

**ESTIMACIÓN NUMÉRICA DE LA EXPLOSIÓN PARA UN
PROBLEMA PARABÓLICO DEBIDO A CONDICIONES DE
FRONTERA NO LINEALES Y TÉRMINO NO LINEAL**

WILSON YAMID OBREGÓN

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2012**

**ESTIMACIÓN NUMÉRICA DE LA EXPLOSIÓN PARA UN
PROBLEMA PARABÓLICO DEBIDO A CONDICIONES DE
FRONTERA NO LINEALES Y TÉRMINO NO LINEAL**

WILSON YAMID OBREGÓN

**Monografía presentada para optar al
título de Licenciado en Matemáticas**

**Director
JULIO CÉSAR CARRILLO ESCOBAR, Ph.D.**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2012**

Agradecimientos

- * A Dios, a mi papá, a mi mamá, a Yesica Paola por su apoyo incondicional.
- * A mi director de Proyecto prof. Julio Carrillo, y al prof. Rafael Castro por sus aportes y colaboración.
- * A todos los profesores de la escuela de matemáticas, quienes me aportaron conocimientos a lo largo de estos 5 años.
- * A todos mis amigos: Nelson Rueda, Mauricio Quiñones, Michael Álvarez, Daniel Bernal, Fabio Ruiz, Jairo Ardila, Diego Badillo, Diego Medina, Laura Bosque, Vladimir Angulo, Jazmin Gomez, Marisol Rueda, Miguel Camacho, Luis Pérez, Pedro Jaimes, Jerson Bautista, Mario Castañeda, Nicolas Torres, Jorge Leal, Sergio Pérez, Michael Rincón,...
y a todos los que quisiera nombrar, pero el espacio de esta hoja no es suficiente para esto...

Gracias!!!

Tabla de Contenido

Introducción	9
1. Preliminares	10
1.1. Fundamentos de análisis funcional lineal	10
1.1.1. Espacios normados y de Banach	10
1.1.2. Espacios de Hilbert	14
1.1.3. Espacios de Lebesgue	14
1.1.4. Espacios de Sóbolev	15
1.1.5. Teoremas de representación de Riesz	16
1.1.6. El Teorema de Lax-Milgram	16
2. Cálculo variacional, problemas de valores en la frontera y el método de los elementos finitos	18
2.1. El problema de valores en la frontera	18
2.1.1. Extremos de funciones y funcionales	18
2.1.2. La ecuación de Euler-Lagrange	26
2.1.3. Condiciones de frontera naturales y esenciales	28
2.1.4. El método de Ritz-Galerkin	40
2.1.5. El método de los elementos finitos	44
3. Explosión de las soluciones para un problema parabólico	60
3.1. La ecuación de calor unidimensional	60
3.1.1. El modelo de la ecuación de calor	60
3.1.2. Formulación variacional del un PVF	64
3.1.3. Formulación variacional discreta generalizada	92
3.2. Análisis de B-explosión semidiscreta	93
Bibliografía	97

RESUMEN

TÍTULO: ESTIMACIÓN NUMÉRICA DE LA EXPLOSIÓN PARA UN PROBLEMA PARABÓLICO DEBIDO A CONDICIONES DE FRONTERA NO LINEALES Y TÉRMINO NO LINEAL*

AUTOR: WILSON YAMID OBREGÓN**

PALABRAS CLAVES: Blow-up, semidiscretización de un problema parabólico.

DESCRIPCIÓN: Esta tesis presenta el análisis hecho sobre un problema de evolución no lineal de tipo parabólico, que describe algunos procesos de la teoría de la combustión, dinámica de gases y termodinámica. Este problema parabólico modela procesos relacionados con difusión y reacción exponencial, el cual bajo algunas restricciones a ciertos parámetros, exhibe un fenómeno conocido como "blow-up".

En el capítulo I presentamos ciertos resultados importantes de análisis funcional, necesarios para abordar los métodos que utilizaremos en el análisis del problema propuesto. Trataremos así algunos espacios abstractos (espacios vectoriales, espacios normados, espacios de Banach, espacios de Hilbert, espacios L^p , entre otros) donde consideramos, que en algunos de ellos, se encuentra la solución (aproximada) a un problema diferencial específico, y posteriormente veremos los teoremas de extensión, que serán de gran utilidad porque estos garantizaran la existencia y la unicidad de la solución en algunos de los espacios mencionados anteriormente.

En el capítulo II consideramos algunos resultados relacionados con la teoría de optimización de funcionales, y luego presentamos el método de los elementos finitos (FEM) como herramienta para calcular soluciones aproximadas a problemas planteados en forma de ecuaciones diferenciales. Abordaremos así la relación existente entre problemas de valores en la frontera y el cálculo variacional, vínculo que será fundamental para establecer la equivalencia de soluciones entre estos dos esquemas.

Finalmente, en el capítulo III estudiaremos los principios de termodinámica, esquematizaremos de manera numérica y semidiscreta el problema propuesto y posteriormente analizaremos las condiciones bajo las cuales ocurre el fenómeno de blow-up para el problema planteado.

*Proyecto de Grado.

**Facultad de Ciencias. Escuela de Matemáticas. Director Julio C. Carrillo E.

ABSTRACT

TITLE: NUMERICAL ESTIMATES OF THE EXPLOSION FOR A PARABOLIC PROBLEM DUE TO NONLINEAR BOUNDARY CONDITIONS AND NONLINEAR TERM*

AUTHOR: WILSON YAMID OBREGÓN**

KEYWORDS: Blow-up, semidiscretization of a parabolic problem.

DESCRIPTION: This thesis presents the analysis done in a non-linear evolution problem of parabolic type, which describes some processes of the theory of combustion, gases dynamics and thermodynamics. This parabolic problem model related processes with exponential diffusion and reaction, which under some restrictions on certain parameters, exhibits a phenomenon known as "blow-up".

In Chapter I we present some important results of functional analysis required to approach the methods we use in analyzing the proposed problem. We'll try some abstract spaces (vector spaces, normed spaces, Banach spaces, Hilbert spaces, L^p spaces, etc.) where we consider that in some of these, is the solution (approximate) to a specific differential problem and after we will see the extension theorems, which will be useful because they guarantee the existence and uniqueness of the solution in some of the spaces mentioned before.

In Chapter II we consider some results related to functional optimization theory, and after we present the finite element method (FEM) as a tool to compute approximate solutions to problems proposed in form of differential equations. We'll approach the relationship between boundary value problems and the variational calculus, link which will be fundamental for establish the equivalence of solutions between these two schemes.

Finally, in Chapter III we'll study the principles of thermodynamics, we'll schematize in numericall and semidiscrete form the proposed problem and after we'll analyze the conditions under which occurs the phenomenon of blow-up for the established problem.

* Graduation project

** Faculty of Sciences. Mathematics School. Julio C. Carrillo E. , Advisor

Introducción

Un sinnúmero de fenómenos físicos del mundo real (difusión de calor, dinámica de gases) requieren el uso de ecuaciones diferenciales en derivadas parciales (EDPs) o sistemas de estas para predecir, justificar y/o manipular el comportamiento de tales fenómenos. En muchas de estas ecuaciones en derivadas parciales se establecen condiciones particulares (como por ejemplo, condiciones de frontera y/o condiciones de absorción, lineales o no, dependiendo del fenómeno). Precisamente aquí analizaremos un problema, conocido como fenómeno de explosión (o Blow-Up), el cual involucra condiciones de tipo no lineal, donde además la solución de este problema se comporta de manera no acotada. En un problema que involucra el fenómeno de explosión, usualmente se presentan comportamientos singulares de la solución, que pueden ser identificados como comportamientos asintóticos para determinados valores del tiempo; esto quiere decir que, después de un periodo de evolución predecible de la solución a la ecuación diferencial, se puede observar que la solución al problema se hace no acotada en tiempo finito, resultado de los efectos acumulados de la no linealidad. En particular, nos interesa estudiar el fenómeno de explosión para el problema de tipo parabólico (1) donde se presenta la explosión de u (bajo ciertas restricciones sobre los parámetros p, q y λ), con la siguiente condición de frontera no lineal y termino de absorción no lineal:

$$(1) \quad \begin{array}{ll} \text{(EDP)} & u_t - u_{xx} = -\lambda u^p \quad (x, t) \in (0, 1) \times [0, T) \\ \text{(CF)} & u_x(0, t) = 0, u_x(1, t) = u(1, t)^q \quad t \in [0, T) \\ \text{(CI)} & u(x, 0) = u_0(x) \quad x \in [0, 1] \end{array}$$

en donde $p > 1$ y $q > 1$, $\lambda > 0$ son parámetros, $T > 0$ y u_0 es una función no negativa. En concreto, nos proponemos estudiar el fenómeno de explosión utilizando dos esquemas de representación: un esquema numérico y un esquema semidiscreto.

Capítulo 1

Preliminares

El propósito de este capítulo es revisar ciertos conceptos y resultados referentes al análisis funcional que serán de gran utilidad para analizar el problema inicialmente planteado. De esta manera se tratarán temas como espacios vectoriales, espacios normados, espacios de Banach, formas lineales, espacios con producto interno, espacios métricos, espacios de Hilbert y espacios de funciones. La bibliografía tratada en este capítulo está fundamentada por [9, 10, 14, 15, 16, 17, 18].

1.1. Fundamentos de análisis funcional lineal

La mayoría de las ciencias exactas utilizan constantemente la idea de aproximación, de tal manera que es importante contar con algunas herramientas que puedan medir *la distancia* entre una solución exacta y una solución aproximada de un problema.

1.1.1. Espacios normados y de Banach

Sea X un espacio vectorial sobre el campo de escalares \mathcal{K} . Siempre asumimos que $\mathcal{K} = \mathbb{R}$ es el campo de los números reales o $\mathcal{K} = \mathbb{C}$ es el campo de los números complejos.

Definición 1.1. Una *norma* en X es una función $x \rightarrow \|x\|$ de X en \mathbb{R} , que cumple las siguientes propiedades:

(N1) para cada $x \in X$ se tiene que $\|x\| \geq 0$ y la igualdad se cumple si y solo si $x = 0$.

(N2) para cada $x \in X$ y cada $\lambda \in \mathcal{K}$ se tiene que $\|\lambda x\| = |\lambda| \|x\|$.

(N3) para cada $x, y \in X$ se cumple que $\|x + y\| \leq \|x\| + \|y\|$.

$\|x\|$ es una medida del tamaño del vector x .

Definición 1.2. Un espacio vectorial X con la norma $\|\cdot\|$ que cumple las propiedades (N1)-(N3) se llama un *espacio normado*.

Definición 1.3. Una norma determina una distancia entre vectores del espacio normado X , es decir, la función

$$d(x, y) = \|x - y\|$$

para cada $x, y \in X$ es conocida como una **métrica** inducida por tal norma en X . Verificamos que, para cada $x, y, z \in X$ se cumplen las tres propiedades de una distancia:

(D1) $d(x, y) \geq 0, d(x, y) = 0$ si y solo si $x = y$

(D2) $d(x, y) = d(y, x)$

(D3) $d(x, z) \leq d(x, y) + d(y, z)$

Definición 1.4. La **bola** de centro en un punto $x \in X$ con radio positivo r se denota como:

$$B(x; r) = \{y \in X \mid \|y - x\| < r\}$$

Si $B(x; r) \subseteq V$, para algún $r > 0$ decimos que el conjunto V es una **vecindad** de x . Equivalentemente, si x es un punto interior de V .

Definición 1.5. Un conjunto $V \subseteq X$ es **abierto** si este es vecindad de cada uno de sus puntos. En otras palabras, si para cada $x \in X$ existe un $r > 0$ tal que $B(x; r) \subseteq V$.

Definición 1.6. Un conjunto $U \subseteq X$ es **cerrado** si su complemento X/U es abierto.

Definición 1.7. Toda función f definida de \mathbb{N} y de valores en X se llama **sucesión** en X . Es costumbre denotar toda sucesión en X de la forma (x_n) .

Definición 1.8. Decimos que una **sucesión es convergente** (que la sucesión (x_n) en X converge a un punto $x \in X$) si

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0$$

Definición 1.9. Dados dos espacios normados X, Y decimos que la función $f : X \rightarrow Y$ es una **función continua** si, para cada $x \in X$ y cada $\epsilon > 0$ existe un $\delta > 0$ tal que

$$\|f(x') - f(x)\| < \epsilon \quad \text{si } x' \in X, \|x' - x\| < \delta$$

Definición 1.10. Una sucesión (x_n) es una **sucesión de Cauchy** si para cada $\epsilon > 0$ uno puede encontrar un entero N lo suficientemente grande tal que

$$\|x_m - x_n\| \leq \epsilon \quad \text{si } m, n \geq N$$

Definición 1.11. Un espacio normado X es un **espacio completo** si cada sucesión de Cauchy en X converge a un punto en X .

Definición 1.12. Un espacio normado completo se llama **espacio de Banach**.

Ejemplos de espacios normados y de Banach

Ejemplo 1.1. El espacio de dimensión finita $\mathbb{R}^n = \{x = (x_1, \dots, x_n), x_i \in \mathbb{R}\}$ con la norma euclidiana

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$

es un espacio de Banach sobre el campo de los números reales.

Ejemplo 1.2. En el espacio vectorial \mathbb{R}^n se pueden considerar también las normas

$$\|x\|_p = (|x_1| + \dots + |x_n|)^{1/p}, \quad \|x\|_\infty = \max\{|x_i|, 1 \leq i \leq n\} \equiv \max_{1 \leq i \leq n} |x_i|.$$

En este caso, $1 \leq p < \infty$. Cada una de estas normas hace de \mathbb{R}^n un espacio de Banach.

Ejemplo 1.3. Para un intervalo dado $[a, b]$, el espacio de las funciones continuas $f : [a, b] \rightarrow \mathbb{R}$ con la norma

$$\|f\|_0 = \max\{|f(x)|, x \in [a, b]\} \equiv \max_{x \in [a, b]} |f(x)|$$

es un espacio de Banach.

Definición 1.13. Sea X un espacio vectorial sobre el campo $\mathbb{K} = \mathbb{R}$ ó \mathbb{C} . Se dice que una función $(\cdot, \cdot) : X \times X \rightarrow \mathbb{K}$ es un **producto interno** en X , si esta función cumple con las siguientes propiedades, para todo $x, y, z \in X$ y $a, b \in \mathbb{K}$:

(PI-1) $(ax + by, z) = a(x, z) + b(y, z)$.

(PI-2) $(x, y) = \overline{(y, x)}$.

(PI-3) $(x, x) > 0$ si y sólo si $x \neq 0$, y $(x, x) = 0$ si y sólo si $x = 0$.

En este caso, si X es un espacio vectorial real, la propiedad (PI-2) se reduce a la condición de simetría

$$(x, y) = (y, x) \quad \text{para cada } x, y \in X.$$

Definición 1.14. Sea V un espacio vectorial. Una **forma bilineal** es una aplicación $f : V \times V \rightarrow \mathbb{R}$, que verifica las condiciones

(1) $f(u + v, h) = f(u, h) + f(v, h), \forall u, v, h \in V$.

(2) $f(\alpha u, v) = f(u, \alpha v) = \alpha f(u, v), \forall u, v \in V, \forall \alpha \in \mathbb{R}$.

(3) $f(u, v) = f(v, u), \forall u, v \in V$.

Definición 1.15. F es definida como una **forma bilineal simétrica** si $F(u, v) = F(v, u)$, para cada $u, v \in X$.

Definición 1.16. Sea F una forma bilineal sobre X . F es definida como **positiva**, si $F(u, u) > 0$ para cada $u \in X$

Definición 1.17. Sea V un espacio vectorial sobre \mathbb{R} , entonces un **funcional lineal** f es una función de V a \mathbb{R} que es lineal, es decir, que cumple:

(1) $f(u + v) = f(u) + f(v), \forall u, v \in V$.

(2) $f(\alpha u) = \alpha f(u), u \in V, \forall \alpha \in \mathbb{R}$.

Definición 1.18. Un funcional lineal $F(\cdot)$ en un espacio normado $(V, \|\cdot\|)$ es un **funcional acotado** si existe una constante $c > 0$ tal que $\forall v \in V$ se cumple que

$$|F(v)| \leq c \|v\|$$

Definición 1.19. Una forma bilineal $b(\cdot, \cdot)$ en un espacio normado $(V, \|\cdot\|)$ es una **forma bilineal coerciva** si existe una constante $\mu > 0$ tal que $\forall u \in V$ se cumple que

$$b(u, u) \geq \mu \|u\|^2$$

Definición 1.20. Una forma bilineal $b(\cdot, \cdot)$ en un espacio normado $(V, \|\cdot\|)$ es una **forma bilineal acotada** si existe una constante $\lambda > 0$ tal que $\forall u, v \in V$ se cumple que

$$|b(u, v)| \leq \lambda \|u\| \|v\|$$

Definición 1.21. Sea V un espacio de Banach y sea $M \subset V, M \neq \emptyset$

1. M es **denso** en V , si cada elemento de V puede ser representado como el límite de una sucesión convergente con elementos de M .
2. V es **separable** si contiene un subconjunto denso enumerable.
3. M es **cerrado** si el límite de toda sucesión convergente en M pertenece a M .
4. M es **acotado** si existe una constante c tal que $\|u\|_V \leq c, \forall u \in M$.
5. M es **compacto** si toda sucesión en M tiene una subsucesión convergente en M (el límite de la subsucesión puede no pertenecer a M).

Teorema 1.1. (Desigualdad de Cauchy-Schwarz) Sea V un espacio vectorial con producto interno $(V, (\cdot, \cdot))$, donde $u, v \in V$, entonces:

$$|(u, v)| \leq (u, u)^{\frac{1}{2}} (v, v)^{\frac{1}{2}} \tag{1.1}$$

Teorema 1.2. (Desigualdad de Hölder) Si $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$ números reales, y $p > 0, q > 0$, tales que $\frac{1}{p} + \frac{1}{q} = 1$, entonces:

$$\sum_{i=1}^n |x_i y_i| \leq \left[\sum_{i=1}^n |x_i|^p \right]^{\frac{1}{p}} \cdot \left[\sum_{i=1}^n |y_i|^q \right]^{\frac{1}{q}} \tag{1.2}$$

Teorema 1.3. (Ley del paralelogramo) Sean $u, v \in V$. Se cumple que

$$\|u + v\|_V^2 + \|u - v\|_V^2 = 2 \left[\|u\|_V^2 + \|v\|_V^2 \right] \quad (1.3)$$

Teorema 1.4. (Desigualdad triangular) Sean $u, v \in V$, entonces

$$\|u + v\|_V \leq \|u\|_V + \|v\|_V$$

Definición 1.22. Sea $f : \Omega \rightarrow \mathbb{R}$, luego el **soporte de f** ($\text{sop}(f)$), se define por

$$\text{sop}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}$$

1.1.2. Espacios de Hilbert

Definición 1.23. X es un **Espacio de Hilbert** si X es un espacio de Banach bajo la norma

$$\|(x, x)\|_X = \sqrt{(x, x)_X} \quad (1.4)$$

1.1.3. Espacios de Lebesgue

Definición 1.24. Sean $1 \leq p < \infty$. Se define **el espacio $L^p[a, b]$** como el conjunto de todas las funciones reales integrables $f(x)$, $-\infty < a \leq x \leq b < \infty$, tales que

$$\int_a^b |f(x)|^p dx < \infty$$

En este caso se considera la integral en el sentido de Riemann.

Definición 1.25. En los espacios L^p se define una norma, conocida como **norma L^p** , $\forall u \in L^p(\Omega)$ dada por el funcional $\|\cdot\|_p$ y determinada por:

$$\|u\|_p = \left[\int_{\Omega} |u(x)|^p dx \right]^{\frac{1}{p}}$$

donde $1 \leq p < \infty$.

Es claro que $\|u\|_p \geq 0$.

Definición 1.26. El **espacio L^2** esta determinado por las funciones cuadrado integrables:

$$L^2(\Omega) = \left[f : \int_a^b |f(x)|^2 dx < \infty \right]$$

Este espacio $L^2(\Omega)$ se puede entender como el conjunto de clases de equivalencia de funciones que son iguales en casi todos los puntos de Ω . Si tenemos dos funciones arbitrarias $f, g \in L^2(\Omega)$, entonces su producto es Lebesgue integrable.

Definición 1.27. Una sucesión u_n en $L^2(\Omega)$ es una **sucesión convergente en media** a $u \in L^2(\Omega)$ si $\lim_{n \rightarrow \infty} \|u_n - u\|_2 = 0$.

El espacio $L^p(\Omega)$ es un espacio de Banach si $1 \leq p < \infty$ (ver[18])

1.1.4. Espacios de Sóbolev

Definición 1.28. El **Espacio de Sóbolev** $W^{k,p}(\Omega)$ se define como el conjunto de todas las funciones $u \in L^p(\Omega)$ tal que para cada multi-índice α con $|\alpha| \leq k$ la derivada parcial $D^\alpha u$ pertenece a $L^p(\Omega)$ (es decir $W^{k,p}(a,b)$ es el conjunto de todas las funciones que tienen hasta k derivadas que son L^p integrables en (a,b)):

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega), \forall |\alpha| \leq k\}$$

En este caso Ω es un conjunto abierto en \mathbb{R}^n y $1 \leq p < \infty$. El número natural k se denomina **orden del espacio de Sóbolev** $W^{k,p}(\Omega)$.

Definición 1.29. En los espacios $W^{k,p}(\Omega)$ se define una norma, conocida como **norma de Sóbolev**, dada por el funcional $\|\cdot\|_{k,p}$ y determinada por:

$$\|u\|_{k,p} = \left[\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right]^{\frac{1}{p}}, \quad 1 \leq p < \infty$$

Es convencional denotar $W^{k,p}(\Omega)$ por $H^k(\Omega)$ ya que este es un espacio de Hilbert con la norma $\|\cdot\|_{W^{k,2}(\Omega)}$

El espacio $W^{m,p}(\Omega)$ es un espacio de Banach si $1 \leq p < \infty$.

Espacio de Sóbolev $W^{1,2}(\Omega)$

Sea $\Omega \subset \mathbb{R}$. El Espacio de Sóbolev $W^{1,2}(\Omega)$ o $H^1(\Omega)$, es el conjunto de todas las funciones $u \in L^2(\Omega)$ para las cuales existe una función $f \in L^2(\Omega)$ donde:

$$\int_{\Omega} u(x)v'(x)dx = - \int_{\Omega} f(x)v(x)dx, \quad v \in C_0^1(\Omega) \quad (1.5)$$

Este conjunto (denotado por $H^1(\Omega)$) es un espacio vectorial, bajo las operaciones usuales de suma de elementos, y producto por un escalar. Consideramos también el espacio $H_0^1(\Omega)$ que consiste en el conjunto de funciones de $H^1(\Omega)$ que tienen soporte compacto:

$$H_0^1(\Omega) = \{v \in H^1(\Omega) \mid v(\partial\Omega) = 0\} \quad (1.6)$$

Definición 1.30. La **norma** $\|f\|_{\infty,\Omega}$ para una función $f \in L^\infty$ esta definida por

$$\|f\|_{\infty,\Omega} = \sup_{x \in \mathbb{R}} |f(x)|$$

De hecho, $L^\infty(\Omega)$ denota el espacio de Banach de funciones acotadas sobre Ω con la norma $\|f\|_{\infty,\Omega}$.

1.1.5. Teoremas de representación de Riesz

Definición 1.31. Dos elementos u y v de un espacio con producto interno V son **ortogonales** si $(u, v) = 0$. Dado un conjunto $A \subset V$, se denota por A^\perp el conjunto de todos los elementos de V que son ortogonales a cada elemento de A :

$$A^\perp = [u \in V : (u, v) = 0, \text{ para cada } v \in A]$$

El siguiente resultado garantiza la existencia de una proyección ortogonal de un elemento de un espacio de Hilbert si tal subespacio es cerrado (ver [10]).

Teorema 1.5. (Teorema de la proyección) Sea A un subespacio cerrado de un espacio de Hilbert V . Entonces $\forall v \in V$ tenemos que $u = v + w, v \in A, w \in A^\perp$.

El Teorema de representación de Riesz

El Teorema de representación de Riesz es una caracterización muy útil para funcionales lineales continuos sobre espacios de Hilbert. Considérese el siguiente resultado (ver [10]).

Teorema 1.6. Para todo funcional lineal continuo f sobre un espacio de Hilbert V , existe un único elemento $u \in V$ tal que $f(v) = (u, v), \forall v \in V$ y $\|f\|_{V'} = \|u\|$

Del teorema de representación de Riesz aparece la definición de la siguiente función:

$$\begin{aligned} T : V &\longrightarrow V' \\ u &\longrightarrow Tu = f \end{aligned}$$

Donde esta es una transformación lineal, y además el teorema de Representación de Riesz muestra que la aplicación es también biyectiva y es además una isometría ($\|Tu\| = \|u\|, \forall u \in V$).

1.1.6. El Teorema de Lax-Milgram

El teorema de Representación de Riesz es importante para el tratamiento de ecuaciones diferenciales que son variacionales (ecuaciones diferenciales que son la integral de la ecuación de Euler-Lagrange). La estructura de dichas ecuaciones necesita de una modificación utilizando el teorema de representación de Riesz, debida a Peter Lax y Arthur Milgram (ver [10]).

Teorema 1.7. Sea $B(\cdot, \cdot)$ una forma bilineal continua y coerciva, sobre un espacio de Hilbert V . Entonces para todo funcional lineal continuo $G : V \longrightarrow \mathbb{R}$ existe un único elemento $u \in V$ tal que $B(u, v) = G(v), \forall v \in V$.

Definición 1.32. Sea Ω un dominio en \mathbb{R}^n . Para algún entero no negativo m , el **espacio vectorial** $C^m(\Omega)$ consiste en todas las funciones ϕ las cuales junto con sus derivadas parciales $D^\alpha \phi$ de orden $|\alpha| \leq m$ son continuas en Ω .

Aplicaciones lineales y la norma dual

Considérese el siguiente resultado (ver [15]).

Teorema 1.8. Sean $n+1$ espacios normados V_1, \dots, V_n, W y sea T una función multilineal (lineal en cada uno de sus argumentos) de $V_1 \times \dots \times V_n$ en W . Entonces T es continua si y solo si es acotada, es decir, si existe un número $c > 0$ tal que para cualquier $(u_1, \dots, u_n) \in V_1 \times \dots \times V_n$

$$\|T(u_1, \dots, u_n)\|_W \leq c \|u_1\|_{V_1} \cdot \|u_2\|_{V_2} \cdots \|u_n\|_{V_n}$$

El espacio de todas las funciones multilineales continuas de $V_1 \times \dots \times V_n$ en W se denota por $\mathcal{L}(V_1, \dots, V_n; W)$.

Definición 1.33. Sean V y W espacios de Banach. Una función $f : V \rightarrow W$ es llamada **función continua** en $u_0 \in V$, si para toda sucesión u_n en V tal que $u_n \rightarrow u_0$ se tiene que $T(u_n) \rightarrow T(u_0)$. La función es llamada continua en V , si es continua en todo punto de V .

Teorema 1.9. Sean V y W espacios de Banach y sea $T : V \rightarrow W$ una transformación lineal. De esta manera, las siguientes propiedades son equivalentes:

1. T es continua
2. T es continua en $0 \in V$
3. T es acotada :

$$\|T\|_{V \rightarrow W} = \sup_{u \in V} \frac{\|Tu\|_W}{\|u\|_V} < \infty$$

Definición 1.34. Sea V un espacio de Banach. El espacio $\mathcal{L}(V, \mathbb{R})$ es llamado el **espacio dual de V** y es denotado por V' . Los elementos de V' son llamados **funcionales lineales continuos**.

Capítulo 2

Cálculo variacional, problemas de valores en la frontera y el método de los elementos finitos

En este capítulo estableceremos la equivalencia entre problemas variacionales y problemas de valores en la frontera (PVF), haciendo uso de las herramientas que proporciona el cálculo variacional. Posteriormente veremos el método de los elementos finitos, que será una herramienta fundamental para calcular soluciones aproximadas a PVF, gracias a ciertos resultados del cálculo variacional. La bibliografía tratada en este capítulo está fundamentada por [1, 3, 7, 8, 9, 10, 12, 14, 15, 16, 17, 18].

2.1. El problema de valores en la frontera

Formulación variacional

El cálculo variacional es una rama matemática que está relacionada con la teoría general de valores extremos de funcionales y funciones. La aplicación del cálculo variacional hace posible la formulación de problemas planteados en función de máximos y mínimos usando la teoría de análisis funcional, la cual puede llevar un PVF a un equivalente variacional y viceversa.

