



UNIVERSIDAD MICHOACANA DE

SAN NICOLÁS DE HIDALGO

INSTITUTO DE FÍSICA Y MATEMÁTICAS

**“ELEMENTOS FINITOS PARA PROBLEMAS ELÍPTICOS E HIPERBÓLICOS
EN UNA DIMENSIÓN ESPACIAL”**

TESIS

PARA OBTENER EL TÍTULO DE
MAESTRO EN CIENCIAS EN EL ÁREA DE FÍSICA

PRESENTA

JUAN PABLO CRUZ PÉREZ

ASESOR

DR. OLIVIER CHARLES-ALBERT SARBACH

Diciembre 2009

Índice general

1. Introducción	3
2. Problema Elíptico	7
2.1. Formulación débil del problema elíptico en una dimensión	8
2.1.1. Caso Particular con $\gamma(x) = 1$	12
2.1.2. Caso de $\gamma(x) > 0$	13
2.1.3. Energía potencial total del sistema	14
2.2. El Método de Galerkin	17
2.2.1. Discretización en el método de Galerkin.	17
2.2.2. Elementos Finitos	19
2.2.3. Aproximación por la interpolante.	20
2.2.4. Convergencia del Método.	26
2.3. Implementación del Método de Galerkin	29
2.3.1. Cálculo de la matriz A	30
2.3.2. Aproximación de L_F	33
2.3.3. Solución al Sistema $Av = J$	35
2.3.4. Condición del sistema	38
2.3.5. Número condicionante	38
2.3.6. Numero condicionante del caso elíptico	39
2.3.7. Ejemplo de la Implementación Numérica	42
3. Problema Hiperbólico	49
3.1. Formulación débil del problema.	49
3.2. Estimación de Energía Apriori	50
3.3. El problema semidiscreto	52
3.4. Método Implícito para la discretización Temporal.	56

3.4.1. Consistencia	58
3.4.2. Estabilidad y Convergencia	60
3.4.3. Solución al método implícito	65
4. Conclusiones	67
A. Apéndices	71
A.0.4. El Espacio de Hilbert H^1	71
A.0.5. Demostración del Lema (2.1.1)	73
A.0.6. Demostración al Lema (2.3.1)	76
Bibliografía	80

Capítulo 1

Introducción

Muchos problemas en la naturaleza son modelados matemáticamente mediante ecuaciones diferenciales, por ejemplo en Mecánica Clásica la segunda ley de Newton da lugar a resolver una ecuación diferencial, o bien en la Mecánica Cuántica, donde tenemos la ecuación de Schrödinger. Para éstos y otros ejemplos, existen soluciones que se obtienen de forma analítica, bajo ciertas suposiciones.

Sin embargo, es posible encontrarse con modelos matemáticos cuyas ecuaciones diferenciales no pueden resolverse de forma analítica. Para estos problemas, es necesario buscar métodos alternativos de solución. Es aquí donde los métodos numéricos toman una gran relevancia para aproximar una solución de estos problemas.

Consideremos, por ejemplo, un sistema de un agujero negro supermasivo que tiene una partícula, tal como una estrella, orbitando alrededor de él. Si las dimensiones de ambos son tales que el cociente entre la masa de la partícula y el objeto supermasivo es pequeño, podemos considerar en primera aproximación que la estrella sigue la trayectoria de una geodésica para una partícula libre alrededor del agujero negro. Ya que el campo gravitacional de la partícula sería mucho menor con respecto al campo generado por el objeto supermasivo.

Entonces en sistemas binario, *agujero negro - estrella*, la presencia de un objeto que se encuentra orbitando alrededor del agujero negro modifica el espacio tiempo en el que se encuentra inmerso. Esto da lugar a la generación de *Ondas Gravitacionales*, por lo que el sistema pierde energía por la radiación de las mismas, modificando así la trayectoria de las partículas provocando que eventualmente la partícula disminuya la distancia que existe entre ella y el objeto supermasivo, produciendo así que sea absorbida por él.

Estas correcciones de la métrica pueden ser calculadas en primera aproximación mediante la teoría de perturbaciones de agujeros negros. Existen diversos trabajos sobre las técnicas empleadas para realizar estas aproximaciones, tal es el caso de formalismo adoptado para los agujeros negros que no tiene rotación, en el trabajo debido a *Regge-Wheeler-Zerilli* [1, 2]. Mientras que del caso de

los agujeros negros rotando, descritos por la métrica de Kerr, el formalismo es debido a los trabajo de *Teulosky* [3]. Para calcular la corrección de la trayectoria debida a la pérdida de energía por la emisión de ondas gravitacionales en un sistema binario, se debe al trabajo de *Gralla-Wald* [4].

El hecho de considerar a una partícula orbitando alrededor del objeto supermasivo, da lugar a un tensor de Energía Momento distributivo de tipo delta de Dirac. Esto lleva a una ecuación tipo onda con una fuente distributiva.

Una de las formas de estudiar este tipo de ecuaciones diferenciales, es cambiando el problema de *Ecuaciones Diferenciales Parciales (EDP)* original con un problema en la *formulación débil*. Esto se hace tomando una función de prueba y se multiplica por la ecuación diferencial del problema, después integramos por partes y bajamos el orden de las derivadas involucradas en el problema original, pero bajo una integración.

Al mismo tiempo que esta nueva formulación débil del problema original nos abre las puertas para implementar el Método de Galerkin para resolver este problema, puesto que esta formulación se adapta de forma exacta para buscar una solución en un conjunto con diferentes características al conjunto de funciones infinitamente diferenciables. Este método encuentra una la solución en un espacio de Hilbert X asegurando su unicidad. Después aproxima dicha solución en un subespacio de X de dimensión finita, digamos $X_n \subset X$. Aquí X_n puede ser un espacio de polinomios (esto lleva a los métodos espectrales) o un espacio de funciones que solamente son distintas de cero en una región pequeña (lo que lleva a los elementos finitos). El método de los elementos finitos ha sido implementado en el problema de una partícula orbitando alrededor de un agujero negro supermasivo para calcular la radiación emitida en forma de ondas gravitacionales [5].

En este trabajo estudiamos dos tipos de problemas en una dimensión espacial: *elíptico* e *hiperbólico*. Llevamos a su formulación débil estos problemas, en donde podremos encontrar una solución en un espacio de Hilbert X , y aproximamos a la solución expandiéndola en términos de un conjunto de funciones lineales a trozos y distintas de cero solamente en un subconjunto pequeño del intervalo de definición, de donde el método toma el nombre de *Elementos Finitos*.

Es tarea importante comprobar si la solución del método que se obtiene al usar elementos finitos converge a la solución exacta del problema débil y a que orden lo hace. Entonces en este trabajo buscamos estimar y controlar el error de la solución del problema débil que se encuentra en X con respecto de la solución aproximada, la cual se encuentra en $X_n \subset X$. Elemento importante en el proceso de la estimación del error, es la introducción de la interpolante de la solución en X para el problema débil. La interpolante es pieza clave y fundamental en la demostración de la convergencia y estimación de los errores en nuestro método numérico.

En el **Capítulo 2** de este trabajo de tesis nos ocupamos del problema elíptico en una dimensión. Una vez definido de forma precisa junto con condiciones de frontera tipo *Dirichlet*, procedemos a formular el problema débil correspondiente

al problema elíptico, en donde la solución para esta formulación se encuentra en un espacio de Hilbert X . Vemos como el producto escalar queda definido mediante los datos que componen el problema elíptico en su forma primitiva, también la fuente del problema original nos ayuda a construir un funcional lineal que posteriormente junto con el *lema de Riesz* [6] nos ayudará a mostrar la existencia y unicidad de la solución en el espacio de Hilbert. Mostramos también que la solución depende de forma continua de la fuente definida en el problema elíptico. En este capítulo mostramos que la formulación del problema débil es equivalente a un problema variacional, donde variamos un funcional definido sobre los elementos de X llamado *Energía del Sistema* [7].

Una vez encontrada la solución en el espacio de Hilbert, procedemos a emplear el Método de Galerkin para buscar una aproximación a la solución en un subespacio de dimensión finita, y que como ya hemos mencionado tenemos un subespacio de funciones lineales a trozos. Calculamos y estimamos los errores de la solución exacta con respecto a la interpolante de la solución, la cual nos ayuda a acotar los errores de la solución exacta que se encuentra sobre X y la solución aproximada que se encuentra sobre el subespacio de dimensión finita X_n .

Al momento de realizar la expansión de la aproximación de la solución en un subespacio de dimensión finita X_n , vemos que para encontrar la solución aproximada $x \in X_n$, el problema se reduce a resolver un sistema de ecuaciones lineales de la forma $Ax = J$, donde A es una matriz tridiagonal de dimensión $n \times n$, siendo n la dimensión de X_n . Cada uno de los elementos de esta matriz tiene relación directa con los datos del problema elíptico, así como J es un vector de dimensión n . Puesto que la matriz A es tridiagonal, se puede invertir de manera eficiente [8, 7].

Una vez calculado $x \in X_n$, un vector de dimensión n , mostramos un ejemplo de la implementación del Método de Elementos Finitos y algunos resultados que tratan de demostrar que las cotas para el error así como el orden de convergencia del método son los correctos.

En el **Capítulo 3**, tomamos un problema hiperbólico en una dimensión espacial, con condiciones iniciales y condiciones de Dirichlet en la frontera. Este problema se formula de forma débil siguiendo la misma idea que para el caso elíptico. La diferencia aquí, es que tenemos dos variables en donde se puede integrar, la temporal t y la espacial x .

Tomamos entonces una función de prueba que depende de x compatible con las condiciones a la frontera, la multiplicamos con la ecuación e integramos sobre la dimensión espacial en donde está definido el problema.

Esto nos lleva a una formulación débil del problema, donde la solución $u(t)$ depende de la variable temporal y por lo cual $u(t) \in X$, tal que $u : [0, \infty) \mapsto X$, donde nuevamente X es un espacio de Hilbert. Definimos la *Energía Total del Sistema*, mediante la cual podemos estimar la solución con respecto del dato del problema hiperbólico. En particular, esta estimación nos permite demostrar la unicidad y la dependencia continua de la solución respecto del dato.

Procedemos mediante el Método de Galerkin, a discretizar el espacio de Hilbert X , en un subespacio de dimensión finita X_n donde buscaremos la aproximación para la solución al problema débil. Sobre esta discretización e inspirados en la Energía Total de Sistema, podemos controlar el error de la solución aproximada encontrada en el subespacio de dimensión finita $X_n \subset X$.

Escogiendo a X_n como un conjunto de funciones lineales a trozos y distintas de cero en subconjunto pequeño del intervalo de definición, podemos expandir $u_n(t) \in X_n$, en términos de la base para X_n . En este caso, la dependencia temporal es asignada a los coeficientes de la expansión.

Así implementando esta expansión en la formulación débil del problema, se reduce a un sistema del tipo $M\ddot{x} + Ax = J$, donde M y A son matrices tridiagonales y positivas definidas. La matriz M es invertible y el problema se reduce a un sistema de ecuaciones diferenciales ordinarias lineales, con su respectiva condición inicial. Este sistema se llama el problema semidiscreto.

Para resolver este problema, usamos un método implícito descrito en [7]. Demostramos primero la consistencia con el problema semidiscreto. A continuación, usando una versión discreta de la Energía del Sistema, mostramos la estabilidad y convergencia del método implícito. Por último encontramos que para implementar este método para la integración en el tiempo y así encontrar una aproximación a la solución del problema débil tenemos dos sistemas del tipo $Bx = b$, donde B es una matriz tridiagonal de dimensión $n \times n$, los cuales funcionan de forma acoplada.

Finalmente, en las **Conclusiones** resumimos los resultados más importantes que se pueden apreciar en este trabajo de tesis, tales como la comprobación de los resultados teóricos de convergencia con los resultados numéricos para el caso elíptico, y los resultados teóricos sobre la convergencia del método semidiscreto y del método implícito para el caso del problema hiperbólico.

En el apéndice mostramos algunos resultados técnicos los cuales eran demasiado extensos para presentarse en el contenido de trabajo de tesis, pero cuyos resultados forman parte del desarrollo del trabajo mismo.

Capítulo 2

Problema Elíptico

En este capítulo, consideramos un problema elíptico lineal. Primero, se reescribe el problema en su forma débil abstracto, el cual consiste de tener un espacio de Hilbert X , junto con su producto escalar (\cdot, \cdot) definido sobre sus elementos, y un funcional lineal f acotado sobre él. De tal forma que se pueda plantear con ayuda del *lema de Riesz* la formulación abstracta del problema débil. Es decir, encontrar una $u \in X$ tal que $(u, v) = f(v)$ para toda $v \in X$.

El método de Galerkin consiste básicamente en aproximar tal solución $u \in X$, mediante una función que se encuentra en un subespacio de dimension finita, $X_n \subset X$. Consideramos que la formulación débil del problema también es válida sobre cualquier subespacio cerrado de X .

La elección del subespacio de dimensión finita X_n , hace que el método de Galerkin adopte diferentes nombres, por ejemplo los *Métodos Espectrales* y *Métodos de Elementos Finitos*. La principal diferencia entre estos métodos radica en que el primero aproxima la solución mediante un conjunto de funciones que son analíticas (polinomios) en todo el intervalo de definición; mientras que el segundo, aproxima la solución mediante un conjunto de funciones que son suaves por trozos y diferentes de cero solamente en un subconjunto pequeño del dominio de definición.

Por esta razón, los Métodos Espectrales son referidos como Métodos de aproximación global, mientras que a los Elementos Finitos los reconocemos como métodos de aproximación local. Nosotros estaremos trabajando con los Elementos Finitos y tomaremos como subconjunto de X funciones que son lineales por trozos.

Demostraremos de manera analítica como con experimentos numéricos que este método converge a la solución exacta del problema cuando aumenta la resolución.

2.1. Formulación débil del problema elíptico en una dimensión

Definimos en esta primera sección el problema elíptico, junto con sus condiciones de frontera. Una vez definido, lo llevamos hasta una formulación débil, este proceso nos da las herramientas para encontrar una solución sobre un espacio de Hilbert, donde el producto interno que está definido sobre él, necesita de algunos datos de la definición del problema original, así como un funcional lineal acotado.

Aplicando el *Lema de Riesz* sobre él, encontramos que la solución existe y es única. Definimos la Energía del Sistema y vemos que el problema en su formulación débil es equivalente a un problema variacional.

Planteamos primero el siguiente problema elíptico en una dimensión, sobre el intervalo $I \equiv [0, 1] \subset \mathbb{R}$.

Definición: 2.1.1 (Definición del problema elíptico.) Sea $I \subset \mathbb{R}$, tal que $I \equiv [0, 1]$. Y sean las funciones γ y F , definidas sobre I de tal forma que tales funciones son al menos clase $C^0(I)$,

$$\gamma, F : I \subset \mathbb{R} \rightarrow \mathbb{R}, \quad (2.1)$$

donde

$$\gamma(x) > 0, \forall x \in I. \quad (2.2)$$

Entonces tenemos que encontrar una u al menos clase $C^2(I)$, tal que satisface la ecuación diferencial

$$\left[-\frac{d^2}{dx^2} + \gamma(x) \right] u(x) = F(x), \quad (2.3)$$

junto con las condiciones a la frontera

$$u(0) = u(1) = 0. \quad (2.4)$$

Dada una $F \notin C^0(I)$, la solución u del problema elíptico no tiene por que ser continua sobre el intervalo de definición. Para tratar de encontrar una solución a este tipo de problemas definimos un conjunto de funciones de prueba.

Definición: 2.1.2 (Conjunto de funciones de prueba.) Sea C_p^∞ , conjunto de funciones ϕ de variable real sobre el intervalo $I \subset \mathbb{R}$ tal que

$$C_p^\infty := \{ \phi \in C^\infty[0, 1] | \phi(0) = \phi(1) = 0 \} \quad (2.5)$$

a este conjunto le llamamos conjunto de funciones de prueba.

2.1. FORMULACIÓN DÉBIL DEL PROBLEMA ELÍPTICO EN UNA DIMENSIÓN

Ahora hacemos el producto de cualquier función de prueba $\phi \in C_p^\infty$ con la ecuación diferencial (2.3) e integramos sobre el dominio de definición

$$-\int_0^1 \phi(x) \frac{d^2 u(x)}{dx^2} dx + \int_0^1 \phi(x) \gamma(x) u(x) dx = \int_0^1 \phi(x) F(x) dx. \quad (2.6)$$

Podemos observar que

$$\frac{d}{dx} \left(\phi(x) \frac{du(x)}{dx} \right) = \frac{d\phi(x)}{dx} \frac{du(x)}{dx} + \phi(x) \frac{d^2 u(x)}{dx^2} \quad (2.7)$$

de donde se obtiene fácilmente

$$-\int_0^1 \phi(x) \frac{d^2 u(x)}{dx^2} dx = \int_0^1 \frac{d\phi(x)}{dx} \frac{du(x)}{dx} dx - \int_0^1 \frac{d}{dx} \left(\phi(x) \frac{du(x)}{dx} \right) dx. \quad (2.8)$$

Ahora veamos el segundo término de la derecha de la ecuación (2.8),

$$\int_0^1 \frac{d(\phi(x) du(x)/dx)}{dx} dx = \phi(x) \frac{du(x)}{dx} \Big|_0^1 = 0, \quad (2.9)$$

ya que ϕ se anula en la frontera debido a que $\phi \in C_p^\infty$.

Por esta razón (2.8) se reduce a la expresión

$$-\int_0^1 \phi(x) \frac{d^2 u(x)}{dx^2} dx = \int_0^1 \frac{d\phi(x)}{dx} \frac{du(x)}{dx} dx, \quad (2.10)$$

y por tanto (2.6) se puede escribir de la forma

$$\boxed{\int_0^1 \frac{d\phi(x)}{dx} \frac{du(x)}{dx} dx + \int_0^1 \phi(x) \gamma(x) u(x) dx = \int_0^1 \phi(x) F(x) dx.} \quad (2.11)$$

La ecuación (2.11) nos da oportunidad de definir la formulación débil del problema elíptico de manera formal de la siguiente manera:

Definición: 2.1.3 (Formulación Débil del Problema Elíptico.) *Sea el problema elíptico dado en la definición (2.1.1), entonces la formulación débil del problema consiste en encontrar una función u definida sobre $I \subset \mathbb{R}$, tal que satisface la ecuación (2.11), $\forall \phi \in C_p^\infty$.*

Podemos observar que la formulación débil, requiere que la derivada de u y ϕ sean de cuadrado integrables, esto motiva a la siguiente definición:

Definición: 2.1.4 (Derivada débil.) Existe la derivada de $u \in L^2(I)$ en el sentido débil, si podemos encontrar $v \in L^2(I)$, tal que

$$\int_0^1 v(x)\phi(x)dx = - \int_0^1 u(x)\phi'(x)dx, \forall \phi \in C_o^\infty[0, 1]. \quad (2.12)$$

Entonces decimos que $u' = v$, en el sentido débil.

Ahora podemos definir el conjunto de funciones donde buscaremos las soluciones como

Definición: 2.1.5 (H^1) Sea H^1 el conjunto de funciones que son de cuadrado integrables sobre el intervalo $I \subset \mathbb{R}$, y que su derivada u' existe en el sentido débil, es decir,

$$H^1 := \{u \in L^2(I) | \text{Existe } u' \in L^2(I), \text{ en el sentido débil.}\}. \quad (2.13)$$

El espacio H^1 es más grande que el espacio $C^1(I)$ de funciones continuamente diferenciables. Por ejemplo, contiene las funciones que son continuas y lineales por trozos.

Si definimos el siguiente producto escalar sobre H^1 ,

$$(u, v)_{H^1} := (u, v)_{L^2} + (u', v')_{L^2}, \quad u, v \in H^1, \quad (2.14)$$

donde $(u, v)_{L^2} := \int_0^1 uv dx < \infty$ para $u, v \in L^2(I)$, entonces H^1 es un espacio de Hilbert, ver el apéndice.

Para buscar soluciones sobre un conjunto que sea compatible con las condiciones de frontera, usaremos el siguiente lema.

