegamos a que

$$\frac{1}{\Delta s^n} s_n \|X^n - X^{n-1}\|^2 + s_n \|X^n\|_a^2 = s_n a(X^n, X^{n-1}).$$

A partir de aquí, después de aplicar la desigualdad de Cauchy y usar (3.26), resulta para $n \geq 2$:

$$\begin{aligned} \Delta s^n s_n \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\|^2 + \frac{s_n}{2} \|X^n\|_a^2 &\leq \frac{s_n}{2} \|X^{n-1}\|_a^2 \\ &\leq \frac{s_{n-1}}{2} \|X^{n-1}\|_a^2 + \frac{\gamma}{2} \Delta s^{n-1} \|X^{n-1}\|_a^2. \end{aligned}$$

Luego,

$$\Delta s^n s_n \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\|^2 + \frac{s_n}{2} \|X^n\|_a^2 - \frac{s_{n-1}}{2} \|X^{n-1}\|_a^2 \leq \frac{\gamma}{2} \Delta s^{n-1} \|X^{n-1}\|_a^2.$$

Sumando esta última desigualdad desde $n = 2$ a $n = N$ y teniendo en cuenta (3.27), se tiene:

$$\begin{aligned} \sum_{n=2}^N \Delta s^n s_n \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\|^2 + \frac{s_N}{2} \|X^N\|_a^2 - \frac{s_1}{2} \|X^1\|_a^2 \\ \leq \frac{\gamma}{2} \sum_{n=1}^{N-1} \Delta s^n \|X^n\|_a^2 \leq \frac{\gamma}{4} \|X^0\|^2, \end{aligned} \quad (3.28)$$

mientras que para $n = 1$ en (3.24), tomando $v = s_1(X^1 - X^0)$ tenemos

$$\Delta s^1 s_1 \left\| \frac{X^1 - X^0}{\Delta s^1} \right\|^2 + s_1 \|X^1\|_a^2 = s_1 a(X^1, X^0) = \Delta s^1 a(X^1, X^0), \quad (3.29)$$

puesto que $\Delta s^1 = s_1$.

Notemos además que en particular, cuando $n = 1$ y $v = X^0$ en (3.24), resulta:

$$\Delta s^1 a(X^1, X^0) = \|X^0\|^2 - \langle X^1, X^0 \rangle \leq \|X^0\|^2 + \frac{\|X^1\|^2}{2} + \frac{\|X^0\|^2}{2}.$$

Pero ya que se cumple (3.23), por el cambio de variables se satisface que $\|X^n\| \leq \|X^{n-1}\|$ para $n = N, \dots, 1$, con lo que $\Delta s^1 a(X^1, X^0) \leq 2\|X^0\|^2$, quedando entonces a partir de (3.29):

$$\Delta s^1 s_1 \left\| \frac{X^1 - X^0}{\Delta s^1} \right\|^2 + s_1 \|X^1\|_a^2 \leq 2\|X^0\|^2. \quad (3.30)$$

Por lo tanto, de la suma de (3.30) y (3.28) se deduce que

$$\sum_{n=1}^N \Delta s^n s_n \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\|^2 \leq C_2^2 \|X^0\|^2, \quad (3.31)$$

donde $C_2 = C_2(\gamma) = \sqrt{\frac{\gamma}{4} + 2}$. Luego, teniendo en cuenta el cambio de variables llegamos a (3.21).

Para probar (3.22), observemos que de la desigualdad de Cauchy - Schwartz

$$\sum_{n=1}^N \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\| \Delta s^n \leq \left(\sum_{n=1}^N \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\|^2 \Delta s^n s_n \right)^{1/2} \left(\sum_{n=1}^N \frac{\Delta s^n}{s_n} \right)^{1/2}.$$

Pero

$$\sum_{n=1}^N \frac{\Delta s^n}{s_n} = 1 + \sum_{n=2}^N \frac{\Delta s^n}{s_n} \leq 1 + \int_{s_1}^{s_N} \frac{1}{s} ds = 1 + \log\left(\frac{s_N}{s_1}\right).$$

Con este resultado y aplicando (3.31) tenemos que

$$\sum_{n=1}^N \left\| \frac{X^n - X^{n-1}}{\Delta s^n} \right\| \Delta s^n \leq C_2 \|X^0\| \left(1 + \log \frac{T}{\Delta s^1}\right)^{1/2}.$$

Luego, el cambio de variables nos permite llegar a (3.22). \square

El lema que acabamos de probar, nos será útil para obtener la estimación del error provocado por la semidiscretización temporal del problema (3.1)–(3.2). Dicha estimación se demuestra en el Teorema 3.6 a continuación.

Denotamos con \overline{f}_n y \overline{U}_n a los promedios de f y u respectivamente, en el intervalo de tiempo (t_{n-1}, t_n) , es decir, $\overline{f}_n = \frac{1}{\Delta t^n} \int_{I_n} f(t) dt$ y $\overline{U}_n = \frac{1}{\Delta t^n} \int_{I_n} u(t) dt$. Claramente $\overline{f}_n, \overline{U}_n \in H_0^1(\Omega)$. Supondremos de ahora en más que $u' \in C(0, T, L^2(\Omega))$.

Teorema 3.6 (Estimación a priori del error en la semidiscretización temporal). *Sean $u^n = u(t_n) \in H_0^1(\Omega)$ soluciones de (3.1)–(3.2) y $U^n \in H_0^1(\Omega)$ soluciones del problema semidiscreto (3.15)–(3.16) y supongamos que vale (3.17). Entonces si $u' \in C(0, T, L^2(\Omega))$, existe una constante $C_2 = C_2(\gamma) > 0$ (que coincide con la constante del Lema 3.5), tal que:*

$$\max_{1 \leq n \leq N} \|u^n - U^n\| \leq C_2 \left(1 + \log \frac{T}{\tau}\right)^{1/2} \max_{1 \leq n \leq N} \Delta t^n \max_{t \in I_n} \|u'(t)\|, \quad (3.32)$$

donde $\tau = \min_{1 \leq n \leq N} \Delta t^n$.

Demostración. Análogamente a la semidiscretización espacial, nos basaremos en la técnica de dualidad. Comenzamos integrando la forma débil (3.1) sobre $I_n = (t_{n-1}, t_n)$:

$$\int_{I_n} \int_{\Omega} u'(t)v dxdt + \int_{I_n} \int_{\Omega} k(x)\nabla u \nabla v dxdt = \int_{I_n} \int_{\Omega} f(t)v dxdt, \quad \forall v \in H_0^1(\Omega).$$

Si en cada integral intercambiamos el orden de integración y usamos el hecho de que $\int_{I_n} \nabla u dt = \nabla \int_{I_n} u dt$ y que v no depende de t , obtenemos para $n = 1, \dots, N$

$$\langle u^n - u^{n-1}, v \rangle + \Delta t^n a(\overline{U}_n, v) = \Delta t^n \langle \overline{f}_n, v \rangle.$$

Por otro lado, multiplicando la ecuación (3.15) por Δt^n :

$$\langle U^n - U^{n-1}, v \rangle + \Delta t^n a(U^n, v) = \Delta t^n \langle \overline{f}_n, v \rangle.$$

Al restar las dos últimas ecuaciones llegamos a que

$$\langle e^n - e^{n-1}, v \rangle + \Delta t^n a(\overline{U}_n - U^n, v) = 0, \quad \forall v \in H_0^1(\Omega). \quad (3.33)$$

Si tomamos ahora en (3.33), $v = Z^{n-1} \in H_0^1(\Omega)$, solución de (3.18)–(3.19) con $1 \leq n \leq N$ y luego sumamos para $n = 1, \dots, N$, nos queda

$$\sum_{n=1}^N \langle e^n - e^{n-1}, Z^{n-1} \rangle + \sum_{n=1}^N \Delta t^n a(\bar{U}_n - U^n, Z^{n-1}) = S_1 + S_2 = 0, \quad (3.34)$$

donde llamamos S_1 y S_2 a la primera y segunda sumatoria respectivamente.

A continuación, analizaremos por separado cada término S_1 y S_2 .

$$\begin{aligned} S_1 &= \sum_{n=0}^{N-1} \langle e^{n+1} - e^n, Z^n \rangle = \langle e^N, Z^{N-1} \rangle - \langle e^0, Z^0 \rangle - \sum_{n=1}^{N-1} \langle Z^n - Z^{n-1}, e^n \rangle \\ &= \langle e^N, Z^{N-1} \rangle - \langle e^0, Z^0 \rangle - \sum_{n=0}^{N-1} \langle Z^{n+1} - Z^n, e^{n+1} \rangle + \langle Z^N - Z^{N-1}, e^N \rangle \\ &= \langle e^N, Z^N \rangle - \langle e^0, Z^0 \rangle - \sum_{n=0}^{N-1} \langle e^{n+1}, Z^{n+1} - Z^n \rangle. \end{aligned} \quad (3.35)$$

Notemos que si hacemos $v = \bar{U}_n - U^n \in H_0^1(\Omega)$ en la ecuación (3.18) y sumamos desde $n = 1$ hasta $n = N$, obtendremos

$$\sum_{n=1}^N \langle Z^{n-1} - Z^n, \bar{U}_n - U^n \rangle + \sum_{n=1}^N \Delta t^n a(Z^{n-1}, \bar{U}_n - U^n) = 0,$$

donde se ve que la segunda sumatoria es S_2 .

De aquí,

$$S_2 = - \sum_{n=1}^N \langle Z^{n-1} - Z^n, \bar{U}_n - U^n \rangle = - \sum_{n=0}^{N-1} \langle Z^{n+1} - Z^n, U^{n+1} - \bar{U}_{n+1} \rangle. \quad (3.36)$$

Reemplazando las expresiones finales de S_1 y S_2 resultantes de (3.35) y (3.36) en (3.34), obtenemos:

$$\langle e^N, Z^N \rangle - \langle e^0, Z^0 \rangle - \sum_{n=0}^{N-1} \langle Z^{n+1} - Z^n, \underbrace{e^{n+1} + U^{n+1}}_{u^{n+1}} - \bar{U}_{n+1} \rangle = 0.$$

Luego, como $Z^N = e^N$ y por la desigualdad de Cauchy - Schwartz:

$$\|e^N\|^2 \leq \|e^0\| \|Z^0\| + \left(\sum_{n=0}^{N-1} \|Z^{n+1} - Z^n\| \right) \max_{0 \leq n \leq N-1} \|u^{n+1} - \bar{U}_{n+1}\|. \quad (3.37)$$

Si utilizamos (3.22) en esta desigualdad, llegamos a

$$\|e^N\| \leq C_2 \left(1 + \log \frac{T}{\Delta t^N}\right)^{1/2} \max_{1 \leq n \leq N} \|u^n - \bar{U}_n\|, \quad (3.38)$$

ya que $\|e^0\| = 0$.

Para terminar, analicemos $\|u^n - \bar{U}_n\|$:

$$\|u^n - \bar{U}_n\|^2 = \int_{\Omega} |u(x, t_n) - \bar{U}_n(x)|^2 dx = \int_{\Omega} \left| \int_{t(x)}^{t_n} \partial_t u(x, t) dt \right|^2 dx,$$

puesto que para cada $x \in \Omega$ existe $t = t(x) \in I_n$ tal que $u(x, t) = \bar{U}_n(x)$, dado que $\bar{U}_n(x) = \frac{1}{\Delta t^n} \int_{I_n} u(x, t) dt$.

De esta manera,

$$\begin{aligned} \|u^n - \bar{U}_n\|^2 &\leq \int_{\Omega} \left[\left(\int_t^{t_n} dt \right)^{1/2} \left(\int_t^{t_n} |\partial_t u(x, t)|^2 dt \right)^{1/2} \right]^2 dx \\ &\leq \Delta t^n \int_{\Omega} \int_{I_n} |\partial_t u(x, t)|^2 dt dx \\ &\leq (\Delta t^n)^2 \max_{t \in I_n} \int_{\Omega} |\partial_t u(x, t)|^2 dx, \end{aligned}$$

después de intercambiar las integrales.

Entonces

$$\|u^n - \bar{U}_n\| \leq \Delta t^n \max_{t \in I_n} \left(\int_{\Omega} |\partial_t u(x, t)|^2 dx \right)^{1/2}, \quad \text{para } n = 1, \dots, N.$$

Con este resultado y dado que la cota del error (3.38) en lugar de t_N también vale para cualquier tiempo t_n de la partición, se obtiene (3.32). \square

3.5. Discretización completa

Consideremos el problema (3.15)–(3.16), correspondiente a la semidiscretización con respecto al tiempo y sea $V_h \subset H_0^1(\Omega)$ el espacio de los elementos finitos lineales a trozos asociados a una partición uniforme del dominio de tamaño h . El planteo de la formulación completamente discreta del problema consiste en:

Para $n = 1, \dots, N$, dado $U_h^{n-1} \in V_h$, encontrar $U_h^n \in V_h$, solución de

$$\left\langle \frac{U_h^n - U_h^{n-1}}{\Delta t^n}, v \right\rangle + a(U_h^n, v) = (\bar{f}_n, v), \quad \forall v \in V_h, \quad (3.39)$$

$$U_h^0 = P_{V_h}(u^0), \quad (3.40)$$

donde $P_{V_h}(u^0)$ es la proyección $L^2(\Omega)$ de u^0 sobre el espacio V_h , es decir que satisface

$$(P_{V_h}(u^0), v) = (u^0, v) \quad \forall v \in V_h. \quad (3.41)$$

Si $\{\phi_i(x)\}_{i=1}^m$ es la base nodal de V_h , escribiremos a $U_h^n(x) \in V_h$ como

$$U_h^n(x) = \sum_{i=1}^m \xi_i(t_n) \phi_i(x),$$

donde $\xi_i(t_n)$ son los valores nodales de U_h^n que serán unívocamente determinados.

Si reemplazamos v por $\phi_j \in V_h$ en (3.39) y (3.41) y luego introducimos la expresión de U_h^n en términos de la base de V_h en (3.39), obtenemos un sistema lineal de ecuaciones que tiene la forma matricial siguiente:

$$\begin{cases} (M + \Delta t^n K)\xi^n = M\xi^{n-1} + \Delta t^n F(t_n), n = 1, \dots, N \\ M\xi^0 = U^0 \end{cases}$$

donde M es la matriz de masa de elementos (ϕ_i, ϕ_j) y K la matriz de rigidez con elementos $a(\phi_i, \phi_j)$, ambas de orden $m \times m$. Los vectores columnas $F(t_n)$ y U^0 tienen su j -ésima componente igual a $(\overline{f_n}, \phi_j)$ y (u^0, ϕ_j) respectivamente y el vector incógnita ξ^n es el correspondiente vector de valores nodales de U_h^n , es decir, $(\xi^n)_j = \xi_j(t_n) = U_h^n(z_j)$, $z_j \in \mathcal{N}_h$, siendo \mathcal{N}_h el conjunto de nodos interiores de la partición de tamaño h .

3.6. Estimación a priori del error en la discretización completa

Recordemos que $u(t) \in H_0^1(\Omega)$ y $U_h^n \in V_h$ son las respectivas soluciones del problema débil (3.1)–(3.2) y su análogo completamente discreto (3.39)–(3.40).

Estamos interesados en encontrar una cota para estimar el error $u(t) - U_h^n$ asociado a la discretización completa del problema (3.1)–(3.2). La demostración de este resultado se encuentra en el Teorema 3.7, en donde se presenta una cota para el máximo entre las normas de valores discretos del error. La idea general se basa en descomponer el máximo de $\|u^n - U_h^n\|$ en función de $\|u^n - u_h(t_n)\|$, que mide el error en el tiempo t_n debido a la semidiscretización espacial y de $\|u_h(t_n) - U_h^n\|$, que mide el error en la semidiscretización temporal sobre el espacio discreto $V_h \subset H_0^1(\Omega)$, para luego usar las acotaciones de estos errores que fueron dadas por teoremas que demostramos en secciones anteriores.

También demostraremos que la acotación del Teorema 3.7 es válida sobre el intervalo completo $[0, T]$ (Corolario 3.8).

Teorema 3.7 (Estimación a priori del error en la discretización completa). Sean $u \in L^2(0, T, H_0^1(\Omega))$ con $u' \in L^2(0, T, H^{-1}(\Omega))$ solución de (3.1)–(3.2) y $\{U_h^n\} \subset V_h$ sucesión de soluciones de (3.39)–(3.40) obtenidas de la discretización completa y suponemos además que vale (3.17). Entonces si $u \in C(0, T, H^{1+r}(\Omega))$ para algún $0 < r < 1/2$ y $u' \in C(0, T, L^2(\Omega))$, existe una constante $C = C(k, \gamma, r, u) > 0$ independiente de h y de Δt^n , tal que:

$$\max_{1 \leq n \leq N} \|u^n - U_h^n\| \leq \|e(0)\| + C(1 + |\log \frac{T}{h^2}| + \log \frac{T}{\tau})(h^{2r} + \max_{1 \leq n \leq N} \Delta t^n), \quad (3.42)$$

donde $\tau = \min_{1 \leq n \leq N} \Delta t^n$.

Demostración. La estimación del error $u^n - U_h^n \in H_0^1(\Omega)$ en la semidiscretización temporal dada en (3.32), fue obtenida para el problema (3.15)–(3.16) planteado sobre $V = H_0^1(\Omega)$.

Si reemplazamos V por V_h , las soluciones U^n de (3.15)–(3.16) son las soluciones U_h^n del problema completamente discreto y el error en la semidiscretización temporal sobre este espacio discreto será $u_h(t_n) - U_h^n$.

3.6. ESTIMACIÓN A PRIORI DEL ERROR EN LA DISCRETIZACIÓN COMPLETA 43

La cota para este error coincide con (3.32) excepto que tendría presente un término más. Dicho término es la norma L^2 del error inicial $e(0) = u_h(0) - u^0$ que para este caso no es cero. Si tenemos en cuenta ésto en la demostración del Teorema 3.6 a partir (3.37), se obtendría como cota final:

$$\max_{1 \leq n \leq N} \|u_h(t_n) - U_h^n\| \leq \|e(0)\| + C_2(1 + \log \frac{T}{\tau})^{1/2} \max_{1 \leq n \leq N} (\Delta t^n \max_{t \in I_n} \|u'(t)\|). \quad (3.43)$$

Por otro lado, probamos en el Teorema 3.3 que el error entre la solución exacta y la solución del problema de semidiscretización espacial satisface

$$\max_{0 \leq t \leq T} \|u(t) - u_h(t)\| \leq C_1(1 + |\log \frac{T}{h^2}|) \max_{0 \leq t \leq T} h^{2r} \sum_{i=1}^2 \sqrt{k_i} |u(t)|_{1+r, \Omega_i}. \quad (3.44)$$

Si descomponemos el máximo de $\|u^n - U_h^n\|$ como

$$\begin{aligned} \max_{1 \leq n \leq N} \|u^n - U_h^n\| &\leq \max_{1 \leq n \leq N} \|u^n - u_h(t_n)\| + \max_{1 \leq n \leq N} \|u_h(t_n) - U_h^n\| \\ &\leq \max_{0 \leq t \leq T} \|u(t) - u_h(t)\| + \max_{1 \leq n \leq N} \|u_h(t_n) - U_h^n\|, \end{aligned}$$

podemos aplicar las estimaciones del error debido a las discretizaciones espacial y temporal (3.44) y (3.43) de la siguiente manera:

$$\begin{aligned} \max_{1 \leq n \leq N} \|u^n - U_h^n\| &\leq C_1(1 + |\log \frac{T}{h^2}|) h^{2r} \max_{0 \leq t \leq T} \sum_{i=1}^2 \sqrt{k_i} |u(t)|_{1+r, \Omega_i} + \\ &\quad \|e(0)\| + C_2(1 + \log \frac{T}{\tau})^{1/2} \max_{1 \leq n \leq N} (\Delta t^n \max_{t \in I_n} \|u'(t)\|) \\ &\leq \|e(0)\| + \tilde{C}_1 \left(1 + |\log \frac{T}{h^2}| + \log \frac{T}{\tau}\right) h^{2r} + \\ &\quad \tilde{C}_2 \left((1 + \log \frac{T}{\tau})^{1/2} + |\log \frac{T}{h^2}|\right) \max_n \Delta t^n, \end{aligned}$$

donde definimos las constantes positivas $\tilde{C}_1 = C_1 \max_{0 \leq t \leq T} \sum_{i=1}^2 \sqrt{k_i} |u(t)|_{1+r, \Omega_i}$ y $\tilde{C}_2 = C_2 \max_{1 \leq n \leq N} \max_{t \in I_n} \|u'(t)\|$, siendo C_1 y C_2 las constantes de los Teoremas (3.3) y (3.6) respectivamente.

Por último, ya que $(1 + \log \frac{T}{\tau})^{1/2} \leq 1 + \log \frac{T}{\tau}$, concluimos la representación a priori del error en la discretización completa (3.42), con $C = \max\{\tilde{C}_1, \tilde{C}_2\}$. \square

Podemos observar que la representación del error que acabamos de obtener en el Teorema 3.7, está dada en términos de la norma L^2 del error $u(t) - U_h^n$ sobre valores discretos del intervalo $[0, T]$. En el siguiente corolario probaremos que introduciendo un interpolante de U_h^n , la cota obtenida para el $\max_n \|u^n - U_h^n\|$ es válida también sobre todo $[0, T]$.

Corolario 3.8. *Supongamos que valen las mismas hipótesis del Teorema 3.7. Entonces si $u \in C(0, T, H^{1+r}(\Omega))$ para algún $0 < r < 1/2$ y $u' \in C(0, T, L^2(\Omega))$, existe una constante $\tilde{C} > 0$ independiente de h y de Δt^n tal que:*

$$\max_{0 \leq t \leq T} \|u(t) - \tilde{U}(t)\| \leq \|e(0)\| + \tilde{C}(1 + |\log \frac{T}{h^2}| + \log \frac{T}{\tau})(h^{2r} + \max_{1 \leq n \leq N} \Delta t^n).$$

donde $\tilde{U}(t)$ es un interpolante constante a trozos de U_h^n .

Demostración. Definimos $\tilde{U}(t)$ como el interpolante constante a trozos de U_h^n dado por $\tilde{U}(t) = U_h^n$ en $t_{n-1} < t \leq t_n$ (es decir $\tilde{U}(t)$ es continua por izquierda). Mostraremos que entonces existe $\tilde{C} > 0$ tal que vale la misma acotación (3.42) para $\max_{0 \leq t \leq T} \|u(t) - \tilde{U}(t)\|$.

En efecto, si $t \in (t_{n-1}, t_n]$, podemos escribir:

$$\tilde{U}(t) - u(t) = U_h^n - u^n + \int_t^{t_n} u'(s) ds.$$

Entonces, llamando $L_{h,\tau} = 1 + |\log \frac{T}{h^2}| + \log \frac{T}{\tau}$ y aplicando (3.42):

$$\begin{aligned} \|(\tilde{U} - u)(t)\| &\leq \|U_h^n - u^n\| + \int_t^{t_n} \|u'(s)\| ds \\ &\leq \|e(0)\| + CL_{h,\tau}(h^{2r} + \max_n \Delta t^n) + \max_n \Delta t^n \max_{1 \leq n \leq N} \max_{t \in (t_{n-1}, t_n]} \|u'(t)\| \\ &\leq \|e(0)\| + CL_{h,\tau}(h^{2r} + \max_n \Delta t^n) + (\max_n \Delta t^n + h^{2r}) \frac{\tilde{C}_2}{C_2}. \end{aligned}$$

donde C , \tilde{C}_2 y C_2 son las constantes definidas en el Teorema 3.7.

Como $\frac{\tilde{C}_2}{C_2} \leq \frac{\tilde{C}_2}{C_2} L_{h,\tau}$ puesto que $L_{h,\tau} > 1$, se puede realizar la siguiente acotación:

$$\|(\tilde{U} - u)(t)\| \leq \|e(0)\| + \tilde{C} L_{h,\tau} (h^{2r} + \max_n \Delta t^n),$$

siendo $\tilde{C} = \max\{C, \frac{\tilde{C}_2}{C_2}\}$ y $t \in (t_{n-1}, t_n]$. Luego, tomando el máximo sobre el intervalo completo $[0, T]$ llegamos a lo que queríamos probar. \square

Es importante observar que, al momento de resolver numéricamente la discretización completa, tradicionalmente se comienza escogiendo una partición del intervalo de tiempo $[0, T]$ en la que cada subintervalo puede tener tamaño Δt^n no necesariamente constante, y una partición uniforme de tamaño h del dominio. Estos parámetros no serán modificados en todo el proceso para calcular los U_h^n , pero puede ocurrir que la elección de Δt^n y de h que permanecieron fijos en el algoritmo, conduzca a resultados no deseados en la cota del error. Por ejemplo, si los parámetros de discretización son muy grandes, los errores también pueden ser grandes y en consecuencia, la aproximación sería mala. O bien, si los parámetros de discretización son muy pequeños, la aproximación sería muy buena pero el costo computacional elevado.

Actualmente se utilizan algoritmos adaptativos. Lo usual es, dados Δt^n y h iniciales, resolver el problema en cada paso de tiempo permitiendo modificaciones en Δt^n y h (localmente) de manera que el error cometido en ese subintervalo de tiempo permanezca menor que una cierta tolerancia. Para esto, hacen falta los estimadores a posteriori del error que presentaremos en el próximo capítulo.

Capítulo 4

Estimaciones a posteriori del error

Cuando en el capítulo anterior discretizamos el problema débil (3.1)–(3.2) en la variable espacio y/o tiempo, hemos podido observar que las cotas obtenidas para el error dependen de $k(x)$, h y Δt y lo que es más importante, requieren del conocimiento de la solución exacta y de su regularidad. De aquí que estas cotas indicadoras del error son llamadas *estimaciones a priori* del error.

El problema es que, en la mayoría de las aplicaciones no es posible satisfacer este requerimiento. Por ello, surge la necesidad de analizar otro tipo de indicador del error, apropiado para el problema que estudiamos y que no requiere conocer la solución exacta.

Por otro lado, como comentamos al final del capítulo anterior, una inadecuada elección de los tamaños Δt^n de los subintervalos de tiempo o el uso de triangulaciones uniformes del dominio, pueden acarrear resultados no deseados. Por ejemplo, en dominios con singularidades, tales como esquinas o ángulos entrantes, el error puede ser muy grande en las cercanías de estas regiones y puede propagarse en aquellas otras que no las contienen. En estos casos, las estimaciones a priori del error que resultarían pesimistas no reflejarían el verdadero comportamiento del error en todo el dominio.