2.1.1. Extremos de funciones y funcionales

Cuando se aplica el método de Ritz y el método de los elementos finitos a un problema de valores en la frontera, muchas veces se trata de resolver un sistema de ecuaciones lineales de la forma $Hx = b$, donde H es una matriz simétrica definida positiva de orden $N \times N$. Este problema (como se verá más adelante) consiste en minimizar el funcional cuadrático:

$$f(x) = \frac{1}{2}x^T Hx - b^T x$$

donde $x \in \mathbb{R}^N$. Por esta razón abordaremos algunas definiciones básicas y teoremas para funcionales definidos en \mathbb{R}^N , y posteriormente analizaremos el comportamiento de ciertos funcionales de tipo cuadrático.

Extremos sin restricciones

Sea f una función definida sobre un conjunto $S \subseteq \mathbb{R}^N$ en \mathbb{R} . Sea un elemento de S es de la forma $x = (x_1, \dots, x_N)^T$, donde $x_i \in \mathbb{R}$. Si se considera a \mathbb{R}^N dotado de la distancia usual generada por la norma

$$\|x\| = \left[\sum_{i=1}^N x_i^2 \right]^{1/2}, x = (x_1, \dots, x_N)$$

se tiene que la vecindad de x de radio ϵ , para $x \in \mathbb{R}^N$ y $\epsilon > 0$, esta dada por

$$B_\epsilon(x) = \{y \in \mathbb{R}^N : 0 \leq \|x - y\| < \epsilon\}$$

Definición 2.1. Sea f una función definida sobre $S \subseteq \mathbb{R}^N$. Entonces $x_0 \in S$ es un **punto minimal local** de f , si existe $\epsilon > 0$ tal que

$$f(x_0) \leq f(x), \forall x \in B_\epsilon(x_0) \cap S$$

Definición 2.2. Sea f una función definida sobre $S \subseteq \mathbb{R}^N$. Entonces $x_0 \in S$ es un **punto minimal local fuerte** de f , si existe $\epsilon > 0$ tal que

$$f(x_0) < f(x), \forall x \in B_\epsilon(x_0) \cap S, x \neq x_0$$

Definición 2.3. Sea f una función definida sobre $S \subseteq \mathbb{R}^N$. Entonces $x_0 \in S$ es un **minimal global** de f , si

$$f(x_0) \leq f(x), \forall x \in S$$

Definición 2.4. Sea f una función definida sobre $S \subseteq \mathbb{R}^N$. Entonces $x_0 \in S$ es un **minimal global fuerte** de f , si

$$f(x_0) < f(x), \forall x \in S, x \neq x_0$$

Los **maximales locales**, **maximales locales fuertes**, **maximales globales** y los **maximales globales fuertes** se definen de manera análoga. Denotaremos a continuación a S como un subconjunto abierto de \mathbb{R}^N .

Definición 2.5. Sea $f \in C^1(S)$. El **gradiente** de f en el punto $x \in S$ es el vector

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_N} \end{bmatrix}$$

Si $f \in C^2(S)$, la **matriz Hessiana** de f en el punto $x \in S$ es la matriz:

$$H(x) = [h_{ij}] \in \mathbb{R}^{N \times N}, \quad h_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

Si tenemos que $f \in C^2(S)$ y como

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$$

entonces la matriz Hessiana de f es simétrica. La expansion de Taylor de f cerca de un punto x se puede expresar en términos de la matriz hessiana y del gradiente. Si $f \in C^1(S)$, entonces para cualquier $x \in S$ se tiene que

$$f(x+h) = f(x) + (\nabla f(x))^T h + o(\|h\|)$$

y si $f \in C^2(S)$ entonces

$$f(x+h) = f(x) + (\nabla f(x))^T h + \frac{1}{2} h^T H(x) h + o(\|h\|)^2$$

Definición 2.6. Sea $f \in C^1(S)$. $x_0 \in S$ es un **punto estacionario** de f si $\nabla f(x_0) = 0$.

Teorema 2.1. Sea $f \in C^1(S)$. Si $x_0 \in S$ es un minimal local de f entonces x_0 es un punto estacionario de f .

Definición 2.7. Sea A una matriz simétrica. Se dice que A es una **matriz definida positiva** si $x^T A x > 0, \forall x \in \mathbb{R}^N, x \neq 0$.

Definición 2.8. Sea A una matriz simétrica. Se dice que A es una **matriz definida negativa** si $x^T A x < 0, \forall x \in \mathbb{R}^N, x \neq 0$.

El siguiente resultado es necesario (ver [20])

Teorema 2.2. (Teorema de Sylvester) La matriz simétrica $A_{k \times k} = (a_{ij})_{k \times k}$ es definida positiva si $\text{Det}(A_{k \times k}) > 0$.

Teorema 2.3. Sea $f \in C^2(S)$ y x_0 un punto estacionario de f . Entonces x_0 es un mínimo local fuerte de f si la matriz Hessiana $H(x_0)$ es definida positiva.

Demostración. Utilizando la expansion de Taylor, y como $\nabla f(x_0) = 0$, entonces

$$f(x_0+h) = f(x_0) + \frac{1}{2} h^T H(x_0) h + o(\|h\|)^2$$

si $H(x_0)$ es definida positiva, entonces existe un numero positivo λ_1 el cual puede ser tomado como el mas pequeño de los valores propios de $H(x_0)$, tal que:

$$h^T H(x_0) h \geq \lambda_1 \|h\|^2, \forall h \in \mathbb{R}^N$$

Así,

$$f(x_0 + h) - f(x_0) \geq \lambda_1 \|h\|^2 + o(\|h\|)^2$$

y como la expresión debe ser positiva para valores suficientemente pequeños de $\|h\|$, esto lleva a que $f(x_0 + h) > f(x_0)$ en alguna vecindad de x_0 , esto es que x_0 es un mínimo local fuerte de f . \square

En el caso de máximos locales de f la situación es análoga.

Teorema 2.4. *Sea $f \in C^2(S)$ y $x_0 \in S$ un punto estacionario de f . Entonces x_0 es un mínimo local fuerte de f si la matriz Hessiana $H(x_0)$ es definida positiva .*

Si x_0 es un máximo local de f entonces x_0 debe ser un punto estacionario de f y si $H(x_0)$ es negativamente definida, entonces x_0 es un máximo local fuerte de f . Además se puede mostrar que si x_0 es un punto estacionario de f , entonces:

1. x_0 es un mínimo o un máximo local, si $H(x_0)$ tiene valores propios positivos o negativos.
2. x_0 puede o no ser un minimal local, si $H(x_0)$ es semidefinida positiva , es decir $x^T H(x_0)x \geq 0$, para todo $x \in \mathbb{R}^N$.
3. x_0 puede o no ser un maximal local, si $H(x_0)$ es semidefinida negativa, es decir $x^T H(x_0)x \leq 0$, para todo $x \in \mathbb{R}^N$.

Definición 2.9. *Sea $f \in C^m(S)$, $x \in S \subseteq \mathbb{R}^N$, $y \in \mathbb{R}^N$, donde $\|y\| = 1$. La **derivada direccional** de orden m de f en el punto x en la dirección y es :*

$$f^{(m)}(x; y) = \left. \frac{d^m f(x + yt)}{dt^m} \right|_{t=0}$$

La derivada direccional de f puede ser calculada mediante la regla de la cadena. Si suponemos que $f \in C^1(S)$, tenemos que

$$f^{(1)}(x; y) = \left. \frac{df}{dt} (x_1 + ty_1, \dots, x_n + ty_n) \right|_{t=0}$$

$$\left. \frac{df}{dt} (x_1 + ty_1, \dots, x_n + ty_n) \right|_{t=0} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} (x_1, \dots, x_n) y_i = ((\nabla f(x)))^T y$$

Si derivamos encontramos que para cualquier $f \in C^2(S)$ tenemos

$$f^{(2)}(x; y) = y^T H(x)y$$

porque se tiene que

$$\begin{aligned}
\frac{d^2 f}{dt^2}(x_1 + ty_1, \dots, x_n + ty_n) &= \frac{d}{dt} \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1 + ty_1, \dots, x_n + ty_n) y_i \right) \\
&= \sum_{i=1}^n \left(\frac{d}{dt} \left(\frac{\partial f}{\partial x_i}(x_1 + ty_1, \dots, x_n + ty_n) \right) \right) y_i \\
&= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1 + ty_1, \dots, x_n + ty_n) y_i y_j \\
&= x^t H(x + ty) x
\end{aligned}$$

Teorema 2.5. Sea $f \in C^1(S)$. Si $x_0 \in S$ es un minimal local de f entonces $f^{(1)}(x_0; y) = 0$ para todas las direcciones y .

Teorema 2.6. Sea $f \in C^2(S)$ y supongamos que para algún $x_0 \in S$ se tiene que $f^{(1)}(x_0; y) = 0$ para todas las direcciones y . Entonces x_0 es un mínimo local fuerte de f si $f^{(2)}(x_0; y) > 0$ para todas las direcciones y .

Definición 2.10. Sea f definida sobre $S \subseteq \mathbb{R}^N$ y sea $k \in \mathbb{R}$ es el recorrido de f . Entonces el conjunto:

$$L_k = \{x \in S : f(x) = k\}$$

Define la **superficie de nivel** de f para el valor de k .

Un hecho bien conocido es el que si x pertenece a la superficie de nivel L_k , entonces el vector gradiente $\nabla f(x)$ es perpendicular a L_k en los x , en que la dirección en que la función f crece mas rápidamente. Así, el gradiente en cercanía de un punto mínimo local fuerte apunta hacia el exterior de la superficie, mientras que en cercanía de un máximo local fuerte apunta hacia el interior de la superficie.

Teoría de funcionales cuadráticos en \mathbb{R}^N

Definición 2.11. Un **funcional cuadrático** de \mathbb{R}^N es una función de la forma

$$f(x) = \frac{1}{2} x^T H x - b^T x + c, x \in \mathbb{R}^N$$

donde H es una matriz simétrica de orden $N \times N$, $b \in \mathbb{R}^N$, y $c \in \mathbb{R}$.

el gradiente y la matriz hessiana de un funcional cuadrático se encuentran determinados respectivamente por

$$\nabla f(x) = Hx - b \quad H(x) = H$$

Observemos que en un funcional cuadrático la matriz Hessiana tiene todas sus entradas constantes. Los funcionales cuadráticos son un ejemplo sencillo de funciones que poseen un mínimo o máximo local fuerte. Un punto x_0 es un punto estacionario de f si su gradiente

se desvanece en x_0 , esto es, si $Hx - b = 0$. Si H es no singular, lo cual se supondrá de aquí en adelante, entonces x_0 es único y está determinado como $x_0 = H^{-1}b$. Luego, un funcional cuadrático puede ser escrito de la forma:

$$f(x) = \frac{1}{2}(x - x_0)^T H(x - x_0) - b^T x + \hat{c}, x \in \mathbb{R}^N$$

donde $\hat{c} = -\frac{1}{2}b^T x_0 + c$. Sean $\{\lambda_i, x_i\}_{i=1}^N$ las soluciones propias de H , es decir son las soluciones de la ecuación:

$$Hx_i = \lambda_i x_i, i = 1, \dots, N$$

Como H es una matriz simétrica, sus valores propios son números reales y pueden ser ordenados $\lambda_1 < \dots < \lambda_N$ y supongamos que los vectores propios satisfacen la condición de ortogonalidad $x_i^T x_j = \delta_{ij}, i = 1, \dots, N$ al definir la matriz diagonal de los valores propios Φ y la matriz V de los vectores propios de H como

$$\Phi = \text{diag}(\lambda_1, \dots, \lambda_N)$$

$$V = [x_1 \dots x_N]$$

se puede observar que V es una matriz ortonormal (es decir $v^{-1} = V^T$) y $HV = V\Phi$. Si introducimos una nueva variable $z = V^T(x - x_0)$ tenemos que:

$$\hat{f}(z) \equiv f(Vz + x_0) = \frac{1}{2}z^T V^T H V z + \hat{c} = \frac{1}{2}z^T \Phi z + \hat{c}$$

o de una manera equivalente:

$$\hat{f}(z) = \frac{1}{2} \sum_{i=1}^N \lambda_i z_i^2 + \hat{c}, z \in \mathbb{R}^N$$

Como $\hat{f}(z) = f(x)$ bajo la transformación z , se puede hacer un análisis de f en \hat{f} . Si H es definida positiva, entonces todos sus valores propios son positivos y el recorrido de \hat{f} es $[\hat{c}, \infty)$. Veamos que $z = 0$ es el mínimo global fuerte de \hat{f} . Si H es definida negativa, entonces el recorrido de \hat{f} es $(-\infty, \hat{c}]$ y $z = 0$ es un máximo global fuerte. Si H tiene valores propios positivos y negativos, entonces el recorrido de \hat{f} es $(-\infty, \infty)$ y \hat{f} tiene un punto estacionario pero no posee mínimo o máximo allí. Las propiedades de los funcionales cuadráticos son comparables para cualquier funcional no cuadrático en $C^2(S)$. Para ver esto, se necesita comparar únicamente la expansión de Taylor

$$f(x + h) = f(x) + (\nabla f(x))^T h + \frac{1}{2}h^T H(x)h$$

Con la expansión de Taylor para un funcional cualquiera en $C^2(S)$

$$f(x + h) = f(x) + (\nabla f(x))^T h + \frac{1}{2}h^T H(x)h + o(\|h\|)^2$$

En una vecindad suficientemente pequeña de x el funcional arbitrario se comporta de manera similar a un funcional cuadrático. Haciendo $h = ty$ donde $\|y\| = 1$ y recordando el calculo hecho para la derivada direccional mediante la regla de la cadena, se tiene cualquier funcional cuadrático la identidad:

$$f(x + ty) = f(x) + tf^1(x; y) + \frac{1}{2}f^2(x; y)$$

Esta expresión establece una conexión entre los funcionales cuadráticos definidos sobre \mathbb{R}^N y los funcionales cuadráticos definidos sobre espacios de funciones (funciones definidas en un espacio de funciones en \mathbb{R}).

Variación de funcionales

Definición 2.12. Sea el funcional $f : V \rightarrow \mathbb{R}$. El conjunto V se llama **conjunto de funciones admisibles de f** , que es algún conjunto de funciones reales continuas definidas en un intervalo cerrado $[a, b]$. Los elementos de V los denotaremos por $u = u(x)$, $a \leq x \leq b$.

Para $u \in V$ y $\epsilon > 0$, la vecindad de u de radio ϵ

$$B_\epsilon(u) = \{v \in V : 0 \leq \|u - v\|_2 < \epsilon\}$$

A continuación, presentaremos algunas definiciones básicas acerca de la teoría de la variación de funcionales.

Definición 2.13. Sea f un funcional definido sobre V . Entonces $\hat{u} \in V$ es un **minimal local** de f , si existe $\epsilon > 0$ tal que

$$f(\hat{u}) \leq f(u), \forall u \in B_\epsilon(\hat{u}) \cap V$$

Definición 2.14. Sea f un funcional definido sobre V . Entonces $\hat{u} \in V$ es un **minimal local fuerte** de f , si existe $\epsilon > 0$ tal que

$$f(\hat{u}) < f(u), \forall u \in B_\epsilon(\hat{u}) \cap V, u \neq \hat{u}$$

Definición 2.15. Sea f un funcional definido sobre V . Entonces $\hat{u} \in V$ es un **minimal global** de f , si existe $\epsilon > 0$ tal que

$$f(\hat{u}) \leq f(u), \forall u \in V$$

Definición 2.16. Sea f un funcional definido sobre V . Entonces $\hat{u} \in V$ es un **minimal global fuerte** de f , si existe $\epsilon > 0$ tal que

$$f(\hat{u}) < f(u), \forall u \in V, u \neq \hat{u}$$

Los **maximales locales**, **maximales locales fuertes**, **maximales globales** y los **maximales globales fuertes** se definen de manera análoga.

Los **extremos locales o globales de un funcional** son los valores máximos o mínimos locales o globales del funcional.

Definición 2.17. Sea el conjunto \tilde{V} definido por

$$\tilde{V} = \{\eta : \eta = u - v, v \in V\}$$

el cual es un espacio vectorial. Este conjunto \tilde{V} es conocido como **conjunto de las funciones de prueba** (o conjunto de las funciones de perturbación).

Así, el conjunto V puede ser determinado en función de \tilde{V} :

$$V = \{v : v = u^* + \eta, \eta \in \tilde{V}\}$$

donde u^* es un elemento arbitrario pero fijo de V . La vecindad $B_\epsilon(u)$ definida anteriormente, también se puede escribir de una forma equivalente:

$$B_\epsilon(u) = \{u \in V : v = u + t\eta, \eta \in \tilde{V}, \|\eta\|_2 = 1, t \in [0, \epsilon]\}$$

Definición 2.18. Sea f un funcional definido sobre V , $u \in V$ y $\eta \in \tilde{V}$ dado, donde $\|\eta\|_2 = 1$, y supongamos que para algún $t_0 > 0$ la función $f(u + t\eta)$, $|t| < t_0$ tiene derivadas continuas hasta de orden m con respecto a t . Luego La **derivada direccional** de orden m de f en la dirección η es

$$f^{(m)}(u; \eta) = \left. \frac{d^m f(u + t\eta)}{dt^m} \right|_{t=0}$$

Definición 2.19. Sea f definida sobre V y supongamos que para algún $\hat{u} \in V$

$$f^{(1)}(\hat{u}; \eta) = 0, \forall \eta \in \tilde{V}, \|\eta\|_2 = 1$$

entonces \hat{u} se conoce como **punto estacionario** de f . También se dice que f tiene un punto estacionario en \hat{u} . Aquí $\|\cdot\|_2$ es una norma en \tilde{V} .

Teorema 2.7. Sea f definida en V y supongamos que para algún $\hat{u} \in V$ la primera derivada $f^{(1)} = (\hat{u}; \eta) = 0$, para todas las direcciones η . si \hat{u} es un minimal local de f , entonces f tiene un punto estacionario en \hat{u} (ver [12]).

Demostración. demostraremos este teorema por contradicción. Supongamos así que :

$$f(\hat{u} + t\eta) = f(\hat{u}) + tf^{(1)}(\hat{u}, \eta) + o(t)$$

para cualquier $\eta \in \tilde{V}$ con $\|\eta\|_2 = 1$. Si \hat{u} no es un punto estacionario de f tenemos que $f^{(1)}(\hat{u}, \eta) \neq 0$ para algún η . Entonces existe $t_0 > 0$ tal que $tf(\hat{u}, \eta) + o(t) < 0$ para todo $t \in (-t_0, 0) \cup (0, t_0)$, dependiendo del signo de $f^{(1)}(\hat{u}, \eta)$. De aquí que en toda vecindad $B_\epsilon(\hat{u})$ de \hat{u} existe un punto $\hat{u} + t\eta$ tal que $f(u) < f(\hat{u})$, y por tanto \hat{u} no puede ser un minimal local. \square

Definición 2.20. Un funcional $f : V \longrightarrow \mathbb{R}$ es un **funcional cuadrático** si satisface la identidad

$$f(u + t\eta) = f(u) + tf^{(1)}(u, \eta) + \frac{1}{2}t^2 f^{(2)}(u, \eta)$$

para todo $u \in V$ y todo $\eta \in \tilde{V}$, con $\|\eta\|_2 = 1$, y todo $t \in \mathbb{R}$.

Supongamos que f es un funcional con la propiedad de que $f(u + t\eta)$ es un polinomio de grado 2 respecto a t . El calculo directo de $f^{(1)}(u, \eta)$ y $f^{(2)}(u, \eta)$ muestra que f es efectivamente un funcional cuadrático.

Teorema 2.8. *Sea f un funcional cuadrático definido en V . Entonces $\hat{u} \in V$ es el único minimal global y local fuerte de f si*

$$\begin{aligned} f^{(1)}(u, \eta) &= 0, \forall \eta \in \tilde{V}, \|\eta\|_2 = 1 \\ f^{(2)}(u, \eta) &> 0, \forall \eta \in \tilde{V}, \|\eta\|_2 = 1 \end{aligned}$$

La prueba de este resultado es evidente a partir de la definición de funcional cuadrático dada anteriormente. Citaremos el siguiente teorema dado que será de gran utilidad para tratar la forma variacional de la ecuación de Euler-Lagrange.

Teorema 2.9. *Sea $\psi(x)$ una función continua en $[a, b]$. Si se cumple que $\forall \eta \in C[a, b]$ se tiene $\int_a^b \psi(x)\eta dx = 0$, entonces $\psi(x)$ es nula en $[a, b]$ (ver [12]).*

Demostración. supongamos que $\psi(x) \neq 0$, en algún punto $x_0 \in (a, b)$ y además que $\int_a^b \psi(x)\eta dx = 0$. Por la continuidad de $\psi(x)$ en x_0 tenemos que existe una vecindad $B_\epsilon(x_0) - \{x_0\}$ con $\epsilon > 0$ en la que $\psi(x)$ no se anula y tiene signo constante. Ahora es fácil encontrar una función $\eta \in \tilde{V}$ que se desvanece fuera de $B_\epsilon(x_0)$ y que sea positiva en $B_\epsilon(x_0)$. Entonces $\int_a^b \psi(x)\eta dx \neq 0$ para este η . □

2.1.2. La ecuación de Euler-Lagrange

Supongamos que $F(x, r, s)$ es un campo escalar (funciones de varias variables) definido en $[a, b] \times \mathbb{R}^2$ con derivadas parciales continuas hasta de orden dos inclusive. Un problema clásico del calculo variacional es el siguiente: Calcular una función v del conjunto $[a, b] \rightarrow \mathbb{R}$ que tenga derivadas continuas hasta de orden 2 inclusive y que cumpla las condiciones de frontera $v(a) = \alpha, v(b) = \beta$, y que además esa función $v \in C^2[a, b]$ sea el mínimo del funcional:

$$f(u) = \int_a^b F(x, u, u') dx$$

Ahora, de acuerdo a la definición de derivada direccional sobre un funcional f , derivamos la función

$$f(u + t\eta) = \int_a^b F(x, u + t\eta, u' + t\eta') dx$$

respecto a t para obtener

$$\frac{df(u + t\eta)}{dt} = \int_a^b \frac{dF(x, u + t\eta, u' + t\eta')}{dt} dx$$

$$\int_a^b \frac{dF(x, u + t\eta, u' + t\eta')}{dt} dx = \int_a^b \nabla F(x, u + t\eta, u' + t\eta') \frac{d(x, u + t\eta, u' + t\eta')}{dt} dx$$

$$\int_a^b \nabla F(x, u + t\eta, u' + t\eta') \frac{d(x, u + t\eta, u' + t\eta')}{dt} dx = \int_a^b \left[\frac{\partial F}{\partial x}, \frac{\partial F}{\partial(u + t\eta)}, \frac{\partial F}{\partial(u' + t\eta')} \right] \cdot (0, \eta, \eta') dx$$

$$\int_a^b \left[\frac{\partial F}{\partial x}, \frac{\partial F}{\partial(u + t\eta)}, \frac{\partial F}{\partial(u' + t\eta')} \right] \cdot (0, \eta, \eta') dx = \int_a^b \left[\frac{\partial F}{\partial(u + t\eta)} \eta + \frac{\partial F}{\partial(u' + t\eta')} \eta' \right] dx$$

Al hacer $t = 0$ tenemos que

$$f^{(1)}(u, \eta) = \int_a^b \left[\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right] dx, \forall \eta \in \tilde{V}$$

Ahora, de la definición de punto estacionario de un funcional, vemos que f tiene un punto estacionario en algún $u \in V$ si y solo si

$$\int_a^b \left[\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right] dx = 0, \forall \eta \in \tilde{V}$$

La suavidad de las funciones u y F es suficiente para permitir integrar por partes el segundo termino y usar así las condiciones de frontera $\eta(a) = \eta(b) = 0$ que cumplen absolutamente todas las funciones de \tilde{V} y obtener:

$$\int_a^b \frac{\partial F}{\partial u'} \eta' dx = \left[\frac{\partial F}{\partial u'} \eta \right]_a^b - \int_a^b \frac{d}{dx} \left[\frac{\partial F}{\partial u'} \right] \eta dx = \int_a^b \frac{d}{dx} \left[\frac{\partial F}{\partial u'} \right] \eta dx$$

Combinando este resultado con

$$\int_a^b \left[\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right] dx = 0$$

obtenemos:

$$f^{(1)}(u, \eta) = \int_a^b \left[\frac{\partial F}{\partial u} - \frac{d}{dx} \left[\frac{\partial F}{\partial u'} \right] \right] \eta dx = 0, \forall \eta \in \tilde{V}$$

como la función del paréntesis es necesariamente continua, por el Teorema 1.20 y utilizando el teorema 1.21 se concluye que $u \in V$ es un punto estacionario de f si y solo si u satisface

$$-\frac{d}{dx} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} = 0, a < x < b$$

Esta ecuación es conocida como **ecuación de Euler-Lagrange** para el funcional

$$f(u) = \int_a^b F(x, u, u') dx$$

Esta ecuación es usualmente una ecuación diferencial ordinaria de segundo orden en u y de esta forma, el problema de hallar una función u que cumpla la ecuación de Euler-Lagrange y las condiciones de frontera $u(a) = \alpha, u(b) = \beta$ se reduce al clásico problema con valores en la frontera

$$-\frac{d}{dx} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} = 0, a < x < b$$

$$u(a) = \alpha, u(b) = \beta$$

Así, podemos afirmar que resolver el problema de valores en la frontera es equivalente a encontrar un punto estacionario de

$$f(u) = \int_a^b F(x, u, u') dx$$

en el espacio V , es decir, una función que cumpla

$$\int_a^b \left[\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right] dx = 0, \forall \eta \in \tilde{V}$$

2.1.3. Condiciones de frontera naturales y esenciales

Desde la elección del espacio de funciones admisibles para f en $V = \{v \in C^2[a, b] : v(a) = \alpha, v(b) = \beta\}$, se ha requerido que cada $u \in V$ cumpla con las condiciones de frontera, lo cual implica que todo punto estacionario de f cumple también estas condiciones. Así, tenemos una pregunta: si no imponemos condiciones de frontera sobre las funciones admisibles, entonces ¿que condiciones debe cumplir un punto estacionario de tal funcional?. Para responder esta pregunta consideremos el funcional

$$f(u) = \int_a^b F(x, u, u') dx, V = C^2[a, b]$$

Tenemos el espacio de funciones de prueba \tilde{V} definido por:

$$\tilde{V} = \{v \in C^2[a, b] : v(a) = v(b) = 0\}$$

con $V = C^2[a, b] = \tilde{V}$ pero de esto no puede afirmarse que $\eta(a) = \eta(b) = 0 \forall \eta \in V$. Como tenemos que

$$\int_a^b \left[\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right] dx = 0, \forall \eta \in \tilde{V}$$

e integrando por partes, se obtiene que la condición para que u sea punto estacionario de f es (Esto se supone que se cumple para alguna función admisible u):

$$\int_a^b \left[\frac{\partial F}{\partial u} + \frac{d}{dx} \left[\frac{\partial F}{\partial u'} \right] \right] \eta dx + \left[\frac{\partial F}{\partial u'} \eta \right]_a^b = 0, \forall \eta \in \tilde{V}$$

Esta expresión se cumple para toda función $\eta \in C^2[a, b]$, pero en realidad debe cumplirse para el subconjunto $C^2[a, b]$ para el cual $\eta(a) = \eta(b) = 0$. Pero como se sabe, esto implica que

$$-\frac{d}{dx} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} = 0, a < x < b$$

y por nuestro resultado

$$\int_a^b \left[\frac{\partial F}{\partial u} + \frac{d}{dt} \left[\frac{\partial F}{\partial u'} \right] \right] \eta dx + \left[\frac{\partial F}{\partial u'} \eta \right]_a^b = 0, \forall \eta \in \tilde{V}$$

tenemos que

$$\left[\frac{\partial F}{\partial u'} \eta \right]_a^b = 0, \forall \eta \in \tilde{V}$$

Pero ahora tenemos que, en este problema hay tres casos acerca de la restricción sobre las funciones admisibles:

1. Cuando no existe ninguna restricción sobre las funciones admisibles, es decir $V = \tilde{V} = C^2[a, b]$. Como no existen restricciones sobre los valores de $\eta(a)$ y $\eta(b)$, cuando η es cualquier función arbitraria en \tilde{V} debe tenerse que (para $x = a, x = b$):

$$\frac{\partial F}{\partial u'} \eta = 0$$

Luego, todo punto estacionario de f debe satisfacer estas condiciones de frontera (además de satisfacer la ecuación de Euler-Lagrange); así, nuestro problema ahora es: calcular $u \in C^2[a, b]$ tal que

$$\begin{aligned} -\frac{d}{dt} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} &= 0 & a < x < b \\ \frac{\partial F}{\partial u'} \eta \Big|_{x=a} &= 0 & \frac{\partial F}{\partial u'} \eta \Big|_{x=b} &= 0 \end{aligned}$$

Veamos que estas condiciones no se imponen sobre las funciones admisibles, sino sobre F , así estas condiciones son llamadas **condiciones de frontera naturales** para

$$f(u) = \int_a^b F(x, u, u') dx$$

Las condiciones $u(a) = \alpha, u(b) = \beta$ que se imponen sobre las funciones admisibles son llamadas **condiciones de frontera esenciales** para f .