Lema: 2.1.1 Sea $u \in H^1$, entonces u posee un representante continuo en $u \in C^o(I)$, y

$$\|u\|_\infty \leq \sqrt{2}\|u\|_{H^1}, \forall u \in H^1. \quad (2.15)$$

Demostración. Para ver una demostración de este lema véase el apéndice (A.0.5).

Ahora bien, podemos implementar las condiciones de frontera (2.4), las cuales piden que la solución se anule en $x = 0$ y $x = 1$.

Definición: 2.1.6 (H_o^1) Sea el conjunto de funciones de variable real definido de la siguiente forma:

$$H_o^1 := \{u \in H^1 | u(0) = u(1) = 0\}. \quad (2.16)$$

y que esta provisto del mismo producto escalar como en H^1 .

2.1. FORMULACIÓN DÉBIL DEL PROBLEMA ELÍPTICO EN UNA DIMENSIÓN 11

Como consecuencia del Lema 2.1.1, H_o^1 es un subespacio cerrado de H^1 , y por lo tanto es completo.

Así que si buscamos las soluciones en este conjunto H_o^1 , podemos definir una operación de producto interno usando $(\cdot, \cdot)_{H^1}$.

Regresando a nuestro problema elíptico, en su formulación débil (2.11), podemos definir de forma natural un producto interno similar a $(\cdot, \cdot)_{H^1}$.

Definición: 2.1.7 (Producto Interno $Q_\gamma(u, v)$) Sean $u, v \in H_o^1$, entonces se define $Q_\gamma(\cdot, \cdot)$ tal que

$$Q_\gamma : H_o^1 \times H_o^1 \longrightarrow \mathbb{R} : \\ (u, v) \longmapsto Q_\gamma(u, v) := \int_0^1 \phi'(x)v'(x)dx + \int_0^1 \gamma(x)\phi(x)v(x)dx. \quad (2.17)$$

Podemos ver que $Q_\gamma(u, v)$ es lineal en la primer entrada, si reemplazamos u por $u = u_1 + \alpha u_2$, con $\alpha \in \mathbb{R}$, $u_1, u_2 \in H_o^1$ vemos que

$$Q_\gamma(u, v) = Q_\gamma(u_1 + \alpha u_2, v) = Q_\gamma(u_1, v) + \alpha Q_\gamma(u_2, v). \quad (2.18)$$

De forma análoga es lineal en la segunda entrada, reemplazando v por $v = v_1 + \beta v_2$ tal que $\beta \in \mathbb{R}$, $u_1, u_2 \in H_o^1$ obtenemos que

$$Q_\gamma(u, v) = Q_\gamma(u, v_1 + \beta v_2) = Q_\gamma(u, v_1) + \beta Q_\gamma(u, v_2). \quad (2.19)$$

Por lo tanto $Q_\gamma(\cdot, \cdot)$, es *bilineal*. Es fácil observar que tiene la propiedad de la *simetría*, $Q_\gamma(u, v) = Q_\gamma(v, u)$, $\forall u, v \in H_o^1$.

También podemos definir el siguiente funcional lineal.

Definición: 2.1.8 ($L_F(v)$) Sean $F \in L^2$ y $\phi \in H_o^1$, entonces podemos definir L_F tal que

$$L_F : H_o^1 \rightarrow \mathbb{R} : \phi \mapsto L_F(\phi) := \int_0^1 F(x)\phi(x)dx. \quad (2.20)$$

Podemos ver que L_F es lineal, $L_F(v_1 + \beta v_2) = L_F(v_1) + \beta L_F(v_2)$ para todo $v_1, v_2 \in H_o^1$ y todo $\beta \in \mathbb{R}$.

Definición: 2.1.9 (Formulación Débil del Problema Elíptico.) Decimos que $u \in H_o^1$ es solución del problema elíptico en su formulación débil si tenemos que

$$Q_\gamma(u, v) = L_F(v), \forall v \in H_o^1. \quad (2.21)$$

Observación: 2.1.1 Podemos notar que si $u \in H_o^2$, es decir, en un conjunto de funciones de variable real donde $u''(x)$ se pueda formular en su forma débil, entonces $u(x)$ también sería solución para la ecuación diferencial (2.3) con las condiciones de frontera dadas por (2.4).

2.1.1. Caso Particular con $\gamma(x) = 1$.

Consideramos primero el caso en el que γ es una constante y en particular

$$\gamma(x) = 1, \forall x \in [0, 1]. \quad (2.22)$$

Con esto tenemos de forma inmediata que el producto interno

$$Q_\gamma(\phi, u) = \int_0^1 u' \phi' dx + \int_0^1 \gamma u \phi dx = \int_0^1 u' \phi' dx + \int_0^1 u \phi dx = (\phi, u)_{H^1},$$

y entonces de forma inmediata se tiene que los dos productos internos Q_γ y $(\cdot, \cdot)_{H^1}$ coinciden en su definición.

Tomemos, por otra parte, el funcional lineal L_F , y busquemos una cota para el

$$|L_F(\phi)| = \left| \int_0^1 F(x) \phi(x) dx \right| \leq \left(\int_0^1 F(x)^2 dx \right)^{1/2} \left(\int_0^1 \phi(x)^2 dx \right)^{1/2} \quad (2.23)$$

$$= \|F\|_{L^2} \|\phi\|_{L^2}. \quad (2.24)$$

Usando que

$$\|\phi\|_{L^2} \leq \|\phi\|_{H^1}, \quad (2.25)$$

lo cual se puede ver directamente de la definición de $\|\cdot\|_{H^1}$

$$\|\phi\|_{H^1}^2 = \|\phi\|_{L^2}^2 + \|\phi'\|_{L^2}^2 \geq \|\phi\|_{L^2}^2.$$

Y usando este resultado obtenemos que

$$|L_F(\phi)| \leq \|F\|_{L^2} \|\phi\|_{H^1}. \quad (2.26)$$

Por otra parte, exigiendo que la F sea tal que $\|F\|_{L^2} < C$, tenemos que L_F es acotado para toda $\phi \in H_o^1$, y como $L_F(\cdot)$ es lineal tenemos que

$$L_F \in (H_o^1)^*. \quad (2.27)$$

Donde $(H_o^1)^*$ es el *dual* de H_o^1 .

Para mostrar la existencia y unicidad de la solución, usaremos el Lema de Riesz, que formulamos a continuación:

2.1. FORMULACIÓN DÉBIL DEL PROBLEMA ELÍPTICO EN UNA DIMENSIÓN 13

Teorema: 2.1.1 (Lema de Riesz) Sea X un espacio de Hilbert con producto escalar definido por (\cdot, \cdot) . Para cada $T \in X^*$ existe una única $y_T \in X$ tal que $T(x) = (y_T, x)$ para toda $x \in X$. Por lo tanto $\|T\|_{X^*} = \|y_T\|_X$.

Para ver la demostración a este lema véase [6].

Tenemos todas las herramientas necesarias para aplicar el lema de Riesz a nuestro problema, ya que trabajamos sobre un espacio de Hilbert H_o^1 , donde tenemos el producto interno definido por $Q_\gamma(\cdot, \cdot)$, además del funcional lineal $L_F(\cdot)$.

Teorema: 2.1.2 (Aplicación del Lema de Riesz) Para cada $L_F \in (H_o^1)^*$ existe una única $u \in H_o^1$ tal que $L_F(\phi) = Q_\gamma(u, \phi)$ para toda $\phi \in H_o^1$. Además tenemos que $\|u\|_{H_o^1} = \|L_F\|_{(H_o^1)^*}$.

Entonces el lema de Riesz, nos asegura la *existencia y unicidad* para el problema elíptico en una dimensión en la formulación débil.

2.1.2. Caso de $\gamma(x) > 0$.

Para tratar el caso de $\gamma(x) \neq 1$ tomemos el producto $Q_\gamma(\cdot, \cdot)$ y procedemos de la forma siguiente. Consideremos que γ es una función que es acotada en el intervalo $[0, 1]$, i.e., que existen $0 < m < M$ tales que $m < \gamma(x) < M$ para todo $x \in [0, 1]$.

Tomando $Q_\gamma(u, u)$, $u \in H_o^1$ tenemos

$$Q_\gamma(u, u) = \int_0^1 u'(x)^2 dx + \int_0^1 \gamma(x) u(x)^2 dx \leq \int_0^1 u'(x)^2 dx + M \int_0^1 u(x)^2 dx. \quad (2.28)$$

Podemos ver que

$$Q_\gamma(u, u) \leq \max\{1, M\} \|u\|_{H^1}^2, \quad \forall u \in H_o^1. \quad (2.29)$$

De forma similiar se obtiene que

$$Q_\gamma(u, u) \geq \min\{m, 1\} \|u\|_{H^1}^2, \quad \forall u \in H_o^1, \quad (2.30)$$

de donde se sigue inmediatamente que existen $\bar{M} = \max\{1, M\}$, $\bar{m} = \min\{m, 1\}$, tal que

$$\bar{m} \|u\|_{H^1}^2 \leq Q_\gamma(u, u) \leq \bar{M} \|u\|_{H^1}^2, \quad \forall u \in H_o^1. \quad (2.31)$$

Entonces Q_γ define un producto escalar sobre H_o^1 con norma inducida equivalente a la norma usual en H_o^1 . Por el lema de Riesz existe un único $u \in H_o^1$ tal que $Q_\gamma(u, \phi) = L_F(\phi)$ para todo $\phi \in H_o^1$, entonces tenemos en particular

$$\bar{m}\|u\|_{H^1}^2 \leq Q_\gamma(u, u) = L_F(u) \leq \|F\|_{L^2}\|u\|_{H^1}, \quad (2.32)$$

de donde obtenemos que

$$\|u\|_{H^1} \leq \frac{1}{\bar{m}}\|F\|_{L^2}. \quad (2.33)$$

Entonces, tenemos existencia, unicidad y dependencia continua de la solución u con respecto al dato F : Sean u_1 y $u_2 \in H_o^1$ soluciones con dato F_1 y $F_2 \in L^2$, respectivamente, entonces la linealidad del problema implica que $u_2 - u_1$ es una solución débil con dato $F_2 - F_1$. Entonces, la estimación (2.33) implica

$$\|u_2 - u_1\|_{H^1} \leq \frac{1}{\bar{m}}\|F_2 - F_1\|_{L^2}, \quad (2.34)$$

lo que demuestra que u_1 converge a u_2 en H_o^1 si F_1 converge a F_2 en L^2 . En particular, si $F_1 = F_2$, obtenemos $u_1 = u_2$ lo que implica la unicidad de las soluciones.

2.1.3. Energía potencial total del sistema

Sea X un espacio de Hilbert, con producto interno definido por Q_γ . Consideremos el siguiente problema de minimización:

- Encontrar $u \in X$, tal que

$$\boxed{F(u) = \min_{v \in X} F(v)}, \quad (2.35)$$

donde $F(v) = \frac{1}{2}Q_\gamma(v, v) - L(v)$.

Y consideremos ahora el problema débil de la forma:

- Encontrar $u \in X$ tal que

$$\boxed{Q_\gamma(u, v) = L_F(v), \forall v \in X.} \quad (2.36)$$

Entonces podemos mostrar que estos dos problemas son equivalentes, es decir, que si $u \in X$ es solución para el problema dado por (2.35) también lo es para (2.36), y viceversa.

- (2.35) \Rightarrow (2.36): Tomemos $u \in X$, solución al problema (2.35), tomemos $v \in X$ de forma arbitraria, y $\epsilon \in \mathbb{R}$. Formemos un nuevo elemento $u + \epsilon v \in X$.

Tenemos entonces que

2.1. FORMULACIÓN DÉBIL DEL PROBLEMA ELÍPTICO EN UNA DIMENSIÓN15

$$F(u) \leq F(u + \epsilon v), \quad (2.37)$$

considerando $g(\epsilon) \equiv F(u + \epsilon v)$, tenemos que el problema se puede escribir de la forma

$$g(0) \leq g(\epsilon), \forall \epsilon \in \mathbb{R}. \quad (2.38)$$

Esto implica que g tiene un mínimo en $\epsilon = 0$ y por lo tanto que la derivada de g respecto de ϵ evaluada en cero es nula, es decir,

$$g'(0) = 0. \quad (2.39)$$

Escribiendo de forma explícita g , tenemos

$$\begin{aligned} g(\epsilon) &= \frac{1}{2}Q_\gamma(u + \epsilon v, u + \epsilon v) - L_F(u + \epsilon v) \\ &= \frac{1}{2}Q_\gamma(u, u) + \epsilon Q_\gamma(u, v) + \frac{\epsilon^2}{2}Q_\gamma(v, v) - L_F(u) - \epsilon L_F(v), \end{aligned} \quad (2.40)$$

esta expresión al ser derivada con respecto de ϵ obtenemos

$$g'(\epsilon) = Q_\gamma(u, v) - L_F(v) + \epsilon Q_\gamma(u, v). \quad (2.41)$$

Entonces concluimos que $g'(0) = 0$ implica que

$$Q_\gamma(u, v) - L_F(v) = 0. \quad (2.42)$$

Puesto que $v \in X$ era arbitrario, hemos mostrado que si $u \in X$ satisface el problema variacional (2.35), también lo hace con (2.36).

- (2.36) \Rightarrow (2.35): Sea $u \in X$ una solución para (2.36), tomemos una $v \in X$ de forma arbitraria, y formemos $w = v - u \in X$. Escribamos $F(v)$ substituyendo $v = u + w$,

$$\begin{aligned} F(v) &= F(u + w) = \frac{1}{2}Q_\gamma(u + w, u + w) - L(u + w) \\ &= \frac{1}{2}Q_\gamma(u, u) + Q_\gamma(u, w) + \frac{1}{2}Q_\gamma(w, w) - L(u) - L(w) \\ &= \left[\frac{1}{2}Q_\gamma(u, u) - L(u) \right] + [Q_\gamma(u, w) - L(w)] + \frac{1}{2}Q_\gamma(w, w), \end{aligned} \quad (2.43)$$

aquí debemos usar que $Q_\gamma(u, w) = L(w), \forall w \in X$, ya que $u \in X$ es solución. Sabemos que $Q_\gamma(w, w) \leq 0, \forall w \in X$. Y reconociendo la igualdad $\frac{1}{2}Q_\gamma(u, u) - L(u) = F(u)$, tenemos que

$$F(v) \leq F(u), \forall v \in X. \quad (2.44)$$

Con esto hemos visto que $u \in X$ es solución para ambas formulaciones del problema.

2.2. El Método de Galerkin

En esta sección describimos el método de Galerkin, donde reemplazamos el espacio de Hilbert $X := H_o^1$ por un subespacio X_n de dimensión finita, y analizamos su convergencia.

Recordamos primero la formulación del problema débil:

Definición: 2.2.1 (Problema débil Abstracto sobre X .) *El problema débil consiste en encontrar una $u \in X$, tal que para toda $v \in X$*

$$Q_\gamma(u, v) = L_F(v), \quad (2.45)$$

donde $Q_\gamma(\cdot, \cdot)$ es un producto escalar en X , y L_F es un operador lineal acotado sobre X .

Para el problema original dado por la definición (2.1.1), hemos visto que al cambiar a la formulación débil describimos un problema débil descrito por la definición (2.1.9).

Una vez establecido el problema débil en el que se va a trabajar procedemos a buscar una aproximación para la solución al problema débil.

2.2.1. Discretización en el método de Galerkin.

Tomemos ahora subconjuntos de X , $X_n \subset X$, tal que X_n son subespacios de dimensión finita ($\dim X_n < \infty$). Entonces el problema débil discreto consiste en encontrar $u_n \in X_n$ tal que

$$Q_\gamma(u_n, \phi_n) = L_F(\phi_n), \forall \phi_n \in X_n. \quad (2.46)$$

El lema de Riesz nos asegura que existe una única solución discreta $u_n \in X_n$.

La aplicación a nuestro problema es directa.

Podemos entonces hacer una estimación del error, suponiendo que $u \in X$ es la *solución débil exacta*

$$Q_\gamma(u, v) = L_F(v), \forall v \in X, \quad (2.47)$$

y tomando ahora una *solución débil discreta* dada por $u_n \in X_n$

$$Q_\gamma(u_n, v) = L_F(v), \forall v \in X_n \subset X. \quad (2.48)$$

Apelando ahora a la bilinealidad de $Q_\gamma(\cdot, \cdot)$, con $v = \phi_n \in X_n$ tenemos que

$$Q_\gamma(u - u_n, \phi_n) = 0, \forall \phi_n \in X_n. \quad (2.49)$$

Esto nos dice que $u - u_n$, el error de la solución aproximada con respecto a la solución exacta, es ortogonal a X_n bajo Q_γ . Además, que $\|u - \phi_n\|_{Q_\gamma} = \text{dist}(u, X_n) = \inf_{v_n \in X_n} \|u - v_n\|_{Q_\gamma}$, para ver esto de forma más clara, hacemos referencia a la *figura (2.1)*.

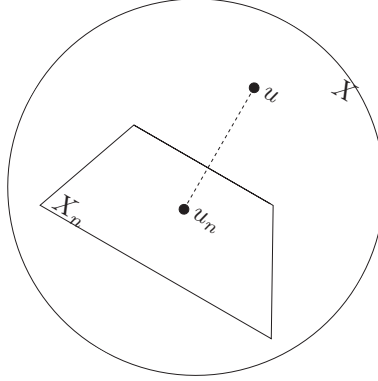


Figura 2.1: Figura donde se muestra que la solución aproximada u_n mediante el Método de Galerkin. u_n es la proyección ortogonal de $u \in X$ sobre X , de acuerdo con la expresión (2.49). El error $\|u - u_n\|_{Q_\gamma}$ es por definición la distancia entre u y X_n .

Ahora tratemos de encontrar una cota para la solución numérica

$$\|u - u_n\|_{H^1}^2 \leq \frac{1}{\bar{m}} Q_\gamma(u - u_n, u - u_n) = \frac{1}{\bar{m}} [Q_\gamma(u - u_n, u - u_n) + Q_\gamma(u - u_n, \phi_n)], \quad (2.50)$$

donde hemos usado la ortogonalidad del error a X_n , $Q_\gamma(u - u_n, \phi_n) = 0$. Definimos $\phi_n = u_n - v_n$, donde claramente $u_n, v_n \in X_n$. Usando ahora la bilinealidad de $Q_\gamma(\cdot, \cdot)$ tenemos que

$$Q_\gamma(u - u_n, u - u_n) + Q_\gamma(u - u_n, \phi_n) = Q_\gamma(u - u_n, u - v_n) \leq \|u - u_n\|_{Q_\gamma} \|u - v_n\|_{Q_\gamma}, \quad (2.51)$$

donde se ha usado la desigualdad de Cauchy-Schwarz. Haciendo uso de $Q_\gamma(u, u) \leq \bar{M} \|u\|_{H^1}^2$, tenemos que

$$\|u - u_n\|_{Q_\gamma} \|u - v_n\|_{Q_\gamma} \leq \bar{M} \|u - u_n\|_{H^1} \|u - v_n\|_{H^1}. \quad (2.52)$$

Y finalmente obtenemos que

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \|u - v_n\|_{H^1}, \forall v_n \in X_n. \quad (2.53)$$

Entonces, el error entre la solución discreta u_n y la solución exacta u satisface

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \inf_{v_n \in X_n} \|u - v_n\|_{H^1}. \quad (2.54)$$

La parte de la derecha de la expresión (2.54) es independiente de la solución numérica. Es proporcional a la distancia entre u y el subespacio X_n . De manera más geométrica, significa que la solución aproximada u_n en el subespacio de dimension finita (que posteriormente será el espacio de elementos finitos), X_n , es la proyección ortogonal de $u \in X$ sobre el espacio X_n , con respecto al producto $Q_\gamma(\cdot, \cdot)$. En particular, vemos que la solución $u_n \in X_n$ es el elemento de X_n que más se aproxima a la solución exacta.

2.2.2. Elementos Finitos

Elegimos el subconjunto X_n de $X = H_o^1$, i.e. el conjunto de *elementos finitos* con el cual vamos a trabajar. En nuestro caso, para el problema elíptico que hemos formulado, será un conjunto de funciones lineales por trozos.