Para reducir el error en la aproximación usaremos la *adaptatividad* en las discretizaciones de espacio y tiempo. El procedimiento adaptativo consiste en agregar nodos sobre la malla donde la solución es menos regular y/o elegir los pasos de tiempo automáticamente de modo de controlar el error para que permanezca menor que una cierta tolerancia y de esta manera quede equidistribuido en todos los elementos del dominio.

El control del error se centra en el uso de *estimadores a posteriori del error* que son cantidades calculables a partir de la solución discreta y de los datos del problema y que miden el error sin tener conocimiento de la solución exacta. En esta tesis, nos apoyamos en el trabajo de S. Berrone [9] para el estudio de los estimadores del error a posteriori apropiados a las características de nuestro problema y uno de estos estimadores involucran al residuo de la solución discreta calculada.

Para obtener los estimadores a posteriori necesitaremos definir un operador de interpolación que tenga en cuenta la *condición de casi monotonía* del coeficiente de difusión. Esta condición sobre k , permitirá que su variación, es decir, el valor del salto de discontinuidad entre los valores máximo y mínimo de k , no se presente en las cotas del error. De esta manera lograremos desarrollar estimadores *robustos*.

4.1. Discretización numérica

Consideremos el problema (2.1) con las mismas hipótesis hechas sobre Ω y sobre los datos $k(x)$ y u^0 y supongamos $f \equiv 0$. La formulación débil de este problema es la dada en (3.1)–(3.2) pero con el lado derecho igual a cero.

A diferencia del capítulo anterior, en esta sección enfatizamos la aproximación mediante discretización completa en lugar de considerar formulaciones semidiscretas, para arribar a estimadores distintos para el error de discretización espacial y el error de discretización temporal.

Consideremos una partición del $[0, T]$ en subintervalos $I_n = (t_{n-1}, t_n)$ de longitud $\Delta t^n = t_n - t_{n-1}$, $n = 1, \dots, N$.

Para cada eslabón $\Omega \times \overline{I_n}$, sea $\{\mathcal{T}_h^n\}$ una familia de particiones de triángulos T sobre $\overline{\Omega}$, admisible, regular y que satisfacen la condición del ángulo mínimo (ver Apéndice, definiciones A.20 y A.21). Los dos últimos requerimientos son equivalentes, por ello lo importante de la condición de regularidad es que no permite ángulos muy pequeños en los triángulos y de esta manera aquellos generados en los sucesivos refinamientos no se degeneran haciéndose arbitrariamente finos.

Asociado a cada triangulación \mathcal{T}_h^n , consideremos el espacio de elementos finitos $V_h^n \subset H_0^1(\Omega)$ de funciones continuas y lineales a trozos que se anulan en $\partial\Omega$, esto es,

$$V_h^n = \{v \in \mathcal{C}(\overline{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h^n\}.$$

Para comenzar definimos $U_h^0 = P_{V_h^0} u^0$, como la proyección $L^2(\Omega)$ de u^0 sobre el espacio inicial V_h^0 . Entonces la formulación completamente discreta del problema (3.1)–(3.2) consiste en, teniendo conocida $U_h^{n-1} \in V_h^{n-1}$, encontrar $U_h^n \in V_h^n$ tal que

$$\left(\frac{U_h^n - U_h^{n-1}}{t_n - t_{n-1}}, v \right) + (k \nabla U_h^n, \nabla v) = 0, \quad \forall v \in V_h^n, \text{ y para } n = 1, \dots, N. \quad (4.1)$$

Notemos que la sucesión de aproximaciones U_h^n pueden pertenecer a subespacios discretos V_h^n distintos, a diferencia de las aproximaciones obtenidas en la sección 3.5 en la que todas pertenecían al mismo espacio discreto V_h .

La dimensión de los subespacios V_h^n es el número de nodos interiores de la triangulación donde está definido dicho espacio y las funciones de base $\{\phi_i\}$ constituyen una base nodal de dicho espacio.

Hacemos la salvedad que durante todo el proceso de adaptatividad con respecto a la variable espacial, sólo trabajaremos con refinamientos y no con engrosamientos de la malla del dominio. Esto implica que el subespacio discreto correspondiente a un paso de tiempo estará contenido en el subespacio discreto del paso de tiempo siguiente, es decir, $V_h^n \subset V_h^{n+1}$. De esta manera, la sucesión de subespacios de dimensión finita finalmente generada por la adaptatividad son anidados.

4.2. Operadores de interpolación

En los textos clásicos de Elementos Finitos la interpolación proporciona una herramienta muy útil para obtener resultados teóricos de aproximación.

En nuestro caso y con el mismo objetivo, analizaremos qué operadores de interpolación usuales serán más ventajosos de usar de acuerdo a las características de nuestro problema.

Previamente introduciremos la siguiente notación. Para cada nivel $\Omega \times \bar{I}_n$ y para un subconjunto $S \subset \bar{\Omega}$, denotaremos con $\mathcal{N}_h^n(S)$ al conjunto de todos los nodos de la triangulación contenidos en S y análogamente con $\mathcal{E}_h^n(S)$ al conjunto de los lados contenidos en S . Todos los triángulos $T \in \mathcal{T}_h^n$ y lados $E \in \mathcal{E}_h^n(\bar{\Omega})$ son conjuntos cerrados. Denotamos con h_T^n al diámetro del elemento T y con h_E^n a la longitud del lado E .

Para un nodo x y un lado E , definimos a los conjuntos w_x y w_E como la unión de todos los triángulos que tienen a x o a E respectivamente.

Al valor del parámetro $k(x)$ sobre cada triángulo T lo simbolizamos con k_T , definimos $k_E = \sum_{T \subset w_E} k_T$ y denotamos con T_E al triángulo de w_E con el máximo k_T .

Escribiremos $a \preceq b$ cuando existe una constante c independiente del tamaño de la malla, paso de tiempo, coeficiente k y salto de discontinuidad de k , tal que $a \leq cb$ e indicaremos con $a \approx b$ cuando $a \preceq b$ y $b \preceq a$.

La hipótesis de regularidad en cada malla (Definición A.20), hace que existan constantes que sólo dependen del ángulo mínimo de la triangulación tales que, para cada $n = 1, \dots, N$:

$$\text{Area}(T) \approx (h_T^n)^2, \quad \forall T \in \mathcal{T}_h^n \quad \text{y} \quad h_T^n \approx h_E^n, \quad \forall E \in \mathcal{E}_h(T).$$

En una familia uniforme de triangulaciones, todos los triángulos pertenecientes a una partición tienen el mismo o aproximadamente el mismo tamaño, dependiendo del modo de particionar cada triángulo. Este tipo de mallas se utilizan en las estimaciones a priori del error (lo vimos en el capítulo anterior), en donde la cota depende de h , entre otros. En cambio, la condición de regularidad en una malla permite particiones en el dominio con triángulos de muy diferentes tamaños. Esto será de gran utilidad en mallas originadas en procedimientos adaptativos, donde los elementos correspondientes a regiones donde la solución no es tan suave, son refinados mientras que otros elementos pueden no haberse refinado nunca durante el procedimiento.

Por el teorema de inmersión de Sobolev (Teorema A.17), las funciones en $H^{1+r}(\Omega)$ con $r > 0$ son continuas en $\bar{\Omega}$. Para esta clase de funciones, se define el operador de interpolación de Lagrange $I_L : H^{1+r}(\Omega) \rightarrow V_h$, de manera que el valor en los nodos de $I_L v$ coincide con la función v en los mismos.

Las estimaciones del error para este operador que están dadas en el Lema A.29 implican estabilidad en H^{1+r} . Sin embargo, el interpolador de Lagrange no es estable en H^1 , dado que en este espacio existen funciones no acotadas.

Para funciones con regularidad H^1 , tanto en problemas elípticos como en parabólicos, es usual utilizar el operador $I_C : H_0^1(\Omega) \rightarrow V_h$ llamado interpolador de Clément (estable en H^1), cuya construcción se detalla en [6] y [5]. Es bien conocido (ver [6]), que las estimaciones locales del error de interpolación usando I_C están dadas por:

$$\begin{aligned} \|v - I_C v\|_T &\preceq h_T |v|_{1, \hat{w}_T}, \quad \forall v \in H_0^1(\Omega), \\ \|v - I_C v\|_E &\preceq h_E^{1/2} |v|_{1, \hat{w}_E}, \quad \forall v \in H_0^1(\Omega), \end{aligned} \tag{4.2}$$

donde \hat{w}_T y \hat{w}_E denotan la unión de triángulos que comparten un nodo con T o E respectivamente. Las constantes intervinientes dependen del ángulo mínimo.

El siguiente lema mostrará que en las estimaciones del error de interpolación con el operador de Clément (ahora en términos de la norma energía de v), aparece un factor que depende de las cotas de k . Más precisamente dependerá del cociente entre los valores extremos de k .

Lema 4.1. Sean $\delta \leq k(x) \leq \delta^{-1}$ para algún $0 < \delta < 1$ y $v(t) \in H_0^1(\Omega)$, $t > 0$ fijo. Para cualquier $T \in \mathcal{T}_h^n$ y $E \in \mathcal{E}_h^n$ valen las siguientes cotas:

$$k_T^{1/2} \|v(t) - I_C v(t)\|_T \leq h_T \delta^{-1} \|v(t)\|_{a, \hat{w}_T}, \quad (4.3)$$

$$k_E^{1/2} \|v(t) - I_C v(t)\|_E \leq h_E^{1/2} \delta^{-1} \|v(t)\|_{a, \hat{w}_E} \quad (4.4)$$

donde $\|\cdot\|_a$ es la norma energía definida en (2.9).

Demostración. Sea $T \in \mathcal{T}_h^n$. Llamemos k_{min} y k_{max} a los valores mínimo y máximo de k sobre \hat{w}_T . Entonces:

$$\begin{aligned} |v(t)|_{1, \hat{w}_T}^2 &= \sum_{T' \subset \hat{w}_T} |v(t)|_{1, T'}^2 \leq \sum_{T' \subset \hat{w}_T} \frac{k_{T'}}{k_{min}} |v(t)|_{1, T'}^2 \\ &= \sum_{T' \subset \hat{w}_T} \frac{1}{k_{min}} \|v(t)\|_{a, T'}^2 \leq k_T^{-1} \frac{k_{max}}{k_{min}} \|v(t)\|_{a, \hat{w}_T}^2. \end{aligned}$$

Si a esta desigualdad obtenida la utilizamos en la estimación local del error con el interpolador de Clément (4.2) y tenemos en cuenta que podemos acotar $\frac{k_{max}}{k_{min}} \leq \delta^{-2}$, llegamos a que

$$\|v(t) - I_C v(t)\|_T \leq h_T |v(t)|_{1, \hat{w}_T} \leq h_T k_T^{-1/2} \delta^{-1} \|v(t)\|_{a, \hat{w}_T},$$

con lo que probamos la desigualdad (4.3). De manera similar se puede probar (4.4). \square

Observamos, de la demostración que acabamos de hacer, que con el operador I_C se puede obtener una cota no deseable cuando el cociente k_{max}/k_{min} tiene un valor muy grande.

Para salvar esta dificultad nos hace falta definir un operador de interpolación apropiado $I_h : H_0^1(\Omega) \rightarrow V_h$ cuyas estimaciones sean de la forma:

$$\begin{aligned} k_T^{1/2} \|v(t) - I_h v(t)\|_T &\leq h_T \|v(t)\|_{a, \tilde{w}_T}, \\ k_E^{1/2} \|v(t) - I_h v(t)\|_E &\leq h_E^{1/2} \|v(t)\|_{a, \tilde{w}_{T_E}}, \end{aligned}$$

donde $\tilde{w}_T, \tilde{w}_{T_E}$ denoten “ciertos” conjuntos de triángulos vecinos a T y a E respectivamente. La característica fundamental de estas estimaciones es que en las cotas no figuran factores adicionales que dependan del parámetro k . Cuando un interpolador satisface este tipo de cotas se llama *robusto*. Del lema anterior se desprende que, para el problema específico que estudiamos, el interpolador de Clément no es apropiado, puesto que no es robusto.

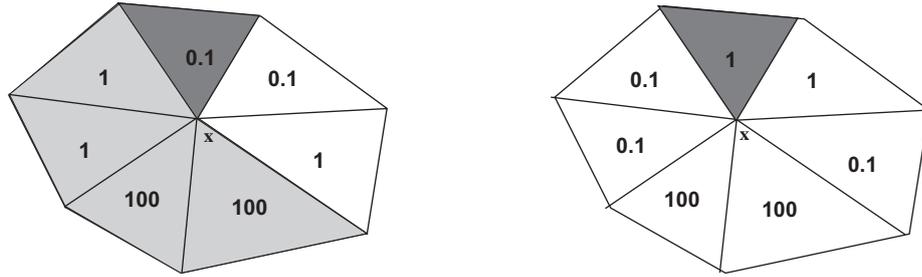


Figura 4.1: La distribución de coeficientes $k_T, T \in w_x$ es casi monótona con respecto al nodo interior x en la gráfica de la izquierda y no lo es en la de la derecha. El triángulo T está coloreado gris oscuro y el conjunto $\tilde{w}_{T,x}$ gris claro en la gráfica de la izquierda.

4.3. La condición de casi monotonía

Acabamos de ver que con el operador de Clément, no se pueden esperar resultados de interpolación robustos para valores arbitrarios de k . Los operadores de interpolación robustos se derivan bajo ciertas restricciones en el coeficiente de difusión.

Siguiendo los trabajos [9], [8], introduciremos el *operador de casi interpolación* $I_h : H_0^1(\Omega) \rightarrow V_h^n$. La definición de este operador de interpolación, requiere de la hipótesis de casi monotonía del parámetro k con respecto a cualquier nodo de la triangulación \mathcal{T}_h^n . Cumpliéndose esta condición sobre k , será posible obtener estimaciones de interpolación robustas.

Definición 4.2 (Casi monotonía de $k(x)$). Sea un nodo $x \in \mathcal{N}_h^n(\overline{\Omega})$. Denotemos con C_x a la unión de triángulos $T \subset w_x$ que alcanzan el máximo valor k_T en w_x . La distribución de coeficientes k_T es casi monótona con respecto al nodo x , si para cada triángulo $T \subset w_x$, existe un conjunto Lipschitz $\tilde{w}_{T,x} \subset w_x$ tal que

- si $x \in \mathcal{N}_h^n(\Omega)$ entonces $T \cup C_x \subseteq \tilde{w}_{T,x}$ y $k_T \leq k_{T'}, \forall T' \subset \tilde{w}_{T,x}$
- si $x \in \mathcal{N}_h^n(\partial\Omega)$ entonces $T \subseteq \tilde{w}_{T,x}$, $|\partial\tilde{w}_{T,x} \cap \partial\Omega| > 0$ y $k_T \leq k_{T'}, \forall T' \subset \tilde{w}_{T,x}$

Diremos que la distribución de coeficientes k_T es casi monótona si es casi monótona con respecto a todos los nodos de $\mathcal{N}_h^n(\overline{\Omega})$.

Observemos que la unión de dos triángulos de \mathcal{T}_h^n que tienen un nodo en común es un dominio Lipschitz, si y sólo si tienen un lado en común (ver Definición A.15).

Para mayor comprensión, ilustramos la condición de casi monotonía en un nodo interior (Figura 4.1) y en un nodo del borde (Figura 4.2), para el caso en que Ω se divida en más de dos subdominios. Seleccionamos un triángulo T pintado en gris oscuro y en color gris claro un ejemplo de conjunto $\tilde{w}_{T,x}$, sólo en los casos donde se cumple la casi monotonía.

La Definición 4.2 es una adaptación de la que se encuentra en [8, Sec 3.5.2], para el problema particular con condición de borde sólo de tipo Dirichlet. Notamos que en nuestra aplicación la casi monotonía es siempre verdadera.

Una manera más intuitiva de formular la definición de casi monotonía con respecto a un nodo, es requerir que:

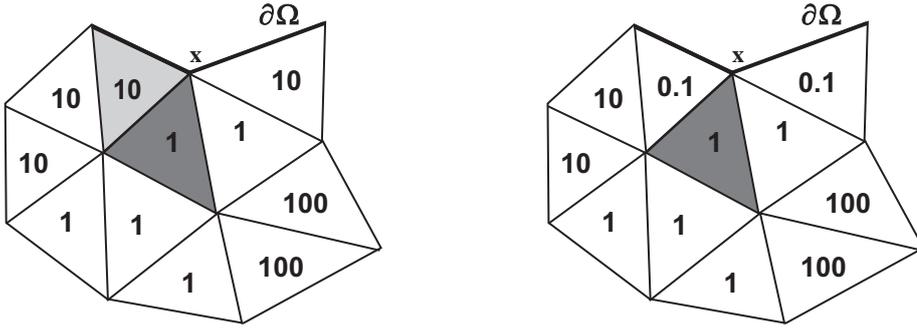


Figura 4.2: La distribución de coeficientes $k_T, T \in w_x$ es casi monótona con respecto al nodo $x \in \mathcal{N}_h^n(\bar{\Omega})$ en la gráfica de la izquierda pero no lo es en la de la derecha. El triángulo T está coloreado gris oscuro y el conjunto $\tilde{w}_{T,x}$ gris claro en la gráfica de la izquierda.

- para un nodo interior, la traza de k sobre un entorno B de x sólo tenga un máximo local
- para un nodo del borde, cada máximo local sea adyacente a $\partial\Omega$.

En este contexto, decimos que un máximo local es alcanzado en $B \cap \Omega_i$ si $k_i > k_j$ para todo subdominio adyacente Ω_j con nodo $x \in \Omega_j$.

También de la Definición 4.2 se desprende que la casi monotonía es independiente de la triangulación y se preserva durante refinamientos sucesivos.

4.4. Un operador de interpolación robusto

En esta sección definiremos un operador de interpolación, que teniendo en cuenta la condición de casi monotonía, logrará que el cociente entre los valores máximo y mínimo de k no se presente en las cotas del error. Este nuevo operador de interpolación (robusto) I_h será necesario para demostrar que los estimadores a posteriori resultan una cota superior robusta del error (Teorema 4.8).

La definición que sigue del nuevo operador es una adaptación de la que se encuentra en [9] y [8].

Definición 4.3 (Operador de casi interpolación). *Sea la distribución de coeficientes $k_T, T \in \mathcal{T}_h^n$ casi monótona. Definimos al operador de casi interpolación $I_h : H_0^1(\Omega) \rightarrow V_h^n$ como*

$$I_h v(x) = \sum_{i=1}^{\dim V_h^n} \phi_i(x) p_{x_i} \text{ con } p_{x_i} = \frac{1}{|C_{x_i}|} \int_{C_{x_i}} v dx, \quad \forall x_i \in \mathcal{N}_h^n(\Omega), \quad (4.5)$$

donde $|C_{x_i}|$ es el área de C_{x_i} , $\phi_i(x)$ son las funciones de base nodal del espacio V_h^n y C_{x_i} como en la Definición 4.2.

Por conveniencia, definimos $p_x = 0$ para los nodos $x \in \partial\Omega$, a fin de que, en la definición de I_h , podamos tomar la sumatoria sobre todos los nodos de $\mathcal{N}_h^n(\bar{\Omega})$.

Los resultados obtenidos para estimar el error de interpolación con el nuevo operador, son presentados en el Lema 4.5. La demostración que exponemos de dicho lema es como la realizada por M. Petzoldt en [8], pero adaptada al caso de dos dimensiones y con condiciones de borde sólo Dirichlet. Antes necesitaremos enunciar el siguiente lema, cuyo resultado no es otra cosa que la desigualdad de trazas escalada.

Lema 4.4. *Sea T un elemento de diámetro h_T perteneciente a una triangulación \mathcal{T}_h^n y sea E cualquier lado de ∂T . Entonces para $v \in H^1(T)$:*

$$\|v\|_E^2 \preceq h_T^{-1} \|v\|_T^2 + h_T \|\nabla v\|_T^2,$$

donde la constante en la cota depende del ángulo mínimo de T .

Demostración. Consideremos el mapeo afín y biyectivo $F_T : T_R \longrightarrow T/F_T(\hat{z}) = z_0 + A_T \hat{z}$, definido en la Sección A.4.2.

Si restringimos F_T sobre cualquier $E \subset \partial T$ de longitud h_E , tenemos:

$$\|v\|_E^2 = h_E \|\hat{v}\|_{\hat{E}}^2 \leq h_T \|\hat{v}\|_{\hat{E}}^2, \quad (4.6)$$

donde $\hat{v} = v \circ F_T \in H^1(T_R)$ y $\hat{E} = F_T^{-1}(E)$.

Por otro lado, por el Teorema de trazas [6, Cap 2, Sec 3.1]:

$$\|\hat{v}\|_{\hat{E}}^2 \preceq \|\hat{v}\|_{T_R}^2 + \|\widehat{\nabla} \hat{v}\|_{T_R}^2, \quad (4.7)$$

donde la constante interviniente depende de T_R .

Entonces, de (4.6), (4.7) y utilizando propiedades del mapeo afín (ver Sección A.4.2):

$$\begin{aligned} \|v\|_E^2 &\preceq h_T \|\hat{v}\|_{T_R}^2 + h_T \|\widehat{\nabla} \hat{v}\|_{T_R}^2 \\ &\preceq h_T |\det A_T^{-1}| \|v\|_T^2 + h_T |\det A_T^{-1}| \int_T \|A_T\|^2 |\nabla u|^2 dz \\ &\preceq h_T h_T^{-2} \|v\|_T^2 + h_T h_T^{-2} h_T^2 \|\nabla u\|_T^2 \\ &\preceq h_T^{-1} \|v\|_T^2 + h_T \|\nabla u\|_T^2. \end{aligned}$$

□

Para una distribución de coeficientes $k_T, T \in \mathcal{T}_h^n$ casi monótona, utilizaremos el conjunto $\tilde{w}_{T,x}$ mencionado en la Definición 4.2 para demostrar estimaciones robustas.

Observemos que si definimos para cada triángulo $T \in \mathcal{T}_h^n$ el siguiente conjunto:

$$\tilde{w}_T = \bigcup_{x \in \mathcal{N}_h^n(T)} \tilde{w}_{T,x},$$

entonces la condición de casi monotonía, implica que

$$k_T \leq k_{T'}, \quad \forall T' \subset \tilde{w}_T. \quad (4.8)$$

Lema 4.5 (Estimaciones del error de interpolación). *Consideremos cualquier $T \in \mathcal{T}_h^n$ y $E \in \mathcal{E}_h^n$. Supongamos se cumple la condición de casi monotonía con respecto a cualquier nodo de T y sea I_h el operador de casi interpolación definido en (4.5). Entonces las siguientes estimaciones valen $\forall v \in H_0^1(\Omega)$:*

$$\|I_h v\|_T \preceq \|v\|_{\tilde{w}_T}, \quad (4.9)$$

$$\|v - I_h v\|_T \preceq h_T^n |v|_{1, \tilde{w}_T} \preceq \frac{h_T^n}{\sqrt{k_T}} \|v\|_{a, \tilde{w}_T}, \quad (4.10)$$

$$\|v - I_h v\|_E \preceq \frac{\sqrt{h_E^n}}{\sqrt{k_{TE}}} \|v\|_{a, \tilde{w}_{TE}}, \quad (4.11)$$

donde las constantes que intervienen sólo dependen del ángulo mínimo en la triangulación.

Demostración. Sea $v \in H_0^1(\Omega)$. Elijamos un triángulo $T \in \mathcal{T}_h^n$ y consideremos sus tres nodos $x_i^T, i = 1, 2, 3$, que por simplicidad llamaremos x_i .

Notemos que por su definición en (4.5), la constante p_{x_i} puede interpretarse como una función de v , es decir $P_{x_i}(v)$, donde $P_{x_i} : L^2(C_{x_i}) \rightarrow \mathbb{R}$ es la proyección L^2 -ortogonal sobre las funciones constantes. De acuerdo a propiedades de esta proyección, podemos decir que para cualquier nodo $x_i \in \mathcal{N}_h^n(T \setminus \partial\Omega)$ y para cualquier $c \in \mathbb{R}$:

$$\begin{aligned} \|p_{x_i} - c\|_T^2 &\preceq \|p_{x_i} - c\|_{C_{x_i}}^2 = \|P_{x_i}(v) - P_{x_i}(c)\|_{C_{x_i}}^2 \\ &= \|P_{x_i}(v - c)\|_{C_{x_i}}^2 \leq \|v - c\|_{C_{x_i}}^2. \end{aligned} \quad (4.12)$$

Sean $\lambda_i(x), i = 1, 2, 3$ las funciones de la base nodal de $P_1(T)$. Dado que $\sum_{i=1}^3 \lambda_i = 1$ y $0 \leq \lambda_i \leq 1$ [7, Cap. 4], escribiremos:

$$v = \sum_{i=1}^3 \lambda_i v \quad \text{y por otro lado} \quad I_h v \Big|_T = \sum_{i=1}^3 \lambda_i p_{x_i}. \quad (4.13)$$

Entonces, de (4.12) con $c = 0$:

$$\begin{aligned} \|I_h v\|_T^2 &= \int_T \left(\sum_{i=1}^3 \lambda_i p_{x_i} \right)^2 \leq \int_T \left(\sum_{i=1}^3 \lambda_i^2 \right) \left(\sum_{i=1}^3 p_{x_i}^2 \right) \leq 3 \sum_{i=1}^3 \|p_{x_i}\|_T^2 \\ &\preceq \sum_{i=1}^3 \|v\|_{C_{x_i}}^2 \preceq \|v\|_{\tilde{w}_T}^2 \end{aligned}$$

y de esta manera probamos (4.9).