2. Cuando existen restricciones sobre las funciones admisibles, es decir cuando

$$V = \{v \in C^2[a, b] : v(a) = \alpha, v(b) = \beta\}$$

$$\tilde{V} = \{\eta \in C^2[a, b] : \eta(a) = \eta(b) = 0\} = C_0^2[a, b]$$

Aquí las funciones de prueba deben cumplir las condiciones de frontera nulas. Para este caso el problema de calcular un punto estacionario de f en V se transforma en el problema de calcular una función $u \in C^2[a, b]$ tal que

$$\begin{aligned} -\frac{d}{dt} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} &= 0 & a < x < b \\ u(a) &= \alpha & u(b) &= \beta \end{aligned}$$

3. Cuando se presenta una condición mixta, es decir cuando alguna de las condiciones de frontera esencial no es impuesta a las funciones admisibles. Para este caso, el punto estacionario de f debe satisfacer la correspondiente condición de frontera natural. Por ejemplo, si únicamente la primera condición de frontera se impone, tenemos que

$$V = \{v \in C^2[a, b] : v(a) = \alpha\}$$

$$\tilde{V} = \{\eta \in C^2[a, b] : \eta(a) = 0\}$$

y además

$$\frac{\partial F}{\partial u'} \eta \Big|_a^b = 0, \forall \eta \in \tilde{V}$$

Esto implica que

$$\eta(b) \frac{\partial F}{\partial u'} \Big|_{x=b} = 0, \forall \eta \in \tilde{V}$$

Esto significa que

$$\frac{\partial F}{\partial u'} \Big|_a^b = 0$$

pues no existen restricciones sobre los valores de $\eta(b)$. De esto concluimos que existe una condición de frontera natural correspondiente a una condición de frontera esencial omitida. Luego, nuestro problema de calcular un punto estacionario de f en V se transforma en el problema de calcular una $u \in C^2[a, b]$ tal que

$$\begin{aligned} -\frac{d}{dt} \left(\frac{\partial F}{\partial u'} \right) + \frac{\partial F}{\partial u} &= 0 & a < x < b \\ u(a) &= \alpha & \frac{\partial F}{\partial u'} \Big|_{x=b} &= 0 \end{aligned}$$

El concepto de Completación

Ya vimos que a partir de un problema variacional podemos obtener su equivalente PVF, pero comúnmente, en aplicaciones de física y de ingeniería se presenta el problema inverso: Dado un PVF, calcular su equivalente problema variacional, así que analizaremos a

continuación este tipo de problema. Deseamos encontrar el funcional asociado al siguiente problema: dada $g \in C^1[a, b]$, tenemos que encontrar una función u que verifique

$$\begin{aligned} -u''(x) + u &= g \\ u(a) = u(b) &= 0 \end{aligned}$$

la solución clásica de este problema es una función de $C^2[a, b]$ que verifica la ecuación diferencial y las condiciones de frontera. Veamos que

$$V = \tilde{V} = \{v \in C^2[a, b]; v(a) = v(b) = 0\} = C_0^2[a, b]$$

al multiplicar la ecuación diferencial

$$-u''(x) + u = g$$

por un $\eta \in C^1[a, b]$, integrando sobre el dominio y usando la integración por partes, para transferir así la derivada de u a η , obtenemos

$$\int_a^b (u'\eta' + u\eta)dx = \int_a^b g\eta dx$$

para todo $\eta \in C^1[a, b]$, tal que $\eta(a) = \eta(b) = 0$, es decir $\eta \in C_0^1[a, b]$. Veamos que η , además de cumplir las condiciones de frontera homogénea, debe también cumplir las condiciones:

$$\int_a^b (u'\eta')dx = - \int_a^b u''\eta dx < \infty$$

$$\int_a^b (u'\eta')dx < \infty$$

Pero estas dos desigualdades no siempre se cumplen. Así, vemos que esto está relacionado con la **completitud del espacio** V y la teoría de **espacios de Sóbolev**. Tal completitud tiene por ventaja la aplicación del teorema de Lax-Milgram, para asegurar así la existencia y unicidad de la solución \hat{u} de la formulación variacional de un PVF con ciertas condiciones. De esta manera, podemos observar a grandes rasgos que el enfoque variacional aplicado a las ecuaciones diferenciales consta de 4 pasos:

1. Se establece la noción de **solución débil**, lo cual implica hacer uso de la teoría referente a los espacios de Sóbolev.
2. Se establece la existencia y unicidad de la solución débil mediante el método variacional, utilizando así el teorema de Lax-Milgram.
3. Posteriormente, se prueba que esta solución es una solución de clase $C^2[a, b]$.
4. Se debe probar que toda solución débil de clase $C^2[a, b]$ es una solución clásica.

Problemas de valores en la frontera: soluciones fuertes-soluciones débiles

Ahora analizaremos el siguiente problema de valores en la frontera: si se tiene una función $h(x) \in C(\Omega)$, o si $h(x) \in L^2(\Omega)$, entonces hay que encontrar una función $u(x)$ que satisfaga el siguiente problema:

$$\begin{aligned} -u''(x) &= h(x) - u(x) \quad , \Omega = (0, 1) \\ u(0) &= u(1) = 0 \end{aligned} \tag{2.1}$$

Definición 2.21. Una *solución clásica* del problema:

$$\begin{aligned} -u''(x) &= h(x) - u(x) \\ u(0) &= u(1) = 0 \end{aligned}$$

con $h(x) \in C(\Omega)$, $\Omega = (0, 1)$ es una función $u(x) \in C^2[0, 1]$ que verifica esta ecuación diferencial en el sentido usual.

Si $u(x)$ es una solución fuerte de este problema donde $v(x)$ es una función tal que $v(0) = v(1) = 0$, y si multiplicamos el término $u''(x)$ por $v(x)$, luego integramos por partes, obtenemos:

$$\int_{\Omega} u''(x)v(x)dx = u'(x)v(x) \Big|_0^1 - \int_{\Omega} u'(x)v'(x)dx$$

como $v(0) = v(1) = 0$ entonces:

$$\int_{\Omega} u''(x)v(x)dx = - \int_{\Omega} u'(x)v'(x)dx$$

Luego, si $u(x)$ es una solución fuerte de este problema y multiplicamos este problema por tal función $v(x)$ e integramos por partes, obtenemos:

$$\int_0^1 u'(x)v'(x)dx + \int_0^1 u(x)v(x)dx = \int_0^1 h(x)v(x)dx$$

Ahora, estas integrales también se pueden expresar mediante:

$$B(u, v) = (u'(x), v'(x)) + (u(x), v(x)) = \int_0^1 u'(x)v'(x)dx + \int_0^1 u(x)v(x)dx$$

$$G(v) = \int_0^1 h(x)v(x)dx$$

Donde (\cdot, \cdot) denota el producto interno en $L^2(0, 1)$

Definición 2.22. Una *solución débil* al problema de valores en la frontera:

$$\begin{aligned} -u''(x) &= h(x) - u(x) \\ u(0) &= u(1) = 0 \end{aligned}$$

con $h(x) \in C(\Omega)$, $\Omega = (0, 1)$, es una función $u(x) \in H^1(0, 1)$ tal que

$$B(u, v) = G(v), \forall v(x) \in H_0^1(0, 1)$$

Luego, el problema de valores de frontera se reduce a encontrar una función $v(x)$ para tal formulación variacional.

Debido a la formulación de integración por partes, se puede concluir que toda solución clásica del problema de valores en la frontera es una solución débil de la formulación. Probaremos que el funcional G es acotado. Se sabe que la forma $B(\cdot, \cdot)$ es bilineal simétrica en $H_0^1(0, 1)$. Como:

$$B(v, v) = (v'(x), v'(x)) + (v(x), v(x)) = \int_0^1 [v'(x)]^2 dx + \int_0^1 [v(x)]^2 dx$$

Donde $v \in H_0^1(0, 1)$. Tenemos entonces que la forma bilineal $B(\cdot, \cdot)$ es coerciva, y como $B(\cdot, \cdot)$ es un producto interno en $H_0^1(0, 1)$ se cumple que la desigualdad de Cauchy-Schwarz en $H_0^1(0, 1)$:

$$|B(u, v)| \leq B(u, u)^{\frac{1}{2}} B(v, v)^{\frac{1}{2}}, \forall u, v \in H_0^1(0, 1)$$

Como

$$B(v, v) = \|v\|_1^2, \forall v \in H_0^1(0, 1)$$

entonces $B(\cdot, \cdot)$ es continua en $H_0^1(0, 1)$, luego

$$|B(u, v)| \leq \|u\|_1 \|v\|_1, \forall u, v \in H_0^1(0, 1)$$

Además, si $v \in H_0^1$ entonces existe un compacto $[a, b]$ en $(0, 1)$ tal que:

$$|G(v)| \leq \int_0^1 |h(x)| |v(x)| dx \leq \sup_{x \in [a, b]} |h(x)| \|v\|_2 \leq \sup_{x \in [a, b]} |h(x)| \|v\|_2 + \|v'\|_2^2 \leq k \|v\|_1^2$$

aquí $k = \max\{\sup_{x \in [a, b]} |h(x)|, 1\}$. Luego el funcional lineal G es acotado en $H_0^1(0, 1)$.

Demostraremos a continuación la existencia y la unicidad de la solución débil.

Teorema 2.10. (Existencia y unicidad de la solución débil). Para cada $h \in L^2(0, 1)$, existe una única solución u de

$$B(u, v) = G(v), \forall v(x) \in H_0^1(0, 1)$$

tal que $u \in H_0^1(0, 1)$. También se tiene que:

$$I(u) = \min_{v \in H_0^1(0, 1)} I(v)$$

donde I es el funcional lineal

$$I(v) = \frac{1}{2} B(v, v) - G(v)$$

Esta última expresión se conoce como Principio de Dirichlet (ver [12]).

Demostración. dado que la forma bilineal $B(\cdot, \cdot)$ es coerciva y acotada, y el funcional $G(\cdot)$ es lineal, así que el resultado se sigue del teorema de Lax-Milgram. La simetría de $B(\cdot, \cdot)$ garantiza que u es el mínimo del funcional $I(\cdot)$ en $H_0^1(0, 1)$ (ver [10]). \square

Ahora presentamos otra versión del Teorema de representación de Riesz y del Teorema de Lax-Milgram.

Formulación variacional discreta

Teorema 2.11. (Teorema de representación de Riesz) Para todo funcional lineal continuo $G : V \rightarrow \mathbb{R}$ existe un único elemento $u \in V$ tal que

$$(u, v) = G(v), \forall v \in V$$

El elemento u es también determinado como la solución del problema variacional:

$$f(v) = \frac{1}{2}(v, v) - G(v) \rightarrow \text{mín}, \forall v \in V$$

o también

$$f(u) = \text{mín}_{v \in V} f(v)$$

Teorema 2.12. (Teorema de Lax-Milgram) Sea $B(\cdot, \cdot)$ una forma bilineal continua y definida positiva sobre un espacio de Hilbert V con producto interno (\cdot, \cdot) . Para todo funcional lineal continuo de V en \mathbb{R} existe un único $v \in V$ tal que

$$B(u, v) = G(v), \forall v \in V$$

Si adicionalmente $B(\cdot, \cdot)$ es simétrica, u está dada como la solución del problema variacional

$$f(v) = \frac{1}{2}B(v, v) - G(v) \rightarrow \text{mín}, \forall v \in V$$

Ahora veremos como solucionar un problema variacional descrito por un problema de valores en la frontera. El problema

$$\begin{aligned} -u''(x) + u &= g, \quad g \in C^1[a, b] \\ u(a) &= u(b) = 0 \end{aligned}$$

tiene la formulación integral :

$$\int_a^b (u'\eta' + u\eta)dx = \int_a^b g\eta dx$$

donde $u \in V$, de tal forma que, integrando por partes sobre esta formulación, obtenemos:

$$\int_a^b (-u'' + u - g)\eta dx = 0, \eta \in C_0^1[a, b]$$

Como $C_0^1[a, b]$ es denso en $L^2[a, b]$ se concluye que $-u'' + u - g = 0$, en todo punto, ya que $u \in C^2[a, b]$. Para comprender el porqué es necesario cumplir estas anteriores etapas y además, para comprender el porqué es necesario extender el conjunto de funciones admisibles hacia el espacio de Sóbolev, consideraremos el funcional

$$f(u) = \int_a^b \left[\frac{1}{2}p(x)(u')^2 + \frac{1}{2}q(x)(u)^2 - g(x)u \right] dx$$

$$u \in V, p(x) \in C^1[a, b], q(x), g(x) \in C[a, b]$$

donde

$$V = \{v \in C^2[a, b] : v(a) = v(b) = 0\} = C_0^2[a, b]$$

El espacio de funciones de prueba es $V = \tilde{V}$. Sea la forma bilineal

$$B : V \times V \longrightarrow \mathbb{R}$$

la cual esta definida como

$$B(u, v) = \int_a^b [p(x)u'v' + q(x)uv]$$

y sea

$$G : V \longrightarrow \mathbb{R}$$

el funcional definido por

$$G(u) = \int_a^b g(x)u dx = (g, u)$$

Veamos que $B(u, v)$ es una forma bilineal que es simétrica, continua y coerciva. Además, el funcional $G(v)$ es un funcional lineal continuo de V en \mathbb{R} . De acuerdo a esto, veamos que f se puede expresar por

$$f(u) = \frac{1}{2}B(u, u) - G(u), \forall u \in V$$

Veamos que

$$B(u + \eta, u + \eta) = B(u, u) + 2B(u, \eta) + B(\eta, \eta)$$

$$G(u, \eta) = G(u) + G(\eta)$$

de lo cual se tiene la identidad

$$f(u + \eta) = f(u) + B(u, \eta) - G(\eta) + \frac{1}{2}B(\eta, \eta)$$

Si reemplazamos en esta identidad a η por $t\eta$ donde $\|\eta\| = 1$ y $t \in \mathbb{R}$ se tiene que

$$f(u + \eta) = f(u) + t[B(u, \eta) - G(\eta)] + \frac{t^2}{2}B(\eta, \eta)$$

cuya primer derivada direccional es

$$f^{(1)}(u; \eta) = B(u, \eta) - G(\eta)$$

Luego, una función $u \in V$ es un punto estacionario de f si $f^{(1)}(u; \eta) = 0, \forall \eta \in V, \|\eta\| = 1$.
Luego

$$B(u, \eta) = G(\eta), \forall \eta \in V$$

Definamos el operador $L : V \rightarrow C[a, b]$ como

$$Lu = -[p(x)u']' + q(x)u$$

el cual es lineal y hace el uso del producto interno en $L^2[a, b]$. Luego podemos decir que

$$G(\eta) = (g, \eta)$$

$$B(u, \eta) = (Lu, \eta)$$

Integrando por partes y como $\eta \in \tilde{V}$ se tiene que

$$B(u, \eta) = \int_a^b [Lu] \eta dx$$

Como sabemos que $B(u, \eta) = G(\eta), \forall \eta \in V$, entonces

$$(Lu - g, \eta) = 0, \forall \eta \in V$$

Esto equivale a $Lu = g$ en (a, b) . Observemos que esta es la ecuación para nuestro funcional $f(u)$, y $B(u, \eta) = G(\eta), \forall \eta \in V$ es la formulación variacional del siguiente PVF: Calcular una función $u \in C^2[a, b]$ con $a < x < b$, y $u(a) = u(b) = 0$, tal que

$$Lu = -[p(x)u']' + q(x)u = g$$

Este problema (según la teoría de ecuaciones diferenciales) tiene una única solución $\hat{u} \in V$. Veamos que $B(u, \eta) = G(\eta), \forall \eta \in V$ se mantiene para $u = \hat{u}$. Una reformulación sencilla de

$$f(u + \eta) = f(u) + B(u, \eta) - G(\eta) + \frac{1}{2}B(\eta, \eta)$$

admite la identidad

$$f(u) = f(\hat{u}) + \frac{1}{2}B(u - \hat{u}, u - \hat{u})$$

y como $B(\eta, \eta) > 0, \forall \eta \in V, \eta \neq 0$ tenemos que

$$\min_{u \in V} f(u) = f(\hat{u})$$

Ahora extenderemos el dominio de f del espacio V a un espacio \bar{V} mas grande, tal que toda función de \bar{V} puede ser aproximada por funciones de V , y además, que la solución \hat{u} del PVF con $a < x < b$, y $u(a) = u(b) = 0$, tal que

$$Lu = - [p(x)u']' + q(x)u = g$$

sea la única solución del siguiente problema variacional: calcular $u \in \bar{V}$ tal que

$$B(u, \eta) = G(\eta), \forall \eta \in \bar{V}$$

y también que

$$\min_{u \in V} f(u) = f(\hat{u}) = \min_{u \in \bar{V}} f(u)$$

Esta ultima condición exige establecer la regularidad de la solución del problema variacional. Recordemos que $L^2[a, b]$ es completo con respecto a la norma $\|\cdot\|$, en el sentido que si $\{u_n\}_{n=1}^{\infty} \subset L^2[a, b]$ es una sucesión de Cauchy en esta norma, es decir, si

$$\lim_{m, n \rightarrow \infty} \|u_m - u_n\| = 0$$

entonces existe $u \in L^2[a, b]$ tal que $\lim_{n \rightarrow \infty} \|u - u_n\| = 0$. Sea

$$(u, v)_L = B(u, v), u, v \in V$$

$$\|u\|_L = (u, u)_L^{\frac{1}{2}} = B(u, u)^{\frac{1}{2}}, u \in V$$

donde $(\cdot, \cdot)_L$ y $\|\cdot\|_L$ definen un producto interno y una norma sobre V . Así, a $(\cdot, \cdot)_L$ se le llama **producto interno energía** y a $\|\cdot\|_L$ se le conoce como **norma energía**, que es equivalente a la norma $L^2[a, b]$. Veamos que tenemos una dificultad: el espacio V no es completo en la norma energía $\|\cdot\|_L$, porque existen sucesiones $\{u_n\}_{n=1}^{\infty}$ tales que

$$\lim_{m, n \rightarrow \infty} \|u_m - u_n\| = 0$$

para las cuales ningún elemento $u \in V$ satisface la condición

$$\lim_{n \rightarrow \infty} \|u - u_n\| = 0$$

Luego, podemos decir que en V con la norma energía existen sucesiones de Cauchy que no convergen en V . A fin de Superar esta dificultad, consideremos en V una sucesión $\{u_n\}_{n=1}^{\infty}$ de Cauchy. Segun [10], tenemos el siguiente resultado: Si p y q cumplen las condiciones

$$f(u) = \int_a^b \left[\frac{1}{2}p(x)(u')^2 + \frac{1}{2}q(x)(u)^2 - g(x)u \right] dx$$

$$u \in V, p(x) \in C^1[a, b], q(x), g(x) \in C[a, b]$$

para todo $u \in V$, donde el conjunto V está determinado por:

$$V = \{v \in C^2[a, b] : v(a) = v(b) = 0\} = C_0^2[a, b]$$

con

$$0 < p_0 \leq p(x) \leq p_1, 0 \leq q(x) \leq q_1, a \leq x \leq b$$

donde p_0, p_1 y q_1 son escalares, entonces

$$\|\cdot\|_L^2 = \rho \|u'\|_2^2 + \rho \|u\|_2^2$$

donde

$$\rho = \rho_0 \min \left\{ \frac{1}{2}, \frac{1}{(b-a)^2} \right\} > 0$$

Luego, por este resultado y considerando el hecho de que $\{u_n\}_{n=1}^\infty$ es una sucesión de Cauchy tal que

$$\lim_{m,n \rightarrow \infty} \|u_m - u_n\|_L = 0$$

$$\lim_{m,n \rightarrow \infty} \|u'_m - u'_n\|_L = 0$$

por la completitud de $L^2[a, b]$ en la norma $\|\cdot\|$, existen $u, \tilde{u} \in L^2[a, b]$ tales que

$$\lim_{n \rightarrow \infty} \|u - u_n\|_2 = 0$$

$$\lim_{n \rightarrow \infty} \|\tilde{u} - u'_n\|_2 = 0$$

Luego, se define como \overline{V} al conjunto de todas las funciones u que pueden ser obtenidas mediante este proceso, y llamamos a \tilde{u} la derivada generalizada de u . Si $u \in C^1[a, b]$ entonces $\tilde{u} = u'$. En contraste con V , \overline{V} es completo con respecto a la norma energía y por esto es un espacio de Hilbert. Según [12], se sabe que

$$\overline{V} \subset \{v \in C[a, b] : v(a) = v(b) = 0\}$$

Ahora, determinamos el siguiente producto interno y la norma:

$$(u, v)_1 = \sum_{i=0}^1 (u^{(i)}, v^{(i)}) = \int_a^b (uv + u'v') dx, u, v \in V$$

$$\|u\|_1 = (u, u)_1^{\frac{1}{2}} = \left[\int_a^b ((u')^2 + u^2) dx \right]^{\frac{1}{2}}, u \in V$$

donde

$$V = \{v \in C^2[a, b] : v(a) = v(b) = 0\} = C_0^2[a, b]$$

De esta forma, se conoce como el **producto interno de Sóbolev de orden 1** a

$$(u, v)_1 = (u, v) + (u', v')dx, u, v \in V$$

Y se denomina **la norma de Sóbolev de orden 1** a

$$\|u\|_1 = \left[\|u\|^2 + \|u'\|^2 \right]^{\frac{1}{2}}, u \in V$$

Se ha mostrado que al completar el espacio V con respecto a la norma energía se obtiene el espacio de Hilbert \overline{V} . Ahora denotaremos a \widehat{V} como la completación de V respecto a la norma de Sóbolev. Para obtener \widehat{V} consideremos aquellas sucesiones $\{u_n\}_{n=1}^{\infty}$ en V y que sean de Cauchy, es decir que

$$\lim_{m, n \rightarrow \infty} \|u_m - u_n\|_1 = 0$$

Tenemos también que

$$\lim_{m, n \rightarrow \infty} \|u_m - u_n\|_2 = 0$$

además la completez de $L^2[a, b]$ con respecto a la norma $\|\cdot\|_2$ determina la existencia de un $u \in L^2[a, b]$ tal que

$$\lim_{n \rightarrow \infty} \|u - u_n\|_2 = 0$$

El espacio \widehat{V} es así el conjunto de todas las funciones que pueden ser obtenidas por tal método. Comparando este proceso con el de la completación de V respecto a la norma energía, se ve que $\widehat{V} = \overline{V}$ si y solo si para toda sucesión $\{u_n\}_{n=1}^{\infty}$ en V se cumple que

$$\lim_{m, n \rightarrow \infty} \|u_m - u_n\|_1 = 0 \iff \lim_{m, n \rightarrow \infty} \|u_m - u_n\|_L = 0$$

La equivalencia de las dos normas implica esta doble implicación, así que $\widehat{V} = \overline{V}$. Veremos ahora una variedad de los espacios de Sóbolev, de los cuales \overline{V} es un caso especial. Sea k un entero no negativo, así definimos

$$C^{k*}(a, b) = \left\{ v \in C^k(a, b) : \int_a^b (v^{(i)})^2 dx < \infty \right\}$$

donde $i = 0, 1, \dots, k$. Sean $u, v \in C^{k*}(a, b)$. Se define el **producto interno de Sóbolev de orden k** como

$$(u, v)_k = \sum_{i=0}^k (u^{(i)}, v^{(i)}) = \int_a^b (uv + u'v' + \dots + u^{(k)}v^{(k)})dx$$

Y además definiendo así la **norma de Sóbolev de orden k**

$$\|u\|_k = (u, u)_k^{\frac{1}{2}} = \left[\int_a^b (u^2 + (u')^2 + \dots + (u^{(k)})^2) dx \right]^{\frac{1}{2}}$$

La completación de $C^{k*}(a, b)$ respecto a la norma $\|\cdot\|_k$ produce el **espacio de Sóbolev de orden k** , $H^k(a, b)$. Esta completación permite extender el dominio de definición de producto interno y norma de $C^{k*}(a, b)$ a todo $H^k(a, b)$. Denotaremos así a $C_0^k(a, b)$ al conjunto de todas las funciones de $C^k(a, b)$ con soporte compacto sobre (a, b) y se define a $H_0^k(a, b)$ como la completación de $C_0^k(a, b)$ en la norma $\|\cdot\|_k$. $H_0^k(a, b)$ es un espacio de Hilbert con respecto al producto interno y a la norma de Sóbolev donde

$$H_0^k(a, b) = \{v \in H^k(a, b) : v^{(i)}(a) = v^{(i)}(b) = 0, i = 0, 1, \dots, k-1\}$$

Aunque $C_0^1(a, b)$ no es igual a $V = C_0^2(a, b)$ la completación de los dos espacios si es igual respecto a la norma $\|\cdot\|_1$, es decir $\bar{V} = H_0^1(a, b)$. Podemos así resumir el concepto de completación de la siguiente manera: Los métodos para solucionar problemas variacionales son realmente útiles, porque por ejemplo en el caso del problema variacional

$$B(u, v) = G(v), \forall v(x) \in H_0^1(0, 1)$$

el método obtiene una sucesión de soluciones aproximadas (la cual es una sucesión de Cauchy), pero lo que garantiza que el límite de esta sucesión es la solución débil del problema variacional es la completitud de $H_0^1(0, 1)$, porque si el problema se trabaja sobre $C_0^1(0, 1)$ se tiene el inconveniente de que este espacio $C_0^1(0, 1) = \{v \in C^1(0, 1) \mid v(0) = v(1) = 0\}$ no es completo en la norma de $L^2(0, 1)$ y por lo tanto tampoco es completo en la norma de Sóbolev. El espacio de $C_0^1(0, 1)$ se puede completar en la norma de Sóbolev de orden 1, obteniendo así el espacio de Sóbolev $H_0^1(0, 1)$ (ver [10]), en el cual existe una solución única y débil del problema de valores en la frontera. A continuación describiremos el método de Ritz-Galerkin para calcular la solución aproximada a un problema variacional.