Consideramos el intervalo $[0, 1] \subset \mathbb{R}$ y realizamos sobre él una partición $\{x_i\}$ con $i = 0, \dots, n+1$ del subconjunto de los reales de tal forma que

$$0 = x_0 < x_1 < x_2 < \dots < x_n < x_{n+1} = 1. \quad (2.55)$$

Tomemos una función $\phi \in H_o^1$, y evaluemos la función en cada uno de los puntos de la partición $\{x_i\}$ de tal forma que

$$\begin{aligned} \phi(x_j) &= \phi_j, \quad j = 1, \dots, n \\ \phi_0 &= \phi_{n+1} = 0. \end{aligned} \quad (2.56)$$

Entonces podemos formar una función, $N_j(x)$, tal que al ser evaluada en x_j es igual a 1 y que sea igual a cero para cualquier otro punto de la partición.

Definamos $N_j \in H_o^1$ como una función que sea lineal a trozos, como se dice a continuación:

Sea $N_j : I_{j-1, j+1} \subset \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$N_j(x) = \begin{cases} \frac{x - x_{j-1}}{\rho_{j-1}} & \text{si } x \in I_{j-1, j} \\ \frac{x_{j+1} - x}{\rho_j} & \text{si } x \in I_{j, j+1} \\ 0 & x \notin I_{j-1, j+1}. \end{cases} \quad (2.57)$$

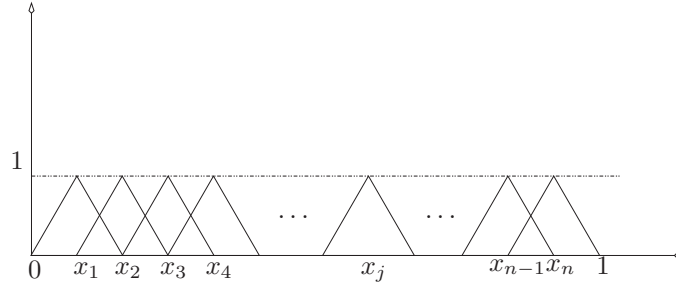


Figura 2.2: Figura que muestra los elementos finitos $N_j(x)$, $j = 1, \dots, n$, centrado en el punto x_j de la partición del intervalo $[0, 1]$.

donde $\rho_j = x_{j+1} - x_j$, con $j = 1, \dots, n$. Los subconjuntos $I_{\alpha, \beta} \subset \mathbb{R}$ son $I_{\alpha, \beta} := [x_\alpha, x_\beta] \subset [0, 1]$.

Entonces con la ayuda de estas funciones podemos aproximar una función $\phi \in H_o^1$ de forma única, mediante una expansión en terminos de (2.57)

$$\phi = \sum_{j=0}^{n+1} \phi_j N_j(x) = \sum_{j=1}^n \phi_j N_j(x), \quad (2.58)$$

donde hemos considerado de forma simple que $N_0 = N_{n+1} = 0$.

Entonces $X_n \subset H_o^1$, el conjunto de funciones lineales a trozos, forma un espacio vectorial de dimensión n ,

$$X_n := \left\{ \sum_{i=1}^n a_i N_i(x) \mid a_i \in \mathbb{R}, i = 1, \dots, n \right\}. \quad (2.59)$$

2.2.3. Aproximación por la interpolante.

Ahora bien, como habíamos visto, la norma H^1 del error discreto $u_n - u$ está acotado por la distancia $\text{dist}(u, X_n) = \inf_{v_n \in X_n} \|u - v_n\|_{Q_\gamma}$ entre u y el subespacio X_n . En términos de las normas H^1 ,

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \inf_{v_n \in X_n} \|u - v_n\|_{H^1}. \quad (2.60)$$

Para estimar $\inf_{v_n \in X_n} \|u - v_n\|_{H^1}$, construimos la *interpolante* $\hat{u}_n \in X_n$ de u , definida por la función lineal por trozos que satisface

$$\hat{u}_n(x_j) = u(x_j), \quad j = 0, \dots, n+1. \quad (2.61)$$

Entonces, llegamos a la estimación

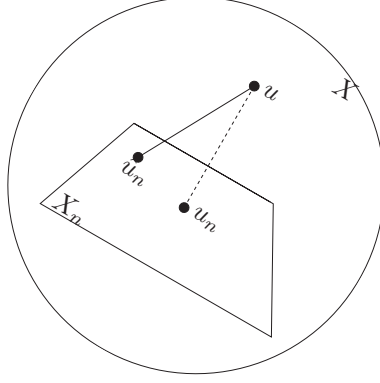


Figura 2.3: Figura donde se ilustra el método para estimar el error $u - u_n$.

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \|u - \hat{u}_n\|_{H^1} \quad (2.62)$$

para la norma H^1 del error. A continuación, acotamos la distancia entre u y su interpolante, demostrando de esta manera que u_n converge a u en el límite $n \rightarrow \infty$.

Para acotar esta distancia, nos basamos en el *teorema del valor medio* que nos dice que si nos situamos en un fragmento de la discretización, $I_{j,j+1} = [x_j, x_{j+1}]$, existe un punto en I , digamos x_o , tal que

$$u'(x_o) = \frac{u(x_{i+1}) - u(x_i)}{\rho_i}, \quad \rho_i = x_{i+1} - x_i. \quad (2.63)$$

Pero como sabemos de la propiedad anterior $\hat{u}_n(x_j) = u(x_j)$ y que \hat{u}_n es lineal en el intervalo $I_{j,j+1}$, concluimos que

$$u'(x_o) = \frac{\hat{u}_n(x_{i+1}) - \hat{u}_n(x_i)}{\rho_i} = \hat{u}'_n(x) \quad \forall x \in I_{j,j+1}. \quad (2.64)$$

Podemos mostrar el siguiente lema que nos ayuda a estimar la diferencia entre la derivada de la función $u \in X$ y la derivada de su interpolante $u_n \in X_n$.

Lema: 2.2.1 *Sea $u \in X$, y sea $\hat{u}_n \in X_n$ la interpolante de u . Entonces tenemos que*

$$\|u' - \hat{u}'_n\|_{\infty} \leq h \|u''\|_{\infty}, \quad (2.65)$$

donde $h := \max_{j=1,\dots,n} \rho_j$ y donde $\|f\|_\infty = \sup\{|f(x)| : \forall x \in I\}$.

Demostración. Situándonos de nuevo en el intervalo $I_{j,j+1}$, tenemos que

$$u'(x) = \int_{x_o}^x u''(z) dz + u'(x_o), \quad (2.66)$$

$$\hat{u}'_n(x) = \int_{x_o}^x \hat{u}''_n(z) dz + \hat{u}'_n(x_o). \quad (2.67)$$

Donde x_o es el punto proveniente del teorema del valor medio y x cualquier punto en $I_{j,j+1}$. Entonces tomando la diferencia de ambas expresiones tenemos que

$$u'(x) - \hat{u}'_n(x) = \int_{x_o}^x [u''(z) - \hat{u}''_n(z)] dz + u'(x_o) - \hat{u}'_n(x_o), \quad (2.68)$$

donde usando el hecho de que $u'(x_o) = \hat{u}'_n(x_o)$, tenemos

$$|u'(x) - \hat{u}'_n(x)| = \left| \int_{x_o}^x [u''(z) - \hat{u}''_n(z)] dz \right| \leq \int_{x_o}^x |u''(z) - \hat{u}''_n(z)| dz. \quad (2.69)$$

Tomemos en cuenta que \hat{u}_n se extrae de X_n y que éste está formado por funciones lineales a trozos por lo que su segunda derivada es igual a cero en todo punto de $I_{j,j+1}$, es decir, $\hat{u}''_n = 0$, $\forall x \in I_{j,j+1}$. Por lo tanto

$$|u'(x) - \hat{u}'_n(x)| \leq \int_{x_o}^x |u''(z)| dz. \quad (2.70)$$

Apelando a que $u \in C^2(I)$ en particular lo es para $I_{j,j+1}$ y entonces $u'' \in C^0(I_{j,j+1})$. Esto implica que existe un punto $z^* \in I_{j,j+1}$ tal que $u''(z^*) \geq u''(z)$ para todo $z \in I_{j,j+1}$. Por lo tanto obtenemos la expresión

$$|u'(x) - \hat{u}'_n(x)| \leq |u''(z^*)| \int_{x_o}^x dz \leq |u''(z^*)| h, \quad (2.71)$$

donde hemos usado el hecho de que $|x - x_o| \leq \rho_i \leq h$. Podemos deducir de forma inmediata la desigualdad que es válida para todo $x \in I_{j,j+1}$,

$$|u'(x) - \hat{u}'_n(x)| \leq \max_{x_j \leq x \leq x_{j+1}} |u''(x)| h, \quad (2.72)$$

podemos apreciar que tenemos una desigualdad como ésta para cada $I_{j,j+1}$ y entonces podemos obtener la norma infinito, tomando el supremo de ambos lados de la desigualdad

$$\|u' - \hat{u}'_n\|_\infty \leq h \|u''\|_\infty. \quad (2.73)$$

□

De forma análoga podemos hacer una estimación para la diferencia entre la solución u y la interpolante \hat{u}_n .

Lema: 2.2.2 *Sea $u \in X$, y sea $\hat{u}_n \in X_n$ la interpolante de u . Entonces tenemos que*

$$\|u - \hat{u}_n\|_\infty \leq h^2 \|u''\|_\infty. \quad (2.74)$$

Demostración. Tomando en cuenta que

$$u(x) = \int_{x_j}^x u'(z) dz + u(x_j), \quad (2.75)$$

$$\hat{u}'_n(x) = \int_{x_j}^x \hat{u}'_n(z) dz + \hat{u}_n(x_j), \quad (2.76)$$

donde usamos el hecho de que $\hat{u}_n(x_j) = u(x_j)$, obtenemos así la diferencia

$$|u(x) - \hat{u}_n(x)| = \left| \int_{x_j}^x [u'(z) - \hat{u}'_n(z)] dz \right| \leq \int_{x_j}^x |u'(z) - \hat{u}'_n(z)| dz. \quad (2.77)$$

Usando la estimación (2.72) encontrada con anterioridad, tenemos que

$$|u(x) - \hat{u}'_n(x)| \leq \max_{x_j \leq x \leq x_{j+1}} |u''(x)| h \left(\int_{x_j}^x dz \right) \leq \max_{x_j \leq x \leq x_{j+1}} |u''(x)| h^2, \quad (2.78)$$

y de forma análoga al caso anterior tenemos

$$\|u - \hat{u}_n\|_\infty \leq h^2 \|u''\|_\infty. \quad (2.79)$$

□

Podemos encontrar las estimaciones correspondientes bajo la norma L^2 .

Lema: 2.2.3 *Sea $u \in X$, y sea $\hat{u}_n \in X_n$ la interpolante de u . Entonces tenemos que*

$$\|u' - \hat{u}'_n\|_{L^2} \leq \frac{h}{\sqrt{2}} \|u''\|_{L^2}, \quad (2.80)$$

donde $\|f\|_{L^2}^2 = \int_0^1 |f(x)|^2 dx$, $f \in L^2(I)$.

Demostración. Tomemos la norma L^2 elevada a la segunda potencia en los puntos del intervalo $I = [x_j, x_{j+1}]$,

$$\|u' - \hat{u}'_n\|_{L^2(x_j, x_{j+1})}^2 = \int_{x_j}^{\xi} |u'(x) - \hat{u}'_n(x)|^2 dx + \int_{\xi}^{x_{j+1}} |u'(x) - \hat{u}'_n(x)|^2 dx \quad (2.81)$$

donde nuevamente $\xi \in I_{j,j+1}$ es el punto intermedio proveniente del teorema del valor medio para cada subconjunto de $(0, 1)$.

Tomamos el valor absoluto de $u'(x) - \hat{u}'_n(x)$ y lo elevamos al cuadrado para acotar la segunda integral de (2.81), donde conocemos que $x > \xi$,

$$|u'(x) - \hat{u}'_n(x)|^2 = \left| \int_{\xi}^x u''(z) dz \right|^2 \leq \int_{\xi}^x |u''(z)|^2 dz (x - \xi) = \|u\|_{L^2(\xi, x)}^2 (x - \xi), \quad (2.82)$$

donde hemos usado la desigualdad de Cauchy-Schwarz. Entonces al realizar la segunda integral de (2.81) tenemos que

$$\begin{aligned} \int_{\xi}^{x_{j+1}} |u'(x) - \hat{u}'_n(x)|^2 dx &\leq \|u\|_{L^2(\xi, x_{j+1})}^2 \int_{\xi}^{x_{j+1}} (x - \xi) dx \\ &= \|u\|_{L^2(\xi, x_{j+1})}^2 \frac{(x_{j+1} - \xi)^2}{2} \leq \|u\|_{L^2(\xi, x_{j+1})}^2 \frac{h^2}{2}. \end{aligned} \quad (2.83)$$

En la primera integral de (2.81) considerando que $x < \xi$

$$|u'(x) - \hat{u}'_n(x)|^2 = \left| \int_x^{\xi} u''(z) dz \right|^2 \leq \int_x^{\xi} |u''(z)|^2 dz (\xi - x) = \|u\|_{L^2(x, \xi)}^2 (\xi - x). \quad (2.84)$$

Entonces al acotar la primera integral encontramos que

$$\begin{aligned} \int_{x_j}^{\xi} |u'(x) - \hat{u}'_n(x)|^2 dx &\leq \|u\|_{L^2(x_j, \xi)}^2 \int_{x_j}^{\xi} (\xi - x) dx \\ &= \|u\|_{L^2(x_j, \xi)}^2 \frac{(\xi - x_j)^2}{2} \leq \|u\|_{L^2(x_j, \xi)}^2 \frac{h^2}{2}. \end{aligned} \quad (2.85)$$

Así al sumar las dos aproximaciones para las integrales encontramos que

$$\|u' - \hat{u}'_n\|_{L^2(x_j, x_{j+1})}^2 \leq \frac{h^2}{2} \left\{ \|u\|_{L^2(x_j, \xi)}^2 + \|u\|_{L^2(\xi, x_{j+1})}^2 \right\} = \frac{h^2}{2} \|u\|_{L^2(x_j, x_{j+1})}^2. \quad (2.86)$$

La afirmación se obtiene de este resultado al sumar sobre los intervalos $I_{j,j+1}$ y tomar la raíz cuadrada.

□

Lema: 2.2.4 *Sea $u \in X$, y sea $\hat{u}_n \in X_n$ la interpolante de u . Entonces tenemos que*

$$\|u - \hat{u}_n\|_{L^2} \leq \frac{h^2}{2} \|u''\|_{L^2}. \quad (2.87)$$

Demostración. Para esto procedemos tomando de la definición, la norma de la diferencia entre la solución y su interpolante sobre el intervalo $I_{j,j+1}$, donde usaremos nuevamente que en los nodos x_j la solución es igual a su interpolante. Tenemos

$$\|u - \hat{u}_n\|_{L^2(x_j, x_{j+1})}^2 = \int_{x_j}^{x_{j+1}} |u(x) - \hat{u}_n(x)|^2 dx. \quad (2.88)$$

Tomemos ahora el integrando

$$\begin{aligned} |u(x) - \hat{u}_n(x)|^2 &= \left| \int_{x_j}^x [u'(z) - \hat{u}'_n(z)] dz \right|^2 \leq \left[\int_{x_j}^x |u'(z) - \hat{u}'_n(z)| dz \right]^2 \\ &\leq \left[\int_{x_j}^{x_{j+1}} |u'(z) - \hat{u}'_n(z)|^2 dz \right] (x - x_j), \end{aligned} \quad (2.89)$$

donde en el último paso hemos usado la desigualdad de Cachy-Schwarz y hemos integrado, así vemos que podemos reescribir el resultado anterior como

$$|u(x) - \hat{u}_n(x)|^2 \leq \|u' - \hat{u}'_n\|_{L^2(x_j, x_{j+1})}^2 (x - x_j). \quad (2.90)$$

Ahora procedemos a implementar esta desigualdad en la ecuación (2.88), e integramos

$$\begin{aligned} \|u - \hat{u}_n\|_{L^2(x_j, x_{j+1})}^2 &\leq \|u' - \hat{u}'_n\|_{L^2(x_j, x_{j+1})}^2 \int_{x_j}^{x_{j+1}} (x - x_j) dx \\ &= \frac{h^2}{2} \|u' - \hat{u}'_n\|_{L^2(x_j, x_{j+1})}^2. \end{aligned} \quad (2.91)$$

$\ \cdot\ _\infty$	$\ \cdot\ _{L^2}$
$\ u - \hat{u}_n\ _\infty \leq h^2 \ u''\ _\infty$	$\ u - \hat{u}_n\ _{L^2} \leq \frac{h^2}{2} \ u''\ _{L^2}$
$\ u' - \hat{u}'_n\ _\infty \leq h \ u''\ _\infty$	$\ u' - \hat{u}'_n\ _{L^2} \leq \frac{h}{\sqrt{2}} \ u''\ _{L^2}$

Cuadro 2.1: Estimaciones de la diferencia entre la solución y su interpolante.

Sumando sobre todos los intervalos obtenemos

$$\|u - \hat{u}_n\|_{L^2}^2 \leq \frac{h^2}{2} \|u' - \hat{u}'_n\|_{L^2}^2, \quad (2.92)$$

y usando el resultado del lema (2.2.3)

$$\|u - \hat{u}_n\|_{L^2} \leq \frac{h^2}{2} \|u''\|_{L^2}. \quad (2.93)$$

□

2.2.4. Convergencia del Método.

Ya que tenemos las estimaciones de la diferencia entre $u \in H_o^1$, y la interpolante $\hat{u}_n \in X_n$, tomamos a $v_n = \hat{u}_n$ en la estimación bajo la norma en H^1 , dada por la desigualdad (2.54), tenemos que

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \|u - \hat{u}_n\|_{H^1}. \quad (2.94)$$

Aquí usamos las estimaciones del Cuadro (2.1), de donde encontramos que

$$\|u - \hat{u}_n\|_{H^1}^2 = \|u - \hat{u}_n\|_{L^2}^2 + \|u' - \hat{u}'_n\|_{L^2}^2 \leq h^2 \|u''\|_{L^2}^2. \quad (2.95)$$

Usando este resultado junto con (2.94) obtenemos una estimación para la norma del error en la norma H^1 , de la solución con respecto a la solución aproximada

$$\|u - u_n\|_{H^1} \leq \frac{\bar{M}}{\bar{m}} \|u - \hat{u}_n\|_{H^1} \leq h \frac{\bar{M}}{\bar{m}} \|u''\|_{L^2}. \quad (2.96)$$

o bien

$$\boxed{\|u - u_n\|_{H^1} \leq Ch \|u''\|_{L^2}, \quad C = \frac{\bar{M}}{\bar{m}}.} \quad (2.97)$$

Esto nos dice que la norma H^1 del error converge a cero por lo menos a orden h .

Ahora vamos a ver si podemos obtener una mejor cota para la norma L^2 del error.

Tomemos el error $e = u - u_n \in X = H_o^1$, sabemos que

$$Q_\gamma(e, v_n) = 0, \forall v_n \in X_n. \quad (2.98)$$

Ahora estimemos la norma sobre L_2 de $e \in H_o^1$, $\|e\|_{L_2}$, usando un argumento de dualidad.