Para demostrar (4.10), comenzamos observando que a partir de (4.13) obtenemos:

$$\|v - I_h v\|_T^2 = \left\| \sum_{i=1}^3 \lambda_i v - \sum_{i=1}^3 \lambda_i p_{x_i} \right\|_T^2 \preceq \sum_{i=1}^3 \|\lambda_i (v - p_{x_i})\|_T^2 \leq \sum_{i=1}^3 \|v - p_{x_i}\|_T^2. \quad (4.14)$$

Si aplicamos la desigualdad (4.12) a cualquier nodo $x_i \in \mathcal{N}_h^n(T \setminus \partial\Omega)$, es cierto que para todo $c \in \mathbb{R}$:

$$\|v - p_{x_i}\|_T^2 = \|v - c - p_{x_i} + c\|_T^2 \leq 2\|v - c\|_T^2 + 2\|p_{x_i} - c\|_T^2 \preceq \|v - c\|_T^2 + \|v - c\|_{C_{x_i}}^2.$$

Por otro lado, recordemos que de la definición 4.2, $T \cup C_{x_i} \subset \tilde{\omega}_{T,x_i} \subset \tilde{\omega}_T$. Entonces, si sumamos sobre los nodos x_i de T que no están sobre el borde $\partial\Omega$ y aplicamos la desigualdad de Poincaré sobre el conjunto Lipschitz $\tilde{\omega}_T$, para cualquier constante c vale:

$$\begin{aligned} \sum_{x_i \in \mathcal{N}_h^n(T \setminus \partial\Omega)} \|v - p_{x_i}\|_T^2 &\leq \|v - c\|_T^2 + \sum_{x_i \in \mathcal{N}_h^n(T \setminus \partial\Omega)} \|v - c\|_{C_{x_i}}^2 \\ &\leq \|v - c\|_{\tilde{\omega}_T}^2 \leq (h_T^n)^2 |v - c|_{1, \tilde{\omega}_T}^2 = (h_T^n)^2 |v|_{1, \tilde{\omega}_T}^2. \end{aligned} \quad (4.15)$$

Para los nodos de T que también están sobre $\partial\Omega$ usaremos el hecho de que $p_{x_i} = 0$ y que v se anula sobre $\partial\tilde{\omega}_{T,x_i} \cap \partial\Omega$. Nuevamente, por la desigualdad de Poincaré:

$$\|v - p_{x_i}\|_T^2 = \|v\|_T^2 \leq \|v\|_{\tilde{\omega}_{T,x_i}}^2 \leq (h_T^n)^2 |v|_{1, \tilde{\omega}_{T,x_i}}^2. \quad (4.16)$$

Si retomamos (4.14) y aplicamos lo obtenido en (4.15) y (4.16) se tiene:

$$\begin{aligned} \|v - I_h v\|_T^2 &\leq \sum_{x_i \in \mathcal{N}_h^n(T \setminus \partial\Omega)} \|v - p_{x_i}\|_T^2 + \sum_{x_i \in \mathcal{N}_h^n(T \cap \partial\Omega)} \|v - p_{x_i}\|_T^2 \\ &\leq h_T^2 \left(|v|_{1, \tilde{\omega}_T}^2 + \sum_{x_i \in \mathcal{N}_h^n(T \cap \partial\Omega)} |v|_{1, \tilde{\omega}_{T,x_i}}^2 \right) \leq (h_T^n)^2 |v|_{1, \tilde{\omega}_T}^2. \end{aligned} \quad (4.17)$$

Así llegamos a la primera desigualdad de (4.10).

Para acotar $|v|_{1, \tilde{\omega}_T}^2$ usamos (4.8):

$$|v|_{1, \tilde{\omega}_T}^2 = \sum_{T' \subset \tilde{\omega}_T} \frac{1}{k_{T'}} \|\sqrt{k_{T'}} \nabla v\|_{T'}^2 = \sum_{T' \subset \tilde{\omega}_T} \frac{1}{k_{T'}} \|v\|_{a, T'}^2 \leq \frac{1}{k_T} \|v\|_{a, \tilde{\omega}_T}^2, \quad (4.18)$$

y obtenemos la segunda desigualdad de (4.10).

Por último, para probar la estimación que nos falta (4.11), nos será útil primero mostrar que

$$|v - I_h v|_{1, T} \leq |v|_{1, \tilde{\omega}_T}. \quad (4.19)$$

Para ello, comenzamos usando las dos igualdades de (4.13):

$$|v - I_h v|_{1, T}^2 = \left| \sum_{i=1}^3 \lambda_i (v - p_{x_i}) \right|_{1, T}^2 \leq \sum_{i=1}^3 |\lambda_i (v - p_{x_i})|_{1, T}^2. \quad (4.20)$$

Mediante la transformación afín definida en la Sección A.4.2, se puede ver que $\|\lambda_i\|_{L^\infty(T)}^2 = 1$ y $\|\nabla \lambda_i\|_{L^\infty(T)}^2 \leq (h_T^n)^{-2}$. Estas propiedades de λ_i implican que:

$$\begin{aligned} |\lambda_i (v - p_{x_i})|_{1, T}^2 &= \|\nabla [\lambda_i (v - p_{x_i})]\|_T^2 = \int_T |(\nabla \lambda_i)(v - p_{x_i}) + \lambda_i \nabla (v - p_{x_i})|^2 \\ &\leq \int_T |(\nabla \lambda_i)(v - p_{x_i})|^2 + \int_T |\lambda_i \nabla (v - p_{x_i})|^2 \\ &\leq \|\nabla \lambda_i\|_{L^\infty(T)}^2 \|v - p_{x_i}\|_T^2 + \|\lambda_i\|_{L^\infty(T)}^2 \|\nabla v\|_T^2 \\ &\preceq (h_T^n)^{-2} \|v - p_{x_i}\|_T^2 + |v|_{1, T}^2. \end{aligned} \quad (4.21)$$

Si tenemos en cuenta las cotas obtenidas en (4.15) y (4.16) para cualquier nodo de T que se encuentre sobre $\partial\Omega$ o no, podemos afirmar que

$$\sum_{i=1}^3 \|v - p_{x_i}\|_T^2 \preceq (h_T^n)^2 |v|_{1, \tilde{\omega}_T}^2. \quad (4.22)$$

Luego, combinando (4.20), (4.21) y (4.22) llegamos a la desigualdad que queríamos (4.19):

$$\begin{aligned} |v - I_h v|_{1, T}^2 &\preceq \sum_{i=1}^3 |\lambda_i(v - p_{x_i})|_{1, T}^2 \\ &\preceq (h_T^n)^{-2} \sum_{i=1}^3 \|v - p_{x_i}\|_T^2 + |v|_{1, T}^2 \preceq |v|_{1, \tilde{\omega}_T}^2. \end{aligned}$$

Ahora, para cualquier $E \in \mathcal{E}_h^n$, obtenemos del Lema 4.4 y las desigualdades (4.17) y (4.19) que:

$$\begin{aligned} \|v - I_h v\|_E^2 &\preceq h_{T_E}^{-1} \|v - I_h v\|_{T_E}^2 + h_{T_E} |v - I_h v|_{1, T_E}^2 \\ &\preceq h_{T_E}^{-1} h_{T_E}^2 |v|_{1, \tilde{\omega}_{T_E}}^2 + h_{T_E} |v|_{1, \tilde{\omega}_{T_E}}^2 \preceq h_{T_E} |v|_{1, \tilde{\omega}_{T_E}}^2. \end{aligned}$$

Después de multiplicar por k_{T_E} llegamos a que, como $k_{T_E} \leq k_{T'}$, $T' \subset \tilde{\omega}_{T_E}$ por la casi monotonía y además $h_{T_E} \approx h_E$:

$$\begin{aligned} k_{T_E} \|v - I_h v\|_E^2 &\preceq k_{T_E} h_{T_E} |v|_{1, \tilde{\omega}_{T_E}}^2 = h_{T_E} \sum_{T' \subset \tilde{\omega}_{T_E}} k_{T'} |v|_{1, T'}^2 \\ &\preceq h_{T_E} \|v\|_{a, \tilde{\omega}_{T_E}}^2 \preceq h_E \|v\|_{a, \tilde{\omega}_{T_E}}^2, \end{aligned}$$

quedando finalizada la prueba de (4.11). □

4.5. Estimadores a posteriori del error

El objetivo ahora es encontrar *estimaciones a posteriori del error*, que se obtienen a partir de la sucesión de soluciones del problema débil (3.1)–(3.2) planteado con discretización completa.

Dado que de la resolución de (4.1) se conocen las soluciones $U_h^n(x) \in V_h^n$ sólo en un número finito de valores del intervalo de tiempo $[0, T]$, para medir el error de aproximación en el intervalo completo será necesario introducir un interpolante de la sucesión $\{U_h^n\}$, $n = 0, \dots, N$, continuo y lineal a trozos en el tiempo. A dicho interpolante lo denotamos con $U_I(x, t)$ y en cada eslabón $\bar{\Omega} \times [t_{n-1}, t_n]$ lo definimos como:

$$U_I(x, t) = \frac{t - t_{n-1}}{t_n - t_{n-1}} U_h^n(x) + \frac{t_n - t}{t_n - t_{n-1}} U_h^{n-1}(x), \quad n = 1, \dots, N. \quad (4.23)$$

Es importante notar de la definición, que:

$$\frac{\partial U_I}{\partial t} = \frac{U_h^n - U_h^{n-1}}{t_n - t_{n-1}}, \quad \text{en cada } \Omega \times (t_{n-1}, t_n).$$

Definiremos al error de aproximación como $e = u - U_I$, donde u es la solución de (3.1)–(3.2), con $f \equiv 0$.

Siguiendo el trabajo de S. Berrone [9], lo que haremos en esta sección es presentar un estimador a posteriori del error que resulta una cota superior del error. Este estimador es una suma de distintos términos, pero contiene esencialmente dos partes: una debida a la discretización espacial y otra debida a la discretización temporal.

Utilizando con criterio este estimador, desarrollaremos en el próximo capítulo un método adaptativo para la ecuación (3.1)–(3.2).

Los indicadores del error producidos por la discretización espacial involucran a los residuos (interior y de borde) del interpolante U_I , con respecto a la ecuación diferencial de (2.1). Estos residuos se definen de la siguiente manera en cada subintervalo $I_n = (t_{n-1}, t_n)$, $n = 1, \dots, N$:

- $R_T^n(U_I)$ denota el *residuo interior a cada elemento* T de una triangulación de Ω y se define por:

$$R_T^n(U_I) := \frac{\partial U_I}{\partial t} - k_T \Delta U_h^n(x), \quad x \in T.$$

En particular, ya que $U_h^n(x)$ es lineal a trozos sobre la triangulación, resulta $R_T^n(U_I) = \frac{\partial U_I}{\partial t}$.

- Sean $E \in \mathcal{E}_h(\Omega)$, T_1, T_2 los dos triángulos que forman a w_E y n_E el vector unitario normal a E que apunta de T_1 a T_2 . Denotamos con $J_E^n(U_I)$ al *residuo en el borde interelemento* E , que es el salto de flujo de U_h^n a través de la normal n_E , definido por:

$$J_E^n(U_I) := \left[k_T \frac{\partial U_h^n}{\partial n_E} \right]_E = k_{T_1} \frac{\partial U_h^n}{\partial n_E} - k_{T_2} \frac{\partial U_h^n}{\partial n_E}.$$

El residuo interior servirá para indicar cuán lejos está la aproximación $U_I(x, t)$ de satisfacer la ecuación diferencial y el residuo en el borde para medir el error debido a la discontinuidad de la derivada normal de U_h^n sobre el borde de cada elemento.

Habiendo ya definido los dos residuos $R_T^n(U_I)$ y $J_E^n(U_I)$, presentamos a continuación los estimadores a posteriori locales y globales, que indicarán el error producido por las discretizaciones de las variables espacio y tiempo.

Definición 4.6 (Estimadores del error a posteriori). *Definimos los siguientes estimadores locales en el espacio y locales en el tiempo:*

$$\begin{aligned} (\eta_{R,T}^n)^2 &= \Delta t^n \left((h_T^n)^2 \left\| \frac{1}{\sqrt{k_T}} R_T^n(U_I) \right\|_T^2 + \frac{1}{2} \sum_{E \in E_h^n(T) \cap E_h^n(\Omega)} h_E^n \left\| \frac{1}{\sqrt{k_{T_E}}} J_E^n(U_I) \right\|_E^2 \right), \\ (\eta_{\nabla,T}^n)^2 &= \Delta t^n \left\| \sqrt{k_T} \nabla (U_h^n - U_h^{n-1}) \right\|_T^2. \end{aligned}$$

La suma de las contribuciones de los indicadores del error sobre cada T definen los siguientes estimadores globales en el espacio y locales en el tiempo:

$$(\eta_R^n)^2 = \sum_{T \in \mathcal{T}_h^n} (\eta_{R,T}^n)^2, \quad (\eta_{\nabla}^n)^2 = \sum_{T \in \mathcal{T}_h^n} (\eta_{\nabla,T}^n)^2.$$

Finalmente, definimos los estimadores globales en el espacio y globales en el tiempo por:

$$\eta_R^2 = \sum_{n=1}^N (\eta_R^n)^2, \quad \eta_{\nabla}^2 = \sum_{n=1}^N (\eta_{\nabla}^n)^2.$$

El estimador η_R^n es el indicador del error debido a la discretización espacial y entonces está relacionado con la calidad de la triangulación \mathcal{T}_h^n , mientras que η_{∇}^n da información del error debido a la discretización temporal.

Para el caso general donde $f \neq 0$, en [9] se define además el estimador global η_f^2 que servirá para indicar el error debido a la aproximación en el dato f .

Observación 4.7. *Supongamos que adicionamos al problema (2.1) una condición de borde tipo Neumann de la forma*

$$k(x) \frac{\partial u}{\partial n} = g \quad \text{en } \Gamma_N \times (0, T),$$

con n el vector normal unitario exterior a Γ_N y $g \in L^2(0, T, L^2(\Gamma_N))$.

Aquí estamos considerando que $\partial\Omega$ se puede descomponer como unión de partes correspondientes a la condición de borde tipo Dirichlet y Neumann, es decir, $\partial\Omega = \Gamma_D \cup \Gamma_N$, $\Gamma_D \cap \Gamma_N = \emptyset$ y Γ_D tiene medida positiva.

En [8] pueden verse las definiciones de casi monotonía de k y del operador de interpolación para este caso, de condiciones mixtas en el borde.

En particular, el estimador global en el espacio y local en el tiempo se define por:

$$(\eta_R^n)^2 = \Delta t^n \left(\sum_{T \in \mathcal{T}_h^n} \Delta t^n (h_T^n)^2 \left\| \frac{1}{\sqrt{k_T}} R_T^n(U_I) \right\|_T^2 + \sum_{E \in E_h^n(\bar{\Omega}) \setminus E_h^n(\Gamma_D)} h_E^n \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n(U_I) \right\|_E^2 \right),$$

donde

$$J_E^n(U_I) = \begin{cases} [k_T \frac{\partial U_h^n}{\partial n_E}]_E & \text{si } E \in \Omega \\ g|_E - k_T \frac{\partial U_h^n}{\partial n_E} & \text{si } E \in \Gamma_N \end{cases}$$

4.6. Cota superior global

A partir de la aproximación por elementos finitos $U_I(x, t)$ definida en (4.23), vamos a presentar dos estimaciones a posteriori del error que se determinan a través de dos cotas superiores: una cota superior *local* (Teorema 4.8) y una cota superior *global* del error (Corolario 4.9).

Estas dos acotaciones contienen a los estimadores dados en la Definición 4.6 y son muy importantes porque dan información sobre el tamaño y la distribución del error en cada nivel $\Omega \times [t_{n-1}, t_n]$ (en el caso de la cota local) y en $\Omega \times [0, T]$ (en el caso de la cota global).

Teorema 4.8 (Cota superior local del error). *Sean u la solución de (3.1)–(3.2) para $f \equiv 0$ y $\{U_h^n\}_{n=0}^N$ la sucesión de soluciones de la formulación completamente discreta (4.1) y supongamos que la distribución de coeficientes k es casi monótona. Entonces, para cada $n = 1, \dots, N$, existe una constante \tilde{C}_n que sólo depende del ángulo mínimo de la triangulación regular \mathcal{T}_h^n , tal que*

$$\|u^n - U_h^n\|^2 + \int_{t_{n-1}}^{t_n} \|\sqrt{k} \nabla(u - U_I)\|^2 dt \leq \|u^{n-1} - U_h^{n-1}\|^2 + \tilde{C}_n ((\eta_R^n)^2 + (\eta_\nabla^n)^2), \quad (4.24)$$

siendo $u^n = u(t_n)$.

Demostración. Observamos que como $e = u - U_I$, para cualquier $v \in H_0^1(\Omega)$:

$$\left\langle \frac{\partial e}{\partial t}, v \right\rangle + a(e, v) = -\left\langle \frac{\partial U_I}{\partial t}, v \right\rangle - a(U_I, v), \quad t \in I_n.$$

En particular cuando $v = e$, si usamos (2.21) y la expresión que resulta en (4.1) después de reemplazar v por $I_h v \in V_h^n$, siendo I_h el operador de interpolación que definimos en (4.5), se tiene para $t \in (t_{n-1}, t_n)$:

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial t} \|e\|^2 + \|e\|_a^2 &= -\left\langle \frac{\partial U_I}{\partial t}, e \right\rangle - a(U_I, e) \\ &= \left\langle \frac{\partial U_I}{\partial t}, I_h e - e \right\rangle + a(U_h^n, I_h e - e) + a(U_h^n - U_I, e). \end{aligned} \quad (4.25)$$

Llamemos I_1 , I_2 e I_3 respectivamente, a cada término resultante en la última igualdad. Procederemos a acotar por separado cada integral de I_i , $i = 1, 2, 3$.

Para I_1 , aplicamos la desigualdad (4.10) del Lema 4.5:

$$I_1 = \sum_{T \in \mathcal{T}_h^n} \int_T \frac{\partial U_I}{\partial t} (I_h e - e) \, dx \preceq \sum_{T \in \mathcal{T}_h^n} \frac{h_T^n}{\sqrt{k_T}} \|R_T^n(U_I)\|_T \|e\|_{a, \tilde{w}_T}.$$

El término I_2 se puede acotar usando integración por partes, la desigualdad (4.11) del Lema 4.5 y el hecho de que ΔU_h^n se anula sobre cada $T \in \mathcal{T}_h^n$:

$$\begin{aligned} I_2 &= \sum_{T \in \mathcal{T}_h^n} \int_T k_T \nabla U_h^n \cdot \nabla (I_h e - e) \, dx \\ &= - \sum_{T \in \mathcal{T}_h^n} \int_T (I_h e - e) k_T \Delta U_h^n \, dx + \sum_{E \in \mathcal{E}_h(\Omega)} \int_E \left[k_T \frac{\partial U_h^n}{\partial n_E} \right]_E (I_h e - e) \, ds \\ &\preceq \sum_{E \in \mathcal{E}_h(\Omega)} \frac{\sqrt{h_E^n}}{\sqrt{k_{T_E}}} \|J_E^n\|_E \|e\|_{a, \tilde{w}_{T_E}}. \end{aligned}$$

Para I_3 , tendremos en cuenta que

$$U_h^n - U_I = \frac{t_n - t}{t_n - t_{n-1}} (U_h^n - U_h^{n-1}),$$

entonces, aplicando la desigualdad de Cauchy- Schwartz:

$$I_3 = \int_{\Omega} k(x) \nabla(U_h^n - U_I) \cdot \nabla e \, dx \leq \frac{t_n - t}{\Delta t^n} \|\sqrt{k} \nabla(U_h^n - U_h^{n-1})\| \|\sqrt{k} \nabla e\|.$$

Con las estimaciones obtenidas de I_1 , I_2 e I_3 , cuando integramos (4.25) sobre (t_{n-1}, t_n) , llegamos a que:

$$\begin{aligned} \frac{1}{2} \|e^n\|^2 - \frac{1}{2} \|e^{n-1}\|^2 + \int_{t_{n-1}}^{t_n} \|e\|_a^2 \, dt &\leq \int_{t_{n-1}}^{t_n} \sum_{T \in \mathcal{T}_h^n} h_T^n \left\| \frac{1}{\sqrt{k_T}} R_T^n \right\|_T \|e\|_{a, \tilde{w}_T} \, dt \\ &\quad + \int_{t_{n-1}}^{t_n} \sum_{E \in \mathcal{E}_h(\Omega)} \sqrt{h_E^n} \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n \right\|_E \|e\|_{a, \tilde{w}_{TE}} \, dt \\ &\quad + \int_{t_{n-1}}^{t_n} \frac{t_n - t}{\Delta t^n} \|\sqrt{k} \nabla(U_h^n - U_h^{n-1})\| \|\sqrt{k} \nabla e\| \, dt, \end{aligned}$$

donde $e^n = u^n - U_h^n$.

La constante interviniente en esta desigualdad está en función de las constantes del Lema 4.5, que sólo dependen del ángulo mínimo de la triangulación.

Si ahora aplicamos la desigualdad de Young (2.22) en los integrandos de cada término del lado derecho y sabiendo que

$$\int_{t_{n-1}}^{t_n} \left(\frac{t_n - t}{\Delta t^n} \right)^2 \, dt = \frac{\Delta t^n}{3},$$

se tiene:

$$\begin{aligned} \|e^n\|^2 - \|e^{n-1}\|^2 + 2 \int_{t_{n-1}}^{t_n} \|e\|_a^2 \, dt &\leq \int_{t_{n-1}}^{t_n} \left(\sum_{T \in \mathcal{T}_h^n} \frac{(h_T^n)^2}{4\epsilon} \left\| \frac{1}{\sqrt{k_T}} R_T^n \right\|_T^2 + \sum_{T \in \mathcal{T}_h^n} \|e\|_{a, \tilde{w}_T}^2 \epsilon \right) \, dt \\ &\quad + \int_{t_{n-1}}^{t_n} \left(\sum_{E \in \mathcal{E}_h(\Omega)} \frac{h_E^n}{4\epsilon} \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n \right\|_E^2 + \sum_{E \in \mathcal{E}_h(\Omega)} \|e\|_{a, \tilde{w}_{TE}}^2 \epsilon \right) \, dt \\ &\quad + \int_{t_{n-1}}^{t_n} \left[\left(\frac{t_n - t}{\Delta t^n} \right)^2 \frac{\|\sqrt{k} \nabla(U_h^n - U_h^{n-1})\|^2}{4\epsilon} + \|\sqrt{k} \nabla e\|^2 \epsilon \right] \, dt \\ &\leq \frac{1}{\epsilon} \Delta t_n \sum_{T \in \mathcal{T}_h^n} (h_T^n)^2 \left\| \frac{1}{\sqrt{k_T}} R_T^n \right\|_T^2 + \frac{1}{\epsilon} \Delta t_n \sum_{E \in \mathcal{E}_h(\Omega)} h_E^n \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n \right\|_E^2 \\ &\quad + \frac{1}{\epsilon} \frac{\Delta t_n}{3} \|\sqrt{k} \nabla(U_h^n - U_h^{n-1})\|^2 + \epsilon \int_{t_{n-1}}^{t_n} \|e\|_a^2 \, dt \end{aligned}$$

donde las constantes en la cota dependen del ángulo mínimo.

Luego, eligiendo ϵ apropiadamente concluimos que existirá la constante $\tilde{C}_n = \tilde{C}_n(\epsilon)$ que sólo depende del ángulo mínimo de la triangulación, tal que podemos escribir la última desigualdad como:

$$\begin{aligned} \|e^n\|^2 + \int_{t_{n-1}}^{t_n} \|e\|_a^2 dt &\leq \|e^{n-1}\|^2 \\ &+ \tilde{C}_n \left[\Delta t^n \left(\sum_{T \in \mathcal{T}_h^n} (h_T^n)^2 \left\| \frac{1}{\sqrt{k_T}} R_T^n \right\|_T^2 + \sum_{E \in \mathcal{E}_h(\Omega)} h_E^n \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n \right\|_E^2 \right) \right. \\ &\left. + \Delta t^n \sum_{T \in \mathcal{T}_h^n} \|U_h^n - U_h^{n-1}\|_{a,T}^2 \right] \\ &= \|u^{n-1} - U_h^{n-1}\|^2 + \tilde{C}_n \left((\eta_R^n)^2 + (\eta_{\nabla}^n)^2 \right). \end{aligned}$$

□

La cota superior que acabamos de demostrar, es local en el tiempo pero global en el espacio. En base a ella, obtendremos una estimación para el error que resulta global tanto en el espacio como en el tiempo, que es la que presentamos en el siguiente corolario:

Corolario 4.9 (Cota superior global del error). *Bajo las mismas hipótesis del Teorema 4.8, existe una constante \tilde{C} que sólo depende del ángulo mínimo de la familia regular de triangulaciones $\{\mathcal{T}_h^n\}$, tal que se cumple:*

$$\max_{1 \leq n \leq N} \|u^n - U_h^n\|^2 + \int_0^T \|\sqrt{k} \nabla(u - U_I)\|^2 dt \leq 2 \|u^0 - U_h^0\|^2 + \tilde{C}(\eta_R^2 + \eta_{\nabla}^2), \quad (4.26)$$

siendo $\tilde{C} = 2 \max_n \tilde{C}_n$.

Demostración. Si tomamos las N cotas superiores locales (4.24) obtenidas al considerar $n = 1, \dots, N$ y luego sumamos todas ellas, es inmediata la expresión de la cota superior global del error (4.26). □

Gracias a los estimadores a posteriori del error, se hace posible mejorar o controlar el error de aproximación. Este hecho es la esencia de los métodos adaptativos, los que a través de modificaciones en la malla o del tamaño del paso de tiempo, permiten reducir y controlar numéricamente el error.

El capítulo que sigue, se centrará en el desarrollo e implementación de un algoritmo adaptativo, que con el estimador $\eta_R^2 + \eta_{\nabla}^2$ y la estimación obtenida (4.26) calculará una muy buena aproximación $U_I(x, t)$ del problema (3.1)–(3.2) a bajo costo computacional.

4.7. Un ejemplo numérico

En esta sección ilustramos a través de un ejemplo, el comportamiento de los estimadores a posteriori del error obtenidos en el Teorema 4.8 y Corolario 4.9.

Pretendemos comparar los errores exacto y estimado representados por los valores obtenidos de cada lado de la desigualdad (4.24) en cada paso de tiempo y por los valores obtenidos de cada lado de la desigualdad (4.26) en todo el intervalo $[0, T]$. Para ello, consideramos un problema con solución conocida, la cual para simplificar dependerá de una variable espacial.

Consideremos un dominio rectangular Ω dividido en dos subdominios $\Omega_1 = (0, 1) \times (0, 1)$ y $\Omega_2 = (1, \beta) \times (0, 1)$, donde $\beta > 1$ es un parámetro que depende de los valores $k_i > 0$ del coeficiente de difusión en cada Ω_i , $i = 1, 2$, los cuales son conocidos.