2.1.4. El método de Ritz-Galerkin

Método de Ritz

Sea $V_N \subset H_0^1$ un espacio vectorial de dimension finita. Una **solución de Ritz** al problema con valores en la frontera descrito anteriormente es una función $u_N \in V_N$ tal que:

$$B(u_N, v) = (h, v), \forall v(x) \in V_N$$

De esta manera el problema de aproximación de Ritz sobre el problema de valores en la frontera, consiste en encontrar una solución $u_N \in V_N$ del problema $B(u_N, v) = (h, v), \forall v(x) \in V_N$. Aquí se utilizara el Teorema de Lax-Milgram para obtener una solución del problema variacional en el espacio de dimension finita V_N de $H_0^1(0, 1)$. Sabemos que V_N es también un espacio de Hilbert, así que el Teorema de Lax-Milgram garantiza que la solución del problema variacional en V_N existe y es única. Supondremos ahora que existe una forma

bilineal, continua y coerciva $B(\cdot, \cdot)$ y un funcional lineal continuo $G(\cdot)$, ambos definidos sobre un espacio de Hilbert V tal que la solución del problema de valores en la frontera es idéntica al elemento $\hat{u} \in V$, donde la existencia de \hat{u} es garantizada por el Teorema de Lax-Milgram. De tal forma, que este elemento satisface la condición

$$B(\hat{u}, v) = G(v), \forall v \in V$$

y si $B(\cdot, \cdot)$ es simétrica entonces

$$\min_{u \in V} f(u) = f(\hat{u})$$

donde tenemos que

$$f(u) = \frac{1}{2}B(u, u) - G(u), \forall u \in V$$

De esta forma, la idea central del método de Ritz consiste en que, al determinar el extremo del funcional $f(u)$ se consideran de las funciones admisibles solo aquellas que se puedan expresar como la combinación lineal

$$u_N(x) = \sum_{i=1}^n \alpha_i \phi_i(x)$$

donde los α_i son constantes y el sistema $\{\phi_i(x)\}$ esta formado por funciones linealmente independientes. Usualmente los métodos para obtener soluciones aproximadas requieren de algún tipo de discretización, de tal forma que para el método de Ritz (para considerar el teorema de Lax-Milgram) la discretización exige la selección de un subespacio V_N de V de dimension N . Como V_N es un espacio de Hilbert las condiciones del teorema de Lax-Milgram se satisfacen cuando V es reemplazado por V_N . Así, procederemos a calcular la aproximación de \hat{u} que determinaremos por \hat{u}_N , la cual obtendremos por un procedimiento sobre V_N . Al suponer que $B(\cdot, \cdot)$ es simétrica y que $\{\phi_i\}_{i=1}^N$ es una base para V_N (N funciones linealmente independientes de V generan V_N). Así, cada $u \in V_N$ tiene una representación única de la forma

$$u = \sum_{i=1}^n \alpha_i \phi_i$$

donde los α_i son escalares. A continuación calcularemos (haciendo uso de esta base para V_N) la representación matricial de

$$f(u) = \frac{1}{2}B(u, u) - G(u), \forall u \in V$$

Teorema 2.13. (*Existencia y unicidad de la solución de Ritz*) Para cada $h \in L^2(0, 1)$ existe una única solución u_N de

$$B(u_N, v) = (h, v), \forall v(x) \in V_N$$

También se cumple que

$$I(u_N) = \min_{v \in V_N} I(v)$$

donde se sabe que I es el funcional lineal

$$I(v) = \frac{1}{2}B(v, v) - G(v)$$

Este teorema no nos da una procedimiento para calcular u_N , pero como se sabe que V_N es un espacio vectorial de dimension finita, entonces todo elemento de V_N cuenta con una representación única en función de los elementos de la base, y dicha base la denominaremos por ϕ_N . Así existen reales $\alpha_1, \dots, \alpha_N$ que satisfacen:

$$u_N(x) = \sum_{i=1}^N \alpha_i \phi_i(x)$$

De tal manera que si reemplazamos esta combinación lineal en $I(v) = \frac{1}{2}B(v, v) - G(v)$, haciendo uso de la propiedad de la forma lineal $B(\cdot, \cdot)$ y la propiedad de linealidad de $G(\cdot)$, tenemos que

$$I(u_N) = \frac{1}{2}B\left(\sum_{i=1}^N \alpha_i \phi_i, \sum_{i=1}^N \alpha_i \phi_i\right) - G\left(\sum_{i=1}^N \alpha_i \phi_i\right)$$

$$I(u_N) = \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j B(\phi_i, \phi_j) - \sum_{i=1}^N \alpha_i G(\phi_i)$$

$$I(u_N) = \frac{1}{2} \alpha^T M \alpha - \alpha^T G = I(\alpha) = J(\alpha)$$

donde

$$M = B(\phi_i, \phi_j)_{N \times N} = (\phi'_i, \phi'_j), \quad \alpha = (\alpha_i)_{N \times 1}, \quad G = G(\phi_i)_{N \times 1}$$

Aquí la simetría de $B(\cdot, \cdot)$ implica que M es simétrica. Como $B(\cdot, \cdot)$ es coerciva, tenemos que

$$\alpha^T M \alpha = B(v, v) \geq \rho \|v\|_V^2 \geq 0$$

Como

$$\min_{u \in V_N} I(u) = \min_{\alpha \in \mathbb{R}^N} J(\alpha)$$

y además de que J es un funcional cuadrático sobre \mathbb{R}^N con matriz Hessiana definida positiva, entonces tenemos que J es minimizado por el único α para el cual

$$M\alpha = G$$

Así, el mínimo de I sobre V_N , \widehat{u}_N esta relacionado con este α por

$$u = \sum_{i=1}^N \alpha_i \phi_i$$

De tal manera, podemos concluir que el procedimiento para calcular \widehat{u}_N mediante el método de Ritz es el siguiente:

1. elegir la base $\{\phi_i\}_{i=1}^N$ para V_N .
2. construir las matrices M y G .
3. Calcular $\alpha \in \mathbb{R}^N$ que minimiza a J en $M\alpha = G$.
4. construcción de $\widehat{u}_N = \sum_{i=1}^N \alpha_i \phi_i$.

analizaremos a continuación el caso para el cual $B(\cdot, \cdot)$ no es necesariamente simétrica.

Método de Galerkin

Ahora supondremos que la forma bilineal $B(\cdot, \cdot)$ no es necesariamente simétrica pero que satisface las condiciones restantes del teorema de Lax-Milgram. Como V_N es un espacio de Hilbert, y V_N es un subespacio de V , entonces este espacio V puede ser reemplazado por V_N pues las hipótesis del teorema de Lax-Milgram son también así satisfechas. Así podemos obtener un procedimiento para calcular \widehat{u}_N . Si $\{\phi_i\}_{i=1}^N$ es una base para V_N y si $\{\alpha_i\}_{i=1}^N$ es un conjunto único de escalares tales que

$$\widehat{u}_N = \sum_{i=1}^N \alpha_i \phi_i$$

entonces por el teorema de Lax-Milgram (para V , reemplazado por V_N) se tiene que

$$\sum_{j=1}^N \alpha_j B(\phi_j, v) = G(v), \forall v \in V_N$$

En particular si tenemos que $v = \phi_i$, obtenemos el siguiente sistema de ecuaciones lineales, determinado por las incógnitas α_j :

$$\sum_{j=1}^N \alpha_j B(\phi_j, \phi_i) = G(\phi_i), i = 1, \dots, N$$

equivalente a

$$M\alpha = G$$

El siguiente resultado muestra que este sistema tiene solución única (ver libro Axelsson Barker)

Teorema 2.14. Sea $B(\cdot, \cdot)$ una forma bilineal coerciva. Entonces la matriz M en

$$M\alpha = G$$

resultante de la aplicación del método de Galerkin, no es singular si y solo si $\{\phi_i\}_{i=1}^N$ es una base de V_N .

Según este resultado, podemos calcular α resolviendo

$$M\alpha = G$$

y luego determinar \widehat{u}_N en

$$\widehat{u}_N = \sum_{i=1}^N \alpha_i \phi_i$$

Este es el *método de Galerkin* para calcular \widehat{u}_N . Este método, como se había mencionado, es idéntico al método de Ritz cuando $B(\cdot, \cdot)$ es simétrica. Veamos que si $B(\cdot, \cdot)$ no es simétrica, entonces M tampoco lo es. Luego, podemos afirmar así que la resolución de un problema variacional cuya forma bilineal $B(\cdot, \cdot)$ no es simétrica, exige mas tiempo en términos de computo para resolver el sistema, que la resolución de un problema variacional cuya forma bilineal $B(\cdot, \cdot)$ es simétrica. A continuación describiremos de una forma muy detallada el método de los elementos finitos (implementando el método de Ritz y el método de Galerkin) que nos será de gran utilidad dado que también deseamos encontrar la solución aproximada al problema

$$\begin{array}{lll} \text{(EDP)} & u_t - u_{xx} = -\lambda u^p & (x, t) \in (0, 1) \times [0, T) \\ \text{(CF)} & u_x(0, t) = 0, u_x(1, t) = u(1, t)^q & t \in [0, T) \\ \text{(CI)} & u(x, 0) = u_0(x) & x \in [0, 1] \end{array}$$

en donde $p, q > 1$, $\lambda > 0$ son parámetros, $T > 0$ y u_0 es una función no negativa.

2.1.5. El método de los elementos finitos

El **método de los elementos finitos** (MEF)(ver [12]) es un método numérico general para la aproximación de soluciones de ecuaciones diferenciales parciales y ordinarias muy utilizado en ingeniería y ciencia. La utilidad de este método reside en que muy a menudo se encuentran problemas cuyas condiciones particulares no tienen solución analítica, o cuando la solución analítica es aún más difícil de aplicar que una solución numérica: así se justifica el uso de este método de aproximación a soluciones de ecuaciones diferenciales parciales mediante elementos finitos y diferencias finitas. El método de los elementos finitos permite obtener una solución numérica aproximada sobre un cuerpo, estructura o dominio (medio continuo) sobre el que están definidas ciertas ecuaciones diferenciales en forma débil o integral que caracterizan el comportamiento físico del problema- dividiéndolo en un gran número de subdominios, denominados "*elementos finitos*". El conjunto de elementos finitos forma una partición del dominio, denominada **discretización**. Dentro de cada elemento

se distinguen una serie de puntos representativos llamados *nodos*. Dos nodos son adyacentes si pertenecen al mismo elemento finito; además, un nodo sobre la frontera de un elemento finito puede pertenecer a varios elementos. El conjunto de nodos considerando sus relaciones de adyacencia se llama *malla*. Los cálculos se realizan sobre tal malla, definida por los nodos, que sirven a su vez de base para discretización del dominio en elementos finitos. De acuerdo con estas relaciones de adyacencia o conectividad se relaciona el valor de un conjunto de variables incógnitas definidas en cada nodo y denominadas grados de libertad. El conjunto de relaciones entre el valor de una determinada variable y entre los nodos se puede escribir en forma de sistema de ecuaciones lineales (o ecuaciones linealizadas). La matriz de dicho sistema de ecuaciones se llama *matriz de rigidez* del sistema. El número de ecuaciones de dicho sistema es proporcional al número de nodos. Las consideraciones que se realizan acerca de los intervalos de espacio Δx y los intervalos de tiempo Δt simplifica considerablemente el sistema, ya que en lugar de seguir una función suave a un número infinito de puntos, con este método uno sólo se ocupa de un número finito de valores, por ejemplo valores de temperatura (para el caso de la ecuación de calor) en un número finito de espacios y tiempos. Entre las grandes ventajas de los elementos finitos frente a otros métodos se pueden destacar el hecho de que Las geometrías complicadas y las condiciones frontera pueden manipularse con mayor facilidad. La eficacia de un método en elementos finitos se basa en la medida que la solución aproximada a el problema se acerca o converge a su solución exacta. Esta convergencia depende de dos factores: de la consistencia y de la estabilidad del método. La *consistencia del método numérico* hace referencia a su coherencia a la hora de aproximar la ED. Se trata simplemente de comprobar si el esquema numérico utilizado es un esquema razonable para aproximar la ED en cuestión o, si por el contrario, corre el riesgo de aproximar a otra ED. Lo más habitual es comprobar la consistencia mediante el desarrollo de Taylor. El problema se reduce entonces a verificar si, cuando el paso del mallado tiende a cero [$h \rightarrow 0$], las soluciones regulares del problema continuo son soluciones aproximadas del problema discreto en la medida en que el error tiende a cero. La *estabilidad del método numérico* consiste en asegurarse de que los esquemas discretos o semi-discretos, en su evolución temporal (discreta o continua), no amplifiquen los errores iniciales o, al menos, no lo hagan de manera creciente y descontrolada a medida que el paso del mallado tiende a cero. Vamos a considerar dos técnicas para la construcción de aproximaciones mediante elementos finitos: el principio de Rayleigh-Ritz y el principio de Galerkin. Para utilizar el Método de elementos finitos (MEF), se debe contar con un conjunto $\Omega \subset R^2$, donde Ω debe ser abierto y simplemente conexo. Los elementos finitos mas comúnmente usados son triángulos o cuadriláteros. Por facilidad se asume que todos los elementos de la malla son del mismo tipo y tienen así el mismo numero de lados y vertices. La tarea de generar una malla de elementos finitos apropiados no es una labor fácil, porque se deben tener en cuenta muchos factores, entre ellos los que influyen en el error, los cuales determinan ciertas reglas:

1. Ángulos muy grandes o muy pequeños en los elementos finitos se deben evitar.
2. Los elementos deben ser colocados con mayor densidad en regiones donde se sospecha que la solución del problema y sus derivadas varían rápidamente
3. La alta precision requiere de una malla fina o de muchos nodos por elemento (esta

ultima condición lleva a una precisión alta, solo si la solución es suficientemente suave)

Así, el método de elementos finitos se puede ver como el uso de el método de Ritz con funciones de base de elementos finitos. Tal tipo de funciones base están constituidas por polinomios a trozos y que tienen soporte local, entendiendo esto como el hecho de que estas funciones toman el valor cero fuera de un intervalo pequeño de $[a, b]$. Para nuestro problema unidimensional, las funciones estarán definidas con respecto a una partición del intervalo $[0, 1]$. A continuación veremos como calcular la solución aproximada del problema variacional

$$B(u, v) = G(v), \forall v(x) \in H_0^1(0, 1)$$

utilizando las funciones de base global y los elementos finitos.

Formulación del método de elementos finitos

Consideraremos el problema (ver [1, 12, 3]) de encontrar una solución para:

$$\begin{aligned} -\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] + q(x)u &= g(x) \quad , \Omega = (a, b) \\ u(a) &= \alpha, u(b) = \beta \end{aligned}$$

El cual es el caso general para el problema que se había tratado anteriormente como ejemplo (donde $p = q = 1, g(x) = h(x)$ y $\Omega = (0, 1)$):

$$\begin{aligned} -u''(x) &= h(x) - u(x) \quad , \Omega = (0, 1) \\ u(0) &= u(1) = 0 \end{aligned}$$

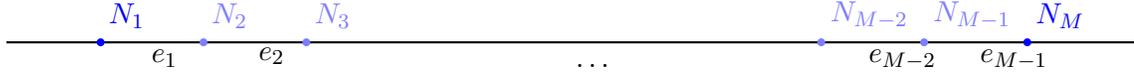
Discretización del dominio - Mallado en elementos finitos

Notación (ver [12, 3])

$e_l, l = 1, \dots, M$	Representa los M elementos de una malla
$N_i, i = 1, \dots, \widetilde{M}$	Representa los \widetilde{M} nodos globales de la malla
$N_r^{(l)}, r = 1, \dots, T$	Representa los T nodos locales del elemento e_l de la malla
$\Phi_i, i = 1, \dots, \widetilde{M}$	Representa las funciones de base global definidas sobre $\overline{\Omega}$
ψ_r^l	Representa las funciones de base local definidas sobre e_l
\tilde{e}	Representa el elemento estandar
$\tilde{\psi}_r$	Representa las funciones de base global estandar definidas sobre $\overline{\Omega}$

Sea $\Omega = (a, b)$ un abierto acotado. Luego se procede a dividir a $\overline{\Omega}$ en un número finito de subintervalos cerrados, llamados *elementos de la malla* $\{e_l\}_{l=1}^M$. Si los elementos tienen igual longitud entonces se dice que la malla es uniforme. Luego, seleccionamos en $\overline{\Omega}$ un número de puntos o *nodos globales* $\{N_i = x_i\}_{i=1}^{\widetilde{M}}$. Cada nodo puede encontrarse dentro del elemento o estar en la intersección de dos elementos. Estos nodos se llaman *nodos locales* tal

que, para un elemento e_l sus nodos se denotan por $\{N_r^l = x_r^l\}_{r=1}^{\tilde{T}}$. Mediante esta elección de elementos y nodos definimos la malla de elementos finitos. La precisión aumenta con el incremento de los nodos. De esta forma se ha construido el mallado global sobre nuestro problema, representado los nodos y los elementos en el siguiente esquema



Funciones base de elementos finitos

Para una malla en elementos finitos (ver [12]) asociamos a cada nodo N_i una función ϕ_i definida sobre $\bar{\Omega}$ con las propiedades

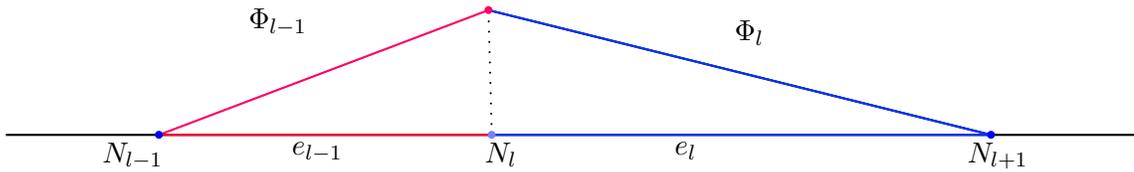
1. ϕ_i sobre e_l es un polinomio de la forma

$$\phi_i = \sum_{s=1}^T c_{il}^{(s)} x^{p_s}, x \in e_l$$

donde T representa el número de funciones de base local sobre e_l .

2. $\phi_i(N_j) = \delta_{ij}$, $i, j = 1, 2, \dots, M$
3. ϕ_i se determina por su valor sobre un nodo determinado.

Así construimos el espacio V_M que es el *espacio de funciones base de elementos finitos*, y las ϕ_i son las *funciones base de elementos finitos*



Los coeficientes de

$$\phi_i = \sum_{s=1}^T c_{il}^{(s)} x^{p_s}, x \in e_l$$

son obtenidos por $\phi_i(N_j) = \delta_{ij}$ al construir el sistema lineal:

$$\sum_{s=1}^T c_{il}^{(s)} (N_r^{(l)})^{p_s} = \begin{cases} 1, & r = i \\ 0, & \text{en cualquier otro caso} \end{cases}$$

El sistema se puede expresar matricialmente como $S_i c_{il} = d_{il}$, donde

$$S_i = \begin{bmatrix} (N_1^l)^{p_1} & (N_1^l)^{p_2} & (N_1^l)^{p_3} \dots & (N_1^l)^{p_{T-2}} & (N_1^l)^{p_{T-1}} & (N_1^l)^{p_T} \\ (N_2^l)^{p_1} & (N_2^l)^{p_2} & (N_2^l)^{p_3} \dots & (N_2^l)^{p_{T-2}} & (N_2^l)^{p_{T-1}} & (N_2^l)^{p_T} \\ (N_3^l)^{p_1} & (N_3^l)^{p_2} & (N_3^l)^{p_3} \dots & (N_3^l)^{p_{T-2}} & (N_3^l)^{p_{T-1}} & (N_3^l)^{p_T} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ (N_{T-2}^l)^{p_1} & (N_{T-2}^l)^{p_2} & (N_{T-2}^l)^{p_3} \dots & (N_{T-2}^l)^{p_{T-2}} & (N_{T-2}^l)^{p_{T-1}} & (N_{T-2}^l)^{p_T} \\ (N_{T-1}^l)^{p_1} & (N_{T-1}^l)^{p_2} & (N_{T-1}^l)^{p_3} \dots & (N_{T-1}^l)^{p_{T-2}} & (N_{T-1}^l)^{p_{T-1}} & (N_{T-1}^l)^{p_T} \\ (N_T^l)^{p_1} & (N_T^l)^{p_2} & (N_T^l)^{p_3} \dots & (N_T^l)^{p_{T-2}} & (N_T^l)^{p_{T-1}} & (N_T^l)^{p_T} \end{bmatrix}$$

como ϕ_i se determina por su valor sobre un nodo determinado, entonces ϕ_i es continua sobre los nodos frontera del elemento e_l , y así, por la continuidad de los polinomios se tiene que $\phi_i \in C(\overline{\Omega})$. Supongamos así que un nodo N_i no pertenece a un elemento e_l entonces $d_{il} = 0$. Así podemos concluir que la función ϕ_i es cero en aquellos elementos a los que el nodo N_i no pertenece. Por definición una función u pertenece a $V_{\widetilde{M}}$ si y solo si puede ser expresada por

$$u(x) = \sum_{i=1}^{\widetilde{M}} c_i \phi_i(x), x \in \overline{\Omega}$$

de tal forma que, para cualquier $u \in C(\overline{\Omega})$, se tiene u_I (que pertenece a $V_{\widetilde{M}}$) definida por

$$u_I(x) = \sum_{i=1}^{\widetilde{M}} u(N_i) \phi_i(x), x \in \overline{\Omega}$$

Funciones base local de elementos finitos

Definimos las funciones de base local (ver [12]) sobre e_l como

$$\psi_r^l(x) = \phi_{N_r^l}(x), x \in e_l, r = 1, 2, \dots, T$$

Así podemos afirmar que una *función de base local es una restricción de una base global a el elemento e_l* . Ahora, si V_T^l es el espacio de los polinomios del tipo:

$$\phi_i = \sum_{s=1}^T c_{il}^{(s)} x^{p_s}, x \in e_l$$

definidos sobre e_l , entonces V_T^l es el espacio generado por las funciones $\left\{ \psi_r^l \right\}_{r=1}^T$. Luego

$$\psi_r^l(N_s^{(l)}) = \delta_{rs}, r, s = 1, 2, \dots, T$$

Usando las bases locales se puede escribir a

$$u_I(x) = \sum_{i=1}^{\widetilde{M}} u(N_i) \phi_i(x)$$

como

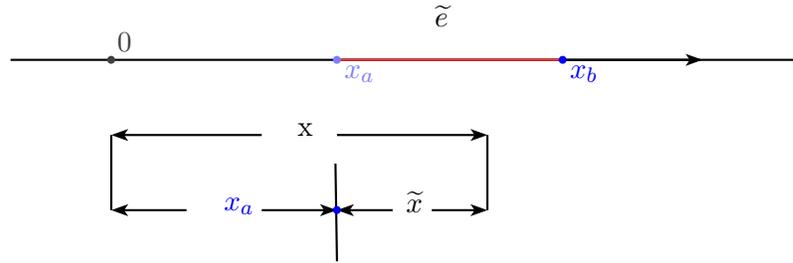
$$u_I(x) = \sum_{i=1}^{\widetilde{M}} u(N_i) \phi_i(x) = \sum_{r=1}^T u(N_r^{(l)}) \psi_r^{(l)}(x), x \in e_l$$

Observemos así que las funciones de base global se constituyen por funciones de base local. Se puede también elegir un elemento estándar $\tilde{e} = (x_a, x_b)$ de la recta, con las funciones de base global $\{\tilde{\psi}_r\}_{r=1}^T$ de tipo

$$\phi_i = \sum_{s=1}^T c_{il}^{(s)} x^{p_s}, x \in e_l$$

y hallar para todo elemento e_l la transformación

$$x = \tilde{x} + x_a, \tilde{x} \in \tilde{e}$$



dependiendo de l tal que $x = \tilde{x} + x_a, \tilde{x} \in \tilde{e}$ transforma a \tilde{e} en e_l y

$$\tilde{\psi}_r(\tilde{x}) = \tilde{\psi}_r^l(x(\tilde{x})) = \psi_r^{(l)}(\tilde{x} + x_a)$$

si denotamos la inversa por

$$\tilde{x} = \tilde{x}(x) = x - x_a, x \in e_l$$

Luego la expresión

$$\tilde{\psi}_r(\tilde{x}) = \tilde{\psi}_r^l(x(\tilde{x})) = \psi_r^{(l)}(\tilde{x} + x_a)$$

se puede ver como

$$\psi_r^{(l)} = \tilde{\psi}_r(\tilde{x})$$

así podemos afirmar que todas las funciones de base local $\psi_r^{(l)}$ con $l = 1, \dots, M$ se pueden reducir mediante la transformación

$$x = \tilde{x} + x_a, \tilde{x} \in \tilde{e}$$

a una función simple $\tilde{\psi}_r$. Pero las funciones de base global están compuestas por funciones de base locales, entonces las operaciones sobre funciones de base global se pueden reducir a operaciones sobre funciones de base local.

Formulación variacional sobre un elemento estándar

Como el problema

$$\begin{aligned} -\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] + q(x)u &= g(x) \quad , \Omega = (a, b) \\ u(a) &= \alpha, u(b) = \beta \end{aligned}$$

es válido sobre el dominio (a, b) también entonces es válido sobre un elemento estándar $\tilde{e} = (x_a, x_b)$ (ver [12]), luego, si seguimos el procedimiento descrito, obtenemos la formulación variacional:

$$\int_{x_a}^{x_b} (-(pu')' + qu - g)\eta dx = \int_{x_a}^{x_b} (pu'\eta' + qu\eta - g\eta) dx + [(pu')\eta] \Big|_{x_a}^{x_b} = 0$$

Ahora con la notación

$$\begin{aligned} u(x_a) &= u_1^{(l)} & u(x_b) &= u_2^{(l)} \\ (-pu') \Big|_{x_a} &= P_1^{(l)} & (-pu') \Big|_{x_b} &= P_2^{(l)} \end{aligned}$$

reemplazamos en el esquema variacional anterior para obtener

$$\int_{x_a}^{x_b} (pu'\eta' + qu\eta - g\eta) dx - P_1^{(l)}\eta(x_a) - P_2^{(l)}\eta(x_b) = 0$$

Observamos así que esto es

$$B(u, \eta) = G(\eta)$$

donde

$$B(u, \eta) = \int_{x_a}^{x_b} (pu'\eta' + qu\eta) dx$$

$$G(\eta) = \int_{x_a}^{x_b} g\eta dx + P_1^{(l)}\eta(x_a) + P_2^{(l)}\eta(x_b)$$

El funcional cuadrático asociado es

$$f(u) = \frac{1}{2}B(u, \eta) - G(\eta) = \frac{1}{2} \int_{x_a}^{x_b} (pu'\eta' + qu\eta) dx - \int_{x_a}^{x_b} g\eta dx - P_1^{(l)}\eta(x_a) - P_2^{(l)}\eta(x_b)$$

Ahora, busquemos las soluciones aproximadas del problema

$$\int_{x_a}^{x_b} (pu'\eta' + qu\eta - g\eta)dx - P_1^{(l)}\eta(x_a) - P_2^{(l)}\eta(x_b)dx = 0$$

con las condiciones

$$u(x_a) = u_1^{(l)}, u(x_b) = u_2^{(l)}$$

Sea u_l la aproximación de Ritz de u sobre el elemento e_l de la forma:

$$u_l(x) = \sum_{r=1}^T \alpha_r^{(l)} \psi_r^{(l)}(x)$$

donde los $\alpha_r^{(l)}$ son los parámetros a ser determinados, y $\psi_r^{(l)}(x)$ son las funciones de aproximación a ser construidas. Si tomamos por $u = u_l(x) = \sum_{r=1}^T \alpha_r^{(l)} \psi_r^{(l)}(x)$ y por $\eta = \psi_r^{(l)}(x)$ en el esquema variacional

$$B(u, \eta) = G(\eta)$$

obtenemos el sistema lineal

$$K^{(l)} \alpha^{(l)} = G^{(l)}$$

donde las matrices $K^{(l)}$ y $G^{(l)}$ están representadas por

$$K_{rs}^{(l)} = \int_{x_a}^{x_b} (p(\psi_r^{(l)})'(\psi_s^{(l)})' + q(\psi_r^{(l)})(\psi_s^{(l)}))dx$$

$$G_r^{(l)} = \int_{x_a}^{x_b} g\psi_r^{(l)} dx + P_1^{(l)}\psi_r^{(l)}(x_a) + P_2^{(l)}\psi_r^{(l)}(x_b)$$

Ahora construiremos las funciones de aproximación $\psi_r^{(l)}$ para el elemento. Elijamos una aproximación lineal (esto se obtiene con $T = 2$) de la forma

$$u(x) = c_1 + c_2x = c_1\phi_1 + c_2\phi_2$$

donde $\phi_1 = 1$ y ϕ_2 son linealmente independientes. Así, para el elemento $e_l = (x_l, x_{l+1})$, tenemos

$$\begin{aligned} u(x_l) = u_1^{(l)} &= c_1 + c_2x_l \\ u(x_{l+1}) = u_2^{(l)} &= c_1 + c_2x_{l+1} \end{aligned}$$

su representación matricial es

$$\begin{bmatrix} u_1^{(l)} \\ u_2^{(l)} \end{bmatrix} = \begin{bmatrix} 1 & x_l \\ 1 & x_{l+1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

Este sistema tiene por solución

$$c_1 = \frac{u_1^{(l)} x_{l+1} - u_2^{(l)} x_l}{x_{l+1} - x_l} \quad c_2 = \frac{u_2^{(l)} - u_1^{(l)}}{x_{l+1} - x_l}$$

luego, tenemos así que

$$u_l(x) \equiv u(x) = \frac{u_1^{(l)} x_{l+1} - u_2^{(l)} x_l}{x_{l+1} - x_l} + \frac{u_2^{(l)} - u_1^{(l)}}{x_{l+1} - x_l} x = \sum_{r=1}^2 u_r^{(l)} \psi_1^{(l)}$$

Aquí las funciones $\psi_i^{(l)}$ están representadas como

$$\psi_1^{(l)} = \frac{x_{l+1} - x}{x_{l+1} - x_l} \quad \psi_2^{(l)} = \frac{x - x_l}{x_{l+1} - x_l} \quad x \in e_l = (x_l, x_{l+1})$$

de la ecuación

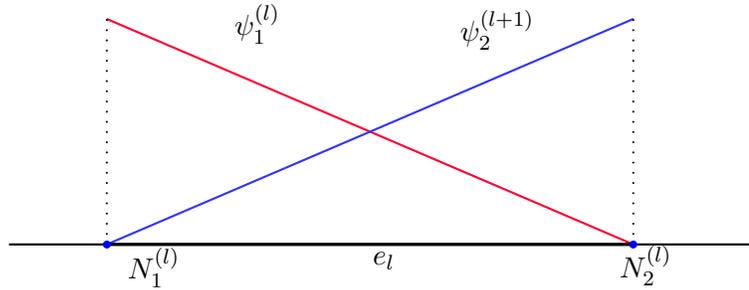
$$K^{(l)} \alpha^{(l)} = G^{(l)}$$

y de la aproximación $u_l(x)$ hecha sobre $u(x)$, utilizando las propiedades de las funciones de aproximación (donde para los elementos lineales tenemos que $x_a = x_l$ y $x_b = x_{l+1}$) y si hacemos (por simplicidad) $q = 0$ tenemos

$$K_{rs}^{(l)} = \int_{x_l}^{x_{l+1}} p(\psi_r^{(l)})' (\psi_s^{(l)})' dx$$

$$G_r^{(l)} = \int_{x_l}^{x_{l+1}} g \psi_r^{(l)} dx + P_r^{(l)}$$

Veamos que si el grado del polinomio $\psi_i^{(l)}$ es n , entonces $K^{(l)}$ es una matriz de orden $(n+1) \times (n+1) = T \times T$, observemos que $n+1$ es el número de nodos necesarios en e_l para obtener los polinomios $\psi_i^{(l)}$