Entonces consideramos el siguiente problema dual auxiliar,

$$-\phi'' + \gamma\phi = e, \forall x \in I, \quad (2.99)$$

$$\phi(0) = \phi(1) = 0, \quad (2.100)$$

donde $\phi \in H_o^2$ es la solución al problema dual auxiliar. La desigualdad (2.33) implica que $\|\phi\|_{H^1} \leq \bar{m}^{-1}\|e\|_{L^2}$. Por lo tanto, $\|\phi''\|_{L^2} \leq \bar{M}\|\phi\|_{L^2} + \|e\|_{L^2} \leq (1 + \bar{M}/\bar{m})\|e\|_{L^2}$. Entonces,

$$\|\phi\|_{H^2} \leq c_o\|e\|_{L^2}. \quad (2.101)$$

Por lo tanto, usando (2.99) en la norma L^2 del error

$$\|e\|_{L^2}^2 = (e, e)_{L^2} = (e, -\phi'' + \gamma\phi) = Q_\gamma(e, \phi). \quad (2.102)$$

Ahora tomemos la interpolante $\Pi_n\phi \in X_n \subset H_o^1$ y observamos que $Q_\gamma(e, \Pi_n\phi) = 0$. Usando este argumento de ortogonalidad y el hecho de que $Q_\gamma(\cdot, \cdot)$ es bilineal tenemos

$$\|e\|_{L^2}^2 = Q_\gamma(e, \phi - \Pi_n\phi) \leq c_1\|e\|_{H_1}\|\phi - \Pi_n\phi\|_{H_1}. \quad (2.103)$$

Implementando en la desigualdad anterior la estimación por la interpolante y después el resultado (2.101), es decir,

$$\|e\|_{L^2}^2 \leq c_1c_2h\|e\|_{H_1}\|\phi\|_{H^2} \leq c_oc_1c_2h\|e\|_{H_1}\|e\|_{L^2}, \quad (2.104)$$

al dividir por $\|e\|_{L^2}$ obtenemos

$$\boxed{\|e\|_{L^2} \leq C'h\|e\|_{H_1}}. \quad (2.105)$$

Usando la estimación dada por (2.97), es decir,

$$\|e\|_{H_1} \leq Ch\|u''\|_{L^2}, \quad (2.106)$$

y la desigualdad (2.105), obtenemos

$$\boxed{\|e\|_{L^2} \leq CC'h^2\|u''\|_{L^2}}. \quad (2.107)$$

Con este resultado podemos deducir que el error debido a la diferencia entre la solución exacta al problema débil $u \in X$ y la solución aproximada $u_n \in X_n$ converge a cero como h^2 bajo la norma $L^2(0, 1)$.

Esto nos dice que si tomamos una resolución para una distribución uniforme de tamaño $\rho_j^{(1)} = h$ para todo $j = 1, \dots, n$ y después tomamos otra resolución $\rho_j^{(2)} = h/2$, entonces esperamos obtener el siguiente factor de convergencia definido por el cociente entre el error en la primera resolución y el de la segunda resolución, es decir

$$\frac{\|e^{(1)}\|_{L^2}}{\|e^{(2)}\|_{L^2}} = \frac{h^2}{(h/2)^2} = 4. \quad (2.108)$$

A éste resultado se le llama *factor de convergencia*, y es el que esperamos reproducir al momento de nuestra implementación numérica.

2.3. Implementación del Método de Galerkin

Tenemos todos los elementos necesarios para aproximar $u \in H_o^1$, mediante una expansión de los elementos de X_n .

Así necesitamos encontrar $u_n \in X_n$ tal que

$$Q_\gamma(u_n, \phi_n) = L_F(\phi_n), \quad \forall \phi_n \in X_n. \quad (2.109)$$

Definimos entonces $u_n \in X_n$, mediante una expansión en los elementos de X_n

$$u_n(x) = \sum_{j=1}^n x_j N_j(x), \quad (2.110)$$

$$\phi_n(x) = \sum_{l=1}^n b_l N_l(x), \quad (2.111)$$

tal que $x_j, b_l \in \mathbb{R}$. Aplicando estos elementos de X_n a (2.109), y usando la bilinealidad de $Q_\gamma(\cdot, \cdot)$, así como la linealidad de $L_F(\cdot)$ obtenemos como resultado

$$\sum_{l=1}^n \left(\sum_{j=1}^n x_j Q_\gamma(N_j, N_l) - L_F(N_l) \right) b_l = 0. \quad (2.112)$$

Como ésta debe ser válida para todo $\phi_n \in X_n$, debe serlo también para todo $b_l \in \mathbb{R}$, por lo tanto

$$\sum_{j=1}^n x_j A_{jl} = L_F(N_l), \quad (2.113)$$

donde hemos definido

$$A_{ij} = Q_\gamma(N_i, N_j) = \int_0^1 \gamma(x) N_i(x) N_j(x) dx + \int_0^1 N_i'(x) N_j'(x) dx. \quad (2.114)$$

Con el propósito de expresar de forma matricial (2.113) definimos el vector columna v , tal que

$$v^t = (x_1, x_2, \dots, x_n), \quad (2.115)$$

tenemos que (2.113) se puede escribir como

$$A \cdot v = J, \quad (2.116)$$

donde J es el vector columna definido por

$$J^t = (L_F(N_1), L_F(N_2), \dots, L_F(N_n)). \quad (2.117)$$

2.3.1. Cálculo de la matriz A .

No necesitamos calcular cada uno de los $n \times n$ componentes de la matriz, esto lo podemos apreciar de las propiedades de la matriz A :

- La matriz es simétrica, debido a que

$$A_{ij} = Q_\gamma(N_i, N_j) = Q_\gamma(N_j, N_i) = A_{ji}. \quad (2.118)$$

- Vemos que A_{ij} es distinta de cero, si y sólo si, $|i - j| \leq 1$. Por lo tanto A_{ij} es una matriz tridiagonal.

El siguiente lema, nos ayuda a calcular los elementos de la matriz A .

Lema: 2.3.1 *Sea $g \in C^2(0, 1)$ una función de x que es dos veces continuamente diferenciable y sea $0 = x_0 < x_1 < x_2 < \dots < x_n = 1$, una partición de $[0, 1]$. Entonces*

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \frac{1}{2} [g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (2.119)$$

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) \left(\frac{x_{j+1} - x}{x_{j+1} - x_j} \right)^2 dx = \frac{1}{12} [3g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (2.120)$$

$$\frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx = \frac{1}{12} [3g(x_j) + g(x_{j-1})] + O(\rho^2), \quad (2.121)$$

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) \frac{x_{j+1} - x}{x_{j+1} - x_j} \frac{x - x_j}{x_{j+1} - x_j} dx = \frac{1}{12} [g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (2.122)$$

donde $\rho_j := x_{j+1} - x_j$ y $\rho = \max_{j=0,1,\dots,n} \rho_j$.

Demostración. Ver el Apéndice (A.0.6). □

Elementos de la diagonal de A .

Comenzaremos primero con los elementos de la diagonal

$$A_{l,l} = \int_0^1 \gamma(x) N_l(x)^2 dx + \int_0^1 N'_l(x)^2 dx. \quad (2.123)$$

Resolvamos para la primera de las dos integrales

$$\int_0^1 \gamma(x) N_l(x)^2 dx = \int_{x_{l-1}}^{x_l} \gamma(x) \left(\frac{x - x_{l-1}}{\rho_{l-1}} \right)^2 dx + \int_{x_l}^{x_{l+1}} \gamma(x) \left(\frac{x_{l+1} - x}{\rho_l} \right)^2 dx, \quad (2.124)$$

usando (2.121) y (2.120) para la primera y segunda intergral respectivamente tenemos

$$\begin{aligned} \int_0^1 \gamma(x) N_l(x)^2 dx &= \frac{\rho_{l-1}}{12} [3\gamma(x_l) + \gamma(x_{l-1})] + \frac{\rho_l}{12} [3\gamma(x_l) + \gamma(x_{l+1})] + O(\rho^2) \\ &= \frac{1}{4}(\rho_{l+1} + \rho_l)\gamma(x_l) + \frac{1}{12} [\rho_{l-1}\gamma(x_{l-1}) + \rho_l\gamma(x_{l+1})] + O(\rho^3). \end{aligned} \quad (2.125)$$

o bien,

$$\int_0^1 \gamma(x) N_l(x)^2 dx = \frac{1}{4}(\rho_{l+1} + \rho_l)\gamma(x_l) + \frac{1}{12} [\rho_{l-1}\gamma(x_{l-1}) + \rho_l\gamma(x_{l+1})] + O(\rho^3). \quad (2.126)$$

La segunda integral es calculada de forma sencilla

$$\int_0^1 N'_l(x)^2 dx = \int_{x_{l-1}}^{x_l} \frac{dx}{\rho_{l-1}^2} + \int_{x_l}^{x_{l+1}} \frac{dx}{\rho_l^2} \quad (2.127)$$

y se tiene que

$$\int_0^1 N'_l(x)^2 dx = \frac{1}{\rho_{l-1}} + \frac{1}{\rho_l}. \quad (2.128)$$

Tenemos que la aproximación para los elementos de la diagonal a orden $O(\rho^2)$ es

$$A_{l,l} = \frac{1}{4}\gamma(x_l)(x_{l+1} - x_{l-1}) + \frac{1}{12} [\rho_{l-1}\gamma(x_{l-1}) + \rho_l\gamma(x_{l+1})] + \frac{1}{\rho_{l-1}} + \frac{1}{\rho_l} + O(\rho^3), l = 1, \dots, n.$$

(2.129)

El cual para particiones uniformes del intervalo donde $\rho_l = x_{l+1} - x_l = \rho$, con $l = 1, \dots, n$, se reduce a la siguiente expresión

$$A_{l,l} = \frac{2}{3}\gamma(x_l)\rho + \frac{2}{\rho} + O(\rho^3). \quad (2.130)$$

Elementos fuera de la diagonal de A .

Procedemos a calcular los elementos fuera de la diagonal $A_{l,l+1} = A_{l+1,l}$,

$$A_{l,l+1} = \int_0^1 \gamma(x)N_l(x)N_{l+1}(x)dx + \int_0^1 N'_l(x)N'_{l+1}(x)dx, \quad (2.131)$$

para esto formamos los productos de N_l y N'_l , que son distintos de cero.

Primero

$$N_l(x)N_{l+1}(x) = \frac{1}{(x_{l+1} - x_l)^2}(x_{l+1} - x)(x - x_l), \quad (2.132)$$

la cual es distinta de cero solamente en $I_{l,l+1}$. Por otro lado tenemos

$$N'_l(x)N'_{l+1}(x) = -\frac{1}{(x_{l+1} - x_l)^2}. \quad (2.133)$$

Calculemos la integral

$$\int_0^1 \gamma(x)N_l(x)N_{l+1}(x)dx = \int_{x_l}^{x_{l+1}} \gamma(x) \frac{x_{l+1} - x}{x_{l+1} - x_l} \frac{x - x_l}{x_{l+1} - x_l} dx, \quad (2.134)$$

donde usaremos la aproximación del lema (2.122) y obtenemos

$$\int_0^1 \gamma(x)N_l(x)N_{l+1}(x)dx = \frac{\rho_{l-1}}{12} [\gamma(x_{l-1}) + \gamma(x_l)] + O(\rho^3). \quad (2.135)$$

La segunda integral del elemento fuera de la diagonal es

$$\int_0^1 N'_{l-1}N'_l dx = -\int_0^1 \frac{dx}{\rho_{l-1}^2}, \quad (2.136)$$

y por lo tanto

$$\int_0^1 N'_{l-1}N'_l dx = -\frac{1}{\rho_{l-1}}. \quad (2.137)$$

Entonces podemos escribir los elementos fuera de la diagonal sumando (2.135) y (2.137)

$$\boxed{A_{l,l+1} = A_{l+1,l} = \frac{\rho_{l-1}}{12} [\gamma(x_{l-1}) + \gamma(x_l)] - \frac{1}{\rho_{l-1}} + O(\rho^3), l = 1, \dots, n-1.} \quad (2.138)$$

Para particiones uniformes $\rho_l = x_{l+1} - x_l = \rho$, se pueden expresar de la forma

$$A_{l,l+1} = \gamma(x_l + \rho/2) \frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3). \quad (2.139)$$

En este trabajo de tesis tomaremos una distribución uniforme por lo que presentamos a continuación los elementos de la matriz tridiagonal:

$$A_{l,l} = \frac{2}{3} \gamma(x_l) \rho + \frac{2}{\rho} + O(\rho^3), l = 1, \dots, n, \quad (2.140)$$

$$A_{l,l+1} = A_{l+1,l} = \gamma(x_l + \rho/2) \frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3), l = 1, \dots, n-1, \quad (2.141)$$

$$A_{i,j} = 0, \text{ para todos los demás casos.} \quad (2.142)$$

2.3.2. Aproximación de L_F

Para hacer una aproximación del elemento $L_F(N_j)$, con $j = 1, \dots, n$ tenemos que

$$\frac{1}{\rho} L_F(N_j) = \int_{x_{j-1}}^{x_j} \frac{x - x_{j-1}}{x_j - x_{j-1}} F(x) dx + \int_{x_j}^{x_{j+1}} \frac{x_{j+1} - x}{x_{j+1} - x_j} F(x) dx, \quad (2.143)$$

donde como sabemos $\rho_j = x_{j-1} - x_j$ con $j = 1, \dots, n$, es el tamaño de la discretización.

Ahora bien

$$L_F(N_j) = \frac{1}{\rho_{j-1}} \int_j^x (x - x_{j-1}) F(x) dx + \frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} (x_{j+1} - x) F(x) dx. \quad (2.144)$$

Resolvemos la primera integral la derecha de la ecuación (2.144), haciendo el cambio de variable $x = x_j - w$ y obtenemos la integral

$$\frac{1}{\rho_{j-1}} \int_j^x (x - x_{j-1}) F(x) dx = \frac{1}{\rho_{j-1}} \int_0^{\rho_{j-1}} (\rho_{j-1} - w) F(x_j - w) dw. \quad (2.145)$$

Para la segunda integral del lado derecho de la ecuación (2.144) hacemos $x = x_j + w$ para obtener

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} (x_{j+1} - x)F(x)dx = \frac{1}{\rho_j} \int_0^{\rho_j} (\rho_j - w)F(x_j + w)dw. \quad (2.146)$$

Entonces al sustituir (2.145) y (2.146) en (2.144) tenemos

$$L_F(N_j) = \frac{1}{\rho_{j-1}} \int_0^{\rho_{j-1}} (\rho_{j-1} - w)F(x_j - w)dw + \frac{1}{\rho_j} \int_0^{\rho_j} (\rho_j - w)F(x_j + w)dw. \quad (2.147)$$

En este momento realizamos una expansión de Taylor de la función F alrededor de x_j , de donde encontramos de forma simple que

$$F(x_j \pm w) = F(x_j) \pm F'(x_j)w + \frac{1}{2}F''(x_j)w^2 \pm \dots \quad (2.148)$$

Sustituyendo esta aproximación para F en (2.147) e integrando, obtenemos

$$L_F(N_j) = F(x_j) \frac{x_{j+1} - x_{j-1}}{2} + \frac{x_{j+1} - x_{j-1}}{2} \frac{\rho_j - \rho_{j-1}}{3} F'(x_j) + O(\rho^2). \quad (2.149)$$

Para discretizaciones uniformes donde $\rho_j = \rho$ con $j = 1, \dots, n$, reducimos la expresión anterior a

$$\frac{1}{\rho} L_F(N_j) = F(x_j) + O(\rho^2). \quad (2.150)$$

Observación: 2.3.1 Una vez calculadas todas las aproximaciones de los elementos de la matriz y el vector de la derecha, podemos ver después de algunos cálculos, que la j -ésima entrada del vector Av es

$$(Av)_j = (A_{jm}v^m) = A_{j,j-1}v^{j-1} + A_{j,j}v^j + A_{j,j+1}v^{j+1} = J_j = L_F(N_j). \quad (2.151)$$

Recordando que los elementos $A_{j-1,j}$ y $A_{j,j}$ de forma explícita son

$$A_{j-1,j} = \gamma(x_{j-1} + \rho/2) \frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3), \quad A_{j,j} = \frac{2}{3}\rho\gamma(x_j) + \frac{2}{\rho} + O(\rho^3). \quad (2.152)$$

Podemos escribir $(Av)_j = L_F(N_j)$ de la siguiente forma

$$\left[\gamma(x_{j-1} + \rho/2) \frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3) \right] v^{j-1} + \left[\frac{2}{3}\rho\gamma(x_j) + \frac{2}{\rho} + O(\rho^3) \right] v^j + \left[\gamma(x_j + \rho/2) \frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3) \right] v^{j+1} = L_F(N_j). \quad (2.153)$$

Si dividimos ambos lados de la ecuación anterior entre ρ , y agrupamos en potencias de ρ obtenemos

$$\begin{aligned} \frac{1}{6} [\gamma(x_{j-1} + \rho/2)v^{j-1} + 4\gamma(x_j)v^j + \gamma(x_j + \rho/2)v^{j+1}] - \\ \frac{1}{\rho^2} [v^{j-1} - 2v^j + v^{j+1}] + O(\rho^2) = F(x_j) + O(\rho^2). \end{aligned} \quad (2.154)$$

El primer término de ésta ecuación, cuando hacemos que $\rho \rightarrow 0$, es un promedio de $\gamma(x_j)v_j$, mientras que el segundo término de lado izquierdo de la ecuación es el operador diferencial de segundo orden discretizado que se usa en diferencias finitas. Ésta aproximación es consistente con el problema elíptico inicial.

2.3.3. Solución al Sistema $Av = J$

Como vimos el problema de elementos finitos aplicado al problema elíptico nos da como resultado un sistema de ecuaciones que se puede escribir de forma matricial como

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \quad (2.155)$$

donde sabemos que \mathbf{A} es una matriz cuadrada ($n \times n$), tridiagonal y simétrica, \mathbf{b} es un vector de n dimensiones. Entonces el objetivo consiste en encontrar \mathbf{x} un vector de n dimensiones tal que satisfaga a (2.155).

La matriz tridiagonal luce de la siguiente forma,

$$\begin{pmatrix} a_{11} & a_{12} & & & & \\ a_{21} & a_{22} & a_{23} & & & \\ & a_{32} & a_{33} & a_{34} & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{j,j-1} & a_{jj} & a_{j,j+1} \\ & & & & \ddots & \ddots \\ & & & & & a_{n,n-1} & a_{nn} \end{pmatrix} \quad (2.156)$$

Así podemos formar la matriz aumentada

$$\bar{\mathbf{A}} = (\mathbf{A}|\mathbf{b}), \quad (2.157)$$

$$\begin{pmatrix} a_{11} & a_{12} & & & & & b_1 \\ a_{21} & a_{22} & a_{23} & & & & b_2 \\ & a_{32} & a_{33} & a_{34} & & & b_3 \\ & & \ddots & \ddots & \ddots & & \vdots \\ & & & a_{j,j-1} & a_{jj} & a_{j,j+1} & b_j \\ & & & & \ddots & \ddots & \vdots \\ & & & & & a_{n,n-1} & a_{nn} & b_n \end{pmatrix} \quad (2.158)$$

donde $\bar{\mathbf{A}} \in \mathbf{M}_{n \times n+1}$.

Con el fin de llevar a la matrix \mathbf{A} a la matriz identidad, realizamos operaciones con los renglones, las cuales básicamente son multiplicar por un escalar, sumar y restar entre los renglones. Las operaciones que se realizan sobre \mathbf{A} , esto cambia las componentes b_j , lo cual nos da la solución al sistema (2.155).

De forma específica: el *Método de Thomas* (ver, por ejemplo, [9]) consiste mediante operaciones de los renglones de la matriz aumentada, hacer ceros la diagonal inferior, para despues hacer ceros en la diagonal superior y por último normalizar los elementos de la diagonal.

Siguiendo el *Método de Thomas* diagonalizamos la matriz \mathbf{A} , para esto definimos un vector por cada renglón de la forma

$$\vec{r}_i = (0, \dots, a_{i,i-1}, a_{i,i}, a_{i,i+1}, 0, \dots, 0), \quad (2.159)$$

a su vez podemos tomar $b_j = J(N_j)$ y así cada renglón de la matriz aumentada es

$$\vec{v}_i = (\vec{r}_i, b_i). \quad (2.160)$$

Entonces usando el método de eliminación gaussiana podemos hacer cero los elementos de la diagonal inferior usando la siguiente relación entre los renglones consecutivos

$$\vec{v}'_{i+1} = \vec{v}_{i+1} - \frac{\vec{v}_i}{a'_{i,i}} a_{i+1,i}, \quad i = 1, 2, \dots, n-1. \quad (2.161)$$

Al tomar la i -ésima componente para los vectores que componen (2.161) tenemos que

- $a'_{i+1,i} = 0$, para $i = 1, \dots, n-1$.
- $a'_{i+1,i} = a_{i+1,i} - \frac{a_{ii}}{a'_{i,i}} a_{i+1,i}$, para $i = 1, \dots, n-1$, donde usamos $a'_{11} = a_{11}$.
- Los elementos de la diagonal superior no sufren cambio alguno.