Definiendo las constantes:

$$\begin{aligned}\alpha &= \sqrt{\frac{k_1}{k_2}} \frac{\pi}{4}, \\ C &= -\frac{1}{\alpha} \frac{k_1}{k_2} \frac{\pi}{4} \frac{\sqrt{2}}{2}, \\ A &= \frac{\sqrt{2}}{2} \sin \alpha + C \cos \alpha \quad \text{y} \\ B &= \frac{\sqrt{2}}{2} \cos \alpha - C \sin \alpha\end{aligned}$$

se puede probar que la función

$$u(x, t) = \begin{cases} u_1(x, t) := e^{-k_1(\frac{\pi}{4})^2 t} \cos(\frac{\pi}{4} x_1), & x = (x_1, x_2) \in \Omega_1, t > 0, \\ u_2(x, t) := e^{-k_2 \alpha^2 t} (A \sin(\alpha x_1) + B \cos(\alpha x_1)), & x = (x_1, x_2) \in \Omega_2, t > 0, \end{cases}$$

es solución del problema

$$\left\{ \begin{array}{ll} \frac{\partial u}{\partial t} - k_i \Delta u = 0 & \text{en } \Omega_i \times (0, T), \\ u(x, t) = 0 & \text{en } \{x_1 = \beta\} \times (0, T), \\ u(x, 0) = u^0(x) & \text{en } \Omega, \\ u_1(x, t) = u_2(x, t) & \text{en } \{x_1 = 1\} \times (0, T), \\ -k_1 \nabla u_1 \cdot n_1 = k_2 \nabla u_2 \cdot n_2 & \text{en } \{x_1 = 1\} \times (0, T), \\ k_i \frac{\partial u}{\partial n_i}(x, t) = 0 & \text{en } (\{x_1 = 0\} \cup \{x_2 = 0\} \cup \{x_2 = 1\}) \times (0, T), \end{array} \right.$$

siempre que

$$\beta = \frac{1}{\alpha} \arcsin \left(\sqrt{\frac{B^2}{A^2 + B^2}} \right),$$

para cualquier $T > 0$. En las pruebas que presentaremos a continuación tomaremos $T = 3$.

De la definición, se puede notar que β depende de la elección que hagamos de los valores de k_1 y k_2 .

Para verificar la robustez de nuestro estimador, realizamos pruebas considerando un valor fijo para k_1 y diferentes valores para el coeficiente k_2 . Así, tomando $k_1 = 1$ (fijo) en Ω_1 y $k_2 = 4, 25, 49$ y 100 en Ω_2 , y aplicando el Algoritmo 2 de la Sección 5.3, calculamos

los términos que componen ambos lados de cada acotación, con el fin de visualizar y comparar los valores que toma cada miembro que conforma la desigualdad de la cota superior local del error dada por el Teorema 4.8 y la cota superior global del error dada por el Corolario 4.9.

Cuando $k_1 = 1$ y $k_2 = 4$, aproximadamente $\beta = 3,8193$. Para este caso, observamos en la Figura 4.3 una malla inicial del dominio con 182 nodos y una malla final de 350 nodos generadas por el algoritmo adaptativo. También mostramos los gráficos de la solución exacta y el error ($|u - U_h^n|$) obtenidos por el algoritmo en el último instante de tiempo $t = 3,02$.

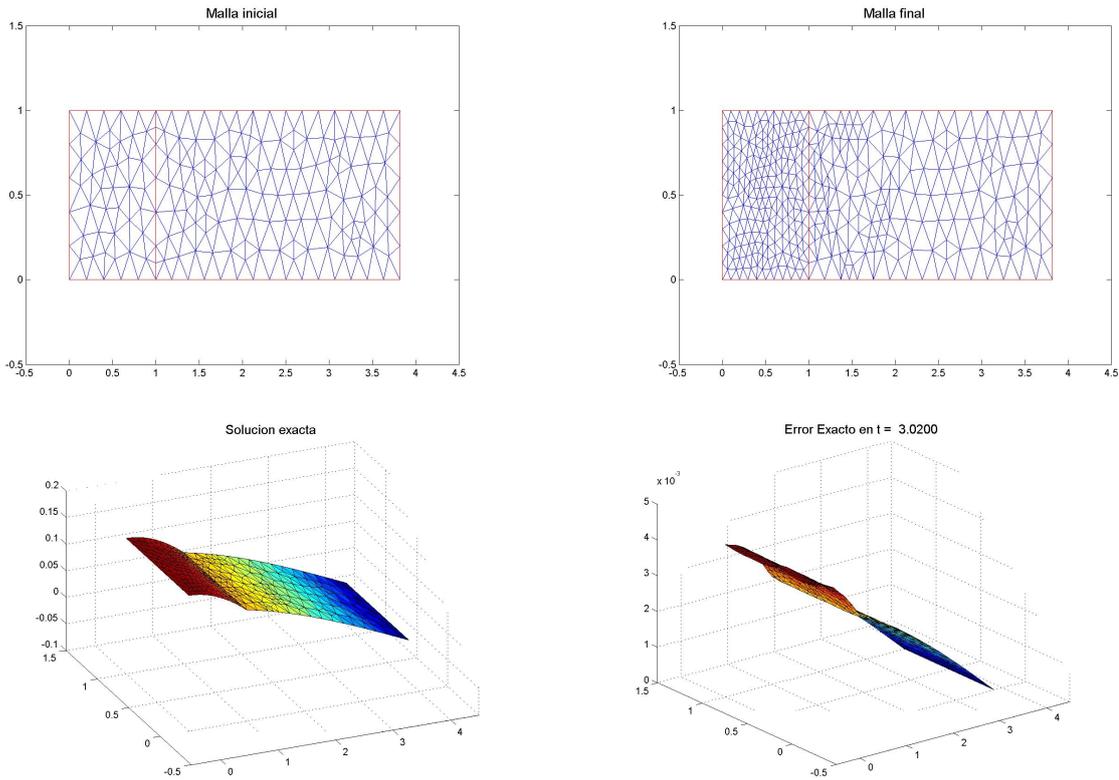


Figura 4.3: Arriba: Malla inicial y final del dominio del problema para el caso en que $k_1 = 1$ y $k_2 = 4$ ($\beta = 3,8193$). Abajo: Solución exacta y Error exacto en $t = 3,02$.

En la Figura 4.4 comparamos, para los distintos valores de k_2 , el lado izquierdo y derecho de las desigualdades (4.24) y (4.26). Las constantes \tilde{C}_n de las cotas locales (4.24) no se pueden determinar exactamente. Para los ejemplos que realizamos, hemos utilizado $\tilde{C}_n = 0,3$ en cada paso n .

En la Figura 4.5 se muestra el comportamiento del cociente de ambos lados de cada desigualdad, en un gráfico para cada valor de k_2 . Puede percibirse que después de una etapa inicial de transición, el cociente entre ambos lados de la desigualdad (4.24) se aproxima a 1, y el cociente entre ambos lados de la acotación global (4.26) se mantiene menor que una constante aproximadamente igual a 3, que no varía fuertemente con el valor de k_2 . Esto indica que el estimador es robusto con respecto al cociente k_2/k_1 , que

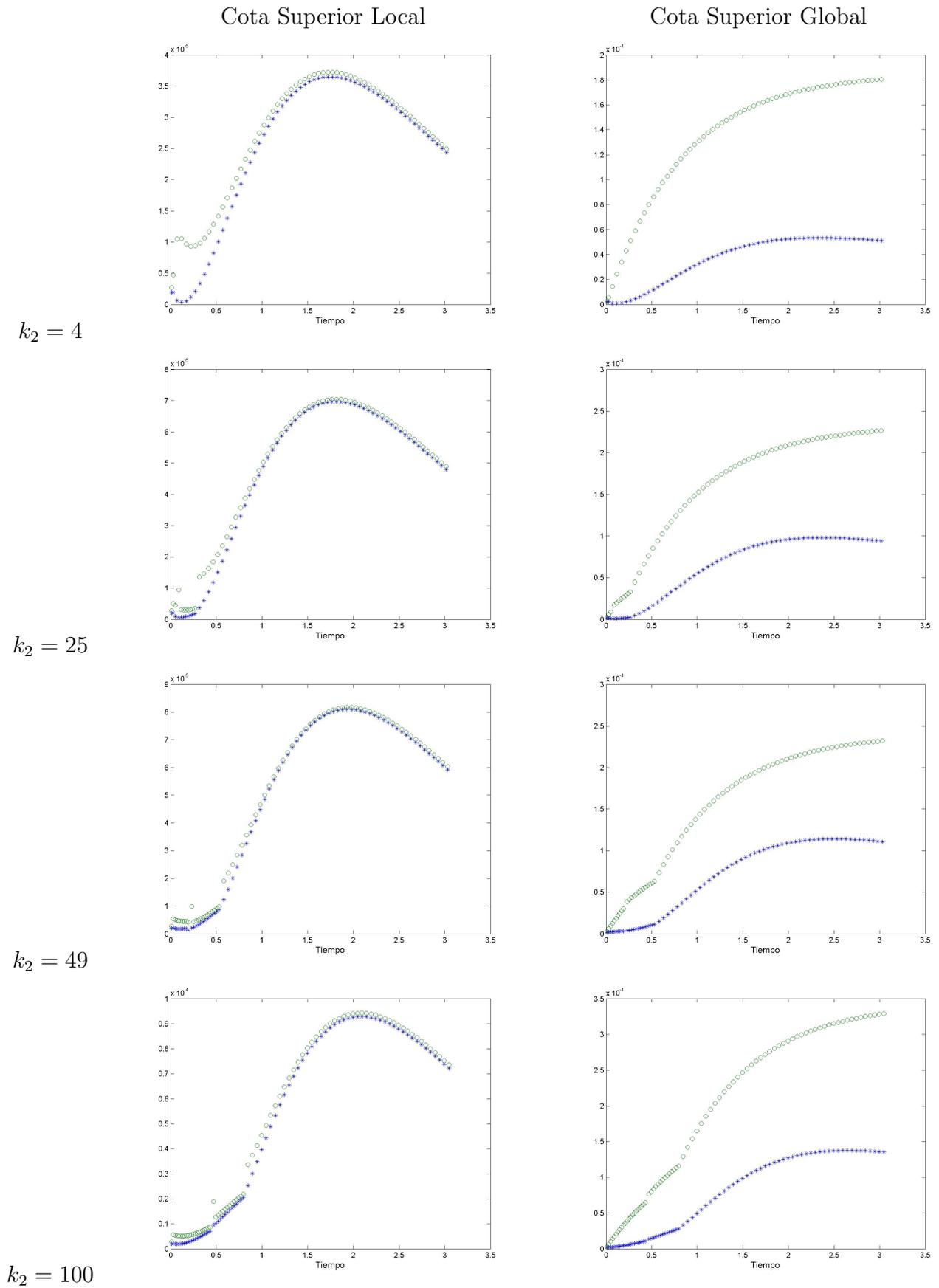


Figura 4.4: Comparación entre el error exacto y estimado. Se muestra en cada gráfico el lado izquierdo (azul) y derecho (verde) de la cota (4.24) (columna izquierda) y (4.26) (columna derecha). El resultado corresponde para ejemplos con $k_1 = 1$ y $k_2 = 4, 25, 49, 100$.

es lo que se pretendía.

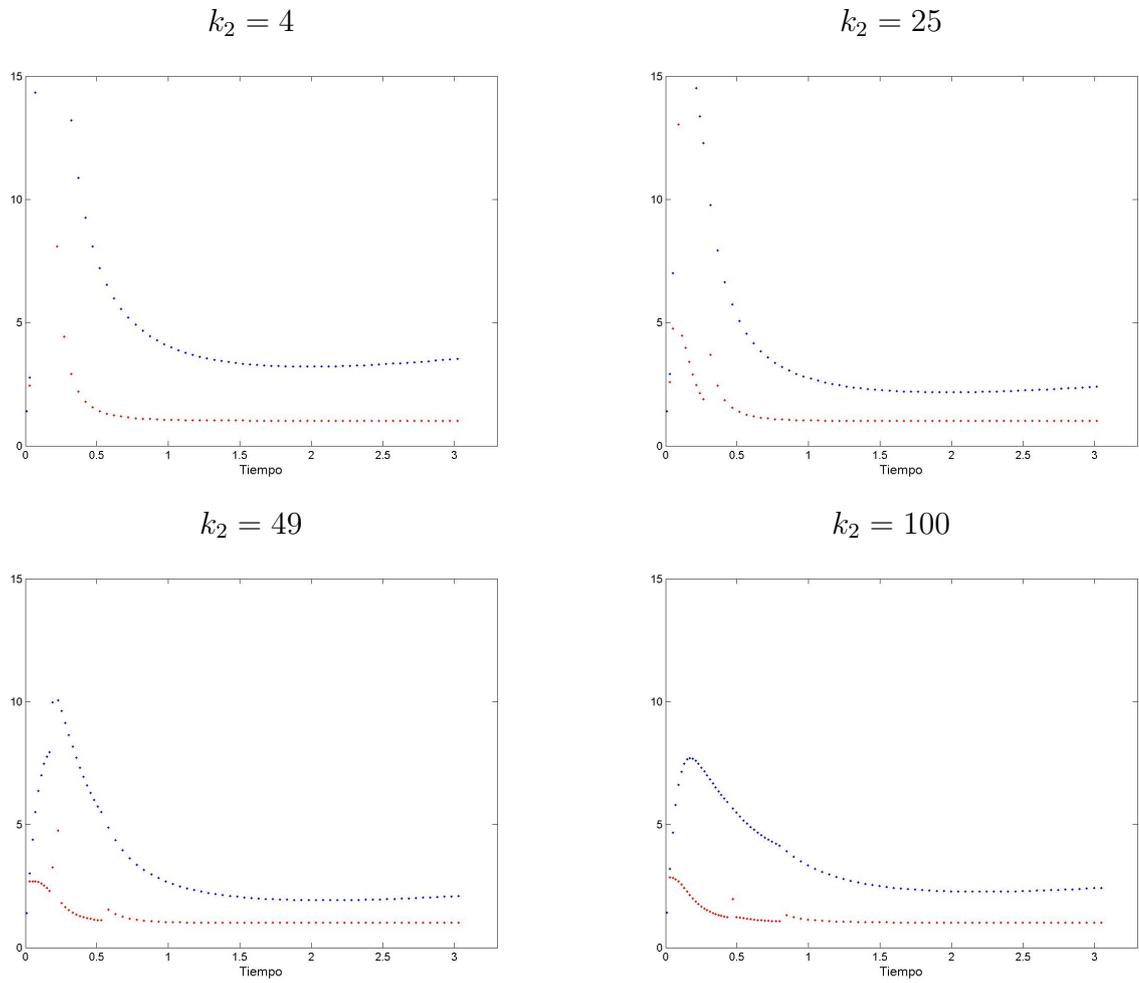


Figura 4.5: Cocientes entre ambos lados de las acotaciones locales (rojo) y globales (azul) dados por la desigualdad (4.24) y (4.26) respectivamente, para $k_1 = 1$ y distintos valores de k_2 .

Capítulo 5

Algoritmo adaptativo

5.1. Introducción

El problema de construir métodos de elementos finitos adaptativos es de gran importancia práctica. En los últimos años, ha sido de interés en numerosas investigaciones, el desarrollo y el análisis de métodos o algoritmos adaptativos para diversos problemas de tipo elíptico o parabólico.

Un algoritmo adaptativo es un procedimiento computacional capaz de construir la solución obtenida mediante discretización por elementos finitos de un problema dado, sin requerir de ninguna información a priori de la solución exacta. El objetivo principal de estos algoritmos es que el error en la aproximación medido en una norma dada, se encuentre acotado por una tolerancia establecida y el número de grados de libertad sea mínimo.

Los estimadores a posteriori del error son el ingrediente esencial en este tipo de procedimientos. Ellos son cantidades calculables en términos de la solución discreta y de los datos del problema, que ofrecen información sobre la calidad de la aproximación y pueden entonces ser usados para realizar con criterio, modificaciones en la malla o en los tamaños de pasos de tiempo. El propósito de estas modificaciones será equidistribuir el error de aproximación y en consecuencia, economizar el esfuerzo computacional.

Es bien conocido que la suavidad de las soluciones de problemas parabólicos puede variar considerablemente en la fase inicial donde oscilaciones pronunciadas de la solución decaen rápidamente. Por lo tanto, los métodos computacionales eficientes para problemas parabólicos, requieren de discretizaciones sobre las dos variables, espacio y tiempo, con pasos pequeños en la fase inicial y pasos más grandes cuando la solución se hace más suave.

En este capítulo, construiremos un algoritmo adaptativo eficiente para resolver el problema (3.1)–(3.2). La construcción del algoritmo se apoyará en el trabajo de Z.Chen y J.Feng [10], el cual se basa en los estimadores a posteriori del error definidos en el capítulo anterior (Definición 4.6), que indican los errores debido a las discretizaciones espacial y temporal.

5.2. Aspectos generales de la implementación

En base al planteo de la forma completamente discreta de (3.1)-(3.2) desarrollaremos un algoritmo adaptativo capaz de resolver aproximadamente dicho problema.

Consideremos el intervalo de tiempo $[0, T]$ y una partición en subintervalos $I_n = (t_{n-1}, t_n)$, $n = 1, \dots, N$. Sea $\Delta t^n = t_n - t_{n-1}$ el n -ésimo tamaño de paso de tiempo, entonces podemos decir que $t_n = \sum_{i=1}^n \Delta t^i$.

Para cada $\Omega \times \bar{I}_n$, sean \mathcal{T}_h^n una triangulación admisible, regular sobre $\bar{\Omega}$ y que satisface la condición del ángulo mínimo (Definiciones A.20 y A.21) y $V_h^n \subset H_0^1(\Omega)$ el espacio de funciones continuas y lineales a trozos sobre la triangulación.

Consideremos $U_h^0 = P_0 u^0$, donde $P_0 : L^2(\Omega) \rightarrow V_h^0$ es el operador proyección $L^2(\Omega)$ sobre el espacio V_h^0 definido en una malla inicial \mathcal{T}_h^0 .

Como ya lo mencionamos en la Sección 3.5, la formulación completamente discreta de (3.1)-(3.2) consiste en: para $n = 1, 2, \dots$ encontrar $U_h^n \in V_h^n$ tal que

$$\left(\frac{U_h^n - U_h^{n-1}}{\Delta t^n}, v \right) + (k \nabla U_h^n, \nabla v) = 0, \quad \forall v \in V_h^n. \quad (5.1)$$

Estamos considerando aquí $f \equiv 0$. Si ésto no ocurriera, el lado derecho de (5.1) sería igual a (\bar{f}_n, v) , donde \bar{f}_n es la aproximación constante a trozos de f sobre $[t_{n-1}, t_n]$ definida como el valor promedio de f en dicho intervalo.

La interpretación del esquema (5.1) es la siguiente: calculado $U_h^{n-1} \in V_h^{n-1}$, sus correspondientes malla \mathcal{T}_h^{n-1} y tamaño de paso de tiempo Δt^{n-1} serán modificados para dar origen a nuevos \mathcal{T}_h^n y Δt^n con los que se calculará $U_h^n \in V_h^n$.

La malla \mathcal{T}_h^n será obtenida a partir de \mathcal{T}_h^{n-1} usando procedimientos de refinamiento, donde la idea general es refinar aquellos triángulos en los que los estimadores del error $(\eta_{R,T}^n)^2$ y $(\eta_{\nabla,T}^n)^2$ sean considerablemente grandes. En esta tesis, no consideraremos el procedimiento de coarsening o engrosamiento sobre ninguna malla, de manera que los espacios de elementos finitos serán anidados ($V_h^{n-1} \subset V_h^n$).

Los refinamientos están basados en el algoritmo de bisección, que conduce a mallas consecutivas compatibles cuyos ángulos mínimos quedan acotados uniformemente lejos de cero. Diremos que dos mallas son *compatibles* si una es el refinamiento local de la otra.

Sobre la base de los estimadores a posteriori locales del error que definimos en el capítulo anterior (Def. 4.6), el algoritmo adaptativo que resuelve nuestro problema, sigue la siguiente secuencia para cada paso de tiempo n :

Resolver \longrightarrow Estimar \longrightarrow Adaptar

Es en la etapa “Adaptar” donde se incluyen las posibles modificaciones en la malla y también en el tamaño del paso de tiempo.

Para cada n , seguiremos el siguiente algoritmo que es propuesto en [10] y permite el control adaptativo de un tamaño del paso de tiempo Δt^n y una malla \mathcal{T}_h^n , comenzando con un tamaño inicial $\Delta t^{n,0}$ y una malla inicial $\mathcal{T}_h^{n,0} := \mathcal{T}_h^{n-1}$:

- Refinar el tamaño de paso de tiempo $\Delta t^{n,0}$ y la malla $\mathcal{T}_h^{n,0}$, de modo que, dada la solución $U_h^n \in V_h^n$ obtenida de (5.1), que está definida sobre la malla final \mathcal{T}_h^n y con tamaño final de paso de tiempo Δt^n , los indicadores del error asociados a las

discretizaciones de las variables espacio y tiempo sean ambos menores que ciertas tolerancias prescriptas.

- En el caso en que el indicador del error de discretización temporal sea *mucho menor* que alguna tolerancia prescripta, agrandar el tamaño de paso de tiempo inicial $\Delta t^{n+1,0}$ para el siguiente paso de tiempo $n + 1$.

En la siguiente sección, formalizamos en detalle este algoritmo a través de un pseudocódigo, que es el dado en el Algoritmo 2.

Si permitiéramos el engrosamiento de las mallas, sería necesario medir el error de coarsening en la diferencia $U_h^n - U_h^{n-1}$ que se presenta en el término $\|\frac{1}{\sqrt{k_T}} R_T^n(U_I)\|_T^2$ de la definición del estimador local espacial $(\eta_{R,T}^n)^2$ y en la definición del estimador local temporal $(\eta_{\nabla,T}^n)^2$ (Def. 4.6). Para ello, se debería introducir $P_n : L^2(\Omega) \rightarrow V_h^n$ el operador proyección $L^2(\Omega)$ y medir el error $P_n U_h^{n-1} - U_h^{n-1}$. Para más detalles, en [10] se puede encontrar el desarrollo y análisis de un algoritmo completo con ambas estrategias de refinamiento y coarsening, junto con la propuesta de un indicador del error asociado al mismo.

5.3. Diseño de un algoritmo adaptativo

Presentamos en esta sección un algoritmo adaptativo que se basa en la cota superior global (4.26), obtenida para estimar el error originado por la discretización completa (5.1). El objetivo del algoritmo será lograr que la estimación del error quede por debajo de una tolerancia dada, reduciendo el tamaño del paso de tiempo y refinando la malla adecuadamente.

Observemos que el lado derecho de dicha cota lo conforman tres términos: uno correspondiente al error de aproximación del dato inicial, y los otros dos correspondientes a los errores generados por las discretizaciones espacial y temporal. Requeriremos entonces que todo el miembro derecho sea menor o igual que una tolerancia TOL establecida previamente y que descomponemos como

$$TOL = TOL_{inicial} + TOL_{espacio} + TOL_{tiempo},$$

siendo $TOL_{inicial}$, $TOL_{espacio}$ y TOL_{tiempo} las tolerancias correspondientes a cada término del lado derecho de (4.26). Es decir, el objetivo del método adaptativo es lograr que:

$$\|u^0 - U_h^0\|^2 \leq TOL_{inicial}, \quad \tilde{C}(\eta_R)^2 \leq TOL_{espacio}, \quad \tilde{C}(\eta_{\nabla})^2 \leq TOL_{tiempo}.$$

La primera desigualdad será el criterio de parada de refinamientos sucesivos a partir de una malla inicial \mathcal{T}_h^0 . Los triángulos marcados para refinar en cada malla son aquellos en donde

$$\|u^0 - U_h^0\|_T^2 > \frac{TOL_{inicial}}{\#\mathcal{T}_h^0},$$

siendo $\#\mathcal{T}_h^0$ la cantidad de triángulos en la malla inicial.

Con respecto a las otras dos desigualdades, los ajustes del tamaño del paso de tiempo y las modificaciones sobre la malla, se realizan en base a la estrategia de equidistribución.

Presentaremos primero el Algoritmo 1 que controla el tamaño del paso de tiempo y luego el Algoritmo 2 que realiza la adaptatividad en las dos variables, es decir controla los tamaños de los pasos de tiempo y permite modificaciones en la malla.

El ajuste del tamaño del paso de tiempo se realiza en base a la estrategia de equidistribución, de tal manera que el error de discretización temporal sea igualmente distribuido en los intervalos de tiempo (t_{n-1}, t_n) , $n = 1, \dots, N$, donde N es el número total de pasos y es tal que $t_N \geq T$. Vale la pena destacar que como el algoritmo adaptativo prevé modificaciones en los tamaños de pasos de tiempos, no se puede a priori determinar el valor de N .

Siendo TOL_{tiempo} la tolerancia establecida para controlar el error debido a la discretización temporal, pediremos que

$$\eta_{\nabla}^2 = \sum_{n=1}^N (\eta_{\nabla}^n)^2 = \sum_{n=1}^N \Delta t^n \left(\sum_{T \in \mathcal{T}_h^n} \|\sqrt{k_T} \nabla (U_h^n - U_h^{n-1})\|_T^2 \right) \leq TOL_{tiempo}. \quad (5.2)$$

Para que ésto se cumpla, modificaremos el tamaño del paso de tiempo Δt^n de manera que se satisfaga

$$\eta_t^n := \sum_{T \in \mathcal{T}_h^n} \|\sqrt{k_T} \nabla (U_h^n - U_h^{n-1})\|_T^2 \leq \frac{TOL_{tiempo}}{T}.$$

Cuando el indicador η_t^n sea demasiado pequeño con respecto a la tolerancia, se permitirá un aumento en el paso de tiempo Δt^{n+1} , con el objetivo de reducir el esfuerzo computacional.

A continuación mostramos el algoritmo que usaremos para el control del tamaño del paso de tiempo en cada instante n :

Algoritmo 1(Controla el tamaño del paso de tiempo):

Dado un Δt^n inicial, suponemos que son conocidos: la tolerancia TOL_{tiempo} y los parámetros $\delta_1 \in (0, 1)$, $\delta_2 > 1$ y $\theta_{tiempo} \in (0, 1)$.

1. Resolver el problema de discretización temporal y estimar el error.
2. Mientras $\eta_t^n > \frac{TOL_{tiempo}}{T}$ hacer

$$\Delta t^n := \delta_1 \Delta t^n$$
 Resolver el problema de discretización temporal y estimar el error.
 fin mientras.
3. Si $\eta_t^n \leq \theta_{tiempo} \frac{TOL_{tiempo}}{T}$
 hacer $\Delta t^{n+1} := \delta_2 \Delta t^n$
 sino hacer $\Delta t^{n+1} := \Delta t^n$
 fin si.

En la práctica, una buena elección de los parámetros en el Algoritmo 1 es tomar $\delta_1 = 0,5$, $\delta_2 = 2$ y $\theta_{tiempo} = 0,5$. Como el algoritmo anterior sólo controla el tamaño del paso de tiempo, aún nos falta combinarlo con otro algoritmo que realice adaptatividad sobre la malla.