Es posible considerar un sistema de coordenadas locales en $e_l = (x_l, x_{l+1})$. Sea \tilde{x} la coordenada local que tiene por origen el extremo izquierdo del elemento. Esta coordenada está relacionada con la coordenada global x mediante la transformación lineal de traslación $x = \tilde{x} + x_l, \tilde{x} \in \tilde{e}$ que determina

$$x = \tilde{x} + x_l$$

En el sistema de coordenadas locales, las ecuaciones

$$u(x) = c_1 + c_2x = c_1\phi_1 + c_2\phi_2$$

y

$$\begin{aligned} u(x_l) = u_1^{(l)} &= c_1 + c_2x_l \\ u(x_{l+1}) = u_2^{(l)} &= c_1 + c_2x_{l+1} \end{aligned}$$

toman así la forma

$$\begin{aligned} u(\tilde{x}) &= \hat{c}_1 + \hat{c}_2\tilde{x} \\ u_1^{(l)} &= u(0) = \hat{c}_1 \\ u_2^{(l)} &= u(x_{l+1} - x_l) = \hat{c}_1 + \hat{c}_2u_1^{(l)}\tilde{x} \end{aligned}$$

con $h_l = x_{l+1} - x_l$ tenemos

$$\begin{bmatrix} u_1^{(l)} \\ u_2^{(l)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & x_{l+1} - x_l \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

con

$$\begin{aligned} \hat{c}_1 &= u_1^{(l)} \\ \hat{c}_2 &= \frac{u_2^{(l)} - u_1^{(l)}}{h_l} \end{aligned}$$

Así obtenemos

$$u_l = \sum_{r=1}^2 u_r^{(l)} \psi_r^{(l)}(\tilde{x})$$

determinando las funciones $\tilde{\psi}_r(\tilde{x})$ por

$$\tilde{\psi}_1(\tilde{x}) = 1 - \frac{\tilde{x}}{h_l}, \quad \tilde{\psi}_2(\tilde{x}) = \frac{\tilde{x}}{h_l}, \quad 0 \leq \tilde{x} \leq h_l$$

los sistemas de coordenadas locales son también convenientes en la evaluación numérica de integrales. Por ejemplo, las componentes de la matriz K^l y el vector columna G^l en

$$K_{rs}^{(l)} = \int_{x_l}^{x_{l+1}} p(\psi_r^{(l)})'(\psi_s^{(l)})' dx$$

$$G_r^{(l)} = \int_{x_l}^{x_{l+1}} g\psi_r^{(l)} dx + P_r^{(l)}$$

pueden ser escritos en términos de coordenadas locales de la forma

$$K_{rs}^{(l)} = \int_{x_l}^{x_{l+1}} \hat{p} \left[\frac{d\widetilde{\psi}_r}{d\widetilde{x}} \right]' \left[\frac{d\widetilde{\psi}_s}{d\widetilde{x}} \right]' \widetilde{dx}$$

$$G_r^{(l)} = \int_{x_l}^{x_{l+1}} \widehat{g}\widetilde{\psi}_r \widetilde{dx} + P_r^{(l)}$$

Aquí $\widehat{p} = p(x_l + \widetilde{x})$, $\widehat{g} = g(x_l + \widetilde{x})$. Si tenemos funciones de aproximación lineales, se tiene que

$$K^{(l)} = \frac{a_l}{h_l} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$G^{(l)} = \frac{g_l h_l}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} P_1^{(l)} \\ P_2^{(l)} \end{bmatrix}$$

Ahora, si suponemos que $\overline{\Omega} = [a, b]$ es dividido en 3 elementos e_l de igual longitud (o de longitud distinta) estos elementos están unidos en los nodos globales N_2 y N_3 y u es continua, el valor de $u_2^{(l)}$ en el elemento e_l debe ser el mismo de $u_1^{(l+1)}$ en el elemento e_{l+1} , con $l = 1, 2$. Para expresar matemáticamente esta correspondencia, se identifican los valores de u en el nodo global N_i , $i = 1, \dots, M$ como U_i , esto quiere decir que $u(N_i) = U_i$. Así obtenemos la siguiente correspondencia entre nodos globales y locales:

$$u_1^{(1)} = u_1, u_2^{(1)} = u_2, u_2^{(3)} = u_4$$

e_l





Estas relaciones son conocidas como *condiciones de continuidad entre elementos*. Como se estableció que $q = 0$, obtenemos que las ecuaciones en términos de los valores de los nodos globales U_i para cada elemento están determinadas por:

1. **Elemento e_1 :** determinado por los nodos U_1 y U_2 :

$$\frac{\alpha_1}{h_1} \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{g_1 h_1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} P_1^{(l)} \\ P_2^{(l)} \\ 0 \\ 0 \end{bmatrix}$$

2. **Elemento e_2 :** determinado por los nodos U_2 y U_3 :

$$\frac{\alpha_2}{h_2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{g_2 h_2}{2} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ P_1^{(2)} \\ P_2^{(2)} \\ 0 \end{bmatrix}$$

3. **Elemento e_3 :** determinado por los nodos U_3 y U_4 :

$$\frac{\alpha_3}{h_3} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{g_2 h_2}{2} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ P_1^{(3)} \\ P_2^{(3)} \end{bmatrix}$$

Observemos que ninguno de estos sistemas tiene solución porque en el lado derecho de cada uno se desconocen los valores de $P_r^{(l)}$, para $r = 1, 2$ y $l = 1, 2, 3$. De esta manera, la solución se obtiene superponiendo las ecuaciones para los tres elementos, obteniendo así el nuevo sistema:

$$\begin{bmatrix} \frac{\alpha_1}{h_1} & -\frac{\alpha_1}{h_1} & 0 & 0 \\ -\frac{\alpha_1}{h_1} & \frac{\alpha_1}{h_1} + \frac{\alpha_2}{h_2} & -\frac{\alpha_2}{h_2} & 0 \\ 0 & -\frac{\alpha_2}{h_2} & \frac{\alpha_2}{h_2} + \frac{\alpha_3}{h_3} & -\frac{\alpha_3}{h_3} \\ 0 & 0 & -\frac{\alpha_3}{h_3} & \frac{\alpha_3}{h_3} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} g_1 h_1 \\ g_1 h_1 + g_2 h_2 \\ g_2 h_2 + g_3 h_3 \\ g_3 h_3 \end{bmatrix} + \begin{bmatrix} P_1^{(1)} \\ P_2^{(1)} + P_1^{(2)} \\ P_2^{(2)} + P_1^{(3)} \\ P_2^{(3)} \end{bmatrix}$$

Esta ecuación representa la matriz global para el método de elementos finitos del problema

$$-\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] + q(x)u = g(x) \quad , \Omega = (a, b)$$

$$u(a) = \alpha, u(b) = \beta$$

Ahora, debemos imponer las condiciones de frontera. Las condiciones de frontera naturales están incluidas en el vector de los $P_i^{(l)}$. Veamos que la ecuación diferencial tiene condiciones de frontera esenciales en $x = a$ y $x = b$, luego se tiene que

$$U_1 = u_1^1 = u(x_1) = u(a) = \alpha,$$

$$U_4 = u_2^3 = u(x_4) = u(b) = \beta$$

Ahora, si observamos el problema diferencial con $q = 0$, sobre el elemento $e_1 = (x_1, x_2)$ tenemos:

$$-\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] = g(x) \quad , \text{ en } e_l$$

$$u(x_1) = \alpha$$

Como no se imponen condiciones de frontera esenciales en x_2 , entonces la correspondiente condición de frontera natural en este punto es:

$$P_2^{(1)} = -p(x) \frac{du}{dx} \Big|_{x=x_2} = 0$$

Al tener la ecuación diferencial con una condición de frontera esencial en $x = a$ tenemos que

$$P_1^{(1)} = -p(x) \frac{du}{dx} \Big|_{x=x_1=a}$$

donde tal $P_1^{(1)}$ tiene un valor arbitrario. De igual forma, para e_2 tenemos que

$$P_1^{(3)} = 0 \quad P_2^{(3)} = -p(x) \frac{du}{dx} \Big|_{x=x_2}$$

De tal forma que la **matriz condiciones de frontera sobre los elementos** esta determinada por

$$\begin{bmatrix} P_1^{(1)} \\ P_2^{(1)} + P_1^{(2)} \\ P_2^{(2)} + P_1^{(3)} \\ P_2^{(3)} \end{bmatrix} = \begin{bmatrix} P_1^{(1)} = -p(x) \frac{du}{dx} \Big|_{x=x_1=a} \\ 0 \\ 0 \\ P_2^{(3)} = -p(x) \frac{du}{dx} \Big|_{x=b} \end{bmatrix}$$

tenemos que

$$u = \sum_{l=1}^3 u_l = \sum_{l=1}^3 \sum_{r=1}^2 u_r^{(l)} \psi_r^{(l)} = \sum_{i=1}^4 U_i \phi_i$$

o lo que es equivalente a

$$u = \begin{cases} U_1\phi_1^{(l)} + U_2\phi_2^{(l)} & , x \in e_1 \\ U_2\phi_1^{(2)} + U_3\phi_2^{(2)} & , x \in e_2 \\ U_3\phi_1^{(3)} + U_4\phi_2^{(3)} & , x \in e_2 \end{cases}$$

Así, obtenemos

$$u = U_1\phi_1^{(l)} + U_2 [\phi_2^{(l)} + \phi_1^{(2)}] + U_3 [\phi_2^{(2)} + \phi_1^{(3)}] + U_4\phi_2^{(3)}$$

y como $U_1 = \alpha$ y $U_4 = \beta$, el sistema

$$\begin{bmatrix} \frac{\alpha_1}{h_1} & -\frac{\alpha_1}{h_1} & 0 & 0 \\ -\frac{\alpha_1}{h_1} & \frac{\alpha_1}{h_1} + \frac{\alpha_2}{h_2} & -\frac{\alpha_2}{h_2} & 0 \\ 0 & -\frac{\alpha_2}{h_2} & \frac{\alpha_2}{h_2} + \frac{\alpha_3}{h_3} & -\frac{\alpha_3}{h_3} \\ 0 & 0 & -\frac{\alpha_3}{h_3} & \frac{\alpha_3}{h_3} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} g_1h_1 \\ g_1h_1 + g_2h_2 \\ g_2h_2 + g_3h_3 \\ g_3h_3 \end{bmatrix} + \begin{bmatrix} P_1^{(1)} \\ P_2^{(1)} + P_1^{(2)} \\ P_2^{(2)} + P_1^{(3)} \\ P_2^{(3)} \end{bmatrix}$$

se reduce a resolver

$$\begin{bmatrix} \frac{\alpha_1}{h_1} + \frac{\alpha_2}{h_2} & -\frac{\alpha_2}{h_2} \\ -\frac{\alpha_2}{h_2} & \frac{\alpha_2}{h_2} + \frac{\alpha_3}{h_3} \end{bmatrix} \begin{bmatrix} U_2 \\ U_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} g_1h_1 + g_2h_2 \\ g_2h_2 + g_3h_3 \end{bmatrix}$$

Y así se establece el sistema para calcular U_2 y U_3 . Con estos valores se encuentra a u utilizando

$$u = \sum_{l=1}^3 u_l = \sum_{l=1}^3 \sum_{r=1}^2 u_r^{(l)} \psi_r^{(l)} = \sum_{i=1}^4 U_i \phi_i$$

o también por

$$u = U_1\phi_1^{(l)} + U_2 [\phi_2^{(l)} + \phi_1^{(2)}] + U_3 [\phi_2^{(2)} + \phi_1^{(3)}] + U_4\phi_2^{(3)}$$

que al ser derivada una vez y multiplicada por la función p se puede evaluar en $x = a$ y $x = b$ para encontrar así los valores de $P_1^{(1)}$ y $P_2^{(3)}$. Observemos que cuando determinamos a u en

$$u = \sum_{l=1}^3 u_l = \sum_{l=1}^3 \sum_{r=1}^2 u_r^{(l)} \psi_r^{(l)} = \sum_{i=1}^4 U_i \phi_i$$

o su equivalente

$$u = U_1\phi_1^{(l)} + U_2 [\phi_2^{(l)} + \phi_1^{(2)}] + U_3 [\phi_2^{(2)} + \phi_1^{(3)}] + U_4\phi_2^{(3)}$$

Esto se puede expresar también por

$$u = U_1\phi_1 + U_2\phi_2 + U_3\phi_3 + U_4\phi_4$$

las funciones ϕ_i están dadas por

$$\phi_1 = \phi_1^{(l)}$$

$$\phi_2 = \phi_2^{(l)} + \phi_1^{(2)}$$

$$\phi_3 = \phi_2^{(2)} + \phi_1^{(3)}$$

$$\phi_4 = \phi_2^{(3)}$$

Etapas del proceso de los métodos de los elementos finitos y de las diferencias finitas

Como ya se había comentado en un inicio, La idea general de cualquier método numérico es discretizar: si en las diferencias finitas se reemplazan las derivadas por cocientes en diferencias, en el método de elementos finitos el factor esencial es que la integral de una función medible se pueda escribir como suma de integrales en dominios disjuntos cuya unión es el dominio original, entonces se puede hacer un análisis local del problema y haciendo los dominios suficientemente pequeños los polinomios serán adecuados para una buena representación del comportamiento de la solución. Luego, para resolver un problema diferencial usando elementos finitos se deben realizar las siguiente etapas:

1. Discretizar el dominio dado en una colección de elementos finitos
 - (a) Construir la malla de elementos finitos
 - (b) Numerar los nodos en el elemento
 - (c) Generar las propiedades necesarias para el problema
2. Obtener las ecuaciones sobre un elemento de la malla.
 - (a) Construir la formulación variacional de la ecuación diferencial sobre un elemento estándar.
 - (b) Expresar la solución u como una combinación lineal, a fin de generar el esquema matricial que representa la formulación variacional para tal elemento.
 - (c) seleccionar las funciones ψ_r en los elementos, y calcular las matrices para cada uno de ellos.
3. Ensamblar las ecuaciones de los elementos para obtener la ecuación de el problema global.
 - (a) Identificar las condiciones entre los elementos para las variables involucradas en la especificación de las condiciones de frontera esenciales.

- (b) Identificar las condiciones entre las variables involucradas en la especificación de las condiciones de frontera naturales.
 - (c) Ensamblar las ecuaciones en cada elemento, utilizando para esto los pasos 3(a) y 3(b).
4. Solucionar la matriz ensamblada, la cual ha sido generada a partir de las ecuaciones.
 5. procesar los resultados (usualmente esto implica hacer un análisis sobre las soluciones al sistema global, obtenidas a través de un Computador).

Capítulo 3

Explosión de las soluciones para un problema parabólico

En este capítulo estudiaremos inicialmente el caso homogéneo del problema propuesto (es decir, estudiaremos la ecuación $u_t = u_{xx}$, bajo ciertas condiciones iniciales y de frontera, la teoría referente a la concepción física de la ecuación de calor (modelada de hecho por una ecuación de tipo parabólico) y las Leyes de termodinámica, para comprender el sentido físico del problema parabólico presentado. Luego estudiaremos la representación variacional de nuestro problema inicialmente planteado y por último analizaremos el fenómeno de blow-Up. La bibliografía tratada en este capítulo está fundamentada por [2, 4, 6, 8, 11, 13, 19].

3.1. La ecuación de calor unidimensional

3.1.1. El modelo de la ecuación de calor

La ecuación que describe la conducción de calor (ver [2, 7, 8, 11, 13]) en los sólidos es una herramienta importante para analizar el movimiento dinámico del calor, así como para solucionar una enorme variedad de problemas de tipo difusión en las ciencias físicas, biológicas, entre otras. El proceso de conducción de calor, descrito por una ecuación diferencial parcial fue planteado por primera vez en la historia por Joseph Fourier (1768-1830), quien decidió publicar los resultados de sus investigaciones y observaciones físicas acerca de termodinámica y teoría de potencial en su obra *Teoría analítica del calor*, en 1807. Dicha ecuación de calor es una ecuación diferencial parcial que describe la distribución de calor (o variación de la temperatura) en una región, y cuya distribución calorífica está determinada por el tiempo.

Las soluciones de la ecuación del calor se caracterizan por una progresiva suavización de la distribución inicial de la temperatura, pues el flujo de calor va de la región de mayor temperatura a aquellas que tienen menor temperatura en el objeto. En general, diferentes estados y condiciones de partida de la ecuación de calor tienden hacia el equilibrio estable del mismo (es decir, la temperatura, trascurrido un tiempo considerable tiende hacia un determinado valor en el cuerpo en el cual se distribuye). La ecuación del calor es un ejemplo

de ecuación diferencial parcial de tipo parabólico, dado que una ecuación parabólica en derivadas parciales es una ecuación diferencial parcial de segundo orden del tipo:

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + F = 0$$

donde $AC - B^2 = 0$. Otro ejemplo de ecuación diferencial parcial parabólica es la ecuación de Schrödinger. La ecuación que modela la difusión del calor rige también otros procesos de difusión, tales como la difusión de partículas o la propagación del potencial de acción en las células nerviosas, inclusive está relacionada con la mecánica cuántica, aunque en este último campo existen ciertos fenómenos no difusivos, los cuales se rigen por un modelo matemático análogo de la ecuación del calor. La temperatura representada por la función $u(x, y, z, t)$ consta de tres variables espaciales, donde x, y y z representan las direcciones, y la variable t (temporal) representa el tiempo, determina la ecuación del calor:

$$\frac{\partial u}{\partial t} - \alpha \left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right] = 0$$

aquí podemos ver la representación del operador laplaciano::

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \Delta u$$

así que, La ecuación puede ser escrita como:

$$\frac{\partial u}{\partial t} - \alpha \Delta u = 0$$

$$\alpha = \frac{h}{C_p \rho}$$

donde α es una constante positiva y representa la **difusibilidad** (de hecho, la difusibilidad caracteriza la rapidez con la que varía la temperatura a través de tal material). Generalmente el valor de α se estima a partir de la medición de la conductividad térmica h , capacidad calorífica C_p y densidad ρ . Físicamente, α *indica como fluye el calor por el material*. Cuanto más alta es la difusibilidad térmica de una sustancia, más alto es el ritmo de propagación del perfil de temperatura: Es decir, la difusibilidad relaciona flujo de energía con gradiente de energía.

A continuación haremos ciertas consideraciones físicas sobre el fenómeno de la conducción de calor, citaremos las 3 leyes de la termodinámica, y posteriormente deduciremos dicha ecuación, utilizando el teorema de conservación de energía y la ley de Fourier.

Leyes de la termodinámica, Ley de Fourier y la ecuación de calor

Los fenómenos de conducción de calor se han hecho presentes en el desarrollo de máquinas, y han permitido avances industriales y por ende la evolución de la civilización. James Watt (1736-1819) inventó el primer motor para trabajo relativamente moderno. Este motor funcionaba por el combustible que consumía: el combustible se quemaba y este se transformaba

en calor y, a continuación se transformaba en trabajo. El hecho de que el calor (u otra forma de energía) sea necesaria para producir trabajo en un motor es la esencia de la primera ley de termodinámica. La **primera ley de la termodinámica** es la formulación de una ley más general de la física (la ley de conservación de la energía) para los procesos termodinámicos: **La energía no se crea ni se destruye, sólo puede ser transformada de una forma u otra.**

En el motor sin embargo, no todo el calor se transforma en trabajo. Leonard Carnot (1796-1832), un ingeniero francés pionero en el estudio de la termodinámica, y al cual se le reconoce hoy día como el fundador de la termodinámica demostró que hay un límite para la cantidad de calor que puede ser transformada en trabajo, así el calor restante se pierde en la atmósfera. Este descubrimiento también está relacionado con el hecho conocido de que el calor no puede, por sí solo ser transferido de un organismo de temperatura más baja a uno de temperatura más alta. Esta es la esencia de la **segunda ley de la termodinámica**, la cual también indica la dirección en que se llevan a cabo las transformaciones energéticas: **El flujo espontáneo de calor siempre es unidireccional, desde los cuerpos de temperatura más alta a aquellos de temperatura más baja.**

En esta ley aparece el concepto de *entropía*, la cual se define como la magnitud física que mide la parte de la energía que no puede utilizarse para producir un trabajo. Esto es más fácil de entender con el siguiente ejemplo, en el que se utiliza una máquina térmica: Una fuente de calor es usada para calentar una sustancia de trabajo (por ejemplo, vapor de agua), provocando la expansión de la misma colocada dentro de un pistón a través de una válvula. La expansión mueve el pistón, y por un mecanismo de acoplamiento adecuado, se obtiene trabajo mecánico. El trabajo se da por la diferencia entre el calor final y el inicial. Es imposible la existencia de una máquina térmica que extraiga calor de una fuente y lo convierta totalmente en trabajo, sin enviar nada a la fuente fría. La entropía de un sistema se puede interpretar también como un grado de desorden del mismo.

Más tarde, se comprobó que todos los procesos termodinámicos (como la transferencia de calor, el intercambio de calor, etc) obedecen a estas dos leyes. La **tercera ley de la termodinámica** establecida por Walther Hermann Nernst (1864-1941, físico y químico alemán) mediante el conocido por entonces *teorema del calor* ("la entropía de una materia tiende a anularse cuando su temperatura se aproxima al cero absoluto"), y formalizada en 1923 por Gilbert N. Lewis (1875-1946, físico-químico estadounidense) y Merle Randall (1888- 1950, físico-químico estadounidense), fue el resultado del estudio de las propiedades de materiales a muy bajas temperaturas, en particular de la capacidad calorífica, y esta afirma lo siguiente: **No se puede alcanzar el cero absoluto en un número finito de etapas.** En otras palabras: al llegar al cero absoluto, (0 K), cualquier proceso de un sistema físico se detiene y además, al llegar al cero absoluto la entropía alcanza un valor mínimo y constante.

Esto se da porque a medida que un sistema dado se aproxima al cero absoluto, su entropía tiende a un valor constante específico. A medida que el sistema se acerca al cero absoluto, el intercambio calórico es cada vez menor hasta llegar a ser casi nulo; Ya que el flujo espontáneo de calor es unidireccional, desde los cuerpos de temperatura más alta a los de temperatura más baja (Segunda ley), sería necesario un cuerpo con menor temperatura que el cero absoluto; y esto es imposible.

La ley de la conducción del calor, también conocida como ley de Fourier, indica que el índice del tiempo de traspaso térmico a través de un material es proporcional a el opuesto del gradiente en la temperatura . De esta manera, tenemos que:

$$q = -h\nabla u$$

donde h es la *conductividad térmica* (la cual es una propiedad física de los materiales, que mide la capacidad de conducción de calor. En otras palabras, la conductividad térmica es también la capacidad de una sustancia de transferir la energía cinética de sus moléculas a otras moléculas adyacentes o a sustancias con las que está en contacto) y u es la temperatura. En una dimensión, el gradiente es una derivada espacial ordinaria, entonces:

$$q = -hu_x$$

donde $u_x = \frac{\partial u}{\partial x}$. Un cambio en el interior de la energía por unidad de volumen en el material, ΔQ , es proporcional al cambio de temperatura Δu , lo que significa que:

$$\Delta Q = C_p \rho \Delta u$$

donde C_p es la capacidad calorífica (La *capacidad calorífica* de un cuerpo es el cociente entre la cantidad de energía calorífica transferida a un cuerpo o sistema en un proceso cualquiera y el cambio de temperatura que experimenta) y ρ es la densidad de masa del material. Aquí el símbolo Δ denota el operador diferencia usual. Si elegimos energía igual a cero, en el cero absoluto de temperatura, esto puede ser escrito como:

$$Q = C_p \rho u$$

Luego, el aumento de la energía ζ en una pequeña región espacial de los materiales

$$x - \Delta x \leq \zeta \leq x + \Delta x$$

en un intervalo de un periodo de tiempo τ :

$$t - \Delta t \leq \tau \leq t + \Delta t$$

esta dada por:

$$C_p \rho \int_{x-\Delta x}^{x+\Delta x} [u(\zeta, t + \Delta t) - u(\zeta, t - \Delta t)] d\zeta = C_p \rho \int_{x-\Delta x}^{x+\Delta x} \int_{t-\Delta t}^{t+\Delta t} \frac{\partial u}{\partial \tau} d\tau d\zeta$$

donde el teorema fundamental del cálculo fue utilizado. Además, si no hay trabajo y en ausencia de fuentes de calor o sumideros, el cambio en el interior de la energía en el intervalo $[x - \Delta x, x + \Delta x]$ se explica en su totalidad por el flujo de calor a través de las fronteras. Por la ley de Fourier, esto es

$$h \int_{t-\Delta t}^{t+\Delta t} \left[\frac{\partial u}{\partial x}(x + \Delta x, \tau) - \frac{\partial u}{\partial x}(x - \Delta x, \tau) \right] d\tau = h \int_{t-\Delta t}^{t+\Delta t} \int_{x-\Delta x}^{x+\Delta x} \frac{\partial^2 u}{\partial \zeta^2} d\zeta d\tau$$

de nuevo por el teorema fundamental del cálculo. Por el teorema de la conservación de la energía, obtenemos que:

$$\int_{t-\Delta t}^{t+\Delta t} \int_{x-\Delta x}^{x+\Delta x} [C_p \rho u_\tau - h u_{\zeta\zeta}] d\zeta d\tau = 0$$

Esto es cierto para cualquier rectángulo $[t - \Delta t, t + \Delta t] \times [x - \Delta x, x + \Delta x]$. En consecuencia, el integrando debe anularse:

$$C_p \rho u_t - h u_{xx} = 0$$

este resultado también puede ser escrito como:

$$u_t = \frac{h}{C_p \rho} u_{xx}$$

y con $\alpha = \frac{h}{C_p \rho}$, obtenemos finalmente la ecuación buscada:

$$\frac{\partial u}{\partial t} - \alpha \Delta u = 0$$

3.1.2. Formulación variacional del un PVF

Consideremos el espacio de Hilbert V :

$$V = \{v \in H^1 | v(0) = 0\}$$

Sabemos que $t \in [0, T)$, así que, si multiplicamos la ecuación diferencial del problema planteado

$$u_t(x, t) - u_{xx}(x, t) = -\lambda u^p(x, t)$$

por una función $v \in H_0^1$, y luego integramos con respecto a x , en el intervalo $[0, 1]$, obtenemos

$$\int_0^1 u_t(x, t) v(x) dx - \int_0^1 u_{xx}(x, t) v(x) dx = -\lambda \int_0^1 u^p(x, t) v(x) dx$$

Luego, si consideramos la segunda integral

$$\int_0^1 u_{xx}(x, t) v(x) dx$$

y posteriormente integramos por partes

$$\int_0^1 u_{xx}(x, t) v(x) dx = u_x(x, t) v(x) \Big|_{x=0}^{x=1} - \int_0^1 u_x(x, t) v'(x) dx$$

$$\int_0^1 u_{xx}(x,t)v(x)dx = v(1)u_x(1,t) - v(0)u_x(0,t) - \int_0^1 u_x(x,t)'v(x)dx$$

utilizamos las condiciones de frontera y las condiciones sobre v para obtener

$$\int_0^1 u_{xx}(x,t)v(x)dx = v(1)u(1,t)^q - \int_0^1 u_x(x,t)'v(x)dx$$

así, calculamos

$$\int_0^1 u_t(x,t)v(x)dx - v(1)u(1,t)^q + \int_0^1 u_x(x,t)'v(x)dx = -\lambda \int_0^1 u^p(x,t)v(x)dx$$

entonces:

$$\int_0^1 u_t(x,t)v(x)dx = - \int_0^1 u_x(x,t)v'(x)dx + v(1)u(1,t)^q - \lambda \int_0^1 u^p(x,t)v(x)dx$$

Observemos que aquí se tiene un producto interno en V , luego obtenemos:

$$(u_t, v) = -(\nabla u, \nabla v) + u(1,t)^q v(1) - \lambda(u^p, v)$$

De esta manera se obtiene la formulación variacional del problema parabólico planteado. Sea $V_h \subset V$ un espacio de elementos finitos respecto a una partición P_h sobre el intervalo $(0, 1)$. Ahora el objetivo es encontrar una función $u_h(t) = u_h(\cdot, t) \in V_h$ tal que

$$\begin{cases} (u_{h,t}, v) = -(\nabla u_h, \nabla v) + u_h(1,t)^q v(1) - \lambda(u_h^p, v) = H(u, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

Un método comúnmente usado es elegir $u_{o_h}(x) = v_h$ como una interpolación - proyección de v en V_h , para una partición φ_h en $[0, 1]$. Sea $\varphi_j(x) : j = 1, \dots, N_h$ una base para V_h . Consideremos en esta formulación variacional a $u_h(x, t)$ como una combinación lineal de esas funciones, y $v = \varphi_i$, para $i = 1, \dots, N_h$. Entonces nuestro problema semidiscreto puede ser establecido como sigue: encontrar los coeficientes $u_j(t)$ en

$$u_h(x, t) = \sum_{i=1}^{N_h} u_i(t) \varphi_i(x)$$

donde dicha aproximación pertenece a $H^1(0, 1)$ tal que u_h es una aproximación de la solución de nuestro problema parabólico. Así, obtenemos

$$u_h(x, t) = u_1(t)\varphi_1(x) + u_2(t)\varphi_2(x) + \dots + u_{N_h}(t)\varphi_{N_h}(x)$$

Luego, la formulación variacional

$$(u_t, v) = -(\nabla u, \nabla v) + u(1,t)^q v(1) - \lambda(u^p, v)$$

esta estructurada por:

$$\int \sum_{i=1}^N u_i'(t) \varphi_i \varphi_j dx = - \int \sum_{i=1}^N u_i(t) \varphi_i' \varphi_j' dx + \left[\sum_{i=1}^N u_i(t) \varphi_i(1) \right]^q \varphi_j(1) - \lambda \int \left[\sum_{i=1}^N u_i(t) \varphi_i \right]^p \varphi_j dx$$

Luego, como las funciones $u_i(t)$ no dependen de x , entonces podemos dejarlas como factores de las integrales:

$$\sum_{i=1}^N u_i'(t) \int \varphi_i \varphi_j dx = - \sum_{i=1}^N u_i(t) \int \varphi_i' \varphi_j' dx + \left[\sum_{i=1}^N u_i(t) \varphi_i(1) \right]^q \varphi_j(1) - \lambda \int \left[\sum_{i=1}^N u_i(t) \varphi_i \right]^p \varphi_j dx$$

esto equivale a :

$$(\varphi_i, \varphi_j) \sum_{i=1}^N u_i'(t) = -(\varphi_i' \varphi_j') \sum_{i=1}^N u_i(t) + \left[\sum_{i=1}^N u_i(t) \varphi_i(1) \right]^q \varphi_j(1) - \lambda \int \left[\sum_{i=1}^N u_i(t) \varphi_i \right]^p \varphi_j dx$$

Veremos a continuación como esta formulación variacional lleva a un esquema matricial del problema parabólico planteado. Así ,para obtener las soluciones numéricas en los esquemas matriciales necesitamos discretizar la variable tiempo, y así obtener una discretización completa del problema; para ello, trataremos el caso $p = q = 1$, con la restricción $0 < \lambda < 1$, dado que los métodos que utilizaremos necesitan una representación explícita de u , luego:

$$(\varphi_i, \varphi_j) \sum_{i=1}^N u_i'(t) = -(\varphi_i' \varphi_j') \sum_{i=1}^N u_i(t) + \sum_{i=1}^N u_i(t) \varphi_i(1) \varphi_j(1) - \lambda \int \sum_{i=1}^N u_i(t) \varphi_i \varphi_j dx$$

Luego, con

$$M = (\varphi_i, \varphi_j)_{N \times N}, A = (\varphi_i', \varphi_j')_{N \times N}, U' = u_i(t)'_{N \times 1}, U = u_i(t)_{N \times 1}, e_N = (\varphi_i(1) \cdot \varphi_j(1))$$

y $0 < \lambda < 1$ obtenemos el problema matricial:

$$\begin{cases} MU' = -AU + U_N e_N - \lambda MU, t > 0 \\ U(0) = \gamma \end{cases}$$

y como la matriz M es invertible y definida positiva (ver [4]), se tiene una solución única para:

$$U' = -M^{-1}AU + M^{-1}U_N e_N - \lambda U$$

Observemos que este proceso nos conduce a una solución aproximada del problema que ha sido semidiscretizado, así que tal aproximación incluye un error. Las estimaciones sobre los errores se puede consultar en ([19]): De esta forma, aplicaremos 3 métodos muy usados para lograr este propósito: *El método de Euler hacia atrás* , *El método de Euler hacia adelante* y el *método de Cranck- Nicholson*.