- Las componentes del vector b se ven afectadas

$$b'_{i+1} = b_{i+1} - \frac{b_i}{a'_{i,i}} a_{i+1,i}, \quad i = 1, 2, \dots, n-1. \quad (2.162)$$

Después de realizar este proceso desde $i = 1$ hasta $i = n-1$, ya que comenzamos usando que $a'_{11} = a_{11}$, obtenemos una matriz que solamente tiene diagonal superior y diagonal.

$$\begin{pmatrix} a_{11} & a_{12} & & & & & b'_1 \\ & a'_{22} & a_{23} & & & & b'_2 \\ & 0 & a'_{33} & a_{34} & & & b'_3 \\ & & \ddots & \ddots & \ddots & & \vdots \\ & & & 0 & a'_{jj} & a_{j,j+1} & b'_j \\ & & & & \ddots & \ddots & \vdots \\ & & & & & 0 & a'_{nn} & b'_n \end{pmatrix} \quad (2.163)$$

Hasta este momento nos damos cuenta que la forma de encontrar la solución es muy sencilla, ya que tenemos

$$a'_{nn} x_n = b'_n, \quad (2.164)$$

de donde es fácil obtener la solución para x_n ,

$$x_n = \frac{b'_n}{a'_{nn}}. \quad (2.165)$$

De forma general, para $j = n-1, n-2, \dots, 1$ tenemos que

$$a'_{jj} x_j + a_{j,j+1} x_{j+1} = b'_j, \quad (2.166)$$

y así resolvemos para x_j

$$x_j = \frac{b'_j - a_{j,j+1} x_{j+1}}{a'_{jj}}, \quad j = n-1, n-2, \dots, 1. \quad (2.167)$$

Como vemos, el Método de Thomas termina en $2(n-1)$ operaciones.

Existen dos problemas para métodos derivados de la eliminación gaussiana, y son

- la presencia de errores de redondeo,
- sistemas que son *mal-condicionados*.

A continuación explicamos cuando un problema es *bien-condicionad*.

2.3.4. Condición del sistema

Un problema *bien-condicionado* es aquel en el que pequeños cambios en algún elemento del problema causa pequeños cambios en la solución.

Un problema *mal-condicionado* es aquel en el que un pequeño cambio en alguno de los elementos del problema causa grandes cambios en la solución del problema. Es necesario comentar, que debido a que los problemas *mal-condicionados* son demasiado sensibles a pequeños cambios en los elementos del problema también son sensibles a errores de redondeo.

Podemos analizar la matriz \mathbf{A} para saber la *condición* del sistema de ecuaciones (2.155). A continuación describimos como podemos estimar el *número condicionante*.

2.3.5. Número condicionante

El *número condicionante* de un sistema es una medida de lo sensible que el sistema puede ser a pequeños cambios en alguno de sus elementos. Consideremos el sistema algebraico de ecuaciones lineales

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}. \quad (2.168)$$

Para esta ecuación tenemos que gracias a la desigualdad de Schwarz,

$$\|\mathbf{b}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|. \quad (2.169)$$

Si modificamos un poco el problema original cambiando a \mathbf{b} sumando un $\delta\mathbf{b}$, esto causa un pequeño cambio en la solución, que se puede expresar como $\mathbf{x} + \delta\mathbf{x}$, y así

$$\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}. \quad (2.170)$$

Es fácil obtener, restando el sistema original,

$$\mathbf{A}\delta\mathbf{x} = \delta\mathbf{b}. \quad (2.171)$$

Resolviendo para $\delta\mathbf{x}$ tenemos

$$\delta\mathbf{x} = \mathbf{A}^{-1}\delta\mathbf{b}. \quad (2.172)$$

Aquí podemos estimar nuevamente con las normas y haciendo uso de la desigualdad de Schwarz

$$\|\delta\mathbf{x}\| \leq \|\mathbf{A}^{-1}\| \|\delta\mathbf{b}\|. \quad (2.173)$$

Multiplicando (2.169) y (2.173) obtenemos

$$\|\mathbf{b}\| \|\delta \mathbf{x}\| \leq \|\mathbf{A}\| \|\mathbf{x}\| \|\mathbf{A}^{-1}\| \|\delta \mathbf{b}\|. \quad (2.174)$$

Por último dividimos la última desigualdad por $\|\mathbf{b}\| \|\mathbf{x}\|$ lo cual da como resultado

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad (2.175)$$

de donde el *número condicionante* queda definido por

$$C(\mathbf{A}) := \|\mathbf{A}\| \|\mathbf{A}^{-1}\|. \quad (2.176)$$

Entonces podemos retomar esta definición y escribir (2.175) como

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq C(\mathbf{A}) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}. \quad (2.177)$$

Usando la *norma espectral* de A , $\|A\|_2 = \rho(A^*A)$, para el cálculo del número condicionante se puede encontrar

$$C(\mathbf{A}) = \frac{\lambda_{max}}{\lambda_{min}}, \quad (2.178)$$

donde el λ_{max} es el valor máximo de los eigenvalores de A y λ_{min} el menor (ver [10, 11]).

2.3.6. Numero condicionante del caso elíptico

Para estimar el *número condicionante* para la matriz del problema elíptico, calculemos sus eigenvalores. Recordemos que la matriz tridiagonal A está descrita de la siguiente forma:

$$A_{l,l} = \frac{2}{3}\gamma(x_l)\rho + \frac{2}{\rho} + O(\rho^3), \quad l = 1, \dots, n, \quad (2.179)$$

$$A_{l,l+1} = A_{l+1,l} = \gamma(x_l + \rho/2)\frac{\rho}{6} - \frac{1}{\rho} + O(\rho^3), \quad l = 1, \dots, n-1, \quad (2.180)$$

$$A_{i,j} = 0, \text{ para todos los demas casos.} \quad (2.181)$$

Donde $\rho = 1/(n+1)$ es el tamaño de la discretización del intervalo. Definimos la resolución de la discretización como

$$r = n + 1. \quad (2.182)$$

Entonces, cuando $r \rightarrow \infty$ tenemos que $\rho \rightarrow 0$, por lo tanto las componentes de la matriz se comportan de la siguiente forma:

$$A_{l,l} \rightarrow \frac{2}{\rho}, l = 1, \dots, n, \quad (2.183)$$

$$A_{l,l+1} = A_{l+1,l} \rightarrow -\frac{1}{\rho}, l = 1, \dots, n-1, \quad (2.184)$$

$$A_{i,j} = 0, \text{ para todos los demas casos.} \quad (2.185)$$

Para calcular los eigenvalores de esta matriz, resolvemos la ecuación

$$|A - \lambda I| = 0. \quad (2.186)$$

Éste problema es análogo a resolver el determinante de la matriz tridiagonal $T = \frac{1}{\rho}M$, donde M está dada por:

$$M_{l,l} = L, l = 1, \dots, n, \quad (2.187)$$

$$M_{l,l+1} = M_{l+1,l} = -1, l = 1, \dots, n-1, \quad (2.188)$$

$$M_{i,j} = 0, \text{ para todos los demas casos,} \quad (2.189)$$

donde $L = 2 - \lambda\rho$.

Resolver (2.186) es lo mismo que calcular el determinante de M e igualarlo a cero

$$|M| = 0. \quad (2.190)$$

Lo cual da como resultado (ver [11])

$$L_j = 2 \cos(j\theta), \quad (2.191)$$

donde $\theta = \frac{\pi j}{n+1}$, con $j = 1, \dots, n-1$.

Sustituyendo $L = 2 - \lambda\rho$ encontramos

$$\lambda_j = \frac{2}{\rho} [1 - \cos(j\theta)]. \quad (2.192)$$

Por tal razón el número condicionante

$$C(\mathbf{A}) = \frac{\lambda_{max}}{\lambda_{min}}, \quad (2.193)$$

conforme la resolución va aumentando, es decir, $\rho \rightarrow 0$

$$C(\mathbf{A}) \rightarrow \infty, \quad (2.194)$$

Ya que el $\lambda_{max} \rightarrow 1$ y $\lambda_{min} \rightarrow 0$.

Entonces para grandes resoluciones, es decir, cuando $r \rightarrow \infty$ tenemos que el *número condicionante* tiende a infinito. Por lo tanto, las perturbaciones sobre $\frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$ tienen un gran efecto sobre $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}$, de acuerdo con (2.177).

2.3.7. Ejemplo de la Implementación Numérica

Con el fin de mostrar que el método numérico basado en *Elementos Finitos* funciona de acuerdo a la teoría, a continuación planteamos un problema elíptico, y buscamos la aproximación de la solución. En ésta sección mostramos el factor de convergencia así como las estimaciones sobre la norma $L^2(0, 1)$ hechas para el error, así como para la derivada del error.

Tomemos entonces el problema elíptico descrito por:

$$\left[-\frac{d^2}{dx^2} + (2 + x^2) \right] u(x) = F(x), \quad (2.195)$$

imponiendo las condiciones de frontera tipo Dirichlet

$$u(0) = u(1) = 0, \quad (2.196)$$

donde

$$\gamma(x) \equiv 2 + x^2, \gamma(x) > 0 \forall x \in [0, 1], \quad (2.197)$$

y $F(x)$ dada por

$$F(x) \equiv x(x - \frac{1}{2})(x - 1) \left[(x - 3)^2 - \frac{11}{2} \right]. \quad (2.198)$$

Para este problema sabemos que la solución es

$$u(x) = x(x - \frac{1}{2})(x - 1), \forall x \in [0, 1]. \quad (2.199)$$

Para implementar el Método de Elementos Finitos y estar habilitados a construir la matriz A , y el vector J , tomamos una partición de intervalo $I = [0, 1]$ uniforme en r fragmentos de la misma longitud. De aquí obtenemos que $r = n + 1$, donde n es el número de elementos finitos que se tendrán, y por lo tanto la dimensión del subespacio de H_o^1 , es de dimensión n ($\dim X_n = n$).

Es fácil calcular $\|F\|_{L^2(0,1)}$,

$$\|F\|_{L^2(0,1)}^2 = \int_0^1 |F(x)|^2 dx = 367/110880, \quad (2.200)$$

por lo tanto obedece a $\|F\|_{L^2(0,1)} < C$, donde $C \in \mathbb{R}$.

Con todas los ingredientes disponibles podemos implementar el método numérico, y aproximar una solución para el problema elíptico antes descrito. Y aprovechando que conocemos la solución exacta podemos monitorear el error, así como el error residual para ver si el método de Thomas con el cual resolveremos el sistema $Ax = J$ está trabajando de forma adecuada.

Entonces para la implementación numérica, hemos discretizado de forma uniforme el intervalo $[0, 1]$ dividiéndolo en $r = n + 1$ partes iguales. Llamamos a r , la resolución de la discretización. Tenemos entonces que $\rho = 1/r$ y cada punto de la discretización es $x_j = j/r$, donde $j = 0, \dots, n + 1$.

Se pudo comprobar, que el error de redondeo se acumula cuando la resolución comienza a crecer. Para ver esto se calcularon la norma discreta $l_{\Delta x}^2$ la cual esta definida mediante

$$\|u\|_{l_{\Delta x}^2} = \sqrt{\sum_{l=1}^n |u_k|^2 \Delta x}, \quad (2.201)$$

aquí Δx es el tamaño de la discretización, en nuestra implementación numérica es $\Delta x = \rho$.

De esta forma, mediante el método numérico se obtiene una solución numérica u^r a una resolución r . El error se calcula usando la solución exacta y diferenciando con respecto a la solución obtenida mediante el método numérico.

$$e_i^r = u_i^r - u(x_i), \quad (2.202)$$

donde $u(x_i)$ con $i = 0, 1, 2 \dots n, n + 1$, es la solución exacta en cada uno de los puntos de la discretización del intervalo $[0, 1]$.

Calculamos la norma $l_{\Delta x}^2$ del error e^r para diferentes resoluciones. Haciendo referencia a la *Figura (2.4)* podemos observar que el error decrece de forma constante, pero después de cierto punto el error comienza a crecer nuevamente. Ésto es el error de redondeo y coincide con la estimación para el número condicional el cual crece cuando $\rho \rightarrow 0$, y por lo tanto el método numérico se vuelve más sensible a los errores de redondeo.

En la *Figura (2.4)* se muestran las gráficas correspondientes a datos en el código numérico del tipo *float* y del tipo *double* encontrando que para el dato tipo *float* la norma del error empieza a crecer para $1,0e - 17$, mientras que para *double* el error se comienza a acumular de forma notable hasta $1,0e - 11$.

Usando dos resoluciones, podemos calcular el parámetro de convergencia punto a punto $conv_i$, que está dado por el cociente

$$conv_i = \frac{e_i^r}{e_i^{2r}}. \quad (2.203)$$

Para que nuestro método sea consistente con la teoría desarrollada en la sección previa, este parámetro en cada punto debe ser de orden de 4. A continuación mostramos algunas gráficas, ver *Figura (2.5)*, donde se observa que satisface satisfactoriamente dicha condición. En esta figura también se puede apreciar que al rededor del cero de la solución exacta, en $x = 0,5$, existe una dispersión en los datos. Podemos graficar el error para una resolución r y para el doble de la res-

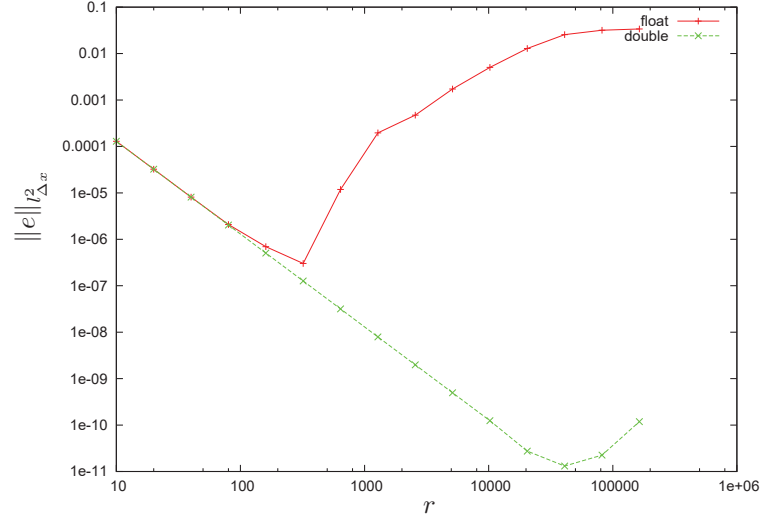


Figura 2.4: En esta gráfica mostramos el error de redondeo que se acumula para altas resoluciones. Mostramos dos gráficas una para tipo *float* donde se muestra que el error se acumula considerablemente para ordenes $1,0e - 07$, mientras que para *double* se observa para $1,0e - 11$.

olución. En la *Figura (2.6)* se observar dos errores y vemos que si aumentamos la resolución al doble, existe un factor 4 entre ambos.

Dependencia del error con la resolución.

A continuación mostramos los resultados gráficos que se obtienen de los datos adquiridos mediante el método numérico. Según la teoría que se tiene en la sección precedente, tenemos que

$$\|e^r\| \sim h^2. \quad (2.204)$$

Se calcula para varias resoluciones la norma l^2 discreta y se grafica contra r , en una escala $\log - \log$. Entonces botenemos una línea recta con pendiente aproximadamente igual a -2. Lo cual concuerda con la teoría.

Los resultados numéricos se pueden observar en la *Figura (2.7)*, donde se ha graficado en escala $\log - \log$ la norma discreta $\|e\|_{l^2_{\Delta x}}$ contra la resolución r .

Dependencia de la derivada del error con la resolución.

También se hace el cálculo para la derivada discreta, mediante la expresión para $u \in H^1_o$ ya que en un subespacio de dimensión finita u se puede representar mediante

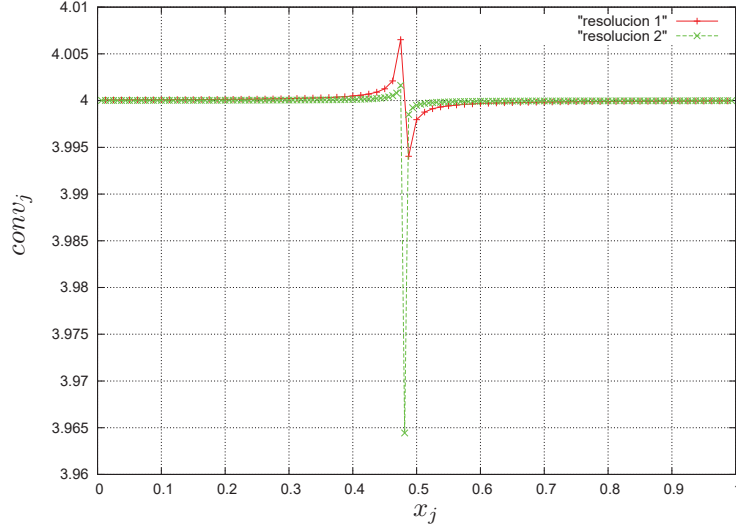


Figura 2.5: En esta grafica mostramos, que para diferentes resoluciones el parámetro de convergencia se apegas de forma muy cercana a 4,0. En la figura se muestras dos gráficas para la convergencia, podemos observar como al rededor de $x = 1/2$, donde la solución exacta se hace cero, la covergencia muestra una ligera dispersión de los datos obtenidos. Sin embargo, el factor de convergencia se acerca a 4,0 cuando aumenta la resolución.

$$u(x) = \sum_{l=0}^{l=r} u^k N_k(x), \quad (2.205)$$

asi la derivada viene dadada por

$$u'(x) = \sum_{l=0}^{l=r} u^k N'_k(x), \quad (2.206)$$

de donde toma la forma explícita siguiente para una distribución uniforme:

$$u'(x) = u_j \left(-\frac{1}{\rho}\right) + u_{j+1} \left(\frac{1}{\rho}\right) = \frac{u_{j+1} - u_j}{\rho}, \quad (2.207)$$

esta ecuación es válida para $x \in I_{j,j+1}$, donde $j = 0, 1, \dots, n, n+1 = r$.

Así pues, con esta aproximación para la derivada, se compara con la derivada de la solución exacta, en los puntos medios entre cada intervalo de discretización del intervalo $(0, 1)$, es decir, $u'(x_j + \rho/2)$.

Al observar la gráfica en la *Figura (2.8)* tenemos que la pendiente de la gráfica es aproximadamente igual -2 . Este resultado se encuentra dentro de la cota

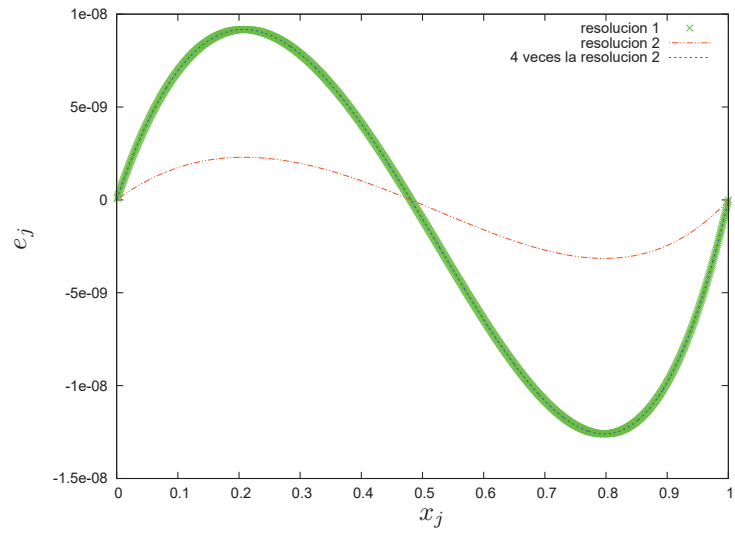


Figura 2.6: Se muestran los errores para dos resoluciones. La *resolución 2* es el doble de la *resolución 1*, entre el error a *resolución 1* y el error a *resolución 2* existe un factor de 4.

impuesta por la teoría, que predice una recta con pendiente, a lo más, igual -1 .