Sea $TOL_{espacio}$ la tolerancia establecida para controlar el error relacionado con la discretización espacial. Entonces, análogamente a lo requerido en (5.2), en cada paso de tiempo n , pediremos que se satisfaga el siguiente criterio de parada para los refinamientos sobre la malla:

$$\eta_e^n := \sum_{T \in \mathcal{T}_h^n} \left[(h_T^n)^2 \left\| \frac{1}{\sqrt{k_T}} R_T^n(U_I) \right\|_T^2 + \frac{1}{2} \sum_{E \in E_h^n(T) \cap E_h^n(\Omega)} h_E^n \left\| \frac{1}{\sqrt{k_{TE}}} J_E^n(U_I) \right\|_E^2 \right] \leq \frac{TOL_{espacio}}{T} \quad (5.3)$$

De esta manera, al cumplirse esta condición, se satisface que

$$\eta_R^2 = \sum_{n=1}^N \Delta t^n \eta_e^n \leq TOL_{espacio}.$$

Si aplicamos las ideas comentadas en esta sección, llegamos al siguiente algoritmo adaptativo, que combina la adaptatividad espacial y temporal, en un determinado paso de tiempo n :

Algoritmo 2 [10](Algoritmo adaptativo en las variables espacio y tiempo):

Suponemos que disponemos de los siguientes datos: las tolerancias TOL_{tiempo} , $TOL_{espacio}$ y los parámetros $\delta_1 \in (0, 1)$, $\delta_2 > 1$ y $\theta_{tiempo} \in (0, 1)$. Además suponemos que en el paso de tiempo anterior fue calculada U_h^{n-1} sobre una malla \mathcal{T}_h^{n-1} y con tamaño de paso de tiempo Δt^{n-1} y que en dicho paso se decidió un valor tentativo para Δt^n .

1. $\mathcal{T}_h^n := \mathcal{T}_h^{n-1}$, $t_n := t_{n-1} + \Delta t^n$.

Resolver el problema discreto para U_h^n sobre \mathcal{T}_h^n usando el dato U_h^{n-1} y calcular la estimación del error sobre \mathcal{T}_h^n .

2. Mientras $\eta_t^n > \frac{TOL_{tiempo}}{T}$ hacer

$$\Delta t^n := \delta_1 \Delta t^n, \quad t_n := t_{n-1} + \Delta t^n.$$

Resolver el problema discreto para U_h^n sobre \mathcal{T}_h^n usando el dato U_h^{n-1} y calcular la estimación del error sobre \mathcal{T}_h^n .

fin mientras

3. Mientras $\eta_e^n > \frac{TOL_{espacio}}{T}$ hacer

refinar la malla \mathcal{T}_h^n generando una malla modificada \mathcal{T}_h^n ,

resolver el problema discreto para U_h^n sobre \mathcal{T}_h^n usando el dato U_h^{n-1} y calcular la estimación del error sobre \mathcal{T}_h^n .

$$\text{Mientras } \eta_t^n > \frac{TOL_{tiempo}}{T} \text{ hacer } \Delta t^n := \delta_1 \Delta t^n, \quad t_n := t_{n-1} + \Delta t^n.$$

Resolver el problema discreto para U_h^n sobre \mathcal{T}_h^n usando el dato U_h^{n-1} y calcular la estimación del error sobre \mathcal{T}_h^n .

```

    fin mientras
fin mientras
4. Si  $\eta_t^n \leq \theta_{tiempo} \frac{TOL_{tiempo}}{T}$ 
    hacer  $\Delta t^{n+1} := \delta_2 \Delta t^n$ 
    sino hacer  $\Delta t^{n+1} := \Delta t^n$ ,
fin si.
```

El propósito de los tres primeros pasos del Algoritmo 2 es reducir el tamaño de paso de tiempo y refinar la malla a fin de que los indicadores del error temporal η_{∇}^n y espacial η_R^n sean menores que las respectivas tolerancias. Logramos este propósito primero reduciendo el tamaño de paso de tiempo (paso 2) para tener la estimación del error temporal por debajo de la tolerancia mientras conservamos la malla sin cambiar. En el paso 3, se refina la malla para asegurar que el error debido a la discretización espacial no supere la tolerancia prescrita, y nuevamente se controla el error temporal. De esta manera, al finalizar el paso 3, ambos errores (espacial y temporal) están debajo de las respectivas tolerancias. En el último paso, se decide cuál es el Δt^{n+1} inicial del ciclo adaptativo. Cuando el indicador del error temporal es mucho más pequeño que la tolerancia, el tamaño del paso se agranda en un factor $\delta_2 > 1$.

En [10] se demuestra que para cualquier tolerancia dada $TOL_{espacio}$ y TOL_{tiempo} el Algoritmo adaptativo 2 termina en un número finito de pasos.

5.4. Implementación

Para el desarrollo del Algoritmo 2 que presentamos en la sección anterior y para su implementación en ejemplos numéricos, utilizamos el Partial Differential Equation (PDE) Toolbox de Matlab [14], herramienta poderosa y flexible para el estudio y resolución numérica de ecuaciones en derivadas parciales en dos dimensiones espaciales y dependientes del tiempo, entre otros.

Comenzamos con la construcción del dominio poligonal Ω dividido en dos subdominios y con la imposición de las condiciones de borde utilizando el entorno gráfico del PDE Toolbox.

Luego, realizamos la programación de los algoritmos. Para ello, invocamos y adaptamos a nuestro problema, determinadas funciones ya incorporadas en el PDE Toolbox, como por ejemplo aquellas para generar o refinar una malla, obtener datos de la geometría, graficar, calcular el flujo interelementos o interpolar datos.

Dada una malla triangular sobre $\bar{\Omega}$ y para cualquier tiempo $t_n \in [0, T]$, hemos visto en la sección 3.3, que la solución por elementos finitos $U_h^n = U_h(x, t_n)$ puede expresarse en términos de $\phi_i, i = 1, \dots, m$, que son las funciones que conforman la base nodal del espacio V_h^n , es decir,

$$U_h(x, t_n) = \sum_{i=1}^m \xi_i(t_n) \phi_i(x)$$

donde $m = \dim V_h^n$, está dado por el número de nodos interiores de la triangulación asociada al eslabón espacio-tiempo $\Omega \times [t_{n-1}, t_n]$.

De aquí que, como ya lo vimos en la Sección 3.5, en cada paso de tiempo n la formulación débil completamente discreta cuando $f \equiv 0$ puede escribirse de manera equivalente en la forma matricial siguiente:

$$\begin{cases} (M^n + \Delta t^n K^n) \xi^n = M^n \xi^{n-1}, & n = 1, \dots, m \\ M^0 \xi^0 = U^0, \end{cases} \quad (5.4)$$

donde M^n es la matriz de masa y K^n la matriz de rigidez correspondientes al paso n .

La incógnita ξ^n del sistema es un vector columna cuya longitud coincide con la dimensión de V_h^n y cada componente i del vector ξ^n es el valor de U_h^n en el nodo x_i , es decir, $\xi_i^n = U_h(x_i, t_n)$.

Si x es un punto interior de un triángulo, el valor de $U_h^n(x)$ se determina por interpolación lineal de los valores nodales.

Cada elemento M_{ij}^n de la matriz M^n contiene integrales en términos de las funciones ϕ_i y ϕ_j de la base nodal y cada elemento K_{ij}^n de la matriz K^n contiene integrales en términos de los gradientes de ϕ_i y ϕ_j . Concretamente,

$$M_{ij}^n = \int_{\Omega} \phi_i(x) \phi_j(x) dx \quad \text{y} \quad K_{ij}^n = \int_{\Omega} k(x) \nabla \phi_i \cdot \nabla \phi_j dx.$$

Dado que por definición $\phi_i(x_j) = \delta_{ij}$, la función de base ϕ_i se anula sobre todos los triángulos que no contienen al nodo x_i . En consecuencia, es suficiente calcular las integrales que aparecen en M_{ij}^n y K_{ij}^n , sólo sobre aquellos triángulos que tienen por vértices a los dos nodos x_i y x_j , ya que en todo otro caso M_{ij}^n y K_{ij}^n son cero. Así, las matrices M^n y K^n son ralas (sparse).

Las integrales de la definición de M_{ij}^n y K_{ij}^n , se calculan de manera exacta utilizando argumentos de escala. Para el cálculo de aquellas que intervienen en la norma L^2 del error $u^0 - U^0$, usamos cuadratura del punto medio y para el cálculo del estimador local $(\eta_{R,T}^n)^2$, cuadratura de orden cúbico.

Para resolver los sistemas lineales que aparecen, utilizamos el comando “\” de Matlab que elige automáticamente el método dependiendo de la estructura y el tamaño de la matriz .

La generación de una nueva malla a través de un refinamiento se realiza mediante una función específica del PDE Toolbox.

Los triángulos marcados para refinar son aquellos para los cuales

$$(\eta_{R,T}^n)^2 > \frac{TOL_{espacio}}{T \cdot \#\mathcal{T}_h^n} \Delta t^n,$$

donde $\#\mathcal{T}_h^n$ indica el número de triángulos sobre la malla. Los triángulos marcados se refinan utilizando la función *refinemesh*. Para más detalles sobre el algoritmo de refinamiento ver [14].

Capítulo 6

Simulación adaptativa de liberación de droga en dispositivos poliméricos

En la industria agroveterinaria, existen dispositivos de liberación de droga por difusión, que son utilizados como un medio para el control de insectos en el ganado.

Como dijimos en la introducción, la motivación inicial de esta tesis es el estudio del efecto de la forma de los dispositivos, en el perfil de liberación de droga hacia el exterior, y el objetivo a alcanzar es una liberación de droga casi constante por unidad de tiempo, de manera de evitar dos problemas fundamentales:

- que la cantidad de droga sea insuficiente para controlar al insecto;
- que la cantidad de droga en el aire sea demasiada y resulte tóxica para el animal.

También es deseable que la cantidad de droga que quede almacenada en el dispositivo al terminar el período sea la menor posible.

En general, el parámetro que interesa a los diseñadores es el flujo por unidad de tiempo hacia fuera del dispositivo, y existen para este parámetro un límite superior de toxicidad y un límite inferior de ineficacia. Si bien es la variable que se pretende controlar, dado que queremos comprender el efecto de la forma del dispositivo, y de eventuales concavidades que pueda tener, entendemos que es igualmente (o más) importante estudiar el flujo de droga hacia fuera de una región un poco más grande que el dispositivo. Esta simulará la región por donde circularán los insectos y/o donde se encuentra la parte del animal susceptible a la droga (ojos, nariz, boca, etc.). Por este motivo, y con el objetivo de comprender mejor el efecto de la forma del dispositivo, estamos estudiando la difusión simultáneamente en su interior y en parte del medio exterior. Para ello consideramos el dominio poligonal Ω particionado en dos subdominios también poligonales Ω_1 y Ω_2 que representan al dispositivo y a parte del ambiente exterior que lo rodea, respectivamente.

En este capítulo, exponemos los resultados numéricos obtenidos de la aplicación del Algoritmo 2 que fue presentado en la sección 5.3, sobre cuatro ejemplos en los que se varía la forma del dominio y las condiciones en la frontera. Comenzaremos con una forma sencilla y con condiciones de borde Dirichlet homogéneas en todo el dominio, y de acuerdo a las curvas de flujo obtenidas que reflejan la cantidad de droga liberada por unidad de tiempo, iremos haciendo modificaciones en dicho dominio hasta encontrar uno cuyas curvas de flujo sean lo más aproximadamente constante.

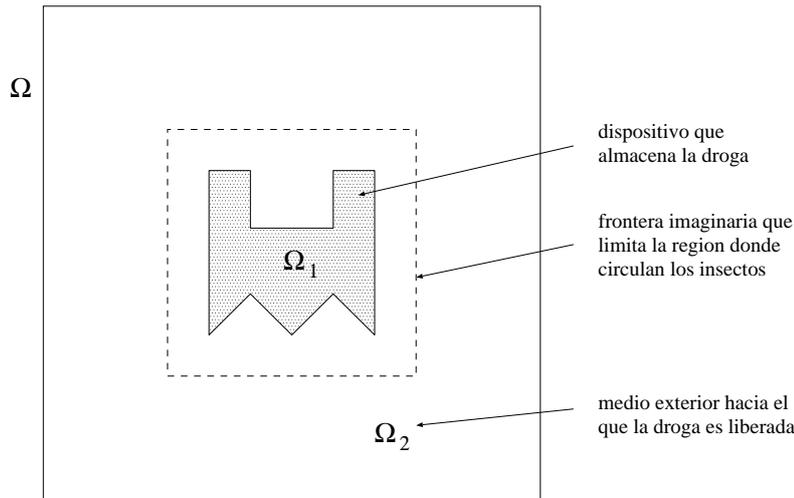


Figura 6.1: Esquema del dominio Ω , subdominios Ω_1 y Ω_2 y la frontera imaginaria

En cada experimento, consideramos como dominio al cuadrado $\Omega = (-3, 3) \times (-3, 3)$ particionado en dos regiones Ω_1 y Ω_2 . El subdominio Ω_1 representará el dispositivo que contiene la droga y en todos los casos será un polígono contenido en el cuadrado $[-1, 1] \times [-1, 1]$. Lo que resta es Ω_2 , representando al medio exterior que rodea al dispositivo, en este caso, el aire (Ver Figura 6.1).

En todos los experimentos analizaremos lo que llamamos flujo *efectivo* de droga hacia el exterior. Para ello, consideraremos una frontera imaginaria dada por el borde del cuadrado $(-1,5, 1,5) \times (-1,5, 1,5)$, como se ilustra con línea punteada en la Figura 6.1, sobre las que se estudiarán las curvas de flujo. Este cuadrado corresponde a la región donde supondremos circulan los insectos que se quieren combatir.

En nuestro problema, el coeficiente de difusión $k(x)$ tiene valores distintos en cada subdominio y posiblemente con órdenes de magnitud muy diferentes. En esta tesis sólo experimentamos con valores ficticios de k y no explicitamos unidades para la variable tiempo ni para la concentración u .

Ya que la difusión en el dispositivo es más lenta que la del aire, consideramos en cada ejemplo, un valor fijo para k_2 (coeficiente de difusión en el aire) y diferentes valores para k_1 (coeficiente de difusión en la matriz). Concretamente experimentamos con $k_2 = 100$ y $k_1 = 0,01, 0,1, 1, 10, 100$.

Con respecto a la condición inicial del problema, suponemos que la función $u^0(x) \equiv 1$ en Ω_1 y $u^0(x) \equiv 0$ en Ω_2 , simulando que inicialmente toda la droga está concentrada en el dispositivo.

En cada ejemplo mostramos:

- la malla inicial del dominio (común para todos los valores de k_1) y una malla final obtenida del procedimiento adaptativo para un valor de k_1 .
- una secuencia de soluciones aproximadas, para dos valores de k_1 . Los valores que toman las soluciones aproximadas se mostrarán por medio de una escala de colores que va de 0 (azul) hasta 1 (rojo). Por ello, en $t = 0$ el dispositivo estará completamente

rojo, indicando que toda la droga está concentrada allí y el resto completamente azul, indicando ausencia de droga. En color blanco podrá verse la forma del dispositivo. A medida que transcurre el tiempo, la distribución de droga en todo el dominio se observará a través de la variación de los colores.

- las curvas de flujo de droga por unidad de tiempo de las cinco elecciones de k_1 , a través del dispositivo y de la frontera imaginaria .

En todos los casos tomaremos $T = 2$ y una tolerancia prefijada TOL que descomponemos en $TOL_{inicial} = TOL/2$ y $TOL_{espacio} = TOL_{tiempo} = TOL/4$. Además, tomamos Δt inicial entre 0,1 y 0,5.

En todos los ejemplos que presentamos a continuación el refinamiento de la malla se debió principalmente a la discontinuidad de la condición inicial. El paso de tiempo resultó muy pequeño al principio, y fue aumentándose a medida que se alejaba de la fase inicial.

6.1. Ejemplo 1: Cuadrado

En este primer ejemplo, consideramos el dispositivo (subdominio Ω_1) de forma cuadrada, específicamente, el cuadrado $(-1, 1) \times (-1, 1)$.

En todas las pruebas realizadas donde fuimos variando los valores de k_1 , comenzamos con una misma partición del dominio, pero al finalizar el proceso adaptativo, se obtuvieron en cada prueba diferentes refinamientos en la malla. Por ejemplo, en la Figura 6.2 mostramos la malla inicial de 210 nodos y la malla final con 6470 nodos obtenida cuando $k_1 = 0,01$ y $TOL = 0,2$.

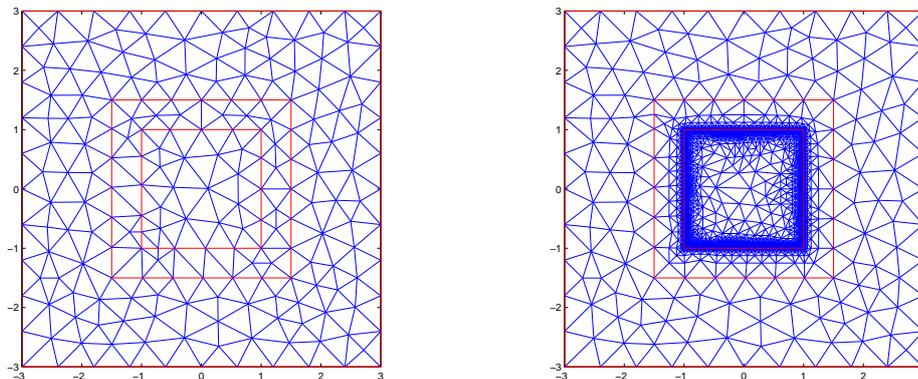


Figura 6.2: Malla inicial y malla final resultante del algoritmo adaptativo cuando el dispositivo es cuadrado y su coeficiente es $k_1 = 0,01$.

En general pudimos observar un gran refinamiento alrededor de la frontera del dispositivo. Esto se debe a que aquí es donde el gradiente de la solución (y el coeficiente de difusión) presenta saltos y pérdida de regularidad.

Los dibujos que presentamos en la Figura 6.3 permiten visualizar y comparar el proceso de difusión de droga desde el dispositivo hacia afuera para dos valores distintos del

coeficiente k_1 . Las imágenes son una selección de soluciones aproximadas que se obtuvieron para $k_1 = 0,1$ y $k_1 = 10$, elegidas de tal manera que los valores de los pasos de tiempo sean los más cercanos posibles entre sí.

Observamos diferencias muy notorias entre las imágenes de las dos columnas. En la primera columna, dado que k_1 es mucho menor que k_2 se observa difusión muy lenta. Entonces, por un lado la poca cantidad de droga que es liberada, rápidamente se expande en el aire (por eso el color azul para todos los valores de t) y por otro lado, al terminar el período de tiempo queda residuo de droga aún en el dispositivo. En cambio, en la segunda columna, donde k_1 es sólo un orden de magnitud menor que k_2 , la difusión es rápida tanto en el dispositivo como en el aire. Esto hace que se libere mucha cantidad de droga en cada t (se observa en los tonos celestes alrededor de la matriz) y que aproximadamente en la mitad del período de tiempo el dispositivo quede prácticamente vacío.

Los gráficos en la Figura 6.4 muestran para los cinco valores de k_1 , el flujo por unidad de tiempo a través de dos fronteras: la frontera del dispositivo y la frontera imaginaria por donde circulan los insectos. Vemos que el flujo sobre esta última difiere sustancialmente —al menos al principio— del flujo hacia fuera del dispositivo.

Para esta forma de dispositivo las dos curvas presentan un decaimiento exponencial, que se acentúa al inicio. Para valores mayores de k_1 el decaimiento es más pronunciado. Este comportamiento resulta indeseable desde el punto de vista práctico porque indica que en muy corto plazo de tiempo toda la droga se diluye en el aire. Sin embargo, cuanto mayor sea k_1 , menor es el resto de droga que queda dentro del dispositivo al finalizar el proceso.

La diferencia entre los dos flujos considerados se percibirá mejor en los siguientes ejemplos donde la forma del dispositivo juega un papel más importante.

6.2. Ejemplo 2: Cuadrado con concavidad

Para comprender el efecto de la forma del dispositivo en el comportamiento de las curvas de flujo, modificamos la geometría del mismo introduciendo una concavidad.

Analizaremos si la concavidad es capaz de producir un retraso o demora en la salida inicial de droga hacia el exterior. Este nuevo Ω_1 tiene la forma de una letra “C” rotada, como se observa en la Figura 6.5. También en esta figura, se muestran la malla inicial de 463 nodos y la malla final con 6370 nodos, obtenida cuando $k_1 = 0,01$ y $TOL = 0,3$.

Hemos notado que, en general, hay un mayor esfuerzo computacional en este ejemplo que en el anterior, en cuanto al número de nodos obtenidos en la mallas finales para cada k_1 . Esto lo atribuimos a las puntas y esquinas originadas por la concavidad de Ω_1 .

Como lo hicimos en el ejemplo anterior, presentamos en la Figura 6.6 una selección de soluciones aproximadas que se obtuvieron para dos valores de k_1 : 0,1 y 10, elegidas de manera que los valores de los pasos de tiempo sean los más cercanos posibles.

Las conclusiones que obtuvimos en el ejemplo anterior cuando describimos y comparamos las imágenes de las dos columnas son similares. Lo único que lo diferencia es que al tener el dispositivo una concavidad, hay mayor perímetro de borde por donde se libera la droga y en consecuencia, la difusión se acelera más para todos los valores de k_1 . Por ello, para $k_1 = 0,1$ casi no se observa residuo de droga dentro del dispositivo en $T = 2$, o para $k_1 = 10$ en $t = 0,1547$, mientras que en el ejemplo anterior ésto no ocurría.

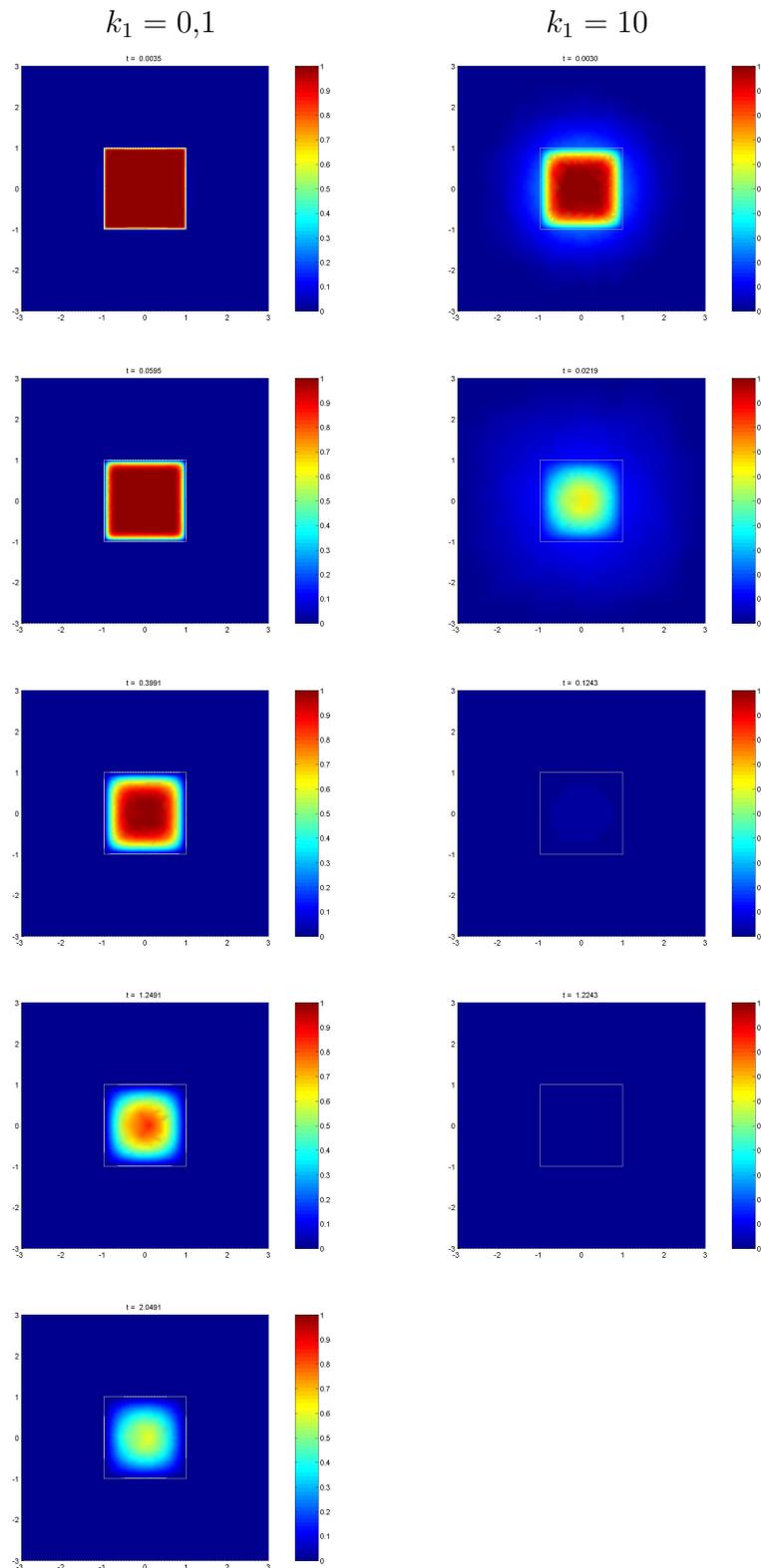


Figura 6.3: Perfil de la concentración de droga para dos valores de k_1 cuando el dispositivo es cuadrado. La primera columna corresponde a $k_1 = 0,1$ y $t = 0,0035$, $t = 0,0595$, $t = 0,3991$, $t = 1,2491$ y $t = 2,0491$. La segunda columna corresponde a $k_1 = 10$ y $t = 0,003$, $t = 0,0219$, $t = 0,1243$ y $t = 1,2243$. Ambas simulaciones fueron realizadas con $k_2 = 100$.