El método de Euler hacia atrás: aplicación a un problema parabólico

El método de Euler hacia atrás esquematiza cualquier ecuación diferencial parcial u ordinaria, bajo la aproximación de la derivada por un cociente en diferencia hacia atrás:

$$\frac{\delta u_h(x, t_n)}{\delta t} = \frac{u_h(x, t_n) - u_h(x, t_{n-1})}{k} \approx \frac{u_h^n - u_h^{n-1}}{k}$$

para nuestro caso, tomaremos la notación:

$$\delta_k u_h^n = \frac{u_h^n - u_h^{n-1}}{k}$$

Luego, reemplazando el valor de u_t por su aproximación $\delta_k u_h^n$ en el esquema variacional del problema

$$\begin{cases} (u_{h,t}, v) = -(\nabla u_h, \nabla v) + u_h(1, t)^q v(1) - \lambda(u_h^p, v) = H(u, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

obtenemos:

$$\begin{cases} \delta_k(u_h^n, v) = H(u, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

Así, con $p = q = 1$ y la restricción $0 < \lambda < 1$:

$$\left(\frac{u_h^n - u_h^{n-1}}{k}, v \right) = -(\nabla u_h^n, \nabla v) + u_h^n(1, t)v(1) - \lambda(u_h^n, v)$$

utilizando las propiedades de las formas bilineales y el producto interno definido, calculamos:

$$\frac{1}{k}(u_h^n, v) - \frac{1}{k}(u_h^{n-1}, v) = -(\nabla u_h^n, \nabla v) + u_h^n(1, t)v(1) - \lambda(u_h^n, v)$$

$$(u_h^n, v) - (u_h^{n-1}, v) = -k(\nabla u_h^n, \nabla v) + k u_h^n(1, t)v(1) - \lambda k(u_h^n, v)$$

Reemplazando la combinación lineal $u_h(x, t) = \sum_{i=1}^{N_h} u_i(t)\varphi_i(x)$ en la anterior expresión, y utilizando las propiedades de las formas bilineales, obtenemos:

$$(\varphi_i, \varphi_j) \sum_{i=1}^N u_i^n(t) - (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^{n-1}(t) = -k(\varphi'_i, \varphi'_j) \sum_{i=1}^N u_i^n(t) + k \sum_{i=1}^N u_i^n(t) [\varphi_i(1) \cdot \varphi_j(1)]$$

$$- \lambda k(\varphi_i, \varphi_j) \sum_{i=1}^N u_i^n(t)$$

Tomando los términos en factor de $\sum_{i=1}^N u_i^n(t)$ y $\sum_{i=1}^N u_i^{n-1}(t)$, y además con

$$M = (\varphi_i, \varphi_j)_{N \times N}, A = (\varphi'_i, \varphi'_j)_{N \times N}, U' = u_i(t)'_{N \times 1}, U = u_i(t)_{N \times 1}, e_N = (\varphi_i(1) \cdot \varphi_j(1))$$

calculamos la matriz:

$$[(1 + \lambda k)A + kB - ke_N]U^n = AU^{n-1}$$

Como la matriz factor de U^n es invertible y definida positiva, tenemos que la solución de U^n es única:

$$U^n = [(1 + \lambda k)A + kB - ke_N]^{-1}AU^{n-1}$$

El método de Crank - Nicolson: aplicación a un problema parabólico

En el análisis numérico, el método de Crank-Nicolson es un método de diferencias finitas utilizado para resolver numéricamente la ecuación de calor y ecuaciones similares en derivadas parciales. Es un método de segundo orden en el tiempo, implícito en el tiempo, y es numéricamente estable. El método de Crank-Nicolson se basa en la diferencia central en el espacio, y la regla del trapecio en el tiempo, dando una convergencia de segundo orden en el tiempo. Supongamos que se tiene la ecuación diferencial parcial (en una dimensión) :

$$\frac{\partial U}{\partial t} = F \left[U, x, t, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2} \right]$$

El método de Crank-Nicolson es el promedio del método de Euler hacia adelante en n , y de el método de Euler hacia atrás en $n + 1$; sin embargo, el método en si mismo no es el simple promedio de estos dos métodos de Euler, pues la ecuación tiene una dependencia implícita de la solución:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = F_i^n \left[U, x, t, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2} \right]$$

Euler Hacia Adelante

$$\frac{U_i^n - U_i^{n-1}}{\Delta t} = F_i^{n+1} \left[U, x, t, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2} \right]$$

Euler Hacia Atrás

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} = \frac{1}{2} \left[F_i^n \left[U, x, t, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2} \right] + F_i^{n+1} \left[U, x, t, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2} \right] \right]$$

Crank-Nicolson

La función F debe ser discretizada espacialmente con una diferencia central. Observemos que se trata de un método implícito: Para obtener el "siguiente" valor de u en el tiempo, un sistema de ecuaciones algebraicas debe ser resuelto. Si el error cometido en una etapa de los cálculos no se amplifica en las etapas posteriores, entonces se dice que el **método es estable**. En el siguiente ejemplo veremos la manera de resolver un problema diferencial, utilizando el método de Crank- Nicolson.

Ejemplo 3.1. *Considere la ecuación de calor :*

$$u_t = \alpha u_{xx} + g(x, t)$$

con $\alpha \in \mathbb{R}$, $t > 0$ en el intervalo $a < x < b$ Obtenga una discretización de esta ecuación parabólica utilizando el método de Crank-Nicolson.

Solución

Utilizando el método de Crank-Nicolson, obtenemos la representación discreta del problema :

$$\frac{u_i^{k+1} - u_i^k}{\Delta t} = \frac{\alpha}{2h^2} \left[u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1} + u_{i-1}^k - 2u_i^k + u_{i+1}^k \right] + g_i^{k+\frac{1}{2}}$$

Así tenemos que k representa el superíndice para la variable tiempo t , y también que i representa el subíndice para la variable espacio x . se tiene que $i = 1, 2, \dots, m - 1$, $g_i^k = g(x_i, t^k)$, u^k es la aproximación a la solución $u(t^k) = u(k\Delta t)$, y $t^k = k\Delta t$. El error global se hace sobre la suma de los dos errores en la variable tiempo y espacio, luego el error es $O(\Delta t^2 + h^2)$. Si disponemos de Los términos para u^{k+1} a la izquierda, y los restantes a la derecha, obtenemos el sistema:

$$\frac{u_i^{k+1}}{\Delta t} - \frac{\alpha}{2h^2} \left[u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1} \right] = \frac{u_i^k}{\Delta t} + \frac{\alpha}{2h^2} \left[u_{i-1}^k - 2u_i^k + u_{i+1}^k \right] + g_i^{k+\frac{1}{2}}$$

De esta manera, el sistema que hay que resolver es el establecido por la matriz

$$MU^{K+1} = NU^K + G^k$$

Cuyas matrices M, N y G^k están dadas por

$$M = \begin{bmatrix} \frac{1}{\Delta t} + \frac{\alpha}{h^2} & \frac{-\alpha}{2h^2} & 0 \cdots & 0 & 0 & 0 \\ \frac{-\alpha}{2h^2} & \frac{1}{\Delta t} + \frac{\alpha}{h^2} & \frac{-\alpha}{2h^2} \cdots & 0 & 0 & 0 \\ 0 & \frac{-\alpha}{2h^2} & \frac{1}{\Delta t} + \frac{\alpha}{h^2} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & \frac{1}{\Delta t} + \frac{\alpha}{h^2} & \frac{-\alpha}{2h^2} & 0 \\ 0 & 0 & 0 \cdots & \frac{-\alpha}{2h^2} & \frac{1}{\Delta t} + \frac{\alpha}{h^2} & \frac{-\alpha}{2h^2} \\ 0 & 0 & 0 \cdots & 0 & \frac{-\alpha}{2h^2} & \frac{1}{\Delta t} + \frac{\alpha}{h^2} \end{bmatrix}$$

$$N = \begin{bmatrix} \frac{1}{\Delta t} - \frac{\alpha}{h^2} & \frac{\alpha}{2h^2} & 0 \cdots & 0 & 0 & 0 \\ \frac{\alpha}{2h^2} & \frac{1}{\Delta t} - \frac{\alpha}{h^2} & \frac{\alpha}{2h^2} \cdots & 0 & 0 & 0 \\ 0 & \frac{\alpha}{2h^2} & \frac{1}{\Delta t} - \frac{\alpha}{h^2} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & \frac{1}{\Delta t} - \frac{\alpha}{h^2} & \frac{\alpha}{2h^2} & 0 \\ 0 & 0 & 0 \cdots & \frac{\alpha}{2h^2} & \frac{1}{\Delta t} - \frac{\alpha}{h^2} & \frac{\alpha}{2h^2} \\ 0 & 0 & 0 \cdots & 0 & \frac{\alpha}{2h^2} & \frac{1}{\Delta t} - \frac{\alpha}{h^2} \end{bmatrix}$$

$$F^k = \begin{pmatrix} f_1^{k+\frac{1}{2}} + \frac{\alpha}{h^2} u(0, t^{k+\frac{1}{2}}) \\ f_2^{k+\frac{1}{2}} \\ f_3^{k+\frac{1}{2}} \\ \vdots \\ f_{m-3}^{k+\frac{1}{2}} \\ f_{m-2}^{k+\frac{1}{2}} \\ f_{m-1}^{k+\frac{1}{2}} + \frac{\alpha}{h^2} u(L, t^{k+\frac{1}{2}}) \end{pmatrix}$$

Si la ecuación diferencial parcial es no lineal, la discretización también será no lineal, así que el avance en el tiempo implica la solución de un sistema de ecuaciones algebraicas no lineales, aunque las linealizaciones son posibles. En muchos problemas, especialmente de difusión lineal, el problema algebraico es tridiagonal y puede ser resuelto de manera eficiente con el *algoritmo de Thomas*, lo que da una solución rápida y directa, a diferencia de los métodos habituales para resolver una matriz completa. El método de Crank-Nicolson es muy utilizado en problemas de difusión, pues esquematiza cualquier ecuación diferencial parcial u ordinaria, bajo la aproximación de la derivada por un cociente en diferencia alrededor del punto $t_{n-\frac{1}{2}}$ para una función f :

$$u(x, t_{n-\frac{1}{2}}) \approx \frac{u(x, t_n) + u(x, t_{n-1})}{2}$$

Existe otra manera equivalente de representar el esquema de Crank-Nicolson, dada por el cociente de diferencia alrededor del punto $t_{n+\frac{1}{2}}$:

$$u(x, t_{n+\frac{1}{2}}) \approx \frac{u(x, t_{n+1}) + u(x, t_n)}{2}$$

Ahora el problema consiste en encontrar una función $u_h^n \in V_h$, la aproximación de $u_{n-\frac{1}{2}}$ tal que:

$$\begin{cases} \delta_k(u_h^n, v) = H \left(\frac{u_h^n + u_h^{n-1}}{2}, v \right), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

con

$$-(\nabla u_h, \nabla v) + u_h(1, t)^q v(1) - \lambda(u_h^p, v) = H(u, v)$$

donde $p = q = 1$ y con la restricción $0 < \lambda < 1$. Reemplazando el valor de $\delta_k u_h^n$, con $b(\varphi_i, \varphi_j)_{N \times N} = (\varphi'_i, \varphi'_j)_{N \times N}$, y utilizando las propiedades de las formas bilineales, tenemos:

$$\left(\frac{u_h^n - u_h^{n-1}}{k}, v \right) = -b \left(\frac{u_h^n + u_h^{n-1}}{2}, v \right) + \left[\frac{u_h^n + u_h^{n-1}}{2} \right] \cdot v(1) - \lambda \left(\frac{u_h^n + u_h^{n-1}}{2}, v \right), \forall n \in \mathbb{N}$$

$$(u_h^n, v) - (u_h^{n-1} k, v) = -\frac{k}{2} b(u_h^n, v) - \frac{k}{2} b(u_h^{n-1}, v) + \frac{k}{2} u_h^n \cdot v(1) + \frac{k}{2} u_h^{n-1} \cdot v(1) - \frac{\lambda k}{2} (u_h^n, v)$$

$$-\frac{\lambda k}{2} (u_h^{n-1}, v), \forall v \in V$$

si utilizamos la combinación lineal $u_h(x, t) = \sum_{i=1}^{N_h} u_i(t) \varphi_i(x)$ en la anterior expresión obtenemos:

$$\begin{aligned} & (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^n(t) - (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^{n-1}(t) = \\ & -\frac{k}{2} (\varphi'_i, \varphi'_j) \sum_{i=1}^N u_i^n(t) - \frac{k}{2} (\varphi'_i, \varphi'_j) \sum_{i=1}^N u_i^{n-1}(t) + \frac{k}{2} [\varphi_i(1) \cdot \varphi_j(1)] \sum_{i=1}^N u_i^n(t) + \\ & \frac{k}{2} [\varphi_i(1) \cdot \varphi_j(1)] \sum_{i=1}^N u_i^{n-1}(t) - \frac{\lambda k}{2} (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^n(t) - \frac{\lambda k}{2} (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^{n-1}(t) \end{aligned}$$

con

$$M = (\varphi_i, \varphi_j)_{N \times N}, A = (\varphi'_i, \varphi'_j)_{N \times N}, U' = u_i(t)'_{N \times 1}, U = u_i(t)_{N \times 1}, e_N = (\varphi_i(1) \cdot \varphi_j(1))$$

y tomando los términos en factor de $\sum_{i=1}^N u_i^n(t)$ y $\sum_{i=1}^N u_i^{n-1}(t)$, obtenemos la matriz:

$$\left[A + \frac{k}{2} B - \frac{k}{2} e_N - \frac{\lambda k}{2} A \right] U^n = \left[A - \frac{k}{2} B + \frac{k}{2} e_N + \frac{\lambda k}{2} A \right] U^{n-1}$$

Como la matriz

$$A + \frac{k}{2} B - \frac{k}{2} e_N - \frac{\lambda k}{2} A$$

es definida positiva e invertible, tenemos que

$$U^n = \left[A + \frac{k}{2}B - \frac{k}{2}e_N - \frac{\lambda k}{2}A \right]^{-1} \left[A - \frac{k}{2}B + \frac{k}{2}e_N + \frac{\lambda k}{2}A \right] U^{n-1}$$

El método de Euler hacia adelante: aplicación a un problema parabólico

El método de Euler hacia adelante esquematiza cualquier ecuación diferencial parcial u ordinaria, bajo la aproximación de la derivada por un cociente en diferencia hacia adelante:

$$\frac{\delta u_h(x, t_n)}{\delta t} = \frac{u_h(x, t_{n+1}) - u_h(x, t_n)}{k} \approx \frac{u_h^{n+1} - u_h^n}{k}$$

para nuestro caso, tomaremos el cociente :

$$\delta_k u_h^n = \frac{u_h^n - u_h^{n-1}}{k}$$

Luego, reemplazando el valor de u_t por su aproximación $\delta_k u_h^n$ en el esquema variacional del problema

$$\begin{cases} (u_{h,t}, v) = H(u_h^n, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

obtenemos:

$$\begin{cases} \delta_k(u_h^n, v) = H(u_h^{n-1}, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

Así , con $p = q = 1$ y la restricción $0 < \lambda < 1$:

$$\left(\frac{u_h^n - u_h^{n-1}}{k}, v \right) = -(\nabla u_h^{n-1}, \nabla v) + u_h^{n-1}(1, t)v(1) - \lambda(u_h^{n-1}, v)$$

utilizando las propiedades de las formas bilineales y el producto interno definido, calculamos:

$$\frac{1}{k}(u_h^n, v) - \frac{1}{k}(u_h^{n-1}, v) = -(\nabla u_h^{n-1}, \nabla v) + u_h^{n-1}(1, t)v(1) - \lambda(u_h^{n-1}, v)$$

$$(u_h^n, v) - (u_h^{n-1}, v) = -k(\nabla u_h^{n-1}, \nabla v) + k u_h^{n-1}(1, t)v(1) - \lambda k(u_h^{n-1}, v)$$

Reemplazando la combinación lineal $u_h(x, t) = \sum_{i=1}^{N_h} u_i(t)\varphi_i(x)$ en la anterior expresión, y utilizando las propiedades de las formas bilineales, obtenemos:

$$(\varphi_i, \varphi_j) \sum_{i=1}^N u_i^n(t) - (\varphi_i, \varphi_j) \sum_{i=1}^N u_i^{n-1}(t) = -k(\varphi_i', \varphi_j') \sum_{i=1}^N u_i^{n-1}(t) + k \sum_{i=1}^N u_i^{n-1}(t)[\varphi_i(1) \cdot \varphi_j(1)] -$$

$$\lambda k(\varphi_i, \varphi_j) \sum_{i=1}^N u_i^{n-1}(t)$$

Tomando los términos en factor de $\sum_{i=1}^N u_i^n(t)$ y $\sum_{i=1}^N u_i^{n-1}(t)$, y además con

$$M = (\varphi_i, \varphi_j)_{N \times N}, A = (\varphi'_i, \varphi'_j)_{N \times N}, U' = u_i(t)'_{N \times 1}, U = u_i(t)_{N \times 1}, e_N = (\varphi_i(1) \cdot \varphi_j(1))$$

calculamos la matriz:

$$AU^n = [(1 - \lambda k)A - kB + ke_N]U^{n-1}$$

Como la matriz factor de U^n es invertible y definida positiva, tenemos que la solución de U^n es única:

$$U^n = A^{-1}[(1 - \lambda k)A - kB + ke_N]U^{n-1}$$

A continuación procederemos con la esquematización numérica del problema.

Esquematización numérica: Modelado de la malla

Sabemos que la formulación integral del problema

$$\begin{array}{lll} \text{(EDP)} & u_t - u_{xx} = -\lambda u^p & (x, t) \in (0, 1) \times [0, T) \\ \text{(CF)} & u_x(0, t) = 0, u_x(1, t) = u(1, t)^q & t \in [0, T) \\ \text{(CI)} & u(x, 0) = u_0(x) & x \in [0, 1] \end{array}$$

es

$$\int_0^1 u_t(x, t)v(x)dx - \int_0^1 u_{xx}(x, t)v(x)dx = -\lambda \int_0^1 u^p(x, t)v(x)dx$$

cuyo equivalente variacional está definido por:

$$(u_t, v) = -(\nabla u, \nabla v) + u(1, t)^q v(1) - \lambda(u^p, v)$$

Donde analizaremos el caso $p = q = 1$ y $0 < \lambda < 1$. Así, el objetivo es aproximar el valor de la solución $u(x, t)$ sobre una grilla determinada por x y t . Consideremos el dominio rectangular en el plano (x, t) con $0 < x < 1$ y $0 < t < T < \infty$, para el caso $p = q = 1$, con la restricción $0 < \lambda < 1$. Sea Δx el espaciado en la dirección de x y Δt el espaciado en la dirección de t . Sea

$$h = \frac{1}{N} = \Delta x, k = \frac{T}{M} = \Delta t$$

donde el intervalo en x tiene N particiones y el intervalo en t tiene M particiones, tal que:

$$\begin{aligned} 0 &= x_0 < x_1 < x_2 < \dots < x_N = 1 \\ 0 &= t_0 < t_1 < t_2 < \dots < t_M = T \end{aligned}$$

Luego tenemos que la partición \wp_h en $[0, 1]$ es

$$x_i = \frac{i}{N} = ih, i = 0, 1, 2, \dots, N$$

y la partición \wp_τ en $[0, T)$ es

$$t_j = \frac{j}{M} = jk, i = 0, 1, 2, \dots, N$$

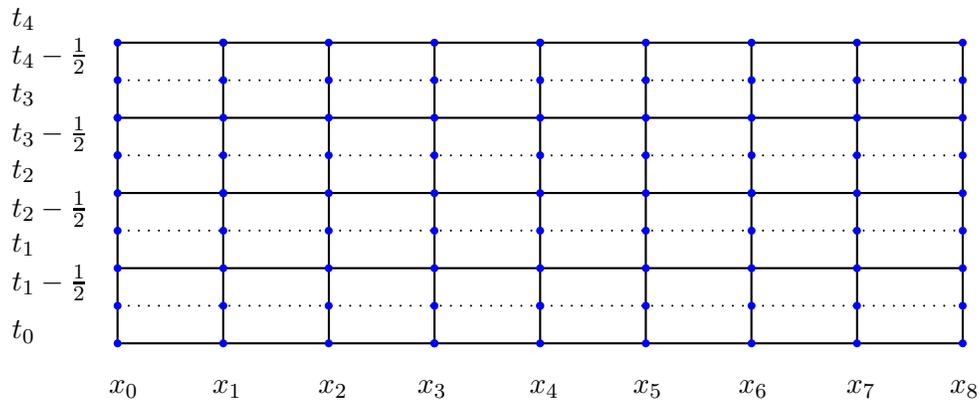
Implementaremos el esquema numérico sobre el esquema obtenido mediante el método de Crank-Nicholson.

Modelado del Esquema numérico para un problema parabólico mediante método Crank-Nicholson

En el paso del tiempo, con $\frac{T}{M} = \Delta t$ tenemos $t_{n-\frac{1}{2}} = t_n - \frac{\Delta t}{2}$ con $n = 1, 2, \dots, M$ Luego, vemos que

$$t_{j-1/2} \in [t_{j-1}, t_j], \quad j = 1, 2, \dots, m.$$

Así, el esquema global tiene que ser resuelto para m niveles de pasos en el tiempo ($n = 1, 2, \dots, m$). La siguiente imagen ilustra la malla para un problema con $N = 8$ y $m = 4$, con $0 \leq i \leq 8$



Cuando $t = t_0$, entonces

$$u_h(x_i, t_j) = v_h(x_i) = \gamma_i$$

para $i = 0, 1, \dots, N$. Esto muestra una solución aproximada en $u_h(x_i, t_j)$ para $i = 1, 2, \dots, N$ y $j = 0, 1, 2, \dots, m$. Cuando $t = t_0$, entonces

$$u_h(x_i, t_j) = v_h(x - i) = \gamma_i$$

para $i = 0, 1, 2, \dots, N$. Sea:

$$V_h = [\varphi : \varphi|_{[x_{i-1}, x_i]}, \text{ es una funcion lineal, con } i = 1, \dots, N, \varphi(0) = 0]$$

Tal espacio de funciones tiene dimension N y sus elementos son llamados **funciones lineales discretas**. Estas funciones, también conocidas como funciones de base global, son la herramienta que utilizaremos en el método de elementos finitos para encontrar una solución aproximada del problema variacional

$$(u_t, v) = -b(u, v) + u(1, t)^q v(1) - (\lambda u^p, v), \forall v \in V_h, p = q = 1, 0 < \lambda < 1$$

definamos también que, si $\varphi_i(x) \in V_h$, entonces $\varphi_i(x_j) = \delta_{ij}$ con $i = 1, 2, \dots, N$ y $j = 0, 1, 2, \dots, N$, donde tales funciones lineales discretas están definidas de la siguiente manera: Para $i = 1, 2, \dots, N - 1$,

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_i}, & x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{h_{i+1}}, & x_i \leq x \leq x_{i+1} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

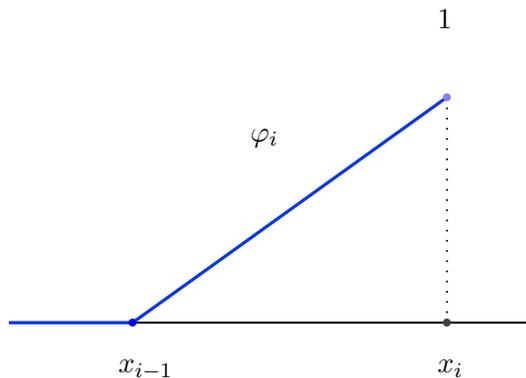
con $h_i = x_i - x_{i-1}$ y $h_{i+1} = x_{i+1} - x_i$. Para estas funciones tenemos que :

1. caso I:

$$\varphi_i(x) = \frac{x - x_{i-1}}{h_i} = \frac{x - x_{i-1}}{x_i - x_{i-1}}$$

Tenemos que $x_{i-1} < x_i$, así $h_i = x_i - x_{i-1} > 0$ Luego, si $x = x_{i-1}$ entonces $\varphi_i(x) = 0$. Esto es una función lineal creciente, pues su pendiente es positiva

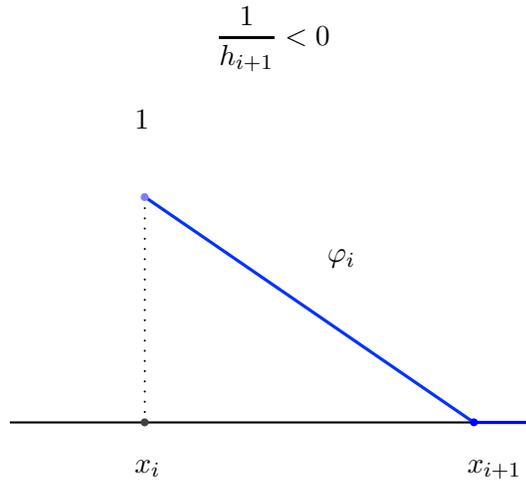
$$\frac{1}{h_i} > 0$$



2. caso II

$$\varphi_i(x) = \frac{x_{i+1} - x}{h_{i+1}}$$

Tenemos que $x_i < x_{i+1}$, así $x_i - x_{i+1} < 0$. Luego, si $\varphi_i(x_i) = 1$ y $\varphi_i(x_{i+1}) = 0$, obtenemos que, para esta función lineal se tiene un comportamiento decreciente, pues su pendiente es negativa



3. caso III

$$i = N$$

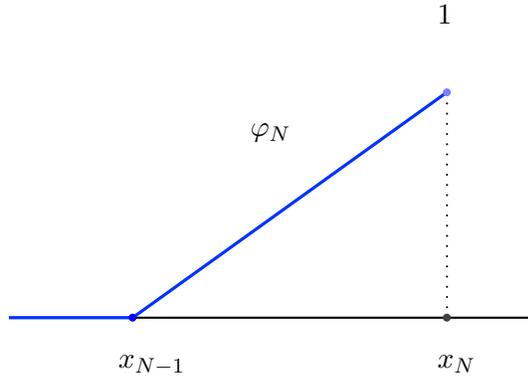
$$\varphi_{N \times N} = \begin{cases} \frac{x - x_{N-1}}{h_N}, & x_{N-1} \leq x \leq x_N \\ 0, & \text{en cualquier otro caso} \end{cases}$$

veamos que

$$h_N = x_N - x_{N-1}$$

para este caso $\varphi_N(x_{N-1}) = 0$, y $\varphi_N(x_N) = 1$. Esto da cuenta de una función lineal creciente, pues su pendiente es positiva

$$\frac{1}{h_N} > 0$$



Ahora podemos ver Gráficamente que si $|i - j| \geq 2$, entonces $Dom\Phi_i \cap Dom\Phi_j = \emptyset$ así que, si $|i - j| \geq 2$ entonces $a_{ij}(\Phi_i, \Phi_j) = 0$. Si $t \in (0, T)$ y $0 < x < 1$ suponemos que $u_h(x, t)$ es la solución aproximada de $u(x, t)$ en V_h . Si tomamos la combinación lineal

$$u_h(x, t) = \sum_{i=1}^{N_h} u_i(t) \varphi_i(x)$$

Luego, el sistema matricial

$$\begin{cases} MU' = -AU + U_N e_N - \lambda MU, t > 0 \\ U(0) = \gamma \end{cases}$$

establecido a través del esquema variacional

$$\begin{cases} (u_{h,t}, v) = -(\nabla u_h, \nabla v) + u_h(1, t)v(1) - \lambda(u_h, v) = H(u, v), \forall v \in V_h, t > 0 \\ u_h(x, 0) = u_{o_h}(x) = v_h(x), x \in (0, 1) \end{cases}$$

será expresado mediante el método de Crank - Nicholson por

$$\left[A + \frac{k}{2}B - \frac{k}{2}e_N - \frac{\lambda k}{2}A \right] U^n = \left[A - \frac{k}{2}B + \frac{k}{2}e_N + \frac{\lambda k}{2}A \right] U^{n-1}$$

con

$$M = (\varphi_i, \varphi_j)_{N \times N}, A = (\varphi'_i, \varphi'_j)_{N \times N}, U' = u_i(t)'_{N \times 1}, U = u_i(t)_{N \times 1}, e_N = (\varphi_i(1) \cdot \varphi_j(1))$$

donde tenemos que:

$$a_{ij} = \int_0^1 \varphi_i(x) \varphi_j(x) dx = (\varphi_i, \varphi_j)$$

$$b_{ij} = \int_0^1 \varphi'_i(x) \varphi'_j(x) dx = (\varphi'_i, \varphi'_j)$$

En estos esquemas matriciales \mathbf{M} es conocida como la *matriz de masa* del sistema, con

$$\mathbf{M} = [(\varphi_i, \varphi_j)]_{N \times N}$$

\mathbf{A} es la *matriz de rigidez*, definida como

$$\mathbf{A} = [b(\varphi_i, \varphi_j)]_{N \times N}$$

con $b(u, v) = (u_x, v')$, y $U = u(t)$ es el vector de los desconocidos $u_i(t)$. A continuación calculamos cada matriz resultante de la aplicación del método de Crank-Nicolson.