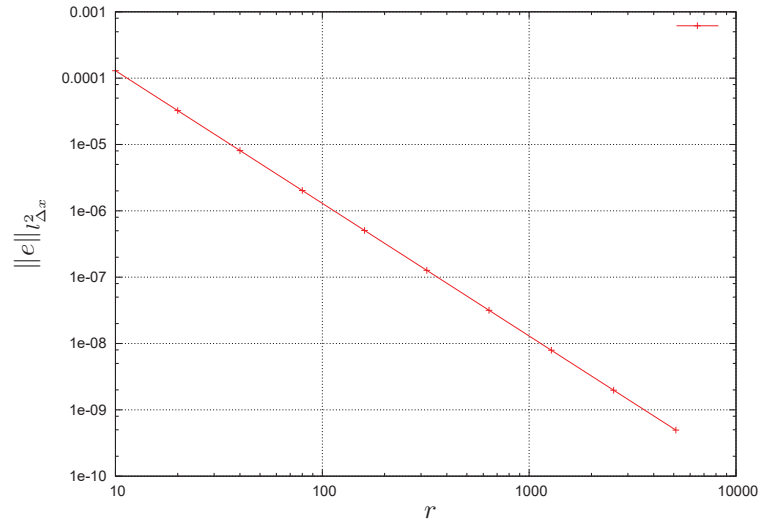


Figura 2.7: Aquí podemos observar que la grafica concuerda con la teoría, ya que forma una recta a simple vista perfecta, cuya pendiente se aproxima a $-2,0$

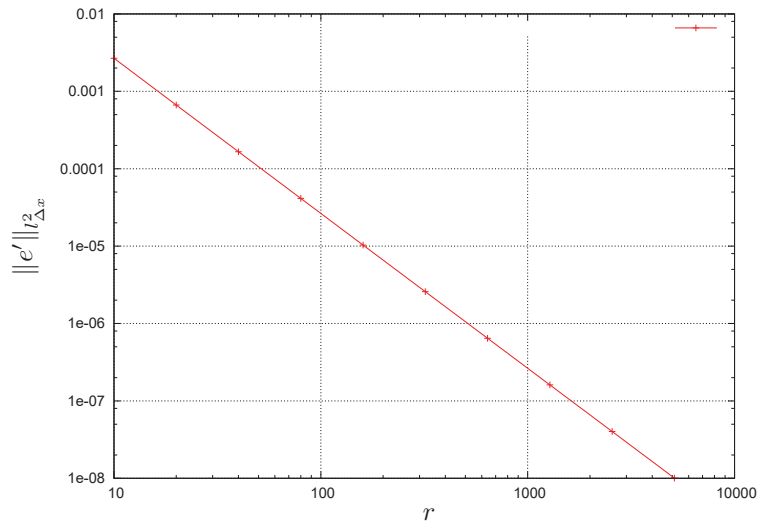


Figura 2.8: En este caso podemos observar que la pendiente de la recta para la derivada es de nuevo aproximadamente $-2,0$.

Capítulo 3

Problema Hiperbólico

En este capítulo trataremos un problema hiperbólico tipo onda en una dimensión espacial. Una vez planteado el problema con sus respectivas condiciones iniciales y condiciones de frontera, que serán del tipo Dirichlet, procedemos a la formulación débil del problema.

Mostramos mediante la definición de la Energía del Sistema, la unicidad y la dependencia continua respecto del dato que tiene que ver con la fuente del problema hiperbólico y los datos iniciales.

Procedemos a discretizar la parte espacial del problema débil y logramos mostrar que el error está controlado usando algunas estimaciones del **Capítulo 2**. Para integrar la parte referente al tiempo, usamos un método implícito [7]. Nos ocupamos de mostrar la consistencia y estabilidad del método.

3.1. Formulación débil del problema.

Definición: 3.1.1 (Problema Hiperbólico) Sea $I \equiv (0, 1) \subset \mathbb{R}$. Consideremos ahora el caso en que tenemos una ecuación diferencial parcial de segundo orden, del tipo hiperbólico definido de la siguiente manera

$$\left[\frac{\partial}{\partial t^2} - \frac{\partial}{\partial x^2} + V(x) \right] u(t, x) = F(x), \quad t > 0, \quad 0 < x < 1, \quad (3.1)$$

donde la función $V \in C[0, 1]$, tal que $0 < m < V(x) < M$ para todo $x \in I$.

Consideremos las siguientes condiciones para u :

- $u(t, 0) = u(t, 1) = 0$, para todo $t \geq 0$, las cuales son condiciones de frontera tipo Dirichlet,
- $u(0, x) = u_o(x)$, $\dot{u}(0, x) = v_o(x)$, con $x \in I$, como condiciones iniciales.

Entonces buscaremos ahora la formulación débil, al igual que en el caso elíptico, multiplicamos por una función de prueba $\phi \in C_o^\infty(I)$ y después integramos en el intervalo de interés, donde después de eliminar los términos de frontera tenemos

$$\begin{aligned} \int_0^1 \ddot{u}(t, x) \phi(x) dx + \int_0^1 u'(t, x) \phi'(x) dx + \\ \int_0^1 V(x) u(t, x) \phi(x) dx = \int_0^1 F(x) \phi(x) dx. \end{aligned} \quad (3.2)$$

Con ayuda de esta formulación débil del problema, podemos definir lo siguiente:

Definición: 3.1.2 ($Q_V(u, v)$) *El producto escalar $Q_V(u, v)$ tal que $\forall u, v \in H_o^1$*

$$Q_V(u, v) = \int_0^1 u' v' dx + \int_0^1 V(x) u v dx. \quad (3.3)$$

Definición: 3.1.3 ($L_F(v)$) *El funcional lineal $L_F(v)$ tal que $\forall v \in H_o^1$*

$$L_F(v) = \int_0^1 F v dx. \quad (3.4)$$

Y por lo tanto la formulación débil del problema hiperbólico puede ser ahora escrita en términos de $Q_V(\cdot, \cdot)$ y $L_F(\cdot)$

Definición: 3.1.4 (Problema Débil Hiperbólico) *El problema débil hiperbólico consiste en encontrar una función*

$$u : [0, \infty) \rightarrow H_o^1, \quad (3.5)$$

con $u \in C^1([0, \infty), H_o^1)$, $\dot{u} \in C^1([0, \infty), L^2(I))$ tal que se cumpla

$$\boxed{(\ddot{u}(t), \phi)_{L^2} + Q_V(u(t), \phi) = L_F(\phi), \forall t \geq 0, \forall \phi \in C_o^\infty(I),} \quad (3.6)$$

y además $u(0) = u_o \in H_o^1$, y $\dot{u}(0) = v_o \in L^2(I)$.

La existencia de soluciones débiles $u : [0, \infty) \rightarrow H_o^1$ se pueden demostrar usando la teoría de semigrupos (ver, por ejemplo, [12]) o usando el método de Galerkin (ver, por ejemplo, [13]).

3.2. Estimación de Energía Apriori

Definimos ahora la energía total del sistema

$$E(t) = \frac{1}{2} \int_0^1 [\dot{u}^2 + u'^2 + V(x)u^2] dx, \quad (3.7)$$

donde podemos apreciar que es igual a

$$E(t) := \frac{1}{2} [(\dot{u}(t), \dot{u}(t))_{L^2} + Q_V(u(t), u(t))]. \quad (3.8)$$

Sea $u(t)$ una solución débil del problema hiperbólico de la definición (3.1.4). Entonces tenemos

$$\frac{\partial E(t)}{\partial t} = (\ddot{u}(t), \dot{u}(t))_{L^2} + Q_V(u(t), \dot{u}(t)) = L_F(\dot{u}(t)), \quad (3.9)$$

donde en la última desigualdad hemos usado (3.6)

Usando la desigualdad de *Cauchy – Schwarz*, en $L_F(\dot{u}(t)) = \int_0^1 \dot{u}(t)F(x)dx$, y usando que $k \cdot \frac{1}{k} = 1$, $k \neq 0$, tenemos que

$$\frac{\partial}{\partial t} E(t) \leq k \|\dot{u}\|_{L^2} \frac{1}{k} \|F(t)\|_{L^2}. \quad (3.10)$$

Usamos ahora que $(a - b)^2 = a^2 + b^2 - 2ab \geq 0$, obteniendo así

$$\frac{\partial}{\partial t} E(t) \leq \frac{k^2}{2} \|\dot{u}\|_{L^2}^2 + \frac{1}{2k^2} \|F(t)\|_{L^2}^2 \leq k^2 E(t) + \frac{1}{2k^2} \|F(t)\|_{L^2}^2, \quad (3.11)$$

siempre y cuando en la última desigualdad se tome en cuenta la definición de $E(t)$.

Puesto que k es un número cualquiera, $k \neq 0$, podemos definirlo como $k^2 = \frac{1}{\tau}$, entonces obtenemos que

$$\dot{E}(t) - \tau^{-1} E(t) \leq \frac{\tau}{2} \|F(t)\|_{L^2}^2, \quad (3.12)$$

de donde podemos expresar esta desigualdad de la siguiente forma

$$\frac{d}{dt} [e^{-t/\tau} E(t)] \leq \frac{\tau}{2} e^{-t/\tau} \|F(t)\|_{L^2}^2, \quad (3.13)$$

y finalmente integrando de 0 hasta un tiempo $T > 0$, obtenemos la condición

$$\boxed{E(T) \leq e^{T/\tau} \left[E(0) + \frac{\tau}{2} \int_0^T e^{-t/\tau} \|F(t)\|_{L^2}^2 dt \right]}. \quad (3.14)$$

A ésta condición se le conoce como *Estimación de Energía Apriori*. Y es de suma importancia para nuestro problema ya que nos da una estimación del error de una solución aproximada con respecto a una solución exacta:

Supongamos que $v(t) \in H_o^1$ es solución al problema débil que corresponde a la ecuación diferencial

$$\ddot{v} - v'' + V(x)v = G(t, x), \quad (3.15)$$

y que $u(t) \in H_o^1$ es la solución exacta al problema débil que corresponde a la ecuación diferencial

$$\ddot{u} - u'' + V(x)u = F(t, x). \quad (3.16)$$

Supongamos que $F(t, x) - G(t, x) = \epsilon(t, x)$ y tomemos la diferencia de ambas ecuaciones diferenciales, de donde obtenemos

$$\ddot{e} - e'' + V(x)e = \epsilon(t, x), \quad (3.17)$$

habiendo definido la función $e(t) \in H_o^1$ como $e(t) = u(t) - v(t) \in H_o^1$.

Aplicando la estimación de (3.14) obtenemos que

$$E^e(T) \leq e^{T/\tau} \left[E^e(0) + \frac{\tau}{2} \int_0^T e^{-\frac{t}{\tau}} \|\epsilon(t)\|_{L^2}^2 dt \right], \quad (3.18)$$

donde $E^e(T)$ es la energía correspondiente al error $e(t) \in H_o^1$.

Como podemos ver $E^e(0)$ se puede calcular con las condiciones iniciales impuestas en el problema. Esta estimación (3.18) implica que para cada $t \in [0, T]$, $u(t)$ converge a $v(t)$ si $u(0)$ converge a $v(0)$, $\dot{u}(0)$ converge a $\dot{v}(0)$ y si ϵ converge a cero, en el sentido de las normas definidas por las expresiones anteriores. Entonces, la solución depende de manera continua de los datos. En particular, si $\epsilon(t) = 0$ para todo $t \in [0, T]$ y $u(0) = v(0)$, $\dot{u}(0) = \dot{v}(0)$, entonces $E^e(0) = 0$ y obtenemos

$$E^e(T) \leq 0. \quad (3.19)$$

Pero sabemos que $E(T) \geq 0$, por lo tanto $E^e(T) = 0$ para todo $T \geq 0$, lo que implica que $u(t) = v(t)$. Entonces tenemos unicidad de las soluciones.

3.3. El problema semidiscreto

Sea $X_n \subset X$ un subespacio de dimensión finita del espacio de Hilbert $X = H_o^1$. El problema débil discreto para el caso hiperbólico consiste en encontrar $u_n \in C^2([0, \infty), X_n)$, tal que

$$(\phi_n, \ddot{u}_n(t))_{L^2} + Q_V(\phi_n, u_n(t)) = L_F(\phi_n), \quad \forall \phi_n \in X_n. \quad (3.20)$$

Y tal que $u_n(0) = u_{on} \in X_n$, $v_n(0) = v_{on} \in X_n$, donde u_{on} y v_{on} aproximan el dato inicial u y v , respectivamente.

Para el caso continuo teníamos

$$(\phi, \ddot{u}(t))_{L^2} + Q_V(\phi, u(t)) = L_F(\phi), \forall \phi \in H_o^1. \quad (3.21)$$

En particular para $\phi = \phi_n \in X_n \subset X$, obtenemos

$$(\phi_n, \ddot{u}(t))_{L^2} + Q_V(\phi_n, u(t)) = L_F(\phi_n). \quad (3.22)$$

Tomando la diferencia entre (3.20) y (3.22) obtenemos

$$(\phi_n, \ddot{e}_n(t))_{L^2} + Q_V(\phi_n, e_n(t)) = 0, \forall \phi_n \in X_n, \quad (3.23)$$

donde se define el *error*

$$e_n(t) = u(t) - u_n(t). \quad (3.24)$$

Podemos tomar $e_n(t)$ como la suma de su parte paralela y ortogonal a X_n

$$e_n(t) = e_n^{\parallel}(t) + e_n^{\perp}(t), \quad (3.25)$$

de acuerdo con el producto definido por Q_V ,

$$Q_V(e_n^{\perp}(t), e_n^{\parallel}(t)) = 0. \quad (3.26)$$

Haciendo uso de ésta propiedad de ortogonalidad, podemos observar que

$$Q_V(\phi_n, e_n(t)) = Q_V(\phi_n, e_n^{\perp}(t) + e_n^{\parallel}(t)) = Q_V(\phi_n, e_n^{\parallel}(t)), \quad (3.27)$$

entonces la ecuación (3.23) da como resultado

$$(\phi_n, \ddot{e}_n(t))_{L^2} + Q_V(\phi_n, e_n^{\parallel}(t)) = 0, \forall \phi_n \in X_n. \quad (3.28)$$

Puesto que $\phi_n \in X_n$ podemos tomar a $\phi_n = \dot{e}_n^{\parallel}$ y por tanto la ecuación anterior resulta

$$(\dot{e}_n^{\parallel}, \ddot{e}_n(t))_{L^2} + Q_V(\dot{e}_n^{\parallel}, e_n^{\parallel}(t)) = 0. \quad (3.29)$$

Ahora bien, inspirada por la Estimación de Energía Apriori (3.14), podemos definir la siguiente cantidad que controla la parte paralela del error e_n^{\parallel}

$$\epsilon_n(t) := \frac{1}{2} \left[(\dot{e}_n^{\parallel}(t), \dot{e}_n^{\parallel}(t))_{L^2} + Q_V(e_n^{\parallel}(t), e_n^{\parallel}(t)) \right], \quad (3.30)$$

y derivando con respecto de t tenemos que

$$\frac{d\epsilon_n(t)}{dt} = (\dot{e}_n^{\parallel}(t), \ddot{e}_n^{\parallel}(t))_{L^2} + Q_V(\dot{e}_n^{\parallel}(t), e_n^{\parallel}(t)) = (\dot{e}_n^{\parallel}(t), -\ddot{e}_n^{\perp}(t))_{L^2}, \quad (3.31)$$

donde se hace uso de la ecuación (3.29). Además, tomando en cuenta que

$$\|\dot{e}_n^{\parallel}\|_{L^2}^2 \leq 2\epsilon_n(t), \quad (3.32)$$

y haciendo uso de la desigualdad de *Cauchy-Schwarz* en (3.31), tenemos que

$$\frac{d\epsilon_n(t)}{dt} \leq \sqrt{2\epsilon_n(t)} \|\ddot{e}_n^{\perp}\|_{L^2} \leq \frac{\sqrt{2\epsilon_n(t)}}{m} \|\ddot{e}_n^{\perp}\|_{Q_V}. \quad (3.33)$$

Por último definamos

$$\frac{1}{2}N^2(t) = \epsilon_n(t), \quad (3.34)$$

entonces al derivar tenemos que

$$\frac{dN(t)}{dt} N(t) \leq \frac{N(t)}{m} \|\ddot{e}_n^{\perp}(t)\|_{Q_V}, \quad (3.35)$$

y eliminado $N(t)$ de ambos lados de la desigualdad e integrando de 0 hasta un cierto $\tau > 0$ tenemos que

$$N(\tau) - N(0) \leq \frac{1}{m} \int_0^{\tau} \|\ddot{e}_n^{\perp}(t)\|_{Q_V} dt. \quad (3.36)$$

o bien

$$N(\tau) \leq N(0) + \frac{1}{m} \int_0^{\tau} \|\ddot{e}_n^{\perp}(t)\|_{Q_V} dt. \quad (3.37)$$

Ahora tomemos la parte ortogonal del error e_n^{\perp} de la siguiente forma

$$e_n^{\perp} = (u - u_n)^{\perp} = u^{\perp} - u_n^{\perp} = u^{\perp}, \quad (3.38)$$

por lo tanto deducimos que

$$\ddot{e}_n^{\perp} = \ddot{u}^{\perp}. \quad (3.39)$$

La magnitud de la parte ortogonal del error depende solamente de la elección de $X_n \subset X = H_o^1$.

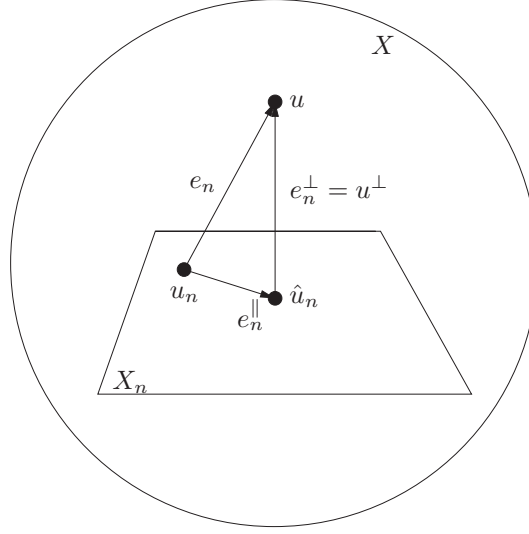


Figura 3.1: Figura donde se muestran la descomposición del error $e_n = e_n^\parallel + e_n^\perp$.

Ahora procedemos a estimar $\|\ddot{u}^\perp\|_{Q_V}$, usando los resultados del **Capítulo 2** tenemos

$$\|\ddot{u}^\perp\|_{Q_V} = \text{dist}_{Q_V}(\ddot{u}, X_n) \leq \frac{\bar{M}}{\bar{m}} \inf_{v_n \in X_n} \|\ddot{u} - v_n\|_{H_o^1}, \quad (3.40)$$

donde usando la aproximación por la interpolante de \ddot{u} tenemos

$$\|\ddot{u}^\perp\|_{Q_V} \leq Ch. \quad (3.41)$$

Entonces tenemos que la expresión para la estimación (3.37)

$$N(\tau) \leq N(0) + Ch\tau, \quad C \in \mathbb{R}. \quad (3.42)$$

Esto significa que si $h \rightarrow 0$ y el dato inicial discreto aproxima el dato inicial exacto, de tal manera que $N(0) \rightarrow 0$ cuando $h \rightarrow 0$, entonces $N(t) \rightarrow 0$, lo que implica que la norma $\epsilon_n(t)$ que controla la parte paralela del error, converge a cero. La parte ortogonal también converge a cero cuando $h \rightarrow 0$ como consecuencia de (3.39, 3.41).