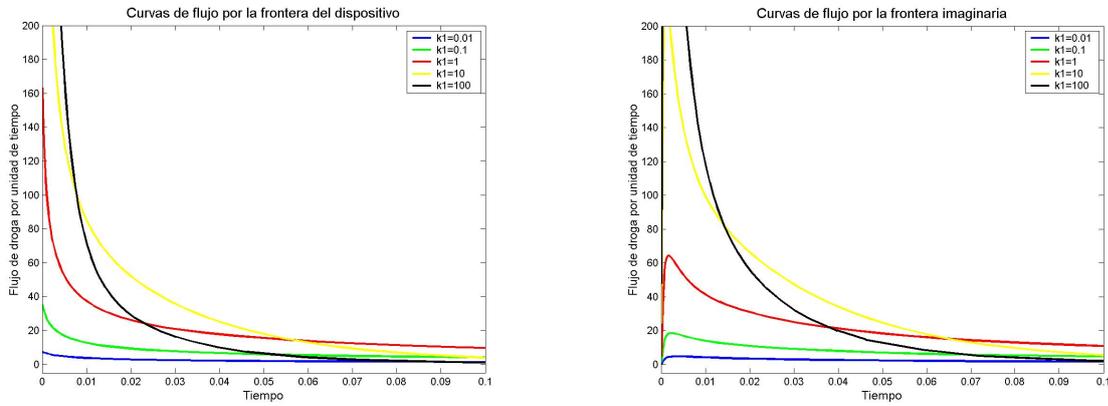


Figura 6.4: Curvas de flujo a través del dispositivo (izquierda) y la frontera imaginaria (derecha) cuando el dispositivo es cuadrado

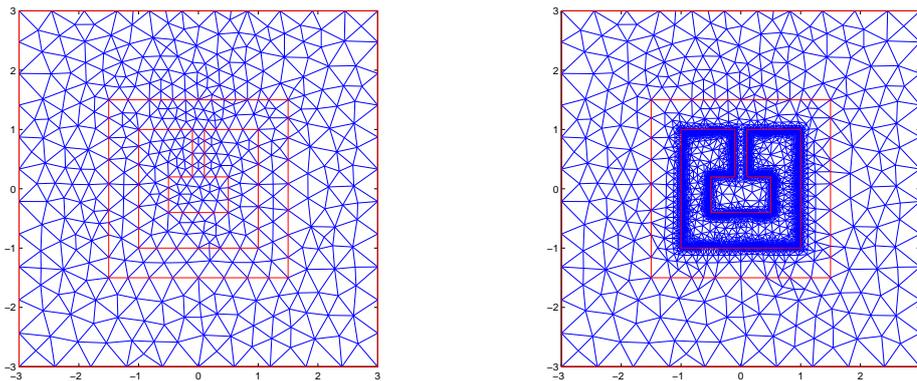


Figura 6.5: Malla inicial y malla final resultante del algoritmo adaptativo cuando el dispositivo es cuadrado con concavidad y coeficiente $k_1 = 0,01$.

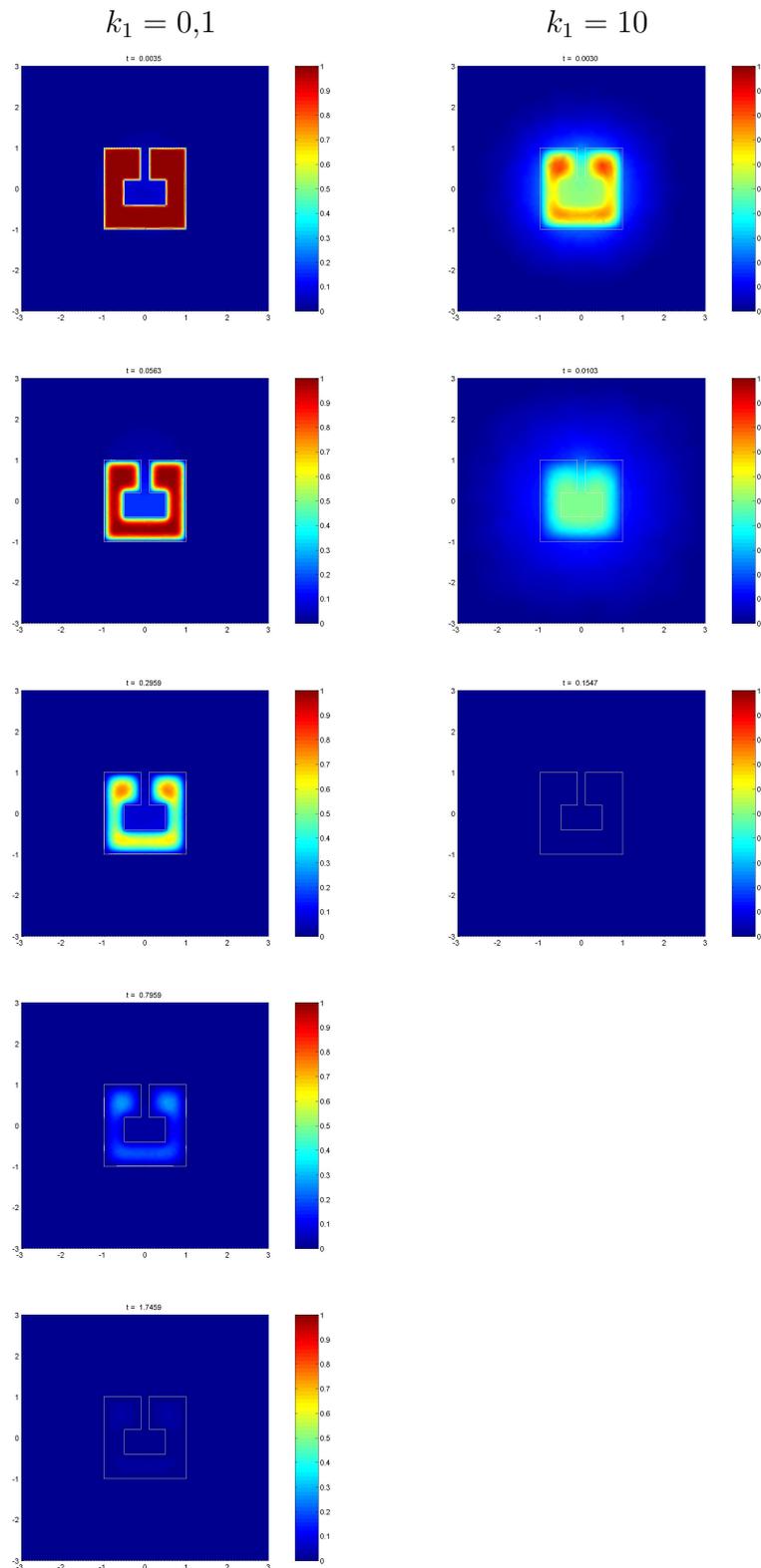


Figura 6.6: Perfil de la concentración de droga para dos valores de k_1 cuando el dispositivo es cuadrado con concauidad. La primera columna corresponde a $k_1 = 0,1$ y $t = 0,0035$, $t = 0,0563$, $t = 0,2959$, $t = 0,7959$ y $t = 1,7459$. La segunda columna corresponde a $k_1 = 10$ y $t = 0,003$, $t = 0,0103$ y $t = 0,1547$. Ambas simulaciones fueron realizadas con $k_2 = 100$.

Para nuestra sorpresa, las curvas de flujo a través de las dos fronteras, como vemos en los gráficos de la Figura 6.7, no presentan diferencias significativas en comparación con las del ejemplo anterior. Estas siguen mostrando un decaimiento exponencial luego de un breve período donde el flujo es creciente. En este ejemplo se observa una diferencia mayor entre el flujo hacia fuera del dispositivo y el flujo hacia fuera del cuadrado $(-1,5, 1,5) \times (-1,5, 1,5)$. Esto se debe a que ahora el dispositivo difiere considerablemente de este cuadrado imaginario, y se produce, por un breve lapso de tiempo, una acumulación de droga en la concavidad.

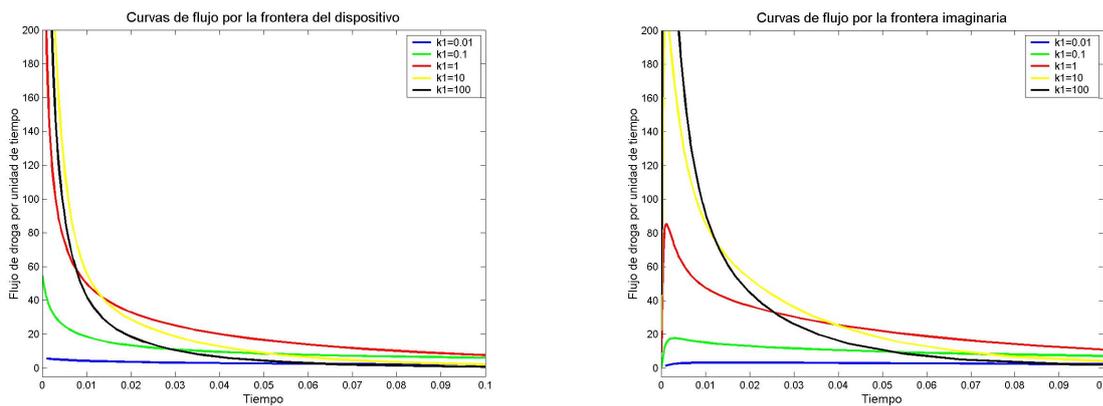


Figura 6.7: Curvas de flujo a través del dispositivo (izquierda) y la frontera imaginaria (derecha) cuando el dispositivo es cuadrado con concavidad

6.3. Ejemplo 3: Cuadrado con parte del borde aislada

Viendo, a partir de los dos ejemplos anteriores, que sólo cambiando la forma del dispositivo no se logra una diferencia sustancial en el perfil de liberación de droga, decidimos realizar cambios de otra naturaleza: aislar parte del dispositivo considerando la difusión sólo a través de una región pequeña del borde. En este caso, dicha región es un segmento contenido en el lado superior del cuadrado Ω_1 del ejemplo 1. Para la implementación, hemos introducido una delgada banda alrededor de la parte aislada del borde del dominio hacia donde no fluye droga, que representa una región de material aislante. Esto significa que en el problema imponemos nuevas condiciones en el borde: sobre esta región se consideran condiciones de tipo Neumann homogéneas.

En la Figura 6.8 mostramos la malla inicial de 244 nodos y la malla final con 1144 nodos, obtenida cuando $k_1 = 0,1$ y $TOL = 0,05$. Notemos la diferencia con el ejemplo 1 en cuanto a la reducción del número de nodos, aún cuando la tolerancia en este caso es menor.

Puede observarse que el refinamiento adaptativo se ha producido principalmente alrededor de la parte del borde del dispositivo a través de la cual la difusión está permitida, y en los vértices de la banda aislante. Esto es razonable dado que allí es donde ocurren los saltos importantes del gradiente, y donde la solución es, por lo tanto, poco regular. Cabe

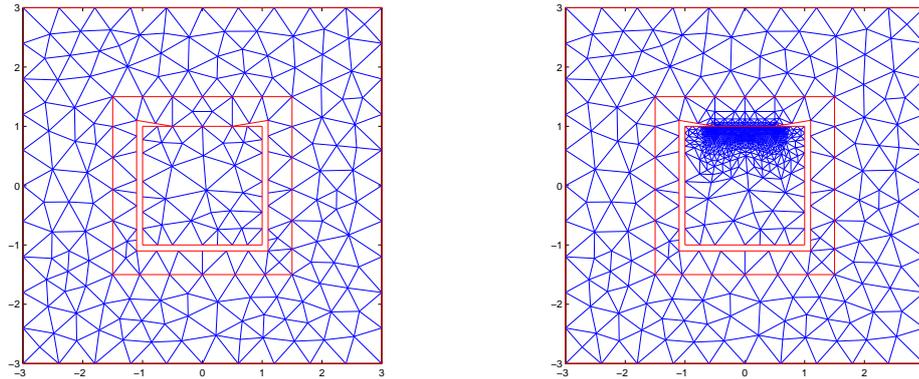


Figura 6.8: Malla inicial y malla final resultante del algoritmo adaptativo cuando el dispositivo es cuadrado y tiene parte aislada en el borde

observar que los vértices de la banda aislante representan concavidades importantes para el dominio computacional, y es sabido que son, en general, responsables de pérdidas de regularidad significativas.

En la Figura 6.9 se observa el resultado de esta nueva simulación y vemos una gran diferencia con los gráficos de la figura 6.3 del ejemplo 1 y los gráficos de la Figura 6.6 del ejemplo 2. Se evidencia para los dos valores de k_1 (y también para los otros k_1) una demora en la difusión.

Las gráficas de la Figura 6.10 muestran las curvas de flujo para los cinco valores del parámetro de difusión. Encontramos ahora diferencias muy significativas respecto de los ejemplos anteriores. Se ve en las mismas que después de una etapa inicial en que ocurre una gran difusión hacia fuera del dispositivo, las curvas presentan un comportamiento aproximadamente constante. Sin embargo, estas curvas no son del todo satisfactorias. Puede observarse que aquella correspondiente al flujo *efectivo*, presenta un pico antes del tiempo $t = 0,005$ que es indeseable pues puede sobrepasar los niveles de toxicidad permitidos.

6.4. Ejemplo 4: Cuadrado con concavidad y parte del borde aislada

Análogamente al ejemplo anterior, consideramos el dispositivo del ejemplo 2 con forma de letra “C”, en donde, aislamos los bordes exteriores de Ω_1 , imponiendo condiciones de borde de tipo Neumann homogéneas, y permitimos el proceso de difusión de droga sobre la parte cóncava de Ω_1 .

Observemos en la Figura 6.11 la malla final de 2830 nodos al usar $TOL = 0,2$, resultado de sucesivos refinamientos adaptativos de la malla inicial de 531 nodos. Análogamente al ejemplo anterior, puede observarse que el refinamiento adaptativo se produce principalmente alrededor de la parte del borde del dispositivo a través de la cual la difusión está permitida, y en los vértices de la banda aislante. Notemos también la diferencia con el ejemplo 2 en cuanto a la reducción del número de nodos, aún cuando la tolerancia sea

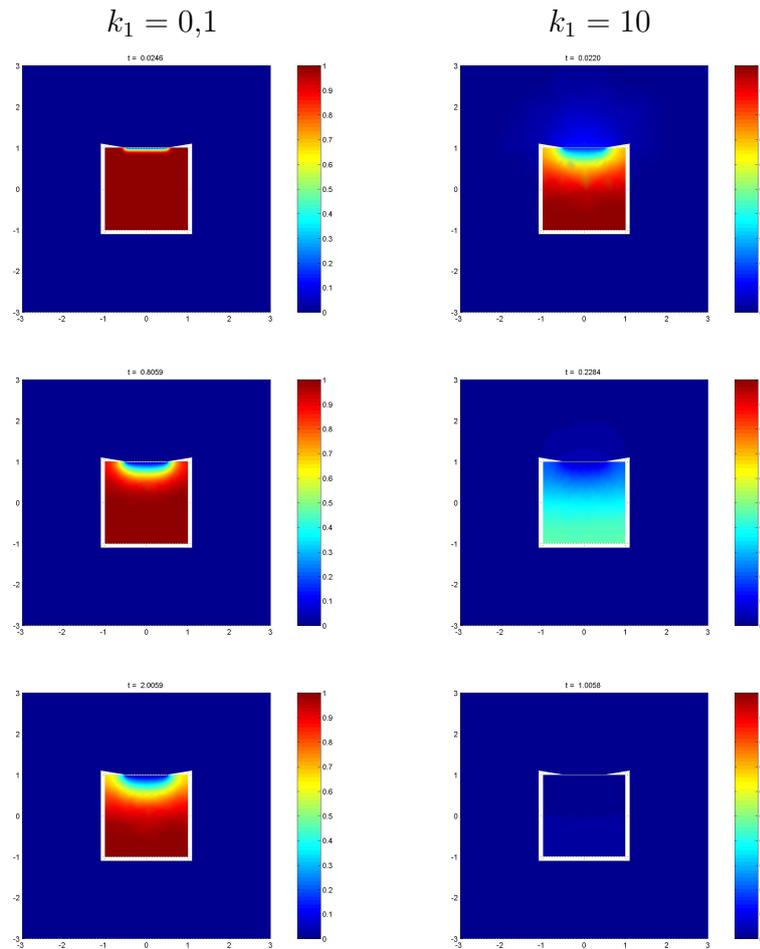


Figura 6.9: Perfil de la concentración de droga para dos valores de k_1 cuando el dispositivo es cuadrado y tiene parte del borde aislado. La primera columna corresponde a $k_1 = 0,1$ y $t = 0,0246$, $t = 0,8059$ y $t = 2,0059$. La segunda columna corresponde a $k_1 = 10$ y $t = 0,0220$, $t = 0,2284$ y $t = 1,0058$. Ambas simulaciones fueron realizadas con $k_2 = 100$.

6.4. EJEMPLO 4: CUADRADO CON CONCAVIDAD Y PARTE DEL BORDE AISLADA 83

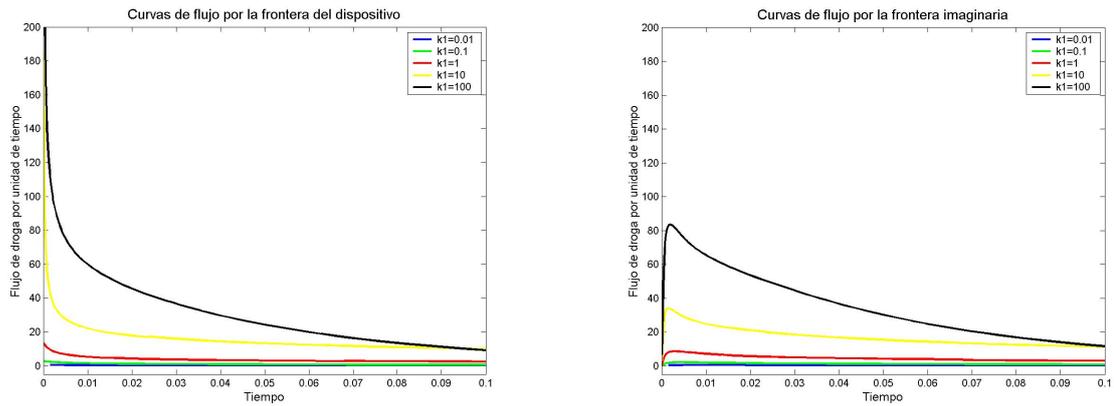


Figura 6.10: Curvas de flujo a través del dispositivo (izquierda) y la frontera imaginaria (derecha) cuando el dispositivo es cuadrado con parte del borde aislado

menor en este caso.

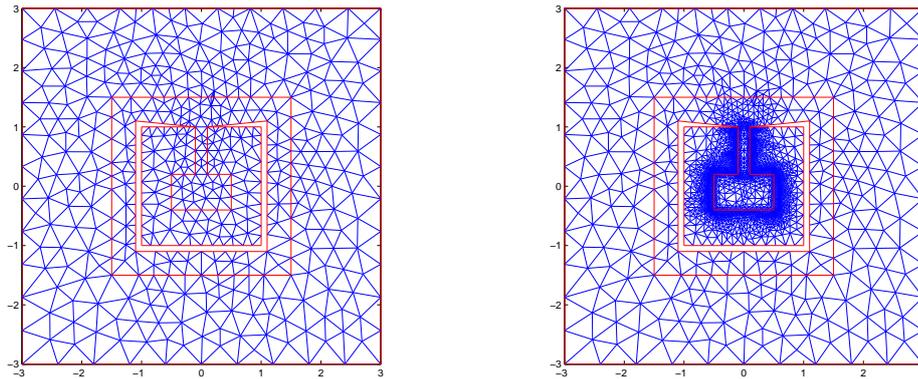


Figura 6.11: Malla inicial y malla final resultante del algoritmo adaptativo cuando el dispositivo es cuadrado, con concavidad y tiene parte aislada en el borde

El resultado de la simulación de liberación de droga para esta nueva forma de dispositivo se observa en la Figura 6.12. Comparando las gráficas del ejemplo 2 en la Figura 6.6 y las del ejemplo 3 en la Figura 6.9, se observa un notorio retraso en la difusión por efecto de la combinación de dos cosas: el hecho de no permitir el paso de droga en casi todo el borde exterior del dispositivo y el hecho de que la difusión sólo se produzca en la concavidad de la matriz provocando que la droga quede retenida por más tiempo.

Encontramos ahora en las curvas de flujo de la Figura 6.13 diferencias muy significativas respecto de los ejemplos anteriores, sobre todo de los dos primeros. Se ve en las mismas que después de una etapa inicial en que ocurre una gran difusión hacia fuera del dispositivo las curvas presentan un comportamiento aproximadamente constante. Puede observarse que la correspondiente al flujo hacia fuera del cuadrado imaginario, que consideramos el flujo *efectivo*, y que resulta la más importante desde el punto de vista de la aplicación, presenta un comportamiento altamente deseable, dado que es *casi constante*

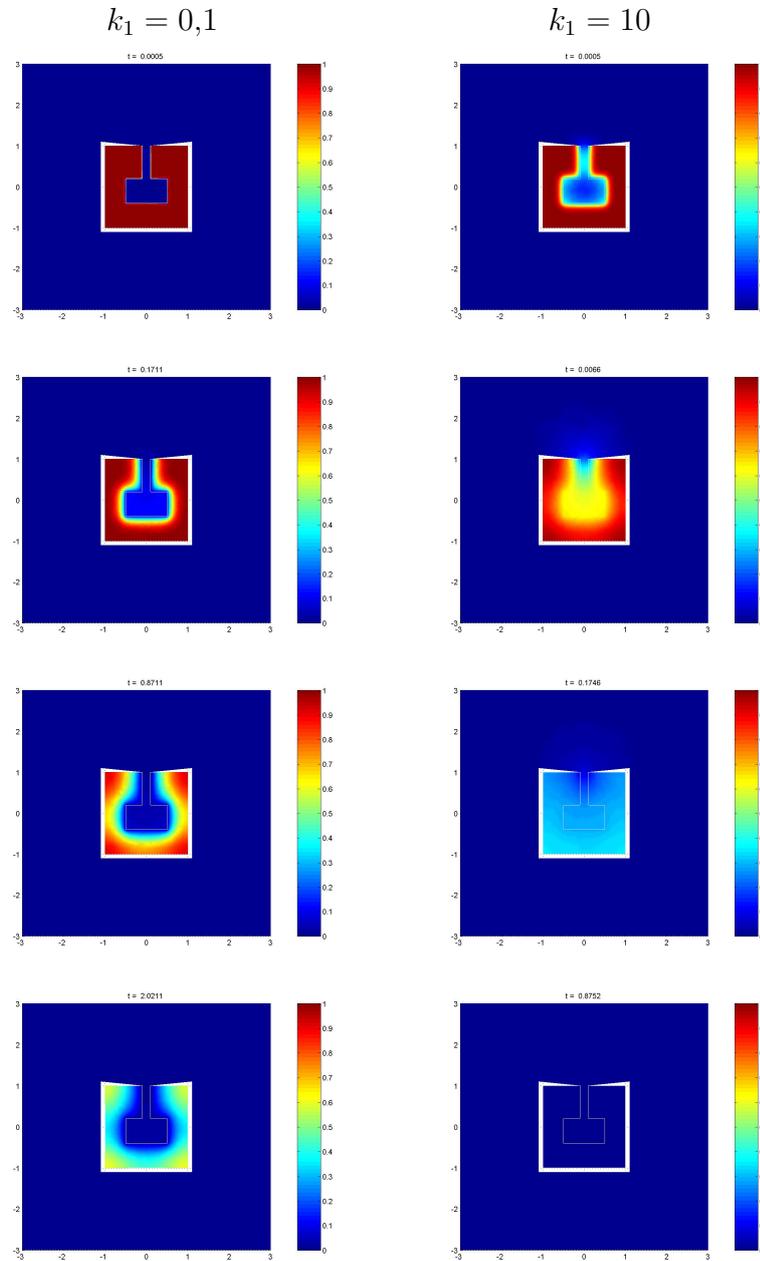


Figura 6.12: Perfil de la concentración de droga para dos valores de k_1 cuando el dispositivo es cuadrado con concavidad y tiene parte aislada en el borde. La primera columna corresponde a $k_1 = 0,1$ y $t = 0,0005$, $t = 0,1711$, $t = 0,8711$ y $t = 2,0211$. La segunda columna corresponde a $k_1 = 10$ y $t = 0,0005$, $t = 0,0066$, $t = 0,1746$ y $t = 0,8752$. Ambas simulaciones fueron realizadas con $k_2 = 100$.

6.4. EJEMPLO 4: CUADRADO CON CONCAVIDAD Y PARTE DEL BORDE AISLADA⁸⁵

luego de un período breve de tiempo, y además no presenta picos que puedan sobrepasar los niveles de toxicidad permitidos.

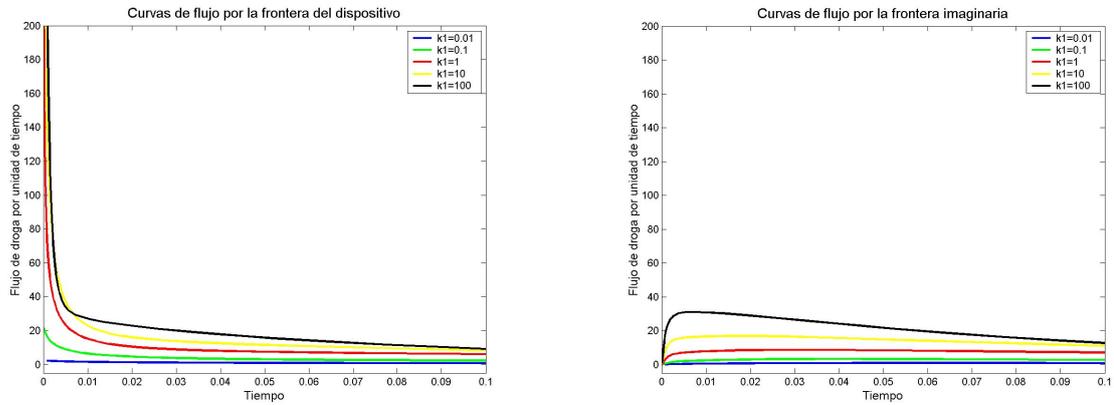


Figura 6.13: Curvas de flujo a través del dispositivo (izquierda) y la frontera imaginaria (derecha) cuando el dispositivo es cuadrado con concauidad y con parte aislada en el borde

Conclusiones

En esta tesis hemos presentado un método de elementos finitos adaptativo para resolver una ecuación de difusión con coeficiente discontinuo.

Esta ecuación surge de una aplicación concreta: el diseño de dispositivos de liberación de droga por difusión. El objetivo de tal diseño es lograr una liberación casi constante de droga por unidad de tiempo. Para lograr una mejor comprensión del fenómeno hemos considerado como dominio computacional al dispositivo junto a parte del medio exterior que lo rodea. De allí resulta una ecuación con coeficiente de difusión discontinuo: un valor para el dispositivo, y otro para el medio exterior.

La adaptatividad se basa en estimadores a posteriori robustos, donde la constante que aparece en la cota superior del error no depende del salto entre el mínimo y el máximo valor del coeficiente de difusión.

El algoritmo fue implementado y utilizado para analizar el efecto de algunos cambios en la forma del dispositivo, sobre el perfil de liberación de droga.

Hemos detallado ejemplos, con diferentes formas de dispositivos y distintas condiciones de borde. A partir de los mismos concluimos que no basta con cambiar la forma del dispositivo, sino que hace falta aislar parte del borde del mismo, de manera que la droga fluya sólo a través de una región. Más aún, a fin de evitar picos indeseables en las curvas de flujo por unidad de tiempo, es conveniente que la parte del borde del dispositivo por donde fluye la droga presente una concavidad. Ésta hace que la cantidad de droga que se libera por unidad de tiempo sea casi constante, y que además no presente picos pronunciados, objetivo altamente deseado desde el punto de vista tecnológico.