Construcción del sistema matricial resultante de la aplicación del Método Crank-Nicolson

Tenemos que si $i - j \geq 2$ o $j - i \geq 2$ entonces

$$a_{ij} = \int_0^1 \varphi_i(x) \varphi_j(x) dx = (\varphi_i, \varphi_j) = 0$$

$$b_{ij} = \int_0^1 \varphi'_i(x) \varphi'_j(x) dx = (\varphi'_i, \varphi'_j) = 0$$

Luego A y B son matrices tridiagonales

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & 0 \cdots & 0 & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 \cdots & 0 & 0 & 0 \\ 0 & a_{3,2} & a_{3,3} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & a_{N-2,N-2} & a_{N-2,N-1} & 0 \\ 0 & 0 & 0 \cdots & a_{N-1,N-2} & a_{N-1,N-1} & a_{N-1,N} \\ 0 & 0 & 0 \cdots & 0 & a_{N,N-1} & a_{N,N} \end{bmatrix}$$

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & 0 \cdots & 0 & 0 & 0 \\ b_{2,1} & b_{2,2} & 0 \cdots & 0 & 0 & 0 \\ 0 & b_{3,2} & b_{3,3} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & b_{N-2,N-2} & b_{N-2,N-1} & 0 \\ 0 & 0 & 0 \cdots & b_{N-1,N-2} & b_{N-1,N-1} & b_{N-1,N} \\ 0 & 0 & 0 \cdots & 0 & b_{N,N-1} & b_{N,N} \end{bmatrix}$$

Observemos que

$$\left[A + \frac{\Delta t}{2} B - \frac{\Delta t}{2} e_N - \frac{\lambda \Delta t}{2} A \right] = \left[A + \frac{k}{2} B - \frac{k}{2} e_N - \frac{\lambda k}{2} A \right]$$

$$\left[A - \frac{\Delta t}{2} B + \frac{\Delta t}{2} e_N + \frac{\lambda \Delta t}{2} A \right] = \left[A - \frac{k}{2} B + \frac{k}{2} e_N + \frac{\lambda k}{2} A \right]$$

De esta manera el esquema matricial

$$\left[A + \frac{k}{2} B - \frac{k}{2} e_N - \frac{\lambda k}{2} A \right] U^n = \left[A - \frac{k}{2} B + \frac{k}{2} e_N + \frac{\lambda k}{2} A \right] U^{n-1}$$

Puede ser representado por:

$$\begin{cases} U^0 = \gamma \\ \tilde{A}U^n = \tilde{F}, \quad n = 1, 2, \dots, m \end{cases}$$

Donde

$$\tilde{A} = \left[A + \frac{\Delta t}{2} B - \frac{\Delta t}{2} e_N - \frac{\lambda \Delta t}{2} A \right]$$

$$\tilde{B} = \left[A - \frac{\Delta t}{2} B + \frac{\Delta t}{2} e_N + \frac{\lambda \Delta t}{2} A \right]$$

$$\tilde{F} = \tilde{B}U^{n-1}$$

Construcción de las matrices \tilde{A} , \tilde{B} Y \tilde{F}

Veamos que A y B son matrices tridiagonales ; luego si representamos a \tilde{A} , \tilde{B} , A Y B por

$$\tilde{A} = \begin{bmatrix} a_1 & b_1 & 0 \cdots & 0 & 0 & 0 \\ c_2 & a_2 & b_2 \cdots & 0 & 0 & 0 \\ 0 & c_3 & a_3 \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & a_{N-2} & b_{N-2} & 0 \\ 0 & 0 & 0 \cdots & c_{N-1} & a_{N-1} & b_{N-1} \\ 0 & 0 & 0 \cdots & 0 & c_N & a_N \end{bmatrix}$$

$$\tilde{B} = \begin{bmatrix} o_1 & p_1 & 0 \cdots & 0 & 0 & 0 \\ q_2 & o_2 & p_2 \cdots & 0 & 0 & 0 \\ 0 & q_3 & o_3 \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & o_{N-2} & p_{N-2} & 0 \\ 0 & 0 & 0 \cdots & q_{N-1} & o_{N-1} & p_{N-1} \\ 0 & 0 & 0 \cdots & 0 & q_N & o_N \end{bmatrix}$$

$$A = \begin{bmatrix} a_{1,1} & b_{1,1} & 0 \cdots & 0 & 0 & 0 \\ c_{1,2} & a_{1,2} & b_{1,2} \cdots & 0 & 0 & 0 \\ 0 & c_{1,3} & a_{1,3} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & a_{1,N-2} & b_{1,N-2} & 0 \\ 0 & 0 & 0 \cdots & c_{1,N-1} & a_{1,N-1} & b_{1,N-1} \\ 0 & 0 & 0 \cdots & 0 & c_{1,N} & a_{1,N} \end{bmatrix}$$

$$B = \begin{bmatrix} a_{2,1} & b_{2,1} & 0 \cdots & 0 & 0 & 0 \\ c_{2,2} & a_{2,2} & b_{2,2} \cdots & 0 & 0 & 0 \\ 0 & c_{2,3} & a_{2,3} \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & a_{2,N-2} & b_{2,N-2} & 0 \\ 0 & 0 & 0 \cdots & c_{2,N-1} & a_{2,N-1} & b_{2,N-1} \\ 0 & 0 & 0 \cdots & 0 & c_{2,N} & a_{2,N} \end{bmatrix}$$

Entonces, las entradas de la matriz \tilde{A} determinadas por a_i, b_i y c_i tienen la estructura

$$a_i = a_{1,i} + \frac{\Delta t}{2} b_{2,i} - \frac{\Delta t}{2} e_N - \frac{\Delta t}{2} a_{1,i}, i = 1, \dots, N$$

$$b_i = a_{1,i} + \frac{\Delta t}{2} b_{2,i} - \frac{\Delta t}{2} e_N - \frac{\Delta t}{2} a_{1,i}, i = 1, \dots, N - 1$$

$$c_i = a_{1,i} + \frac{\Delta t}{2} b_{2,i} - \frac{\Delta t}{2} e_N - \frac{\Delta t}{2} a_{1,i}, i = 2, \dots, N$$

Donde claramente $a_{1,i}, b_{1,i}$ y $c_{1,i}$ son elementos de la matriz A , y los elementos $a_{2,i}, b_{2,i}$ y $c_{2,i}$ son entradas de la matriz B . Para \tilde{B} tenemos :

$$o_i = a_{1,i} - \frac{\Delta t}{2} b_{2,i} + \frac{\Delta t}{2} e_N + \frac{\Delta t}{2} a_{1,i}, i = 1, \dots, N$$

$$p_i = a_{1,i} - \frac{\Delta t}{2} b_{2,i} + \frac{\Delta t}{2} e_N + \frac{\Delta t}{2} a_{1,i}, i = 1, \dots, N - 1$$

$$q_i = a_{1,i} - \frac{\Delta t}{2} b_{2,i} + \frac{\Delta t}{2} e_N + \frac{\Delta t}{2} a_{1,i}, i = 2, \dots, N$$

Calcularemos las entradas de A y B utilizando las funciones de base global.

Matriz A

Elementos de la diagonal

para $i = 1, \dots, N - 1$ con $x_{i-1} \leq x \leq x_{i+1}$, tenemos

$$a_{1,i} = a_{i,i} = \int_0^1 \varphi_i(x)\varphi_i(x)dx = \int_0^1 [\varphi_i(x)]^2 dx$$
$$\int_0^1 [\varphi_i(x)]^2 dx = \int_{x_{i-1}}^{x_i} \left[\frac{x - x_{i-1}}{h} \right]^2 dx + \int_{x_i}^{x_{i+1}} \left[\frac{x_{i+1} - x}{h} \right]^2 dx$$

Por la simetría en el area bajo cada segmento, obtenemos

$$\int_0^1 [\varphi_i(x)]^2 dx = \frac{2}{h^2} \int_{x_{i-1}}^{x_i} \left[\frac{x - x_{i-1}}{h} \right]^2 dx$$

Con el cambio de variable $u = x - x_{i-1}$ obtenemos:

$$a_{1,i} = \frac{2}{h^2} \int_0^h u^2 du = \frac{2}{h^2} \frac{u^3}{3} \Big|_0^h = \frac{2h}{3}$$

para $i = N$, tenemos

$$a_{1,N} = a_{N,N} = \int_{x_{N-1}}^{x_N} \varphi_i(x)\varphi_i(x)dx = \int_{x_{N-1}}^{x_N} [\varphi_i(x)]^2 dx$$

con el cambio de variable $u = x - x_{N-1}$, obtenemos

$$a_{N,N} = \frac{1}{h^2} \int_0^h u^2 du = \frac{h}{3}$$

Elementos por encima de la diagonal para $i = 1, \dots, N - 1$ tenemos

$$b_{1,i} = a_{i,i+1} = \int_{x_i}^{x_{i+1}} \varphi_i(x)\varphi_{i+1}(x)dx$$

$$b_{1,i} = a_{i,i+1} = 2 \int_{x_i}^{\frac{x_i+x_{i+1}}{2}} \left[\frac{x - x_i}{h} \right] \left[\frac{x_{i+1} - x}{h} \right] dx$$

como $x_{i+1} = x_i + h$, tenemos

$$b_{1,i} = a_{i,i+1} = \frac{2}{h^2} \int_{x_i}^{\frac{x_i+x_{i+1}}{2}} - [x_i - x] [(x_i - x) + h] dx$$

$$b_{1,i} = a_{i,i+1} = \frac{2}{h^2} \int_{x_i}^{\frac{x_i+x_{i+1}}{2}} - [(x_i - x)^2 + h(x_i - x)] dx$$

tenemos que $x_{i+1} = x_i + h$ y con el cambio de variable $u = x_i - x$ obtenemos

$$b_{1,i} = a_{i,i+1} = \frac{2}{h^2} \int_0^{\frac{-h}{2}} [u^2 + hu] du$$

$$b_{1,i} = a_{i,i+1} = \frac{2}{h^2} \left[\frac{u^3}{3} + \frac{hu^2}{2} \right] \Big|_0^{\frac{-h}{2}} = \frac{h}{6}$$

Elementos por debajo de la diagonal

para $i = 2, \dots, N$ tenemos

$$c_{1,i} = a_{i,i-1} = \int_{x_{i-1}}^{x_i} \varphi_{i-1}(x) \varphi_i(x) dx$$

$$c_{1,i} = a_{i,i-1} = 2 \int_{x_{i-1}}^{\frac{x_{i-1}+x_i}{2}} \left[\frac{x_i - x}{h} \right] \left[\frac{x - x_{i-1}}{h} \right] dx$$

$$c_{1,i} = a_{i,i-1} = \frac{2}{h^2} \int_{x_i}^{\frac{x_i+x_{i+1}}{2}} \left[\frac{x - x_i}{h} \right] \left[\frac{x_{i+1} - x}{h} \right] dx$$

Así, podemos ver que A es una Matriz Simétrica.

Matriz B

Elementos de la diagonal

para $i = 1, \dots, N - 1$, tenemos

$$a_{2,i} = b_{i,i} = \int_{x_{i-1}}^{x_{i+1}} \varphi'_i(x) \varphi'_i(x) dx = \int_{x_{i-1}}^{x_{i+1}} [\varphi'_i(x)]^2 dx$$

En el intervalo $[x_{i-1}, x_i]$ tenemos

$$\varphi_i(x) = \frac{x - x_{i-1}}{h}$$

En el intervalo $[x_i, x_{i+1}]$ tenemos

$$\varphi_i(x) = \frac{x_{i+1} - x}{h}$$

Luego, la integral la podemos determinar por

$$\int_{x_{i-1}}^{x_{i+1}} [\varphi'_i(x)]^2 dx = \int_{x_{i-1}}^{x_i} [\varphi'_i(x)]^2 dx + \int_{x_i}^{x_{i+1}} [\varphi'_i(x)]^2 dx$$

Calculando las derivadas de funciones lineales, obtenemos

$$\int_{x_{i-1}}^{x_i} [\varphi'_i(x)]^2 dx + \int_{x_i}^{x_{i+1}} [\varphi'_i(x)]^2 dx = \frac{1}{h^2} \int_{x_{i-1}}^{x_i} dx + \frac{1}{h^2} \int_{x_i}^{x_{i+1}} dx$$

$$\frac{1}{h^2} \int_{x_{i-1}}^{x_i} dx + \frac{1}{h^2} \int_{x_i}^{x_{i+1}} dx = \frac{1}{h^2} x \Big|_{x_{i-1}}^{x_i} + \frac{1}{h^2} x \Big|_{x_i}^{x_{i+1}} = \frac{2}{h}$$

para $i = N$, tenemos

$$a_{2,N} = b_{N,N} = \int_{x_{N-1}}^{x_N} \varphi_i(x)'_N \varphi_i(x)'_N dx = \int_{x_{N-1}}^{x_N} [\varphi_i(x)'_N]^2 dx$$

$$\int_{x_{N-1}}^{x_N} [\varphi_i(x)'_N]^2 dx = \frac{1}{h^2} \int_{x_{N-1}}^{x_N} dx$$

$$\frac{1}{h^2} \int_{x_{N-1}}^{x_N} dx = \frac{1}{h^2} x \Big|_{x_{N-1}}^{x_N} = \frac{1}{h^2} x \Big|_{x_{N-1}}^{x_N} = \frac{1}{h}$$

Elementos por encima de la diagonal

para $i = 1, \dots, N-1$ tenemos

$$b_{2,i} = b_{i,i+1} = \int_{x_i}^{x_{i+1}} \varphi_i(x)' \varphi_{i+1}(x)' dx$$

Calculando las derivadas de funciones lineales, obtenemos

$$b_{2,i} = b_{i,i+1} = \frac{1}{h^2} \int_{x_i}^{x_{i+1}} dx = \frac{1}{h^2} x \Big|_{x_i}^{x_{i+1}} = \frac{1}{h^2} x \Big|_{x_i}^{x_{i+1}} = \frac{1}{h}$$

Elementos por debajo de la diagonal

para $i = 2, \dots, N$ tenemos

$$c_{2,i} = b_{i,i-1} = \int_{x_{i-1}}^{x_i} \varphi_{i-1}(x)' \varphi_i(x)' dx$$

$$c_{2,i} = b_{i,i-1} = \frac{1}{h^2} \int_{x_{i-1}}^{x_i} dx$$

$$\frac{1}{h^2} \int_{x_{i-1}}^{x_i} dx = \frac{1}{h^2} x \Big|_{x_{i-1}}^{x_i} = \frac{1}{h^2} x \Big|_{x_{i-1}}^{x_i} = \frac{1}{h}$$

Así, podemos ver que B es una matriz simétrica.

Matriz \tilde{F}

tenemos que

$$\tilde{F} = \tilde{B}U^{n-1}$$

Luego determinaremos cada matriz como se describe a continuación.

Matriz $\tilde{B}U^{n-1}$

Observemos que ya tenemos \tilde{B} donde sus entradas están dadas por:

$$o_i = a_{1,i} - \frac{\Delta t}{2}b_{2,i} + \frac{\Delta t}{2}e_N + \frac{\Delta t}{2}a_{1,i}, i = 1, \dots, N$$

$$p_i = a_{1,i} - \frac{\Delta t}{2}b_{2,i} + \frac{\Delta t}{2}e_N + \frac{\Delta t}{2}a_{1,i}, i = 1, \dots, N - 1$$

$$q_i = a_{1,i} - \frac{\Delta t}{2}b_{2,i} + \frac{\Delta t}{2}e_N + \frac{\Delta t}{2}a_{1,i}, i = 2, \dots, N$$

$$\tilde{B} = \begin{bmatrix} o_1 & p_1 & 0 \cdots & 0 & 0 & 0 \\ q_2 & o_2 & p_2 \cdots & 0 & 0 & 0 \\ 0 & q_3 & o_3 \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & o_{N-2} & p_{N-2} & 0 \\ 0 & 0 & 0 \cdots & q_{N-1} & o_{N-1} & p_{N-1} \\ 0 & 0 & 0 \cdots & 0 & q_N & o_N \end{bmatrix}$$

luego si definimos $\tilde{B}U^{n-1} = C^{n-1}$, tenemos que las entradas de C^{n-1} son

-Para $j = 1$

$$C_1^{n-1} = o_1u_1^{n-1} + p_1u_2^{n-1}$$

-Para $j = 2, \dots, N - 1$

$$C_j^{n-1} = q_ju_{j-1}^{n-1} + o_ju_j^{n-1} + p_ju_{j+1}^{n-1}$$

-Para $j = N$

$$C_N^{n-1} = q_Nu_{N-1}^{n-1} + o_Nu_N^{n-1}$$

Ensamble del vector condición inicial $U^0 = \gamma$

Existen dos maneras de encontrar el vector $U^0 = \gamma$ y la decision final sobre cual de las dos elegir depende de las características de la condición inicial en el problema de la ecuación diferencial parcial. Si $v \in H_0^1(\Omega)$ las dos formas no difieren en mucho. Sin embargo, la primera forma es la mas eficiente pues su calculo no implica resolver un sistema de ecuaciones lineales como veremos.

1. $v \in H_0^1[0, 1]$: sea $I_h : V \rightarrow V_h$ donde I_h es el operador interpolante en V . Supongamos que $U_h(x, 0) = I_h v(x) = v_h(x) = u_{0h}(x)$, luego $v_h(x)$ se puede expresar por:

$$u_{0h}(x) = v_h(x) = \sum_{i=1}^N v(x_i) \varphi_j(x)$$

Esta es la *representación interpolante* de v en V_h . Como tenemos que $\varphi_i(x_j) = \delta_{ij}$ para $i, j = 1, \dots, N$. Luego podemos decir que :

$$\gamma_j = u_{0h}(x_j) = v_h(x_j) = \sum_{i=1}^N v(x_i) \delta_{ij} = v(x_j)$$

Con $j = 1, \dots, N$. Así tenemos que el vector condición inicial tiene la estructura:

$$\gamma_{Nx1} = (v(x_1), v(x_2), v(x_3), \dots, v(x_{N-2}), v(x_{N-1}), v(x_N))^T = \begin{pmatrix} v(x_1) \\ v(x_2) \\ v(x_3) \\ \vdots \\ v(x_{N-2}) \\ v(x_{N-1}) \\ v(x_N) \end{pmatrix}$$

2. $v \in L^2$: Para encontrar las componentes γ_i de la aproximación inicial dada $v_h = u_{0h}$, podemos considerar $u_h(x, 0) = v(x)$. Así, calculamos la representación de $u(x, 0)$ en V_h dada por

$$u_h(x, 0) = \sum_{i=1}^N u_i(0) \varphi_i(x)$$

Entonces, considerando el producto interno en L^2 podemos encontrar los coeficientes $u_i(0) = \gamma_i$ desde la ecuación

$$(u_h(x, 0), \varphi_j(x)) = (v(x), \varphi_j(x)) \quad j = 1, \dots, N$$

Estos significa que la forma iterativa de esta representación es

$$\sum_{i=1}^N u_i(0) (\varphi_i(x), \varphi_j(x)) = (v(x), \varphi_j(x)) \quad j = 1, \dots, N$$

Lo cual equivale matricialmente a $Au(0) = \Phi$. Como A es una matriz diagonal y definida positiva, entonces finalmente encontramos que $u^0 = u(0) = A^{-1}\Phi = \gamma$. En este caso

$$\Phi_j = (v(x), \varphi_j(x))$$

$$\Phi_j = \frac{1}{h} \left[\int_{x_{j-1}}^{x_j} v(x)(x - x_{j-1})dx - \int_{x_j}^{x_{j+1}} v(x)(x - x_{j+1})dx \right]$$

El proceso del refinamiento sobre la malla

Para hacer un análisis numérico de la convergencia para cualquier método tratado, es necesario introducir la idea del **refinamiento** de la partición inicial en la variable espacial. Supongamos así que $h = \frac{1}{N}$ es la longitud de la partición uniforme $\varphi_h : 0 = x_0 < x_1 < \dots < x_N = 1$ en el intervalo $[0, 1]$ de $N+1$ puntos o nodos. El primer paso en el refinamiento sobre φ_h se hace tomando el doble de puntos en $[0, 1]$ lo cual es equivalente a dividir la longitud h en 2:

$$\frac{h}{2} = \frac{1}{2N}$$

En general, el proceso de R-refinamientos sucesivos para una partición uniforme φ_h en $[0, 1]$ se hace al dividir la longitud original h en 2^j , lo cual es equivalente a multiplicar el número de subintervalos N en φ_h por 2^j , para $j = 1, \dots, R$:

$$\frac{h}{2^j} = \frac{1}{2^j N} \quad j = 1, \dots, R$$

De tal forma que es posible denotar $u_h(x, t)$ por $u_N(x, t)$ y en general $u_{\frac{h}{2^j}}(x, t)$ por $u_{2^j N}(x, t)$, lo cual significa que

$$u_{\frac{h}{2^j}}(x, t) \equiv u_{2^j N}(x, t) \equiv u_j(x, t), \quad j = 0, 1, \dots, R$$

Error teórico: caso para la solución desconocida

Si suponemos que la solución a nuestro problema

$$\begin{array}{lll} \text{(EDP)} & u_t - u_{xx} = -\lambda u^p & (x, t) \in (0, 1) \times [0, T) \\ \text{(CF)} & u_x(0, t) = 0, u_x(1, t) = u(1, t)^q & t \in [0, T) \\ \text{(CI)} & u(x, 0) = u_0(x) & x \in [0, 1] \end{array}$$

es desconocida, para examinar error teórico $u_h - u_{\frac{h}{2}}$ consideraremos la norma L^2 y la norma H^1 de la siguiente manera:

$$\left\| u_h - u_{\frac{h}{2}} \right\|_{L^2[0,1],t}^2 = \int_0^1 \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx \quad t = t_n, n = 0, 1, \dots, M$$

$$\left\| u_h - u_{\frac{h}{2}} \right\|_{H^1[0,1]}^2 = \left\| u_h - u_{\frac{h}{2}} \right\|_{L^2[0,1],t}^2 + \left\| u_{h,x} - u_{\frac{h}{2},x} \right\|_{L^2[0,1],t}^2, \quad t = t_n, n = 0, 1, \dots, M$$

Recordemos que si $\varphi_h : 0 = x_0 < x_1 < \dots < x_N = 1$ es una partición sobre $[0, 1]$ de longitud $h = \frac{1}{N}$ entonces $\frac{h}{2N}$ es un refinamiento de φ_h ; este se obtiene tomando el doble de puntos N , o de manera equivalente reduciendo la longitud h a la mitad. Así:

$$u_h(x, t) - u_{\frac{h}{2}}(x, t) = u_N(x, t) - u_{2N}(x, t)$$

$$u_{h,x}(x, t) - u_{\frac{h}{2},x}(x, t) = u_{N,x}(x, t) - u_{2N,x}(x, t)$$

entonces

$$\left\| u_h - u_{\frac{h}{2}} \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \int_{x_{2(i-1)}}^{x_{2i}} \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx$$

$$\left\| u_h - u_{\frac{h}{2}} \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \left[\int_{x_{2(i-1)}}^{x_{2i-1}} \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx + \int_{x_{2i-1}}^{x_{2i}} \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx \right]$$

y también

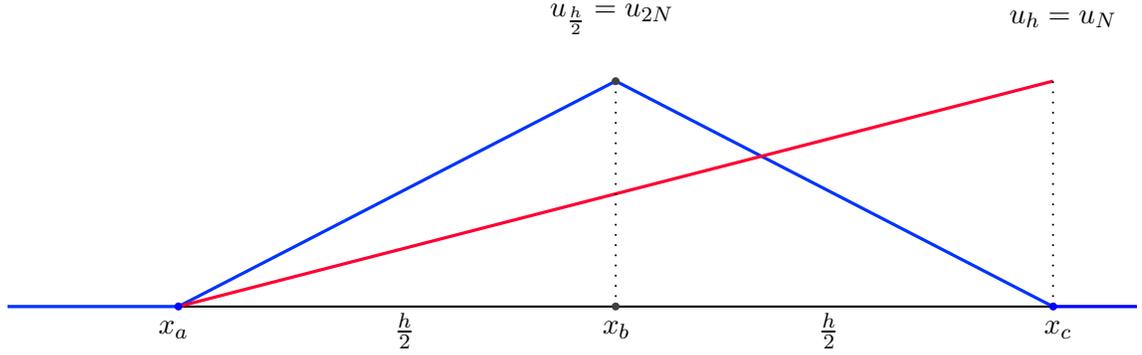
$$\left\| u_{h,x} - u_{\frac{h}{2},x} \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \int_{x_{2(i-1)}}^{x_{2i}} \left[u_{h,x}(x, t) - u_{\frac{h}{2},x}(x, t) \right]^2 dx$$

$$\left\| u_{h,x} - u_{\frac{h}{2},x} \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \left[\int_{x_{2(i-1)}}^{x_{2i-1}} \left[u_{h,x}(x, t) - u_{\frac{h}{2},x}(x, t) \right]^2 dx + \int_{x_{2i-1}}^{x_{2i}} \left[u_{h,x}(x, t) - u_{\frac{h}{2},x}(x, t) \right]^2 dx \right]$$

desde que el intervalo $[x_{2(i-1)}, x_{2i}]$ es la union de los intervalos $[x_{2(i-1)}, x_{2i-1}]$ y $[x_{2i-1}, x_{2i}]$.