3.4. Método Implícito para la discretización Temporal.

Ahora siguiendo el método de Galerkin, discretizamos el espacio H_o^1 en subespacios X_n de dimensión finita n . Y tomamos representantes de ϕ y de $u(t)$ en cada subespacio, $\phi_n \in X_n$ y $u_n(t) \in X_n$ respectivamente, para expandarlos en una base $\{N_k\}_{k=1,\dots,n}$ de X_n

$$\phi_n = \sum_{k=1}^n b_k N_k, \quad u_n(t) = \sum_{l=1}^n x_l(t) N_l, \quad (3.43)$$

donde podemos notar que la dependencia de t en la representación de $u_n(t)$ está presente solamente en los coeficientes $x_l(t)$.

Por esta razón, la segunda derivada temporal de $u_n(t)$ se ve reflejada solamente en los coeficientes $x_l(t)$ de la siguiente forma

$$\ddot{u}(t) = \sum_{l=1}^n \ddot{x}_l(t) N_l. \quad (3.44)$$

Por lo tanto al introducir (3.43) y (3.44) en (3.6) obtenemos

$$\sum_{l=0}^n (\ddot{x}_l(t) M_{lk} + x_l(t) A_{lk}) = J_k, \quad (3.45)$$

donde hemos definido las matrices M de tamaño $n \times n$, tal que

$$M_{lj} = (N_l, N_k)_{L^2}, \quad l, k = 1, 2, \dots, n. \quad (3.46)$$

La matriz A , queda especificada de forma análoga por $Q_V(N_l, N_k)$, mientras que el vector J queda determinado por

$$(J^T) = (L_F(N_1), L_F(N_2), \dots, L_F(N_n)). \quad (3.47)$$

Así, podemos expresar de forma matricial la misma condición

$$M\ddot{x} + Ax = J. \quad (3.48)$$

Definición: 3.4.1 (Matriz definida positiva) Sea M una matriz real simétrica de dimensión $n \times n$, entonces se dice que es definida positiva, si para todo vector de dimensión n , y distinto de cero $x \neq 0$, $x \in \mathbb{R}^n$ tenemos que

$$x^t M x > 0. \quad (3.49)$$

3.4. MÉTODO IMPLÍCITO PARA LA DISCRETIZACIÓN TEMPORAL. 57

Entonces podemos mostrar que M es una matriz definida positiva, ya que para todo $u \neq 0$, donde $u \equiv \sum_i x_i N_i$, podemos hacer la siguiente operación

$$u^t M u = \sum_{i,j} x_i M_{ij} x_j = \sum_{i,j} x_i (N_i, N_j)_{L^2} x_j = \left(\sum_i x_i N_i, \sum_j x_j N_j \right)_{L^2} = (u, u)_{L^2}, \quad (3.50)$$

y por lo tanto tenemos

$$u^t M u = \|u\|_{L^2}^2 > 0. \quad (3.51)$$

Usando que M es *simétrica* y *definida positiva*, sabemos que M es *invertible*. Por lo tanto obtenemos

$$\ddot{x} = -M^{-1} A x + M^{-1} J, \quad (3.52)$$

la cual se reduce a un sistema de ecuaciones diferenciales, que con el dato inicial de problema tenemos que existe una única solución para u .

El método que usaremos para resolver el problema hiperbólico será implícito, para esto recordemos que el sistema viene dado por

$$M \ddot{x} + A x = J \quad (3.53)$$

con las condiciones iniciales dadas por

$$x(0) = \xi_o, \dot{x}(0) = \eta_o. \quad (3.54)$$

Haciendo la transformación siguiente

$$\dot{x} - \eta = 0, \quad (3.55)$$

tenemos que el sistema original toma la siguiente forma

$$M \dot{\eta} + A x = J, \quad (3.56)$$

$$\dot{x} - \eta = 0, \quad (3.57)$$

con las condiciones iniciales

$$\dot{x}(0) = \xi_o, \eta(0) = \eta_o. \quad (3.58)$$

Ahora, discretizamos el intervalo de tiempo $[0, T]$, de tal forma que

$$0 = t_0 < t_1 < t_2 < \cdots < t_N = T, \quad (3.59)$$

de donde tenemos que

$$k_n = t_n - t_{n-1}, \quad n = 1, 2, \dots, N. \quad (3.60)$$

Empleamos el siguiente método implícito [7]

$$M \left(\frac{\eta^{n+1} - \eta^n}{k_n} \right) + \alpha A x^{n+1} + (1 - \alpha) A x^n = \alpha J^{n+1} + (1 - \alpha) J^n, \quad (3.61)$$

$$\left(\frac{x^{n+1} - x^n}{k_n} \right) - (\gamma \eta^{n+1} + (1 - \gamma) \eta^n) = 0, \quad (3.62)$$

donde $\alpha \geq 0$ y $\gamma \geq 0$, con las condiciones iniciales

$$x^o = \xi_o, \quad \eta^o = \eta_o. \quad (3.63)$$

Este método es incondicionalmente estable para $\alpha, \gamma \geq 1/2$, es decir, no existe un factor de courant que relacione el tamaño de la discretización temporal y la discretización espacial para que el método sea estable. Es de segundo orden de precisión si $\alpha = \gamma = 1/2$ [7], como demostramos a continuación.

3.4.1. Consistencia

Tomemos la siguiente definición de consistencia para un esquema de diferencias finitas.[14]:

Supongamos que el esquema de diferencias finitas para el operador diferencial $\mathcal{L}v = F$ está dado por $L_k^n u_k^n = G_k^n$, donde el subíndice k se refiere a la discretización espacial y n se refiere a la discretización temporal.

Definición: 3.4.2 *El esquema de diferencias finitas $L_k^n u_k^n = G_k^n$ es consistente puntalmente con la ecuación diferencial parcial $\mathcal{L}v = F$ en un punto (x, t) si para todas las funciones suaves $\phi = \phi(x, t)$,*

$$(\mathcal{L}\phi - F)_k^n - [L_k^n \phi(k\Delta x, n\Delta t) - G_k^n] \rightarrow 0 \quad (3.64)$$

conforme $\Delta x, \Delta t \rightarrow 0$ y $(k\Delta x, (n+1)\Delta t) \rightarrow (x, t)$.

Sea $\phi = v$, una solución para $\mathcal{L}v = F$, entonces al implementarla en (3.64), tenemos

$$\tau_k^n \equiv [L_k^n v(k\Delta x, n\Delta t) - G_k^n] \rightarrow 0. \quad (3.65)$$

A esta diferencia se le llama error de truncamiento.

Definición: 3.4.3 Se dice que el esquema de diferencias finitas $L_k^n u_k^n = G_k^n$, tiene orden de precisión (p, q) a una ecuación diferencial parcial $\mathcal{L}v = F$ si

$$\|\tau^n\| = O(\Delta x^p) + O(\Delta x^q). \quad (3.66)$$

Ya que el vector η^{n+1} es el vector η evaluado en el nivel de tiempo $n+1$, podemos expandir dicho vector alrededor del tiempo anterior,

$$\eta^{n+1} = \eta^n + (\dot{\eta})^n k_n + (\ddot{\eta})^n \frac{k_n^2}{2} + O(k_n^3), \quad (3.67)$$

y de aquí fácilmente podemos ver que

$$\frac{\eta^{n+1} - \eta^n}{k_n} = (\dot{\eta})^n + (\ddot{\eta})^n \frac{k_n}{2} + O(k_n^2). \quad (3.68)$$

Por otra parte, tenemos que

$$\begin{aligned} \alpha A x^{n+1} + (1 - \alpha) A x^n &= \alpha A (x^{n+1} - x^n) + A x^n = \\ &= \alpha A (\dot{x})^n k_n + \alpha A (\ddot{x})^n \frac{k_n^2}{2} + O(k_n^3) + A x^n. \end{aligned} \quad (3.69)$$

Escribiendo el esquema dado por la ecuación (3.61), tenemos

$$\begin{aligned} M \left[(\dot{\eta})^n + (\ddot{\eta})^n \frac{k_n}{2} + O(k_n^2) \right] + A x^n + \alpha A (\dot{x})^n k_n + O(k_n^2) &= \\ \alpha [J^{n+1} - J^n] + J^n &= \alpha k_n (\dot{J})^n + \frac{\alpha k_n^2}{2} (\ddot{J})^n + O(k_n^3) + J^n. \end{aligned} \quad (3.70)$$

Entonces el esquema se puede reordenar como

$$[M(\dot{\eta})^n + A x^n - J^n] + k_n \left[\frac{M}{2} (\ddot{\eta})^n + \alpha A (\dot{x})^n - \alpha (\dot{J})^n \right] = O(k_n^2). \quad (3.71)$$

De aquí notamos que si $\alpha = \frac{1}{2}$, tenemos

$$[M\dot{\eta} + Ax - J]^n + \frac{k_n}{2} \frac{\partial}{\partial t} [M\dot{\eta} + Ax - J]^n = O(k_n^2) \quad (3.72)$$

y usando la expresión para el esquema de diferencias finitas dado por (3.56), se sigue inmediatamente que el método es consistente, y además de orden de k_n^2 . Si usamos $\alpha \neq 1/2$ el sistema sigue siendo consistente pero a orden de precisión k_n , es decir, el sistema es más eficiente con la elección $\alpha = 1/2$.

Análogamente lo hacemos para

$$\frac{x^{n+1} - x^n}{k_n} = (\dot{x})^n + \frac{k_n}{2}(\ddot{x})^n + O(k_n^2), \quad (3.73)$$

por otra parte tenemos

$$\begin{aligned} \gamma\eta^{n+1} + (1 - \gamma)\eta^n &= \gamma(\eta^{n+1} - \eta^n) + \eta^n = \\ &= \gamma k_n(\dot{\eta})^n + O(k_n^2) + \eta^n, \end{aligned} \quad (3.74)$$

por lo tanto podemos reescribir el esquema (3.62) como

$$\begin{aligned} (\dot{x})^n + \frac{k_n}{2}(\ddot{x})^n + O(k_n^2) - \gamma k_n(\dot{\eta})^n + O(k_n^2) - \eta^n &= \\ (\dot{x})^n - \eta^n + k_n \left[\frac{1}{2}(\ddot{x})^n - \gamma(\dot{\eta})^n \right] + O(k_n^2) \end{aligned} \quad (3.75)$$

podemos usar en este punto la expresión (3.57), y la expresion anterior se reduce a

$$k_n \left[\frac{1}{2}(\ddot{x})^n - \gamma(\dot{\eta})^n \right] = O(k_n^2), \quad (3.76)$$

ahora usamos la misma expresión y obtenemos

$$k_n \left[\frac{1}{2}(\dot{\eta})^n - \gamma(\dot{\eta})^n \right] = O(k_n^2), \quad (3.77)$$

es inmediato que si $\gamma = 1/2$, el esquema es consistente y de orden k_n^2 . Nuevamente para otra elección de $\gamma \neq 1/2$ el método sigue siendo consistente pero a orden k^n .

3.4.2. Estabilidad y Convergencia

Partiendo del esquema de diferencias finitas propuesto en (3.61) y (3.62), multiplicamos la ecuación (3.61) por $(x^{n+1} - x^n)^t$ y la ecuación (3.62) por $(\eta^{n+1} - \eta^n)^t M$ y reduciendo a una ecuación tenemos:

$$\begin{aligned} &\gamma(\eta^{n+1} - \eta^n)^t M(\eta^{n+1}) + (1 - \gamma)(\eta^{n+1} - \eta^n)^t M\eta^n \\ &+ \alpha(x^{n+1} - x^n)^t Ax^{n+1} + (1 - \alpha)(x^{n+1} - x^n)^t Ax^n \\ &= \alpha(x^{n+1} - x^n)^t J^{n+1} + (1 - \alpha)(x^{n+1} - x^n)^t J^n. \end{aligned} \quad (3.78)$$

3.4. MÉTODO IMPLÍCITO PARA LA DISCRETIZACIÓN TEMPORAL. 61

Definiendo $\bar{\alpha} \equiv \alpha - 1/2$ y $\bar{\gamma} \equiv \gamma - 1/2$ podemos reescribir la ecuación anterior como

$$\begin{aligned} & \frac{1}{2}(\eta^{n+1} - \eta^n)^t M(\eta^{n+1} + \eta^n) + \bar{\gamma}(\eta^{n+1} - \eta^n)M(\eta^{n+1} - \eta^n) \\ & + \frac{1}{2}(x^{n+1} - x^n)^t A(x^{n+1} + x^n) + \bar{\alpha}(x^{n+1} - x^n)A(x^{n+1} - x^n) \\ & = \left(\frac{1}{2} + \bar{\alpha}\right)(x^{n+1} - x^n)^t J^{n+1} + \left(\frac{1}{2} - \bar{\alpha}\right)(x^{n+1} - x^n)^t J^n. \end{aligned} \quad (3.79)$$

Podemos notar que

$$(\eta^{n+1} - \eta^n)^t M(\eta^{n+1} + \eta^n) = (\eta^{n+1})^t M \eta^{n+1} - (\eta^n)^t M \eta^n, \quad (3.80)$$

y vemos que la expresión (3.79) se reduce a

$$\begin{aligned} & E^{n+1} - E^n + \bar{\gamma}\|\eta^{n+1} - \eta^n\|_M^2 + \bar{\alpha}\|x^{n+1} - x^n\|_A^2 = \\ & \left(\frac{1}{2} + \bar{\alpha}\right)(x^{n+1} - x^n)^t J^{n+1} + \left(\frac{1}{2} - \bar{\alpha}\right)(x^{n+1} - x^n)^t J^n, \end{aligned} \quad (3.81)$$

donde hemos definido

$$E^{n+1} = \frac{1}{2} [(\eta^{n+1})^t M \eta^{n+1} + (x^{n+1})^t A x^{n+1}], \quad (3.82)$$

y las normas que aparecen están definidas mediante

$$\|\eta^{n+1} - \eta^n\|_M^2 = (\eta^{n+1} - \eta^n, \eta^{n+1} - \eta^n)_M = (\eta^{n+1} - \eta^n)^t M(\eta^{n+1} - \eta^n), \quad (3.83)$$

$$\|x^{n+1} - x^n\|_A^2 = (x^{n+1} - x^n, x^{n+1} - x^n)_A = (x^{n+1} - x^n)^t A(x^{n+1} - x^n) \quad (3.84)$$

Sabiendo que M y A son positivas definidas, obtenemos de la ecuación (3.81) la siguiente desigualdad

$$E^{n+1} - E^n \leq \left(\frac{1}{2} + \bar{\alpha}\right)(x^{n+1} - x^n)^t J^{n+1} + \left(\frac{1}{2} - \bar{\alpha}\right)(x^{n+1} - x^n)^t J^n. \quad (3.85)$$

Tomando en cuenta la definición de E^{n+1} y que $(x^{n+1}, x^{n+1})_A \geq \mu_A \|x^{n+1}\|^2$, tal que $\mu_A \in \mathbb{R}$ es una constante $\mu_A > 0$ tenemos entonces que

$$E^{n+1} \geq \mu_A \|x^{n+1}\|^2, \quad (3.86)$$

de aquí se sigue inmediatamente que

$$\|x^{n+1}\|^2 \leq \frac{1}{\mu_A} E^{n+1}. \quad (3.87)$$

Por otra parte tenemos que

$$(x^{n+1})^t J^{n+1} \leq \|x^{n+1}\| \|J^{n+1}\|, \quad (3.88)$$

gracias a la desigualdad de *Cauchy-Schwarz*, aquí tomamos un $k \in \mathbb{R}$, tal que $k \neq 0$

$$\begin{aligned} x^{n+1} J^{n+1} &\leq \|x^{n+1}\| \|J^{n+1}\| \leq \frac{1}{k} \|x^{n+1}\| k \|J^{n+1}\| \leq \\ \frac{1}{2k^2} \|x^{n+1}\|^2 + \frac{k^2}{2} \|J^{n+1}\|^2 &\leq \frac{1}{2\mu_A k^2} E^{n+1} + \frac{k^2}{2} \|J^{n+1}\|^2, \end{aligned} \quad (3.89)$$

donde hemos usado en la última desigualdad a (3.87). Haciendo uso de éste último resultado tenemos que

$$(x^{n+1} - x^n)^t J^{n+1} \leq \frac{1}{2\mu_A k^2} [E^{n+1} + E^n] + k^2 \|J^{n+1}\|^2, \quad (3.90)$$

$$(x^{n+1} - x^n)^t J^n \leq \frac{1}{2\mu_A k^2} [E^{n+1} + E^n] + k^2 \|J^n\|^2, \quad (3.91)$$

usamos entonces éstas desigualdades en (3.85), para obtener

$$E^{n+1} (2\mu_A k^2 - 1) \leq E^n (2\mu_A k^2 + 1) + 2\mu_A k^4 (\alpha \|J^{n+1}\|^2 + (1 - \alpha) \|J^n\|^2) \quad (3.92)$$

donde imponemos la condición $(2\mu_A k^2 - 1) > 0$ y se sigue entonces que

$$E^{n+1} < \Lambda E^n + \Gamma (\alpha \|J^{n+1}\|^2 + (1 - \alpha) \|J^n\|^2), \quad (3.93)$$

donde hemos definido Λ y Γ de la siguiente forma

$$\Lambda = \frac{2\mu_A k^2 + 1}{2\mu_A k^2 - 1}, \quad (3.94)$$

$$\Gamma = \frac{2\mu_A k^4}{2\mu_A k^2 - 1}. \quad (3.95)$$

Finalmente, podemos usar de forma recursiva la desigualdad (3.93) y obtenemos que

$$E^{n+1} \leq \Lambda^{n+1} E^o + \sum_{j=0}^n \Lambda^j \Gamma (\alpha \|J^{n+1-j}\|^2 + (1-\alpha) \|J^{n-j}\|^2). \quad (3.96)$$

Ahora supongamos que el intervalo temporal $[0, T]$, en n partes iguales, por lo que $\Delta T = T/n$. Entonces por la estimación final tenemos que en un tiempo $t = n\Delta T$

$$E^n \leq \Lambda^n E^o + \sum_{j=0}^{n-1} \Lambda^j \Gamma (\alpha \|J^{n+1-j}\|^2 + (1-\alpha) \|J^{n-j}\|^2). \quad (3.97)$$

Inspirados en la expresión de *Energía Apriori*,

$$\Lambda^n = e^{T/\tau}, \quad (3.98)$$

de donde es fácil obtener que

$$\Lambda = e^{\frac{\Delta T}{\tau}}. \quad (3.99)$$

Implementando el valor de Λ en la estimación para E^n tenemos

$$E^n \leq e^{\frac{T}{\tau}} E^o + \sum_{j=0}^{n-1} e^{\frac{\Delta T j}{\tau}} \Gamma (\alpha \|J^{n+1-j}\|^2 + (1-\alpha) \|J^{n-j}\|^2). \quad (3.100)$$

Factorizamos $e^{\frac{T}{\tau}}$ y obtenemos

$$E^n \leq e^{\frac{T}{\tau}} \left[E^o + \sum_{j=0}^{n-1} e^{-\frac{1}{\tau}(T-\Delta T j)} \Gamma (\alpha \|J^{n+1-j}\|^2 + (1-\alpha) \|J^{n-j}\|^2) \right]. \quad (3.101)$$

Aquí podemos notar que

$$T - \Delta T j > 0, \forall j = 0, 1, \dots, n-1, \quad (3.102)$$

por lo que la exponencial $e^{-\frac{1}{\tau}(T-\Delta T j)}$ es negativa en su argumento. Esta sería una versión discreta para al *Energía Apriori* dada en (3.14). Esto es bueno, ya que con esta estimación logramos mostrar la unicidad de la solución así como la convergencia y dependencia continua del dato.

Ahora definamos, en base a la solución exacta $(x(t), \eta(t))$ al sistema (3.56, 3.57) y la solución numérica (x^n, η^n) dada por el esquema (3.61, 3.62), los errores siguientes

$$\Delta\eta^n \equiv \eta^n - \eta(t^n), \quad (3.103)$$

$$\Delta x^n \equiv x^n - x(t^n), \quad (3.104)$$

donde $t^n = n\Delta t$.

Tomemos las condiciones iniciales que se usan para obtener la solución numérica

$$\eta^o = \eta(0) = \eta_o, \quad (3.105)$$

$$x^o = x(0) = x_o. \quad (3.106)$$

Y por lo tanto tenemos que

$$\Delta\eta^o \equiv \eta^o - \eta(t^o) = 0, \quad (3.107)$$

$$\Delta x^o \equiv x^o - x(t^o) = 0, \quad (3.108)$$

donde hemos usado las condiciones iniciales.