Apéndice A

Apéndice

En este capítulo incluimos por completitud definiciones, comentarios, teoremas y demostraciones que consideramos relevantes para el lector de este trabajo y que fueron citadas en los capítulos precedentes.

A.1. Los espacios L^p

Sea Ω un dominio de \mathbb{R}^d abierto y acotado. Denotamos con $L^p(\Omega)$, $1 \leq p < \infty$ al espacio de las funciones $u : \Omega \rightarrow \mathbb{R}$ medibles Lebesgue para las cuales

$$\|u\|_{L^p(\Omega)} := \left(\int_{\Omega} |u(x)|^p \right)^{1/p} < \infty.$$

En el caso especial $p = 2$, denotaremos con $\|\cdot\|$ a la norma $L^2(\Omega)$ y con $\|\cdot\|_S$ a la norma $L^2(S)$, siendo $S \subset \Omega$ medible.

Para $p = \infty$, $L^\infty(\Omega)$ denota al espacio de las funciones medibles esencialmente acotadas en Ω y la norma se define por

$$\|u\|_{L^\infty(\Omega)} := \sup_{x \in \Omega} |u(x)|.$$

En particular, $L^1_{loc}(\Omega) = \{u : \Omega \rightarrow \mathbb{R} / u \in L^1(K), \forall \text{ compacto } K \subset \Omega\}$. Para $1 \leq p \leq \infty$, se cumple $L^p(\Omega) \subset L^1_{loc}(\Omega)$ ([1, Cor. 2.9]).

A.2. Distribuciones y derivada débil

Sea $x \in \mathbb{R}^d$. Para un multiíndice $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$ con componentes α_i enteras no negativas, definimos $D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$ donde $|\alpha| = \alpha_1 + \dots + \alpha_d$.

Sea Ω un abierto y acotado de \mathbb{R}^d . Denotaremos con $C_0^\infty(\Omega)$ al espacio de las funciones infinitamente diferenciales con soporte compacto en Ω y con $\mathcal{D}(\Omega)$ al espacio $C_0^\infty(\Omega)$ donde la convergencia de sucesiones se define de la sgte. manera:

Una sucesión $\{\phi_j\}$ converge a ϕ en $\mathcal{D}(\Omega)$ si $\text{sop}(\phi_j)$ y $\text{sop}(\phi)$ están contenidos en un compacto fijo $K \subset \Omega$ y para todo multiíndice α , $D^\alpha \phi_j \rightarrow D^\alpha \phi$ uniformemente en \mathbb{R}^d .

Las funciones en $\mathcal{D}(\Omega)$ son conocidas como *funciones de prueba*.

Definición A.1 (El dual $\mathcal{D}'(\Omega)$ de $\mathcal{D}(\Omega)$). [1, Cap.1] El espacio $\mathcal{D}'(\Omega)$, dual de $\mathcal{D}(\Omega)$, es el conjunto de funcionales lineales y continuos $T : \mathcal{D}(\Omega) \longrightarrow \mathbb{R}$. A cada funcional de $\mathcal{D}'(\Omega)$, lo llamaremos distribución en Ω .

En $\mathcal{D}(\Omega)$ la continuidad se define en términos de sucesiones. Es decir, diremos que un funcional lineal T es una distribución, cuando

$$T(\phi_j) \longrightarrow 0 \quad \text{si} \quad \phi_j \longrightarrow 0 \text{ en } \mathcal{D}(\Omega).$$

Lema A.2 (Una distribución particular). [1, Sec. 1.53] Toda función $f \in L^1_{loc}(\Omega)$ tiene asociada una distribución $T_f \in \mathcal{D}'(\Omega)$ definida para toda $\phi \in \mathcal{D}(\Omega)$ por $T_f(\phi) = \int_{\Omega} f(x)\phi(x)dx$.

Lema A.3. [1, Sec. 3.26] Si $v \in L^1_{loc}(\Omega)$ y $\int_{\Omega} v \phi \, dx = 0, \quad \forall \phi \in \mathcal{D}(\Omega)$ entonces $v = 0$ en casi todo punto de Ω .

Definición A.4 (Convergencia débil). [2] Sea X un espacio de Banach. Decimos que una sucesión $\{u^n\}$ converge débilmente a $u \in X$ si

$$F(u^n) \longrightarrow F(u), \quad \forall F \in X' \text{ (dual de } X).$$

Definición A.5 (Derivada débil). [1] Sea $u \in L^1_{loc}(\Omega)$ y sea α un multiíndice.

Si existe una función $g \in L^1_{loc}(\Omega)$ que satisfice

$$\int_{\Omega} u(x) D^{\alpha} \phi(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} g \phi(x) \, dx, \quad \forall \phi \in \mathcal{D}(\Omega),$$

entonces g se llama derivada débil de orden α de u y lo denotaremos por $D^{\alpha}u = g$.

Lema A.6 (Derivada de una distribución). [1, Cap. 1] Sea α un multiíndice. La derivada $D^{\alpha}T$ de una distribución $T \in \mathcal{D}'(\Omega)$ se define por

$$D^{\alpha}T(\phi) = (-1)^{|\alpha|} T(D^{\alpha}\phi), \quad \forall \phi \in \mathcal{D}(\Omega).$$

Como $\phi \in \mathcal{D}(\Omega)$ entonces $D^{\alpha}\phi \in \mathcal{D}(\Omega)$ y $D^{\alpha}T$ es un funcional sobre $\mathcal{D}(\Omega)$. Además es lineal y continuo ([1, Sec. 1.55]). Así $D^{\alpha}T \in \mathcal{D}'(\Omega)$.

A.3. Los espacios de Sobolev H^s

Por el momento supongamos que Ω es un abierto arbitrario de \mathbb{R}^d . Sea s un número real que podemos descomponer como $s = m + \sigma$, donde $m \in \mathbb{Z}$ es su parte entera y $0 \leq \sigma < 1$ su parte fraccionaria.

Definición A.7 (El espacio $H^s(\Omega)$, con $s \in \mathbb{Z}$ y $s \geq 0$). [1, Cap. 3]

Sea s un entero no negativo (aquí $\sigma = 0$). Definimos el espacio $H^s(\Omega)$ al conjunto de funciones $u : \Omega \longrightarrow \mathbb{R}$ tales que, para α multiíndice

$$D^{\alpha}u \in L^2(\Omega), \quad |\alpha| \leq s.$$

Si $u \in H^s(\Omega)$, la seminorma y norma se definen respectivamente

$$\|u\|_s := \left(\sum_{|\alpha|=s} \|D^{\alpha}u\|^2 \right)^{1/2}, \quad \|u\|_s := \left(\sum_{|\alpha| \leq s} \|D^{\alpha}u\|^2 \right)^{1/2}.$$

En particular cuando $s = 0$, $H^0(\Omega) = L^2(\Omega)$. Cuando $s = 1$, $H^1(\Omega) = \{v : v \in L^2(\Omega), \nabla v \in L^2(\Omega)\}$ y

$$|v|_1^2 = \int_{\Omega} |\nabla v|^2 dx, \quad \|v\|_1^2 = \|v\|^2 + |v|_1^2.$$

Definición A.8 (El espacio $H^s(\Omega)$, con $s > 0$ y $s \notin \mathbb{N}$). [4, Sec. 1.2], [1, Sec. 7.48] Sea s un número positivo y no entero (aquí $\sigma \neq 0$) y sean $x, y \in \mathbb{R}^d$. El espacio $H^s(\Omega)$ es el conjunto de todas las funciones $u \in H^m(\Omega)$ tal que la seminorma

$$|u|_s := \left(\sum_{|\alpha|=m} \int_{\Omega} \int_{\Omega} \frac{|D^{\alpha}u(x) - D^{\alpha}u(y)|^2}{|x - y|^{d+2\sigma}} dx dy \right)^{1/2} < \infty.$$

La norma correspondiente a este espacio está dada por

$$\|u\|_s := (\|u\|_m + |u|_s^2)^{1/2}.$$

Si U es un subconjunto medible de Ω , denotaremos con $\|\cdot\|_{s,U}$ a la norma $H^s(U)$ con $s > 0$.

Definición A.9 (El espacio $H_0^s(\Omega)$). [4, Sec. 1.2]

Sea $s \in \mathbb{R}$. Definimos al espacio $H_0^s(\Omega)$ como la clausura de $C_0^\infty(\Omega)$ con la norma de $H^s(\Omega)$. Por lo tanto, $C_0^\infty(\Omega)$ es denso en $H_0^s(\Omega)$ (todo elemento de $H_0^s(\Omega)$ es límite de una sucesión de funciones en $C_0^\infty(\Omega)$).

Definición A.10 (El espacio $H^{-s}(\Omega)$). [4, Sec. 1.2] Sea $s > 0$. Denotamos con $H^{-s}(\Omega)$ al espacio dual de $H_0^s(\Omega)$, es decir $(H_0^s(\Omega))'$, cuya norma se define por:

$$\|F\|_{-s} = \sup_{\varphi \in H_0^s(\Omega), \|\varphi\|_s=1} |F(\varphi)| = \sup_{\varphi \in H_0^s(\Omega), \|\varphi\|_s \neq 0} \frac{|F(\varphi)|}{\|\varphi\|_s}$$

para todo $F : H_0^s(\Omega) \rightarrow \mathbb{R}$ lineal y acotado.

Lema A.11. Sea $s > 0$. Si T es una distribución en Ω tal que para alguna constante positiva C se cumple

$$|T(\phi)| \leq C \|\phi\|_s, \quad \forall \phi \in \mathcal{D}(\Omega) \tag{A.1}$$

entonces T se puede extender a $H_0^s(\Omega)$ de manera continua y diremos que $T \in H^{-s}(\Omega)$ (haciendo abuso de notación y llamando T a la extensión).

Demostración. Consideremos $v \in H_0^s(\Omega)$. Por densidad, elijamos una sucesión $\{\phi_n\} \subset C_0^\infty(\Omega)$ tal que $\phi_n \rightarrow v$ en $H_0^s(\Omega)$. Entonces por linealidad de T y por hipótesis:

$$|T(\phi_n) - T(\phi_m)| = |T(\phi_n - \phi_m)| \leq C \|\phi_n - \phi_m\|_s$$

Pero dado que $\{\phi_n\}$ es de Cauchy en $H_0^s(\Omega)$, tenemos que $\{T(\phi_n)\}$ es de Cauchy en \mathbb{R} , que es completo. Esto implica que $\{T(\phi_n)\}$ es convergente.

Definiremos $T(v) := \lim_{n \rightarrow \infty} T(\phi_n)$. Veamos que esta definición de $T(v)$ no depende de ninguna elección particular de sucesión aproximante.

En efecto, si ϕ_n y ψ_n son dos sucesiones de $C_0^\infty(\Omega)$ que convergen a $v \in H_0^s(\Omega)$, se tiene que $\phi_n - \psi_n \rightarrow 0$ en $H_0^s(\Omega)$. Entonces $T(\phi_n - \psi_n) \rightarrow 0$ por (A.1) y así $\lim_{n \rightarrow \infty} T(\phi_n) = \lim_{n \rightarrow \infty} T(\psi_n)$, con lo que la extensión de T a $H_0^s(\Omega)$ está bien definida.

Además, sigue valiendo la acotación (A.1) para todo $v \in H_0^s(\Omega)$, puesto que

$$|T(v)| = \lim |T(\phi_n)| \leq C \lim \|\phi_n\|_s = C \|v\|_s.$$

Luego, T es un funcional lineal y acotado de $H_0^s(\Omega)$, es decir $T \in (H_0^s(\Omega))'$. \square

Observación A.12. Sean $f \in L^2(\Omega)$ y T_f la distribución asociada a f según el Lema A.2. Como $C_0^\infty(\Omega) \subset H_0^s(\Omega)$ para cualquier $s \geq 0$, se cumple

$$|T_f(\phi)| \leq \|f\| \|\phi\| \leq \|f\| \|\phi\|_s$$

para toda $\phi \in \mathcal{D}(\Omega)$. De acuerdo al Lema A.11, esta cota vale también $\forall \phi \in H_0^s(\Omega)$ lo que implica que $T_f \in H^{-s}(\Omega)$. Además

$$|T_f(\phi)| \leq \|\phi\|_s \|T_f\|_{-s} \leq \|\phi\|_s \|f\|, \quad \forall \phi \in \mathcal{D}(\Omega).$$

Muchas propiedades de los espacios de Sobolev definidas en un dominio Ω , en particular las propiedades de inmersión de estos espacios (*imbedding*), dependen de propiedades de regularidad de Ω . Tal regularidad normalmente se expresa en términos de condiciones geométricas que puede o no satisfacer el dominio. A continuación nombramos dos importantes llamadas *condición de cono* y *condición Lipschitz* en la frontera.

Definición A.13 (Condición de cono interior). [1, Sec. 4.3], [6, Cap. 2, Sec. 1.9] Diremos que un dominio $\Omega \subset \mathbb{R}^d$ satisface la condición de cono, si cada punto x de Ω es el vértice de algún cono finito C_x contenido en Ω .

En el caso de dominios poligonales, esta definición es equivalente a la siguiente:

Definición A.14 (Condición de cono interior). Diremos que un dominio $\Omega \subset \mathbb{R}^d$ satisface la condición de cono, si los ángulos interiores a cada vértice de Ω son positivos y así un cono no trivial se puede posicionar en Ω con su punta en el vértice.

Por lo tanto, todo dominio poligonal satisface la condición de cono interior.

Definición A.15 (Dominio Lipschitz). [6, Cap. 2, Sec. 1.9] Un dominio acotado $\Omega \subset \mathbb{R}^d$ se llama dominio Lipschitz si cumple que $\forall x \in \partial\Omega$, existe un entorno U_x tal que $\partial\Omega \cap U_x$ es el gráfico de una función continua Lipschitz.

En otros libros la definición de dominio Lipschitz es diferente, pero es equivalente a la dada cuando consideramos solamente dominios poligonales.

En la Figura A.1 ilustramos dos ejemplos de dominios poligonales no Lipschitz que satisfacen la condición de cono.

En esta tesis, dado que pediremos que el dominio Ω sea acotado, poligonal y Lipschitz, siempre cumplirá la condición de cono.

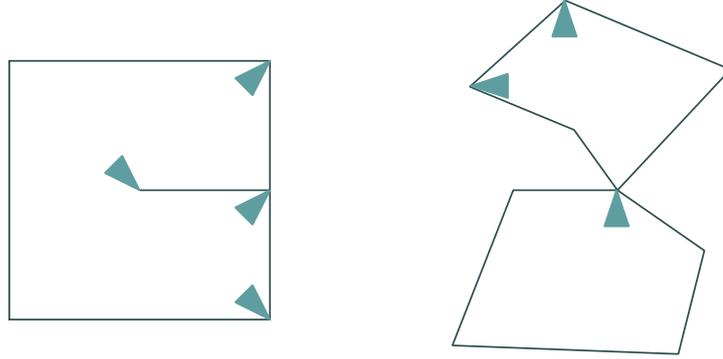


Figura A.1: Dominios no Lipschitz: estos dominios no son Lipschitz, pero sí satisfacen la condición de cono interior.

Definición A.16 (Inmersión continua y compacta). [1, Sec. 1.24] Sean $X \subset Y$ dos espacios normados, con normas respectivamente $\|\cdot\|_X$ y $\|\cdot\|_Y$. Diremos que X está inmerso de manera continua en Y y lo simbolizamos $X \hookrightarrow Y$ si:

$$\exists M \text{ tal que } \|u\|_Y \leq M\|u\|_X, \quad \forall u \in X.$$

Diremos que X está inmerso de manera compacta en Y y lo escribimos $X \hookrightarrow\hookrightarrow Y$ si

- i) $X \hookrightarrow Y$
- ii) toda sucesión acotada en X tiene una subsucesión que converge en Y .

En el siguiente teorema sólo enunciamos las inmersiones más relevantes para esta tesis, que fueron resumidas de [4, Sec. 1.2], [6, Sec. 1.9] y [1, Teo. 6.2].

Teorema A.17 (Teorema de inmersión de Sobolev). Sea Ω un subconjunto abierto de \mathbb{R}^d . Entonces valen las siguientes inmersiones siempre que se cumplan las respectivas restricciones en Ω :

1. Para todo Ω : $H_0^s(\Omega) \hookrightarrow H^s(\Omega) \hookrightarrow L^2(\Omega)$, con $s > 0$.
2. Si Ω es acotado y Lipschitz:
 $H^s(\Omega) \hookrightarrow C^k(\Omega)$, para cualquier entero no negativo $k < s - \frac{d}{2}$.
3. Si Ω es Lipschitz y satisface la condición de cono:
 $H^s(\Omega) \hookrightarrow\hookrightarrow H^t(\Omega)$ para $s > t \geq 0$.
4. Si Ω es acotado y Lipschitz:
 $H^{j+s}(\Omega) \hookrightarrow\hookrightarrow C^j(\overline{\Omega})$ para $j \in \mathbb{N}_0$ y $s \in \mathbb{R}$, $s > \frac{d}{2}$.

A.4. Sobre Elementos Finitos

Teorema A.18 (Integración por partes o Fórmula de Green). *Consideremos Ω y $k(x)$ con las hipótesis descritas en la primera sección del Capítulo 2. Entonces para todo $u, v \in H^1(\Omega)$, por integración por partes en cada $\Omega_i, i = 1, 2$ se tiene:*

$$\int_{\Omega_i} \operatorname{div}(k(x)\nabla u)v \, dx = - \int_{\Omega_i} k(x)\nabla u \cdot \nabla v \, dx + \int_{\partial\Omega_i} k(x)v \frac{\partial u}{\partial n_i} \, d\sigma,$$

donde $\frac{\partial u}{\partial n_i} := \nabla u \cdot n_i$ denota la derivada de u en la dirección n_i unitaria normal exterior a $\partial\Omega_i$.

Teorema A.19 (Unicidad de solución en Sistemas de EDO). *[3] Sean $y_0 \in \mathbb{R}^n$ y $h : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$. Supongamos que h es Lipschitz continua en la variable x y que para cada $R > 0$ existe una función $m_R(t) \in L^1(0, T)$ tal que*

$$|h(t, x)| \leq m_R(t), \quad \forall |x| < R.$$

Entonces existe una única función absolutamente continua $y : [0, T] \rightarrow \mathbb{R}^n$ que satisface

$$\begin{cases} y'(t) = h(t, y(t)), & \text{en casi todo } 0 \leq t \leq T \\ y(0) = y_0. \end{cases}$$

A.4.1. Triangulaciones del dominio

Este apartado es una recopilación de [11, Cap. 3], [6, Cap. 5] y [7, Cap. 4,7].

Sea Ω un dominio poligonal de \mathbb{R}^2 y consideremos una partición *admisibile* \mathcal{T} de Ω en triángulos T que llamaremos *elementos*. *Admisibile* indica que $\bar{\Omega} = \bigcup_{T \in \mathcal{T}} T$ y que la intersección entre dos triángulos de \mathcal{T} es vacía, o bien un vértice común o un lado común.

Definimos para cada elemento $T \in \mathcal{T}$:

$$h_T = \operatorname{diam}(T) = \max\{|x - y|, x, y \in T\} \text{ y}$$

$\rho_T =$ diámetro de la mayor circunferencia contenida en T .

A partir de ahora, denotaremos con \mathcal{T}_h a la triangulación o partición cuyo valor máximo entre todos los diámetros de los elementos es h , esto es, $h = \max_{T \in \mathcal{T}_h} h_T$.

Es trivial que $\rho_T \leq h_T \leq h$.

Definición A.20 (Partición regular y Familia regular de particiones). *Una partición \mathcal{T}_h en Ω se dice regular si existe una constante $\kappa > 0$ tal que*

$$\max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T} \leq \kappa.$$

Diremos que una familia $\{\mathcal{T}_h\}$ de particiones es regular, si el $\max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T}$ está uniformemente acotado para cada partición de la familia, es decir:

$$\max_{\mathcal{T}_h \in \{\mathcal{T}_h\}} \max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T} \leq \kappa. \quad (\text{A.2})$$

En la literatura del método de Elementos Finitos, esta propiedad se menciona frecuentemente como *shape regularity* de una familia de particiones.

Definición A.21 (Condición del ángulo mínimo). Sea $\{\mathcal{T}_h\}$ una familia de triangulaciones de Ω . Diremos que $\{\mathcal{T}_h\}$ satisface la condición del ángulo mínimo si existe una constante $\theta_0 > 0$ tal que

$$\inf_{\mathcal{T}_h} \min_{T \in \mathcal{T}_h} \theta_T \geq \theta_0 \quad (\text{A.3})$$

donde θ_T es el mínimo ángulo interior de T .

Se puede probar que la condición (A.2) es equivalente a la condición (A.3) y que la constante θ_0 depende de la constante de regularidad κ y viceversa.

Definición A.22 (Familia uniforme de particiones). Una sucesión de particiones $\{\mathcal{T}_h\}$ se llama uniforme, si existe una constante $\sigma > 0$ tal que

$$\frac{h}{\rho_T} \leq \sigma, \quad \forall T \in \mathcal{T}_h \text{ y } \forall \mathcal{T}_h \in \{\mathcal{T}_h\}. \quad (\text{A.4})$$

De las definiciones se desprende que una familia de particiones uniforme es también regular pero no viceversa, es decir la condición de uniformidad es más fuerte que la de regularidad.

La condición anterior (A.4) asegura que todos los triángulos T involucrados en la partición \mathcal{T}_h tienen su diámetro h_T del orden de h y en consecuencia su área (que denotamos con $|T|$) del orden de h^2 . Esto se debe a que es posible encontrar constantes α y β para las que $\alpha h^2 \leq |T| \leq \beta h^2$, $\forall T \in \mathcal{T}_h$ y $\forall \mathcal{T}_h$, además de cumplirse la condición del ángulo mínimo.

A.4.2. Un mapeo afín y algunas propiedades

Sea $\Omega \subset \mathbb{R}^2$ y consideremos una partición \mathcal{T}_h de Ω de elementos triangulares. Definimos el subespacio de elementos finitos V_h lineales a trozos asociado con la partición \mathcal{T}_h por:

$$V_h = \{v \in \mathcal{C}(\bar{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h\}.$$

En la implementación y el análisis del método de Elementos Finitos juega un papel importante el uso del *elemento de referencia*, que denotaremos con T_R y se define por el triángulo

$$T_R = \{\hat{z} = (\hat{x}, \hat{y}) / 0 \leq \hat{x} \leq 1, 0 \leq \hat{y} \leq 1 - \hat{x}\}.$$

Cualquier triángulo T de la partición con nodos $z_i = (x_i, y_i)$, $i = 0, 1, 2$ puede pensarse como la imagen del triángulo de referencia bajo la transformación $F_T : T_R \rightarrow T$ afín y biyectiva dada por

$$F_T(\hat{z}) = z_0 + A_T \hat{z} = z$$

donde A_T es la matriz invertible cuyas columnas son $z_i - z_0$, $i = 1, 2$ y z_0 es el nodo de T trasladado del $(0, 0)$ a través de F_T .

Propiedades

Los mapeos F_T y F_T^{-1} son utilizados como una herramienta estándar en los llamados *argumentos de escala*, en los que se tienen en cuenta ciertas propiedades. Mencionaremos sólo aquellas propiedades que necesitaremos en algún momento y que se deducen simplemente de la definición o se demuestran en [6, Cap. 6]:

- $F_T(\hat{z}_i) = z_i$, con \hat{z}_i nodo de T_R .
- Sea $S \subset \Omega$ y $\hat{S} = F_T^{-1}(S)$. Para cualquier $v \in H^1(S)$, definimos $\hat{v} : \hat{S} \rightarrow \mathbb{R}$ por $\hat{v}(\hat{z}) = (v \circ F_T)(\hat{z}) = v(z_0 + A_T \hat{z}) = v(z)$. Luego $\hat{v} \in H^1(\hat{S})$.
- $|T| = |\det A_T| \cdot |T_R| = |\det A_T|/2$
- $\det \mathcal{J}(F_T) = \det A_T$, donde $\mathcal{J}(F_T)$ es el Jacobiano de F_T
- $\|A_T\| \leq \frac{h_T}{\rho_{T_R}}$ y $\|A_T^{-1}\| \leq \frac{h_{T_R}}{\rho_T}$, donde $\|\cdot\|$ es una norma matricial
- $\hat{\nabla} \hat{u} = A_T^t \nabla u$ y $|\hat{\nabla} \hat{u}| \leq \|A_T\| |\nabla u|$
- Si $\{\mathcal{T}_h\}$ es una familia regular de particiones, existe una constante $c_1 = c_1(\kappa)$ tal que

$$\max_{T \in \mathcal{T}_h} \|A_T\| \|A_T^{-1}\| \leq c_1, \quad \forall \mathcal{T}_h$$

A.5. El problema elíptico con coeficiente discontinuo

El interés de esta sección es presentar las estimaciones a priori del error por Elementos Finitos correspondientes al problema (2.1) en estado estacionario, es decir no dependiente del tiempo. Estas estimaciones se utilizan para establecer las estimaciones del error para las soluciones por Elementos Finitos del problema parabólico (2.1). Para ello, consideramos necesario hacer una breve síntesis de lemas y teoremas (donde muchos de ellos demostraremos) que van a ir conduciendo al Teorema principal A.31 que demuestra las acotaciones del error que nos interesan, en distintas normas.

A.5.1. Forma continua y débil

Supongamos que valen las mismas hipótesis que en la sección 2.1 acerca del dominio Ω y el coeficiente de difusión $k(x)$ y consideremos $f(x) \in L^2(\Omega)$.

El problema (2.1) independiente del tiempo (elíptico), con condiciones de Dirichlet tiene la forma

$$\begin{cases} -\operatorname{div}(k(x)\nabla u) = f(x) & \text{en } \Omega_1 \cup \Omega_2, \\ u(x) = 0 & \text{en } \partial\Omega, \\ u_1 = u_2 & \text{en } \Gamma, \\ -k_1 \nabla u_1 \cdot n_1 = k_2 \nabla u_2 \cdot n_2 & \text{en } \Gamma. \end{cases}$$

Sea $V = H_0^1(\Omega)$. La formulación débil de este problema consiste en encontrar $u \in V$ tal que

$$\int_{\Omega} k(x) \nabla u(x) \nabla v(x) \, dx = \int_{\Omega} f(x) v(x) \, dx, \quad \forall v \in V.$$

o equivalentemente

$$a(u, v) = (f, v), \quad \forall v \in V, \tag{A.5}$$

donde a es la forma bilineal definida en (2.6) y (\cdot, \cdot) es el producto usual en $L^2(\Omega)$, ambos independientes del tiempo.