Ahora reducimos el análisis a los intervalos $[x_a, x_b]$ y $[x_b, x_c]$ con longitud igual a $\frac{h}{2}$ y donde asumimos la siguiente correspondencia

$$\begin{array}{lll} x_a \longrightarrow x_{2(i-1)} & u_{h,2(i-1)} \longrightarrow v_a & u_{\frac{h}{2},2(i-1)} \longrightarrow u_a^n \\ x_b \longrightarrow x_{2i-1} & u_{h,2i-1} \longrightarrow v_b & u_{\frac{h}{2},2i-1} \longrightarrow u_b^n \\ x_c \longrightarrow x_{2i} & u_{h,2i} \longrightarrow v_c & u_{\frac{h}{2},2i} \longrightarrow u_c^n \end{array}$$



Donde estos u_a^n, u_b^n, u_c^n son los $u_i(t)^n$. Consideremos a $u_h = u_N$ y $u_{\frac{h}{2}} = u_{2N}$ que son definidos en $[x_a, x_c]$ de la siguiente forma:

$$u_h |_{[x_a, x_c]} = v_a \frac{x_c - x}{h} + v_c \frac{x - x_a}{h}$$

y

$$u_{\frac{h}{2}} |_{[x_a, x_c]} = \begin{cases} u_a^n \frac{x_b - x}{\frac{h}{2}} + u_b^n \frac{x - x_a}{\frac{h}{2}}, & x \in [x_a, x_b] \\ u_b^n \frac{x_c - x}{\frac{h}{2}} + u_c^n \frac{x - x_b}{\frac{h}{2}}, & x \in [x_b, x_c] \end{cases}$$

$$u_{\frac{h}{2}} |_{[x_a, x_c]} = \begin{cases} 2u_a^n \frac{x_b - x}{h} + 2u_b^n \frac{x - x_a}{h}, & x \in [x_a, x_b] \\ 2u_b^n \frac{x_c - x}{h} + 2u_c^n \frac{x - x_b}{h}, & x \in [x_b, x_c] \end{cases}$$

así tenemos que

$$u_h - u_{\frac{h}{2}} |_{[x_a, x_b]} = (v_c - 2u_b^n) \frac{x - x_a}{h} + 2u_a^n \frac{x - x_b}{h} - v_a \frac{x - x_c}{h}$$

$$u_h - u_{\frac{h}{2}} |_{[x_b, x_c]} = v_c \frac{x - x_a}{h} + 2u_c^n \frac{x - x_b}{h} - (v_a - 2u_b^n) \frac{x - x_c}{h}$$

y también

$$u_{h,x} - u_{\frac{h}{2},x} |_{[x_a, x_b]} = \frac{2(u_a^n - u_b^n) - (v_a - v_c)}{h}$$

$$u_{h,x} - u_{\frac{h}{2},x} |_{[x_b, x_c]} = \frac{2(u_b^n - u_c^n) - (v_a - v_c)}{h}$$

Ahora, como $u_h - u_{\frac{h}{2}}$ es una función lineal entonces su derivada $u_{h,x} - u_{\frac{h}{2},x}$ es una constante.

También $x_b - x_a = x_c - x_b = \frac{h}{2}$, $x_c - x_a = h$, así integrando por sustitución obtenemos las siguientes integrales:

$$\begin{aligned} & \int_{x_a}^{x_b} \left[u_h - u_{\frac{h}{2}} \right]^2 dx = \\ &= \frac{h}{3 \left[2(u_a^n - u_b^n) - (v_a - v_c) \right]} \left[\left[(v_c - 2u_b^n) \frac{x_b - x_a}{h} - v_a \frac{x_b - x_c}{h} \right]^3 - \left[2u_a^n \frac{x_a - x_b}{h} - v_a \frac{x_a - x_c}{h} \right]^3 \right] \\ & \int_{x_a}^{x_b} \left[u_h - u_{\frac{h}{2}} \right]^2 dx = \frac{h}{24} \left[\frac{(v_a + v_c - 2u_b^n)^3 - 8(v_a - u_a^n)^3}{2(u_a^n - u_b^n) - (v_a - v_c)} \right] \end{aligned}$$

y

$$\begin{aligned} & \int_{x_b}^{x_c} \left[u_h - u_{\frac{h}{2}} \right]^2 dx = \\ &= \frac{h}{3 \left[2(u_b^n - u_c^n) - (v_a - v_c) \right]} \left[\left[(v_c \frac{x_c - x_a}{h} - 2u_c^n \frac{x_c - x_b}{h})^3 - \left[v_c \frac{x_b - x_a}{h} - (v_a - 2u_b^n) \frac{x_b - x_c}{h} \right]^3 \right] \right] \\ & \int_{x_a}^{x_b} \left[u_h - u_{\frac{h}{2}} \right]^2 dx = \frac{h}{24} \left[\frac{-(u_a^n + 2v_c - 2u_b^n)^3 + 8(v_c - u_c^n)^3}{2(u_b^n) - (v_a - v_c)} \right] \end{aligned}$$

Finalmente tenemos que

$$\int_{x_a}^{x_b} \left[u_{h,x} - u_{\frac{h}{2},x} \right]^2 dx = \left[\frac{2(u_a^n - u_b^n) - (v_a - v_c)}{h} \right] (x_b - x_a) = \frac{2(u_a^n - u_b^n) - (v_a - v_c)}{2}$$

y también

$$\int_{x_b}^{x_c} \left[u_{h,x} - u_{\frac{h}{2},x} \right]^2 dx = \left[\frac{2(u_b^n - u_c^n) - (v_a - v_c)}{h} \right] (x_c - x_b) = \frac{2(u_b^n - u_c^n) - (v_a - v_c)}{2}$$

Esto implica

$$\begin{aligned} & \int_{x_{2(i-1)}}^{x_{2i-1}} \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx = \frac{h}{24} \left[\frac{(v_{2(i-1)} + v_{2i} - 2u_{2i-1}^n)^3 - 8(v_{2(i-1)} - u_{2(i-1)}^n)^3}{2(u_{2(i-1)}^n - u_{2i-1}^n) - (v_{2(i-1)} - v_{2i})} \right] \\ & \int_{x_{2i-1}}^{x_{2i}} \left[u_h(x, t) - u_{\frac{h}{2}}(x, t) \right]^2 dx = \frac{h}{24} \left[\frac{-(u_{2(i-1)}^n + 2v_{2i} - 2u_{2i-1}^n)^3 + 8(v_{2i} - u_{2i}^n)^3}{2(u_{2i-1}^n - u_{2i}^n) - (v_{2(i-1)} - v_{2i})} \right] \end{aligned}$$

y también

$$\int_{x_{2(i-1)}}^{x_{2i-1}} \left[u_{h,x} - u_{\frac{h}{2},x} \right]^2 dx = \frac{2(u_{2(i-1)}^n - u_{2i-1}^n) - (v_{2(i-1)} - v_{2i})}{2}$$

$$\int_{x_{2i-1}}^{x_{2i}} \left[u_{h,x} - u_{\frac{h}{2},x} \right]^2 dx = \frac{2(u_{2i-1}^n - u_{2i}^n) - (v_{2(i-1)} - v_{2i})}{2}$$

De esta manera, obtenemos

$$\left\| u_h(x, t) - u_{\frac{h}{2}}(x, t) \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \left[\frac{h}{24} \left[\frac{(v_{2(i-1)} + v_{2i} - 2u_{2i-1}^n)^3 - 8(v_{2(i-1)} - u_{2(i-1)}^n)^3}{2(u_{2(i-1)}^n - u_{2i-1}^n) - (v_{2(i-1)} - v_{2i})} \right] \right] +$$

$$\sum_{i=1}^N \left[\frac{h}{24} \left[\frac{-(u_{2(i-1)}^n + 2v_{2i} - 2u_{2i-1}^n)^3 + 8(v_{2i} - u_{2i}^n)^3}{2(u_{2i-1}^n - u_{2i}^n) - (v_{2(i-1)} - v_{2i})} \right] \right]$$

$$\left\| u_{h,x} - u_{\frac{h}{2},x} \right\|_{L^2[0,1],t}^2 = \sum_{i=1}^N \left[\frac{2(u_{2(i-1)}^n - u_{2i-1}^n) - (v_{2(i-1)} - v_{2i})}{2} \right] +$$

$$\sum_{i=1}^N \left[\frac{2(u_{2i-1}^n - u_{2i}^n) - (v_{2(i-1)} - v_{2i})}{2} \right]$$

Abordaremos ahora el fenómeno de explosión en EDPs, conocido como *blow-up*.

Análisis del problema parabólico: Fenómeno de explosión (Blow-Up)

Consideremos el problema parabólico inicialmente planteado con condición de frontera no lineal y con termino de absorción no lineal (ver [4]):

$$\begin{aligned} u_t &= u_{xx} - \lambda u^p, & \text{en } (0, 1) \times [0, T) \\ u_x(1, t) &= u(1, t)^q, & \text{en } [0, T) \\ u_x(0, t) &= 0, & \text{en } [0, T) \\ u(x, 0) &= u_0(x) \geq 0 & \text{en } [0, 1] \end{aligned} \tag{3.1}$$

en donde $p, q > 1$, $\lambda > 0$ son parámetros, $T > 0$ y u_0 es una función no negativa. Aquí, $u_x(1, t) = u(1, t)^q$ en $[0, T)$ indica que el cambio de temperatura respecto a la posición en la barra, en el extremo $x = 1$, para cualquier instante de tiempo t , viene dado por $u(1, t)^q$. la condición $u_x(0, t) = 0$, en $[0, T)$ determina que el cambio de temperatura respecto a la posición en la barra es cero en el extremo izquierdo, para cualquier tiempo t . La condición inicial $u(x, 0) = u_0(x) \geq 0$ en $[0, 1]$ implica que la temperatura inicial, para $t = 0$, para cualquier posición x en la barra, es la misma, y es positiva o nula. Para comprender el comportamiento de las soluciones para este problema no homogéneo, primero analizaremos

el comportamiento de las soluciones de un problema similar pero con un menor grado de complejidad. Aunque sabemos que las soluciones de una ecuación de calor homogénea tienden a suavizarse con el transcurrir del tiempo, una pregunta natural es ¿las soluciones u de este problema podrán en algún momento tomar valores inmensurables para ciertos valores de t o de x ? Las soluciones para ciertos problemas de reacción-difusión se pueden hacer ilimitadas (es decir, no acotadas), fenómeno conocido como **explosión** o **blow-up**.

Definición 3.1. Se dice que un punto (\hat{x}, t_b) es un **punto de blow up** (o de *b-explosion*) de una función $u(x, t)$, si existe una sucesión de puntos $(x_k, t_k) \in (0, 1) \times (0, T]$, $T > 0$ donde

$$u(x_k, t_k) \longrightarrow \infty, \text{ cuando } (x_k, t_k) \longrightarrow (\hat{x}, t_b^-)$$

Definición 3.2. Un fenómeno que describe Blow-up puede tener más de un punto de explosión. Así, si el blow-up se da en todo el intervalo $[0, 1]$, el fenómeno se conoce como **blow-up global**. Si este blow-up ocurre en un subintervalo $[\ell, 1]$ con $0 < \ell < 1$ esto se conoce como **blow-up regional**.

El fenómeno de blow-up se puede dar tanto en EDPs como en EDOs. El caso más simple de Blow-up se da en EDOs, con el siguiente problema de valor inicial:

$$u_t = u^p, p > 1 \quad u = u(t) \quad u(0) = a > 0$$

Aplicando el método de separación de variables, obtenemos:

$$\int u^{-p} du = \int dt$$

$$\frac{u^{-p+1}}{-p+1} + C = t$$

con la condición inicial $u(0) = a$, obtenemos $C = \frac{-a^{-p+1}}{-p+1}$, y así calculamos u :

$$u^{1-p} - a^{1-p} = t(1-p)$$

$$u = [a^{1-p} + t(1-p)]^{\frac{1}{1-p}}$$

Elijamos los valores de p como números enteros, con restricción $p > 1$: para $p = 2$, obtenemos la ecuación

$$u = \left[\frac{1}{a} - t \right]^{-1}$$

Es claro que la solución es única para un dato $a > 0$ en el intervalo $0 < t < \frac{1}{a}$. Luego, si tomamos $T = \frac{1}{a}$, y tomamos el límite sobre u cuando $t \rightarrow T^-$, tenemos que este límite va hacia infinito. De esta forma

$$\lim_{t \rightarrow T^-} u(t) = \lim_{t \rightarrow T^-} [T - t]^{-1} = \infty$$

podemos observar el *comportamiento asintótico* de u , por la izquierda de T . Decimos así que u **explota en** $t = T = \frac{1}{a}$. Esto nos permite afirmar que el blow-up es un fenómeno, que desde un enfoque matemático permite ver como, *para ciertos valores de las variables independientes* (el blow-up también se puede dar en una cantidad finita de puntos del dominio Ω de la función), *la variable dependiente va hacia infinito*. Aquí vimos que para cierto valor de t , u va hacia infinito, lo cual puede ser escrito también como

$$\sup_{x \in \Omega} |u(x, t)| \longrightarrow \infty, \text{ cuando } t \longrightarrow T^-$$

La mayoría de fenómenos de reacción-difusión toman la estructura

$$u_t = \Lambda(u) + f(u)$$

en Donde Λ es un operador elíptico de segundo orden, el cual representa la **difusión** (en el problema planteado esta representado por el laplaciano), y $f(u)$ es normalmente una función superlineal de u , que representa la **reacción** (en el problema planteado esta representado por la función $f(u) = -\lambda u^p$). Supongamos que se tiene una solución $u(x, t)$ de $[1,1]$ en $(0, 1) \times [0, T)$ luego, el **conjunto de explosión** (o **conjunto de blow-up**) se define como:

$$B(u_0) = [x \in (0, 1) : \exists [x_n, t_n] \subset (0, 1) \times [0, T), t_n \longrightarrow T^-, x_n \longrightarrow x, u(x_n, t_n) \longrightarrow \infty]$$

luego, el conjunto explosion es no vacío si por lo menos existe un punto $x \in (0, 1)$ tal que para algún $t_n \in [0, T)$, cuando t_n se acerca por la izquierda a T , entonces ocurre que $u(x_n, t_n) \longrightarrow \infty$.

3.1.3. Formulación variacional discreta generalizada

A continuación trataremos el problema determinado por (3.1) desde el blow - up, el cual cuenta con un sistema matricial

$$\begin{cases} MU' = -AU + u_N^q e_N - \lambda MU^p \\ U(0) = u_0^I \end{cases} \quad (3.2)$$

Cuya ecuación, escrita en forma de sistema, se representa por

$$\begin{cases} u'_1 = \frac{2}{h^2}(u_2 - u_1) - \lambda u_1^p \\ u'_k = \frac{1}{h^2}(u_{k+1} - 2u_k + u_{k-1}) - \lambda u_k^p, & 2 \leq k \leq N-1 \\ u'_N = \frac{2}{h^2}(u_{N-1} - u_N) - \lambda u_N^p + \frac{2}{h} u_N^q \\ u_i(0) = u_0(x_i) \end{cases}, \quad 1 \leq i \leq N \quad (3.3)$$

En nuestro problema se tiene un termino de reacción en la frontera y un termino de absorción en la ecuación. Estos dos términos compiten y el fenómeno de blow-up ocurre si y solo si $p < 2q - 1$ o $p = 2q - 1$ con $\lambda < q$.

Teorema 3.1. *Explosion dependiente de parámetros p, q y λ para el problema (3.1)*

1. *Suponga que $p < 2q - 1$ o $p = 2q - 1$ con $\lambda < q$, si $u_0 > v$, donde v es alguna solución máxima estacionaria. Entonces u explota en tiempo finito.*
2. *Suponga que $p > 2q - 1$ o $p = 2q - 1$ con $\lambda \geq q$, entonces toda solución positiva es global*

Teorema 3.2. *Sea $u \in C^{2,1}([0, 1] \times [0, T_1])$ una solución de (3.1) y u_h su aproximación semidiscreta. Entonces existe una constante C dependiente en T_1 y u tal que, para un h lo suficientemente pequeño se cumple que*

$$\|u - u_h\|_{L^\infty, ([0,1] \times [0, T_1])} \leq Ch^{\frac{3}{2}}$$

Describiremos a continuación cuando ocurre el fenómeno de blow-up para el esquema semidiscreto presentado por (3.3), en términos de los parámetros p y q .

3.2. Análisis de B-explosión semidiscreta

Teorema 3.3. *Sea $U = (u_1, \dots, u_N)$ sea una solución positiva de (3.3). Entonces:*

1. *si $p \leq q$ y y el dato inicial es lo suficientemente grande, entonces $U = (u_1, \dots, u_N)$ tiene tiempo de explosion finito.*
2. *Si $p > q$ entonces $U = (u_1, \dots, u_N)$ es global.*

Demostración. 1. caso $p \leq q$. Sea

$$\phi(u) \equiv \int_0^1 \frac{(u_x)^2}{2} + \lambda \int_0^1 \frac{u^{p+1}}{p+1} - \frac{u^{q+1}(1, t)}{q+1}$$

El ϕ es el funcional de Lyapunov para (3.1). Necesitamos observar que si u_0 verifica que $\phi(u_0) < 0$ entonces u tiene un tiempo de explosion finito. El análogo discreto de ϕ es

$$\phi_h(U) \equiv \frac{1}{2} \langle A^{\frac{1}{2}} U; A^{\frac{1}{2}} U \rangle + \frac{\lambda}{p+1} \langle MU^p; U \rangle - \frac{U_N^{q+1}}{q+1}$$

Aquí ϕ_h es el funcional de Lyapunov de (3.2). Ahora sea $W = (w_1, \dots, w_N)$ una solución estacionaria de de (3.2). Entonces tenemos

$$0 = -AW - \lambda MW^p + w_N^q \cdot e_N \tag{3.4}$$

Multiplicando (3.4) por W y como $1 < p \leq q$, tenemos:

$$0 = -\frac{1}{2} \langle A^{\frac{1}{2}} U; A^{\frac{1}{2}} U \rangle - \lambda \sum_{i=1}^N m_{ii} \frac{w_i^{p+1}}{2} + \frac{w_N^{q+1}}{2}$$

$$\geq -\frac{1}{2}\langle A^{\frac{1}{2}}U; A^{\frac{1}{2}}U \rangle - \lambda \sum_{i=1}^N m_{ii} \frac{w_i^{p+1}}{p+1} + \frac{w_N^{q+1}}{q+1} = -\phi_h(W)$$

Así toda solución positiva estacionaria de (3.2) tiene energía positiva (esto significa que $-\phi_h(W) \geq 0$) y entonces si U_0 satisface $\phi_h(U_0) < 0$ cuando toda solución global debe converger a una solución estacionaria, entonces esto explota. Ahora es fácil ver que $\phi_h(u_0^I) \rightarrow \phi(u_0)$ y así concluimos que si $\phi(u_0) < 0$ entonces u y u_h explotan para todo valor pequeño de h .

2. Caso $p > q$. En este caso, tenemos que toda solución esta definida globalmente. Supondremos que no es así, entonces

$$\lim_{t \nearrow T} \max_{k=1,2,\dots,N} |u_k(t)| = \infty$$

Entonces existe $t_0 < T$ y $1 \leq j \leq N$ tal que

$$\max_{k=1,2,\dots,N} \sup_{t \in [0, t_1]} u_k(t) = u_j(t_0) > M$$

Entonces $u'_j(t_0) \geq 0$. Ahora asumamos que $j = N$, porque si no es así, desde las ecuaciones (3.3) y $u'_j(t_0) \geq 0$, deducimos que

$$u_{j+1} = u_{j-1} = u_j$$

así $U(t_0)$ es constante. Pero con $j = N$ (como $p \geq q$)

$$u'_N(t_0) = \frac{2(u_{N-1}(t_0) - u_N(t_0))}{h^2} - \lambda u_N^q(t_0) + \frac{2}{h} u_N^q(t_0) < 0$$

□

Soluciones espurias constantes

Teorema 3.4. *Asumamos que $q < p < 2q - 1$, entonces existe una solución espuria constante $W = (w_1, \dots, w_N)$ de (3.3) la cual es atractiva y verifica que*

$$|w_N| \sim \frac{1}{h^{\frac{1}{p-q}}}$$

Así que si el dato inicial es lo suficientemente grande, la solución es global y converge (cuando $t \rightarrow \infty$) hacia una solución estacionaria y espuria para la cual w_N es de orden $\frac{1}{h^{\frac{1}{p-q}}}$

Demostración. Asumiremos ahora que $q < p < 2q - 1$. Necesitamos trabajar sobre las soluciones estacionarias de (3.3), es decir, sobre

$$\begin{cases} 0 = \frac{2}{h^2}(w_2 - w_1) - \lambda w_1^p \\ 0 = \frac{1}{h^2}(w_{k+1} - 2w_k + w_{k-1}) - \lambda w_k^p, & 2 \leq k \leq N-1 \\ 0 = \frac{2}{h^2}(w_{N-1} - w_N) - \lambda w_N^p + \frac{2}{h} u_N^q \end{cases} \quad (3.5)$$

De esta ecuación podemos obtener w_2 como una función creciente de w_1

$$w_2 = w_1 + \lambda h^2 w_1^p \equiv F_2(w_1)$$

y de la segunda ecuación podemos obtener w_3 como una función de w_1 y w_2 , y usando la ecuación construida como una función de w_1

$$w_3 = 2w_2 - w_1 + \lambda h^2 w_2^p = 2F_2(w_1) - w_1 + \lambda h^2 (F_2(w_1))^p \equiv F_3(w_1)$$

Observemos que w_3 es decreciente como función de w_1 y veamos también que las diferencias $w_2 - w_1$ y $w_3 - w_2$ son funciones decrecientes de w_1 . Podemos continuar con este procedimiento y obtener una secuencia decreciente de $w_1 < w_2 < \dots < w_k = F_k(w_1) < \dots < w_N$ (donde cada F_k es una función decreciente de w_1) la cual satisface

$$\begin{aligned} w_k &= 2w_{k-1} - w_{k-2} + \lambda h^2 w_{k-1}^p \\ w_k &= 2F_{k-1}(w_1) - F_{k-2}(w_1) + \lambda h^2 (F_{k-1}(w_1))^p \equiv F_k(w_1) \end{aligned}$$

Luego, hemos obtenido que toda solución de (3.5) es creciente. Ahora, si $W = (w_1, w_2, \dots, w_N)$ es una solución de (3.5), entonces las ultimas dos coordenadas $w_{N-1} = F_{N-1}(w_1)$ y $w_N = F_N(w_1)$ tienen que satisfacer la ultima condición

$$0 = 2 \left[\frac{F_{N-1} - F_N}{h^2} \right] - \lambda (F_N)^p + \frac{2F_N^q}{h} \equiv G(w_1) \quad (3.6)$$

De esta manera toda solución positiva de (3.5) da una solución de (3.6). De manera inversa, si tenemos una función positiva w_1 tal que $G(w_1) = 0$ entonces obtenemos así una solución de (3.5) tomando $w_k = F_k(w_1)$. Observemos que como $F_N(w_1)$ es una función continua creciente de w_1 y de rango $[0, +\infty)$, entonces existe x , con

$$a < \left[\frac{2}{\lambda} \right]^{\frac{1}{p-q}}$$

tal que

$$F_N(x) = \frac{a}{h^{\frac{1}{p-q}}}$$

Para que el valor de x

$$G(x) \geq -\frac{2a}{h^{2+\frac{1}{p-q}}} + \frac{2a^q}{h^{1+\frac{1}{p-q}}} - \lambda \frac{a^p}{h^{\frac{p}{p-q}}} \geq 0$$

si h es lo suficientemente pequeño (aquí usamos el hecho de que $p < 2q - 1$). Además existe y tal que

$$F_N(y) = \frac{b}{h^{\frac{1}{p-q}}}$$

y si

$$b \geq \left[\frac{2}{\lambda} \right]^{\frac{1}{p-q}}$$

entonces

$$G(y) \leq \frac{2b^q}{h^{1+\frac{q}{p-q}}} - \lambda \frac{b^p}{h^{\frac{p}{p-q}}} \leq 0$$

Como $G(w_1)$ es continua y satisface que $G(x) \geq 0$ y $G(y) \leq 0$, obtenemos que existe una solución de $G(w_1) = 0$ (solución de (3.5)) que satisface

$$F_N(w_1) = \frac{c}{h^{\frac{1}{p-q}}}$$

con

$$\frac{2}{\lambda} \frac{1}{p-q} - \epsilon < c < \frac{2}{\lambda} \frac{1}{p-q}$$

siempre que h sea lo suficientemente pequeño $h = h(\epsilon)$. Ahora necesitamos mostrar que esta solución espuria constante W es atractiva. Para este propósito, únicamente tenemos que observar que la linealización de (3.2) en W tiene todos los eigenvalores con parte real negativa. La linealización tiene la forma

$$Z' = [-M^{-1}A + B] Z$$

donde A es la matriz de rigidez (y semidefinida positiva), M es la matriz de masa (la cual tiene la diagonal con entradas positivas) y B es la matriz diagonal tiene los siguientes coeficientes:

$$b_{ii} = -p\lambda w_i^{p-1}, 1 \leq i \leq N-1$$

$$b_{NN} = -p\lambda w_N^{p-1} + \frac{2qw_N^{q-1}}{h}, i = N$$

Ahora, si tomamos ϵ tal que

$$\frac{2}{\lambda} \frac{1}{p-q} - \epsilon > \frac{2q}{\lambda p} \frac{1}{p-q}$$

obtenemos

$$w_N > \frac{2q}{p\lambda h} \frac{1}{p-q}$$

entonces la matriz B es definida negativa y así $-M^{-1}A + B$ es definida negativa. Esto prueba finalmente este teorema. \square

Bibliografía

- [1] Kendall E. Atkinson. *An introduction to numerical analysis*. John Wiley & Sons Inc., New York, second edition, 1989.
- [2] Yehuda P. , Jacob Rubinstein. *An Introduction to Partial Differential Equations*. Cambridge University Press, 2005.
- [3] E. Süli , D. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, 2003.
- [4] J. Fernández B, J. D. Rossi . *Blow-up vs Spurious Steady Solutions* . American Mathematical Society , volume 129, number 1 , pages 139-144 , 2000.
- [5] Ezequiel Dratman . *Efficient approximation of the solution of certain nonlinear reaction-diffusion equation I.the case of small absorption*. Instituto de ciencias Universidad nacional de general sarmiento,Buenos Aires, Argentina , 2011.
- [6] James S. Walker. *Fourier Series*. Department of Mathematics , University of Wisconsin - Eau Claire Press.
- [7] Richard L. Burden. *Numerical Analysis*. Thomson - Learning, séptima edición.
- [8] R. Winther , A. Tveito. *Introduction to Partial Differential Equations: a Computational Approach*. Springer-Verlag, New York, 1998.
- [9] Robert G. Bartle. *Introducción al análisis matemático*. Editorial Limusa, México, 1982. Translated from the English by Ma. Cristina Gutiérrez González.
- [10] Häim Brézis. *análisis funcional*. Alianza Editorial S.A., 1984.
- [11] Richard Haberman. *Elementary applied partial differential equations*. Prentice Hall Inc., Englewood Cliffs, NJ, second edition, 1987. With Fourier series and boundary value problems.
- [12] O. Axelson , V. A. Barker. *Finite element solution of boundary value problems*. Society for industrial and applied mathematics, Filadelfia , 2001.
- [13] Hans Weinberger. *Ecuaciones diferenciales en derivadas parciales con métodos de variable compleja y de transformaciones integrales*. Editorial Reverté, España, 1992.

- [14] A. Kolmogorov. *Elementos de la teoría de funciones y del análisis funcional* . Editorial MIR, Moscú, 1960.
- [15] J. Dideudonné. *Foundation of modern analysis* . Academic Press, New York, 1960.
- [16] L. Elsgoltz. *Ecuaciones diferenciales y cálculo variacional*. Editorial MIR, Moscú, 1977.
- [17] E. kreyszig. *Introductory functional analysis with applications* . Editorial Wiley, New York, 1978.
- [18] Robert A. Adams . *Sobolev Spaces* . Academic Press, New York, 1975.
- [19] Julio C. Carrillo. *the finite element method for parabolic problems*. Notes of Advanced numerical analysis, University of Louisiana at Lafayette, 2002.
- [20] G. Strang. *Algebra lineal y sus aplicaciones*. Mexico: Addison Wesley, Iberoamérica, 1986.