A.5.2. Lema de Céa

Sea V_h el subespacio finito de V que consiste en el conjunto de funciones continuas lineales a trozos sobre una triangulación \mathcal{T}_h del dominio y que se anulan en $\partial\Omega$, es decir,

$$V_h = \{v \in \mathcal{C}(\bar{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1(T), \forall T \in \mathcal{T}_h\}.$$

Propiedad de ortogonalidad de Galerkin

Consideremos $u \in V$ solución de (A.5) y $u_h \in V_h$ solución del análogo discreto cuya forma es

$$u_h \in V_h : a(u_h, v) = (f, v), \quad \forall v \in V_h. \quad (\text{A.6})$$

Es fácil probar la siguiente propiedad (conocida como de *ortogonalidad de Galerkin*):

$$a(u - u_h, v) = 0, \quad \forall v \in V_h, \quad (\text{A.7})$$

que indica que el error $u - u_h$ es ortogonal a V_h con respecto al producto escalar energía.

El Lema de Céa es de gran importancia para establecer las cotas para el error por Elementos Finitos en problemas elípticos. La desigualdad del lema está expresada en norma energía y su demostración es muy similar a la que se encuentra en [6, Cap. 2, Sec. 4].

Lema A.23 (Lema de Céa). *Sean $u \in V$ solución de (A.5) y $u_h \in V_h$ solución de (A.6). Entonces*

$$\|u - u_h\|_a \leq \frac{\beta}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_a,$$

donde β y α son las constantes de continuidad y coercividad respectivamente de la forma bilineal $a(\cdot, \cdot)$ que induce la norma $\|\cdot\|_a$. (ver (2.7) a (2.9)).

A.5.3. Regularidad y Estabilidad

El Teorema A.24 es el principal resultado de la tesis doctoral de M. Petzoldt [8] donde se demuestra que la regularidad de la solución del problema débil elíptico (A.5) sólo depende de las cotas del coeficiente k .

Teorema A.24. *Supongamos que $\delta < k(x) < \delta^{-1}, \forall x \in \Omega$ y para un número $0 < \delta < 1$. Entonces la solución del problema (A.5) independiente del tiempo tiene regularidad*

$$u \in H^{1+\delta/(2\pi)}(\Omega).$$

Demostración. Se puede encontrar en [8, Sec. 2.6.2]. □

Diremos que la solución de (A.5) es estable en el sentido que se cumple la desigualdad del siguiente lema:

Lema A.25 (Estabilidad). *Supongamos que $\delta < k(x) < \delta^{-1}$ para todo $x \in \Omega$ y para algún $0 < \delta < 1$ y sea $u \in H_0^1(\Omega)$ la solución del problema (A.5), con $f \in L^2(\Omega)$.*

Entonces $u \in H^{1+r}(\Omega)$ para $r = \frac{\delta}{2\pi}$ y además existe una constante $C = C(k, r)$, independiente de f y de u , tal que

$$\|u\|_{1+r} \leq C \|f\|.$$

A la demostración del lema de estabilidad la desarrollaremos al final de este apartado, dado que antes necesitamos de tres resultados intermedios (lemas A.26, A.27 y A.28) que presentamos a continuación.

Lema A.26. *[4, Teo. 1.2.17] Sea Ω un abierto Lipschitz y acotado de \mathfrak{R}^d y sea $s > 0$ y $s \neq 1/2$. Entonces $\frac{\partial}{\partial x_i}$, $i = 1, \dots, d$ es un operador lineal continuo de $H^s(\Omega)$ en $H^{s-1}(\Omega)$.*

En otras palabras, si $u \in H^s(\Omega)$ con $s > 0$ y $s \neq 1/2$ entonces $\frac{\partial}{\partial x_i} T u = T \frac{\partial u}{\partial x_i} \in H^{s-1}(\Omega)$ y además existe una constante \tilde{C} , independiente de u , tal que $\|\frac{\partial}{\partial x_i} T u\|_{s-1} \leq \tilde{C} \|u\|_s$.

Lema A.27. *Sean $0 < r < 1/2$ y $u \in H_0^{1+r}(\Omega)$. Definimos $T : \mathcal{D}(\Omega) \rightarrow \mathfrak{R}$ como*

$$T(\phi) = \int_{\Omega} \nabla u \cdot \nabla \phi \, dx, \quad \forall \phi \in \mathcal{D}(\Omega).$$

Entonces T se puede extender a todo el espacio $H_0^{1-r}(\Omega)$ y así $T \in (H_0^{1-r}(\Omega))' = H^{-1+r}(\Omega)$. También $\|T\|_{-(1-r)} \leq \tilde{C} \|u\|_{1+r}$, $\forall u \in H_0^{1+r}(\Omega)$, donde \tilde{C} es una constante proveniente del Lema A.26.

Demostración. Consideremos $u \in H_0^{1+r}(\Omega)$. Mostraremos primero que T satisface la cota

$$|T(\phi)| \leq C \|\phi\|_{1-r}, \quad \forall \phi \in \mathcal{D}(\Omega) \tag{A.8}$$

para una cierta constante $C > 0$ y luego aplicaremos el Lema A.11.

En efecto, para cualquier $\phi \in \mathcal{D}(\Omega)$ haciendo integración por partes

$$T(\phi) = \int_{\Omega} \nabla u \cdot \nabla \phi \, dx = \int_{\Omega} u(-\Delta \phi) \, dx := F_{-\Delta \phi}(u). \tag{A.9}$$

Estamos interpretando aquí $F_{-\Delta \phi} : H_0^{1+r}(\Omega) \rightarrow \mathfrak{R}$ como un mapeo lineal y acotado de $H_0^{1+r}(\Omega)$ puesto que $|F_{-\Delta \phi}(u)| \leq \|-\Delta \phi\| \|u\| \leq \|-\Delta \phi\| \|u\|_{1+r}$.

Por definición de $\|\cdot\|_{-(1+r)}$:

$$|F_{-\Delta \phi}(u)| \leq \|u\|_{1+r} \|F_{-\Delta \phi}\|_{-(1+r)}. \tag{A.10}$$

Por otro lado, usando dos veces el Lema A.26

$$\|F_{-\Delta \phi}\|_{-(1+r)} \leq \tilde{C} \|\phi\|_{1-r}. \tag{A.11}$$

Combinando (A.9), (A.10) y (A.11) llegamos a la cota (A.8) que queríamos probar con $C = \tilde{C} \|u\|_{1+r}$.

Como consecuencia del Lema A.11, T se puede extender de manera continua a $H_0^{1-r}(\Omega)$, lo que implica que la cota (A.8) vale para toda $\phi \in H_0^{1-r}(\Omega)$ y por consiguiente $T \in (H_0^{1-r}(\Omega))'$.

Además, de la cota (A.8), $\|T\|_{-(1-r)} \leq \tilde{C} \|u\|_{1+r}$, con lo que el lema queda demostrado. \square

Lema A.28. Sean $|r| < 1/2$ y F un funcional en el dual de $H_0^{1-r}(\Omega)$, es decir $F \in H^{-1+r}(\Omega)$. Entonces existe una única $u \in H_0^{1+r}(\Omega)$ que es solución del problema

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = F(v), \quad \forall v \in \mathcal{D}(\Omega) \quad (\text{A.12})$$

y que para alguna constante C_r , independiente de F y de u , pero dependiente de r , satisface

$$\|u\|_{1+r} \leq C_r \|F\|_{-(1-r)}.$$

Demostración. Se puede encontrar en [13]. □

Demostración. Del Lema de Estabilidad A.25

La regularidad global $u \in H_0^{1+r}(\Omega)$ se cumple para $r = \frac{\delta}{2\pi} < 1/2$, por el Teorema A.24.

Para probar la desigualdad, consideremos la distribución T definida para cualquier $\phi \in \mathcal{D}(\Omega)$ como $T(\phi) = \int_{\Omega} \nabla u \cdot \nabla \phi \, dx$, donde u es la solución de la forma débil (A.5).

Por el Lema A.27, notemos que $T \in (H_0^{1-r}(\Omega))'$. Entonces, u coincide con la solución de (A.12) como consecuencia del Lema A.28, y satisface

$$\|u\|_{1+r} \leq C_r \|T\|_{-(1-r)}. \quad (\text{A.13})$$

Para acotar $\|T\|_{-(1-r)}$, tomamos una $\phi \in \mathcal{D}(\Omega)$ y vemos que

$$\begin{aligned} T(\phi) &= \int_{\Omega} \nabla u \cdot \nabla \phi \, dx \leq \frac{1}{\min k(x)} \int_{\Omega} k(x) \nabla u \cdot \nabla \phi \, dx \\ &= \frac{1}{\min k(x)} \int_{\Omega} f \phi \, dx \leq \frac{1}{\min k(x)} \|f\| \|\phi\|_{1-r}. \end{aligned}$$

La última desigualdad es debido a que $H_0^{1-r}(\Omega) \hookrightarrow L^2(\Omega)$.

De aquí que

$$\|T\|_{-(1-r)} \leq \frac{1}{\min k(x)} \|f\|. \quad (\text{A.14})$$

Finalmente, de (A.13) y (A.14) llegamos a la cota final del lema de estabilidad. □

A.5.4. Estimación del error

Como sabemos $\Omega \subset \mathfrak{R}^2$ es acotado y Lipschitz, entonces debido a resultados de inmersión de Sobolev (Teorema A.17), una función $u \in H^{1+r}(\Omega)$ con $r > 0$ es continua en $\overline{\Omega}$. De esta manera, podemos definir el operador de interpolación $I : H^{1+r}(\Omega) \rightarrow V_h$ de manera que Iu sea lineal a trozos sobre una triangulación \mathcal{T}_h de Ω y cuyo valor coincide con u sobre los nodos, es decir, $Iu(z_i) = u(z_i)$, siendo z_i un vértice de $T \in \mathcal{T}_h$.

La acotación dada en el siguiente lema, representa la estimación del error de interpolación mediante polinomios lineales y será de utilidad para aplicarlo en el Teorema A.30 sobre el triángulo de referencia. La demostración de este lema es una adaptación de la que se encuentra en [6, Lema 6.2].

Lema A.29. Sea $T \in \mathcal{T}_h$ y sea $u \in H^{1+r}(T)$ con $r > 0$. Entonces existe una constante C_T , independiente de u pero que sólo depende de T , tal que

$$\|u - Iu\|_{1+r,T} \leq C_T \|u\|_{1+r,T}.$$

Aquí, Iu es lineal por ser la restricción del operador I sobre un triángulo T .

Demostración. Definiremos una nueva norma en $H^{1+r}(\Omega)$ dada para cada $v \in H^{1+r}(\Omega)$ por

$$|||v||| := |v|_{1+r,T} + \sum_{i=0}^2 |v(z_i)|.$$

Al final probaremos que esta nueva norma y $\|\cdot\|_{1+r}$ son equivalentes. Si esta equivalencia fuera cierta, entonces para alguna constante C_1 :

$$\begin{aligned} \|u - Iu\|_{1+r,T} &\leq C_1 |||u - Iu||| = C_1 (|u - Iu|_{1+r,T} + \sum_{i=0}^2 |(u - Iu)(z_i)|) \\ &= C_1 |u - Iu|_{1+r,T} = C_1 |u|_{1+r,T}, \end{aligned}$$

debido a que Iu coincide con u en los nodos z_i de T y además $|Iu|_{1+r,T} = 0$ puesto que $Iu|_T \in P_1$. Entonces la desigualdad del lema queda probada con $C_T = C_1$.

Nos resta mostrar la equivalencia de las normas. Ya que $H^{1+r}(T) \hookrightarrow \mathcal{C}(T)$,

$$|u(z_i)| \leq \max_{x \in T} |u(x)| \leq C \|u\|_{1+r,T}$$

y por definición de $||\cdot|||$ tenemos $|u|_{1+r,T} \leq |||u|||$. De aquí que

$$|||u||| = |u|_{1+r,T} + \sum_{i=0}^2 |u(z_i)| \leq C_2 \|u\|_{1+r,T}$$

con $C_2 = 1 + 3C$.

Para demostrar que existe una constante C_1 tal que $|||u||| \geq C_1 \|u\|_{1+r,T}$ lo haremos por el absurdo. Supongamos entonces que para todo $n \in \mathbb{N}$ existe $\{u_n\} \subset H^{1+r}(T)$ con $\|u_n\|_{1+r,T} = 1$ y $|||u_n||| < \frac{1}{n}$.

Como $\{u_n\}$ es acotada en $H^{1+r}(T)$, por el Teorema A.17 existe una subsucesión $\{u_{n_k}\}$ convergente en $H^1(T)$.

Por otro lado,

$$\begin{aligned} \|u_{n_k} - u_{n_\ell}\|_{1+r,T}^2 &= \|u_{n_k} - u_{n_\ell}\|_{1,T}^2 + |u_{n_k} - u_{n_\ell}|_{1+r,T}^2 \\ &\leq \|u_{n_k} - u_{n_\ell}\|_{1,T}^2 + (|u_{n_k}|_{1+r,T} + |u_{n_\ell}|_{1+r,T})^2 \\ &\leq \|u_{n_k} - u_{n_\ell}\|_{1,T}^2 + \left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)^2 \end{aligned}$$

Cuando $k, \ell \rightarrow \infty$ los dos términos de la última desigualdad tienden a cero, con lo que $\{u_{n_k}\}$ es de Cauchy en H^{1+r} . Debido a la completitud de este espacio, la sucesión converge necesariamente a $u^* \in H^{1+r}(T)$. Luego, $\|u^*\|_{1+r,T} = 1$.

Pero por otro lado, $|||u^*||| = 0$ lo que implica que $u^* = 0$ y ésto es absurdo. \square

Cuando las funciones se aproximan por polinomios a trozos sobre una malla uniforme de tamaño h , la cota del error involucra a una potencia positiva de h . Dicha potencia depende de la regularidad de la solución y del grado del polinomio con que se aproxima.

Un resultado muy conocido es que si la solución u de (A.5) pertenece a $H^2(\Omega)$ (es el caso cuando $k(x)$ es suave y Ω es suave o poligonal convexo) e Iu es su interpolante lineal a trozos, entonces existe C que depende del ángulo mínimo tal que

$$\|u - Iu\| \leq C h^2 |u|_2.$$

Esta cota no es cierta si u tiene regularidad menor, es decir cuando $u \in H_0^{1+r}(\Omega)$ para algún $0 < r < 1/2$. En este caso, el resultado central está dado por el Teorema A.30, en donde se observa que la potencia de h decrece en uno al aplicar una derivada al error.

Este teorema será la herramienta teórica fundamental para obtener las estimaciones del error por elementos finitos del problema elíptico. La demostración utiliza permanentemente argumentos de escala, que recordamos, se basan en el uso de un mapeo que permite trabajar sobre un triángulo más sencillo (ver mapeo afín y sus propiedades en sección A.4.2).

Teorema A.30 (Estimación del error del interpolación). *Sean Ω un dominio acotado y Lipschitz de \mathbb{R}^2 con borde poligonal y $\{\mathcal{T}_h\}$ una familia uniforme de particiones de Ω . Si $u \in H^{1+r}(\Omega)$ para algún $0 < r < 1/2$ es solución del problema (A.5), entonces existe una constante $C_I = C_I(\Omega, \sigma)$, independiente de u , tal que:*

$$i) \|u - Iu\| \leq C_I h^{1+r} |u|_{1+r}$$

$$ii) \|\nabla(u - Iu)\| \leq C_I h^r |u|_{1+r}$$

La constante σ es la correspondiente a la Definición (A.4).

Demostración. En [12] se puede encontrar la demostración completa de la acotación (i). Sacaremos de ella algunas ideas para demostrar (ii).

A través del mapeo biyectivo $F_T : T_R \rightarrow T$ que involucra a la matriz invertible A_T y usando el Lema A.29 para $\hat{u} = F_T^{-1}(u) \in H^{1+r}(T_R)$, se tiene:

$$\begin{aligned} \|\nabla(u - Iu)\|_T^2 &= \int_T |\nabla(u - Iu)|^2 dx \leq |\det A_T| \|A_T^{-1}\|^2 \|\hat{\nabla}(\hat{u} - \hat{I}\hat{u})\|_{T_R}^2 \\ &\leq |\det A_T| \|A_T^{-1}\|^2 C_{T_R} |\hat{u}|_{1+r, T_R}^2. \end{aligned} \quad (\text{A.15})$$

Trataremos de acotar $|\hat{u}|_{1+r, T_R}^2$ usando argumentos de escala y algunas propiedades del mapeo F_T (sección A.4.2):

$$\begin{aligned} |\hat{u}|_{1+r, T_R}^2 &= \int_{T_R} \int_{T_R} \frac{|\hat{\nabla}\hat{u}(\hat{x}) - \hat{\nabla}\hat{u}(\hat{y})|^2}{|\hat{x} - \hat{y}|^{2+2r}} d\hat{x}d\hat{y} \\ &= |\det A_T|^{-2} \int_T \int_T \frac{|A_T^t(\nabla u(x) - \nabla u(y))|^2}{|x - y|^{2+2r}} \left(\frac{|x - y|}{|A_T^{-1}(x - y)|} \right)^{2+2r} dx dy \\ &\leq \frac{\|A_T\|^2}{|\det A_T|^2} \|A_T\|^{2+2r} \int_T \int_T \frac{|\nabla u(x) - \nabla u(y)|^2}{|x - y|^{2+2r}} dx dy. \end{aligned}$$

Al retomar (A.15) obtenemos:

$$\|\nabla(u - Iu)\|_T^2 \leq C_{T_R} \left(\frac{\|A_T\|^2}{|\det A_T|} \right) (\|A_T\| \|A_T^{-1}\|)^2 \|A_T\|^{2r} \int_T \int_T \frac{|\nabla u(x) - \nabla u(y)|^2}{|x - y|^{2+2r}} dx dy.$$

Como se mencionó en la sección A.4.2, $\|A_T\|^{2r} \leq \left(\frac{h}{\rho_{T_R}}\right)^{2r}$ y debido a que la sucesión de particiones es uniforme, podemos decir que, $\|A_T\| \|A_T^{-1}\| \leq c_1(\sigma)$ y como $\frac{h^2}{|T|} \leq 2\sigma^2$, entonces $\frac{\|A_T\|^2}{|\det A_T|} \leq c_2(\sigma)$.

Con lo que acabamos de decir, podemos acotar de la siguiente forma:

$$\|\nabla(u - Iu)\|_T^2 \leq C_{T_R} c_2 c_1^2 \frac{h^{2r}}{\rho_{T_R}^{2r}} \int_T \int_T \frac{|\nabla u(x) - \nabla u(y)|^2}{|x - y|^{2+2r}} dx dy = C_I^2 h^{2r} |u|_{1+r,T}^2,$$

denotando $C_I = \left(C_{T_R} \frac{c_2 c_1^2}{\rho_{T_R}^{2r}}\right)^{1/2}$.

Con las contribuciones de $\|\nabla(u - Iu)\|$ sobre cada T podremos acotar la misma norma sobre todo el dominio y así culminamos la demostración. En efecto, llamando $\phi(x, y)$ al integrando de $|u|_{1+r}^2$:

$$\begin{aligned} \|\nabla(u - Iu)\|^2 &= \sum_{T \in \mathcal{T}_h} \|\nabla(u - Iu)\|_T^2 \leq C_I^2 h^{2r} \sum_{T \in \mathcal{T}_h} |u|_{1+r,T}^2 \\ &= C_I^2 h^{2r} \sum_{T \in \mathcal{T}_h} \int_T \int_T \phi(x, y) dx dy \leq C_I^2 h^{2r} \sum_{T \in \mathcal{T}_h} \int_\Omega \int_T \phi(x, y) dx dy \\ &= C_I^2 h^{2r} \int_\Omega \sum_{T \in \mathcal{T}_h} \int_T \phi(x, y) dx dy = C_I^2 h^{2r} \int_\Omega \int_\Omega \phi(x, y) dx dy \\ &= C_I^2 h^{2r} |u|_{1+r}^2. \end{aligned}$$

□

Teorema principal

El Teorema A.31 establece las cotas para el error por el método de Elementos Finitos, medido en seminorma H^1 y en norma L^2 . En la última, la demostración requiere de un argumento de dualidad muy conocido como el *truco de Aubin-Nitsche*.

Este teorema es de fundamental importancia para poder obtener las estimaciones a priori del error en las semidiscretizaciones del problema parabólico que estamos estudiando en esta tesis (capítulo 3).

Teorema A.31 (Estimación del error por Elementos Finitos). *Supongamos valen las mismas hipótesis que en la sección 2.1 acerca del dominio Ω y el coeficiente de difusión $k(x)$ y supongamos $\delta < k < \delta^{-1}$ para algún $0 < \delta < 1$. Para $f(x) \in L^2(\Omega)$, sea $u \in H_0^1(\Omega)$ la solución del problema (A.5) y $u_h \in V_h$ la solución discreta de (A.6) asociada a una triangulación uniforme de tamaño h .*

Entonces existen constantes C_1 y C_2 independientes de u , que sólo dependen de k , Ω y la constante σ , tales que:

$$i) |u - u_h|_1 \leq C_1 h^r |u|_{1+r}$$

$$ii) \|u - u_h\| \leq C_2 h^{2r} |u|_{1+r}$$

Demostración. Observemos primero que por el Teorema A.24 y el Lema A.25, $u \in H^{1+r}(\Omega)$ y $|u|_{1+r} \leq C \|f\|$, con $r = \frac{\delta}{2\pi}$.

Para probar (i), usamos la definición y propiedades de la norma energía (sección 2.2.2), el Lema de Céa A.23 y la desigualdad (ii) del Teorema A.30. En efecto,

$$\begin{aligned} |u - u_h|_1 &\leq \frac{1}{\sqrt{\min k_i}} \|u - u_h\|_a \leq \frac{1}{\sqrt{\min k_i}} \frac{\beta}{\alpha} \|u - Iu\|_a \\ &\leq \frac{1}{\sqrt{\min k_i}} \frac{\beta}{\alpha} \left(\sum_{i=1}^2 k_i |u - Iu|_{1,\Omega_i}^2 \right)^{1/2} \leq \frac{1}{\sqrt{\min k_i}} \frac{\beta}{\alpha} \left(\sum_{i=1}^2 k_i h^{2r} C_{I_i} |u|_{1+r,\Omega_i}^2 \right)^{1/2} \\ &\leq C_1 h^r \left(\sum_{i=1}^2 |u|_{1+r,\Omega_i}^2 \right)^{1/2} \end{aligned}$$

siendo $C_1 = \sqrt{\frac{\max k_i}{\min k_i}} \frac{\beta}{\alpha} \max_{i=1,2} \sqrt{C_{I_i}}$ y β, α constantes de continuidad y coercividad de a .

Utilizaremos a continuación un argumento de dualidad para la prueba de (ii). Sea w la solución del problema débil dual, que consiste en encontrar

$$w \in H_0^1(\Omega) \quad \text{tal que} \quad a(w, v) = (u - u_h, v) \quad \forall v \in H_0^1(\Omega).$$

Nuevamente, por Teorema A.24 y Lema A.25,

$$w \in H^{1+r}(\Omega) \quad \text{y} \quad |w|_{1+r} \leq C \|u - u_h\|. \quad (\text{A.16})$$

Tomando $v = u - u_h$ en el problema débil dual e introduciendo el interpolador $Iw \in V_h$, tenemos

$$(u - u_h, u - u_h) = a(w, u - u_h) = a(w - Iw, u - u_h),$$

como consecuencia de la propiedad (A.7).

De aquí, vamos a acotar $(u - u_h, u - u_h)$ haciendo uso de la estimación (ii) del Teorema A.30, la estimación (i) que acabamos de probar correspondiente a este teorema y la desigualdad (A.16):

$$\begin{aligned} \|u - u_h\|^2 &\leq \beta \|w - Iw\|_1 \|u - u_h\|_1 \leq \beta (c(\Omega)^2 + 1) |w - Iw|_1 |u - u_h|_1 \\ &\leq \beta (c(\Omega)^2 + 1) C_I h^r |w|_{1+r} C_1 h^r |u|_{1+r} \\ &\leq C_2 h^{2r} \|u - u_h\| |u|_{1+r}, \end{aligned}$$

donde $C_2 = \beta(c(\Omega)^2 + 1)C_I C C_1$, siendo C_I, C, C_1 las constantes respectivas de los teoremas aplicados y $c(\Omega)$ la constante de Poincaré.

Por último, dividiendo por $\|u - u_h\|$ llegamos a la estimación que queríamos probar. \square

Bibliografía

- [1] R. A. ADAMS, Sobolev Spaces, Academic Press, New York - London, 1975. Pure and Applied Mathematics, Vol.65
- [2] L. C. EVANS, Partial Differential Equations, AMS, Volume 19, 1998.
- [3] E. A. CODDINGTON, N. LEVINSON, Theory of Ordinary Differential Equations, McGraw-Hill, 1955.
- [4] P. GRISVARD, Singularities in boundary value problems, Springer-Verlag, 1992. Research Notes in Applied Mathematics, RMA 22.
- [5] R. VERFÜRTH, A Review of a posteriori error estimation and adaptative mesh-refinement techniques, Wiley- Teubner, 1996.
- [6] D. BRAESS, Finite Elements , Cambridge University Press, 2001.
- [7] C. JOHNSON, Numerical solution of partial differential equations by the finite element method, Cambridge University Press, 1995.
- [8] M. PETZOLDT, Regularity and error estimators for elliptic problems with discontinuous coefficients, Thesis Doctoral, Berlin 2001.
- [9] S. BERRONE, Robust a posteriori error estimates for finite element discretizations of the heat equation with discontinuous coefficients, Rapporto N.20, Politecnico di Torino, Italia, 2005.
- [10] Z. CHEN, J. FENG, An adaptative finite element algorithm with reliable and efficient error control for linear parabolic problems, Mathematics of Computation. 73(2004),no.247, 1167-1193.
- [11] P.G.CIARLET, The Finite Element Method for Elliptic Problems, North- Holland Publishing Company, Amsterdam, 1978.
- [12] T. DUPONT, R. SCOTT, Polynomial Aproximation of Functions in Sobolev Spaces, Mathematics of Computation. 34(1980),no.150, 441-463.
- [13] J. NECAS, Sur la coercivité, des formes sesquilinéaires, elliptiques. Rev. Roumaine de Math.Pure et App.9(1964),no.1, 47-69.
- [14] NATICK MA, Partial Differential Equation Toolbox User's Guide, The Mathworks, 1997.

- [15] Y. ZHOU, X. WU, Finite element analysis of diffusional drug release from complex matrix systems. i. complex geometries and composite structures, *Journal of Controlled Release*. 49(1997), 277-288.