Ahora en base a la energía discreta para $(\Delta x, \Delta \eta)$, tenemos la siguiente energía discreta

$$E_\Delta^n = \frac{1}{2} [(\Delta\eta^n)^t M \Delta\eta^n + (\Delta x^n)^t A \Delta x^n]. \quad (3.109)$$

De donde tenemos claramente que

$$E_\Delta^o = 0. \quad (3.110)$$

Ahora, al implementar en (3.61) el siguiente cambio $\eta^n \rightarrow \Delta\eta^n$ y $x^n \rightarrow \Delta x^n$, tenemos

$$\begin{aligned} & \left[M \left(\frac{\eta^{n+1} - \eta^n}{k_n} \right) + \alpha A x^{n+1} + (1 - \alpha) A x^n \right] \\ & - \left[M \left(\frac{\eta(t^{n+1}) - \eta(t^n)}{k_n} \right) + \alpha A x(t^{n+1}) + (1 - \alpha) A x(t^n) \right] \\ & = \alpha (J^{n+1} - J(t^{n+1})) + (1 - \alpha) (J^n - J(t^n)). \end{aligned} \quad (3.111)$$

Por lo tanto tenemos que como J es dato, hacemos que $J^n \rightarrow J(t^n)$, cuando $\Delta t \rightarrow 0$ entonces

$$\Delta J^n \equiv J^n - J(t^n) \rightarrow 0. \quad (3.112)$$

De aquí, la estimación (3.101) se convierte en

$$E_{\Delta}^n \leq e^{\frac{T}{\tau}} \left[\sum_{j=0}^{n-1} e^{-\frac{1}{\tau}(T-\Delta T j)} \Gamma \left(\alpha \|\Delta J^{n+1-j}\|^2 + (1-\alpha) \|\Delta J^{n-j}\|^2 \right) \right], \quad (3.113)$$

y cuando $\Delta t \rightarrow 0$ tenemos

$$E_{\Delta}^n \leq 0. \quad (3.114)$$

Por lo tanto $E_{\Delta}^n = 0$. Y entonces $\Delta \eta^n \rightarrow 0$ y $\Delta x^n \rightarrow 0$ cuando $\Delta t \rightarrow 0$.

3.4.3. Solución al método implícito

Ahora pasaremos a resolver el método implícito, con el esquema de diferencias finitas mencionado con anterioridad en las ecuaciones (3.61) y (3.62). Así, a partir de estas relaciones obtenemos

$$\alpha k_n A x^{n+1} + A \eta^{n+1} - \alpha k_n J^{n+1} = -k_n(1-\alpha) A x^n + M \eta^n + k_n(1-\alpha) J^n \quad (3.115)$$

$$x^{n+1} - k_n \gamma \eta^{n+1} = x^n + k_n(1-\gamma) \eta^n \quad (3.116)$$

Podemos expresar estas relaciones de una forma más compacta, en forma matricial

$$\begin{pmatrix} \alpha k_n A & M \\ I & -k_n \gamma I \end{pmatrix} \begin{pmatrix} x^{n+1} \\ \eta^{n+1} \end{pmatrix} - \begin{pmatrix} \alpha k_n J^{n+1} \\ 0 \end{pmatrix} = \begin{pmatrix} -k_n(1-\alpha) A & M \\ I & k_n(1-\gamma) I \end{pmatrix} \begin{pmatrix} x^n \\ \eta^n \end{pmatrix} + \begin{pmatrix} k_n(1-\alpha) J^n \\ 0 \end{pmatrix}. \quad (3.117)$$

Es necesario invertir la matriz que multiplica al vector del futuro de la izquierda,

$$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} \begin{pmatrix} \alpha k_n A & M \\ I & -k_n \gamma I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \quad (3.118)$$

de donde obtenemos un sistema de cuatro ecuaciones lineales de matrices

$$\alpha k_n X A + Y = I \quad (3.119)$$

$$\alpha k_n Z A + W = 0 \quad (3.120)$$

$$X M - k_n \gamma Y = 0 \quad (3.121)$$

$$Z M - k_n \gamma W = I \quad (3.122)$$

y al resolver obtenemos que

$$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \begin{pmatrix} k_n \gamma L & LM \\ L & -\alpha k_n LA \end{pmatrix} \quad (3.123)$$

de tal manera que

$$L \equiv (\alpha \gamma k_n^2 A + M)^{-1}. \quad (3.124)$$

Entonces, al multiplicar por la matriz inversa tenemos que (3.117) se convierte en

$$\begin{pmatrix} x^{n+1} \\ \eta^{n+1} \end{pmatrix} = \begin{pmatrix} I - k_n^2 \gamma LA & k_n LM \\ -k_n LA & I - \alpha k_n^2 LA \end{pmatrix} \begin{pmatrix} x^n \\ \eta^n \end{pmatrix} + \begin{pmatrix} k_n^2 \gamma L(\alpha J^{n+1} + (1 - \alpha)J^n) \\ k_n L(\alpha J^{n+1} + (1 - \alpha)J^n) \end{pmatrix} \quad (3.125)$$

Ahora podemos multiplicar la expresión anterior por

$$\begin{pmatrix} L^{-1} & 0 \\ 0 & L^{-1} \end{pmatrix}. \quad (3.126)$$

Lo cual da como resultado el siguiente sistema

$$\begin{pmatrix} L^{-1} x^{n+1} \\ L^{-1} \eta^{n+1} \end{pmatrix} = \begin{pmatrix} M + (\alpha - 1)k_n^2 \gamma A & k_n M \\ -k_n A & M + (\gamma - 1)\alpha k_n^2 A \end{pmatrix} \begin{pmatrix} x^n \\ \eta^n \end{pmatrix} + \begin{pmatrix} k_n^2 \gamma (\alpha J^{n+1} + (1 - \alpha)J^n) \\ k_n (\alpha J^{n+1} + (1 - \alpha)J^n) \end{pmatrix} \quad (3.127)$$

Lo que hemos obtenido aquí es un sistema de ecuaciones del tipo

$$L^{-1}v = J, \quad (3.128)$$

y recordando la expresión para $L^{-1} = \alpha \gamma k_n^2 A + M$, es una combinación lineal de matrices simétricas y tridiagonales. Usamos que las matrices A y M son matrices simétricas y positivas definidas, podemos invertir el sistema de tal forma que

$$v = LJ. \quad (3.129)$$

Además tenemos la ventaja que las matrices son tridiagonales, por lo que podremos usar la eficiencia del *Método de Thomas*.

Capítulo 4

Conclusiones

Para el caso del problema elíptico en una dimensión, con condiciones de Dirichlet en la frontera, se formuló el problema débil abstracto. El cual nos permitió encontrar la solución débil al problema, en un espacio de Hilbert X , cuya existencia y unicidad está respaldada por el Lema de Riesz.

Pudimos ver la equivalencia entre el problema débil abstracto, y un problema variacional, que consiste en encontrar el mínimo de un funcional definido sobre X , al cual llamamos *Energía Potencial del Sistema*. Mostramos que si $u \in X$ es solución al problema débil abstracto, entonces también es un mínimo de la Energía Potencial del Sistema.

Logramos mostrar la dependencia continua del dato, además de una estimación que nos muestra la estabilidad de la solución.

Para aproximar la solución al problema débil, $u \in X$, tomamos un subespacio de dimensión finita $X_n \subset X$. Sobre éste subespacio formulamos el problema débil discreto y aproximamos a $u \in X$ mediante $u_n \in X_n$, y pudimos observar que el error $u - u_n$ es ortogonal a X_n .

Ahora, al implementar los *Elementos Finitos*, construimos $X_n \subset X$ tal que contiene funciones lineales por trozos y distintas de cero en un intervalo pequeño del dominio de definición de X_n . Tomamos una interpolante de la solución $u \in X$, $\hat{u}_n \in X_n$, y mostramos que el error $u - \hat{u}_n$ bajo la norma L^2 es del orden de $O(h^2)$, mientras que la derivada del error es del orden de $O(h)$. Con estas estimaciones pudimos mostrar que el error $u - u_n$ converge a cero con orden $O(h^2)$.

Una vez que nos encontramos trabajando sobre X_n , usamos los elementos de su base para generar a partir del problema débil un sistema de ecuaciones lineales de la forma $Ax = J$, y cuya solución $x \in \mathbb{R}^n$ se resuelve de forma numérica. Aquí A es una matriz tridiagonal, simétrica y positiva definida de dimensión $n \times n$, y J viene dado por el dato del problema elíptico.

Para resolver numéricamente este sistema se contruyó un código numérico a partir del Método de Thomas. Nos dimos cuenta que la solución de este tipo

de sistemas mediante la implementación de algoritmos derivados de la eliminación gaussiana, como es el caso del Método de Thomas, presentan un error de redondeo.

Además debe cuidarse que el sistema que se está tratando esté *bien condicionado* o *mal condicionado*. En nuestro caso logramos ver que al aumentar la resolución en la discretización espacial, lo cual hace aumentar la dimensión del sistema a resolver, el *número condicionante* que determina la condición del sistema también aumenta, lo que hace que nuestro sistema sea *mal condicionado* para resoluciones altas.

Sin embargo en el régimen donde el error de redondeo es despreciable y el sistema es bien condicionado, pudimos comprobar las estimaciones del error que se hicieron teóricamente bajo la norma $L^2(0, 1)$, tomando como una aproximación la norma discreta $l_{\Delta x}^2$, donde Δx es el tamaño de la discretización el cual en nuestro caso al hacer la implementación numérica se escogió uniforme, por ejemplo la norma $l_{\Delta x}^2$ del vector $u \in \mathbb{R}^{n-1}$

$$\|u\|_{l_{\Delta x}^2} = \sqrt{\sum_{l=1}^n |u_k|^2 \Delta x}.$$

Se comprobó que la convergencia del método es de orden $O(h^2)$, y que ésta mejora cuando la resolución de la discretización espacial aumenta (*cuidando que el sistema esté bien condicionado*).

Para el caso del problema hiperbólico, llevamos el problema original a una formulación débil tomando en cuenta la dimensión espacial. Definimos al Energía del Sistema, y ésta nos ayuda a demostrar la unicidad de la solución, así como la dependencia continua del dato y de las condiciones iniciales de problema.

Hicimos la semidiscretización del problema usando el Método de Galerkin. En este caso, el error ya no es ortogonal al subespacio de dimensión finita de X , X_n , que se ha elegido para la aproximación de la solución. Sin embargo, separando el error en una parte paralela y otra ortogonal al subespacio de dimensión finita. Se definió una norma que logra controlar la parte paralela de error, sin embargo, esta norma necesita de una cota para la parte ortogonal, usando algunos resultados para el caso elíptico se mostró la parte ortogonal del error converge a cero a orden $O(h)$.

Esto nos ayudó a completar la norma con la cual probamos de nuevo la dependencia continua del dato y la convergencia.

Después de discretizar la parte espacial, y elegir X_n de nuevo como en el caso elíptico, obtuvimos un sistema del tipo $M\ddot{x} + Ax = J$, donde M y A son matrices tridiagonales y además definidas positivas y simétricas, por lo cual el problema se convierte en encontrar la solución de un sistema $\ddot{x} = -M^{-1}Ax + J$, la cual es única dadas las condiciones iniciales del problema.

Para resolver este sistema, se usó un método implícito [7], para el cual se demostró la consistencia, estabilidad y convergencia, haciendo uso de una versión

discreta de la Energía del Sistema, tal y como se vió en el caso continuo.

Una vez mostradas las conclusiones de este trabajo, podemos decir, que el paso siguiente sería comprobar los resultados teóricos mediante la implementación numérica para el caso hiperbólico. Una vez hecho esto, se podría implementar al problema más complejo tal y como el descrito anteriormente para el sistema binario *agujero negro - estrella* en la Introducción de éste trabajo. Para este tipo de problemas ya existen implementaciones del los Elementos Finitos.

Apéndice A

Apéndices

En el *1er. Apéndice* se demuestra que H^1 , el conjunto de funciones tal que son cuadrado integrables así como sus derivadas en el sentido débil también lo son, es completo.

En el *2o. Apéndice* se demuestra el Lema de Sovolev (2.1.1), que se usa en el **Capítulo 1** y que nos asegura que para cada función en H^1 se puede encontrar un representante continuo.

En el *3er. Apéndice* se demuestra el lema (2.3.1), el cual se usa para obtener una aproximación de los elementos matriciales que serán usados en el **Capítulo 2** referente al Método de Galerkin aplicado al problema elíptico.

A.0.4. El Espacio de Hilbert H^1 .

Es la intención de este apéndice mostrar que

$$H^1 := \{u \in L^2[0, 1] | \text{existe } u' \in L^2[0, 1] \text{ en el sentido débil} \} \quad (\text{A.1})$$

es un espacio de Hilbert, es decir, que es un espacio vectorial completo con el producto definido por

$$(u, v)_{H^1} := (u, v)_{L^2} + (u', v')_{L^2}, \forall u, v \in H^1. \quad (\text{A.2})$$

Podemos definir la norma en H^1 mediante el producto interno

$$\|u\|_{H^1} \equiv (u, u)_{H^1}^{1/2} = (\|u\|_{L^2}^2 + \|u'\|_{L^2}^2)^{1/2}. \quad (\text{A.3})$$

Para mostrar que H^1 es completo tomemos una sucesión de Cauchy en H^1 , $\{u_n\}$. Por lo tanto dado $\epsilon > 0$, existe una $N \in \mathbb{Z}$ tal que $\|u_n - u_m\|_{H^1} < \epsilon$ siempre que $n, m > N$. Usando la norma en H^1 dada por (A.3) tenemos que

$$\|u_n - u_m\|_{H^1}^2 = \|u_n - u_m\|_{L^2}^2 + \|u'_n - u'_m\|_{L^2}^2, \quad (\text{A.4})$$

de donde se obtiene de forma inmediata que

$$\|u_n - u_m\|_{L^2}^2 + \|u'_n - u'_m\|_{L^2}^2 < \epsilon^2, \quad (\text{A.5})$$

y por tanto que

$$\|u_n - u_m\|_{L^2} < \epsilon, \quad (\text{A.6})$$

$$\|u'_n - u'_m\|_{L^2} < \epsilon. \quad (\text{A.7})$$

De aquí se sigue directamente que u_n y u'_n son sucesiones de Cauchy en L^2 . Como L^2 es completo, entonces

$$\begin{aligned} u_n &\rightarrow u \in L^2, \\ u'_n &\rightarrow v \in L^2, \end{aligned} \quad (\text{A.8})$$

en L^2 , cuando $n \rightarrow \infty$.

Nos resta mostrar que $u' = v$ en el sentido débil.

Para ésto, sea $\phi \in C_o^\infty(0, 1)$ y multiplicando por u'_n e integrando de cero a uno

$$\int_0^1 u'_n(x) \phi(x) dx = - \int_0^1 u_n(x) \phi'(x) dx. \quad (\text{A.9})$$

Tomemos la siguiente integral

$$\left| \int_0^1 [u'_n(x) - v(x)] \phi(x) dx \right| \leq \|u'_n - v\|_{L^2} \|\phi\|_{L^2}, \quad (\text{A.10})$$

donde hemos usado la desigualdad de Cauchy-Schwartz. De aquí podemos ver que como $u'_n \rightarrow v$ si $n \rightarrow \infty$ en L^2 , entonces

$$\left| \int_0^1 [u'_n(x) - v(x)] \phi(x) dx \right| \rightarrow 0, \text{ si } n \rightarrow \infty, \quad (\text{A.11})$$

de donde se sigue

$$\int_0^1 u'_n(x) \phi(x) dx \rightarrow \int_0^1 v(x) \phi(x) dx, \text{ si } n \rightarrow \infty. \quad (\text{A.12})$$

De la misma manera, se demuestra que

$$\int_0^1 u_n(x)\phi'(x)dx \rightarrow \int_0^1 u(x)\phi'(x)dx, \text{ si } n \rightarrow \infty, \quad (\text{A.13})$$

de donde finalmente obtenemos que

$$\int_0^1 v(x)\phi(x)dx = - \int_0^1 u\phi'(x)dx. \quad (\text{A.14})$$

Esto demuestra que u es diferenciable en el sentido débil y que $u' = v$.

A.0.5. Demostración del Lema (2.1.1)

Lema: A.0.1 Sean $I = (0, 1) \in \mathbb{R}$ y $u \in H^1(I)$, entonces u posee un representante $\tilde{u} \in C^0(\bar{I})$. Además sabemos que

$$\|\tilde{u}\|_\infty \leq \sqrt{2}\|u\|_{H^1}, \forall u \in H^1. \quad (\text{A.15})$$

Demostración.

Sea primero $u \in C^1(I) \cap H^1(I)$, definamos el promedio de u de la siguiente forma

$$\bar{u} \equiv \int_0^1 u(x)dx. \quad (\text{A.16})$$

Estimemos éste promedio bajo la norma L^2 , de la siguiente forma

$$|\bar{u}| \leq \int_0^1 |u(x)| \cdot 1 dx \leq \|u\|_{L^2} \|1\|_{L^2}, \quad (\text{A.17})$$

donde se ha usado la desigual de Cauchy-Schwarz. Recordando que $u \in C^1(I) \cap H^1(I)$, la estimación anterior está bien definida y tenemos finalmente

$$|\bar{u}| \leq \int_0^1 |u(x)| \cdot 1 dx \leq \|u\|_{L^2}. \quad (\text{A.18})$$

Usando el Teorema del Valor Medio, tenemos que existe $x_o \in I$, tal que

$$\bar{u} = u(x_o). \quad (\text{A.19})$$

Para ver esto claramente definimos

$$F(x) := \int_0^x u(y)dy, \quad (\text{A.20})$$

y tenemos que

$$F'(x) = u(x). \quad (\text{A.21})$$

Entonces, por el Teorema del Valor Medio, existe $x_o \in I$ tal que

$$\frac{F(1) - F(0)}{1} = F'(x_o) = u(x_o). \quad (\text{A.22})$$

Pero por la definición de $F(x)$ dada por (A.20), tenemos que

$$F(0) = 0, F(1) = \int_0^1 u(y)dy = \bar{u}. \quad (\text{A.23})$$

De aquí se sigue inmediatamente que (A.19).

Sea $x \in I$, entonces

$$|u(x) - \bar{u}| = |u(x) - u(x_o)| = \left| \int_{x_o}^x u'(y)dy \right| \leq |x - x_o|^{1/2} \|u'\|_{L^2} \leq \|u'\|_{L^2}, \quad (\text{A.24})$$

donde hemos usado la desigualdad de Cauchy-Schwarz. Usando ahora la desigualdad del triángulo en la expresión anterior tenemos

$$|u(x)| \leq |u(x) - \bar{u}| + |\bar{u}| \leq \|u'\|_{L^2} + \|u\|_{L^2}. \quad (\text{A.25})$$

A partir de ésta desigualdad podemos obtener

$$|u(x)|^2 \leq \|u'\|_{L^2}^2 + \|u\|_{L^2}^2 + 2\|u'\|_{L^2}\|u\|_{L^2} \leq 2\|u\|_{H^1}^2, \quad (\text{A.26})$$

de aquí tomamos el supremo en ambos lados de la desigualdad $|u(x)|^2 \leq 2\|u\|_{H^1}^2$, y la raíz cuadrada para obtener

$$\|u\|_{\infty} \leq \sqrt{2}\|u\|_{H^1}. \quad (\text{A.27})$$

Ahora, sea $u \in H^1(I)$. Dado que $C^1(I) \cap H^1(I)$ es denso en $H^1(I)$ (ver, por ejemplo, el Teorema 5.3.2 en [13]), existe una sucesión de Cauchy $(u_n) \in C^1(I) \cap H^1(I)$, tal que $u_n \rightarrow u \in H^1$ cuando $n \rightarrow \infty$.

Entonces usando (A.27) tenemos que

$$\|u_n - u_m\|_{\infty} \leq \sqrt{2}\|u_n - u_m\|_{H^1}. \quad (\text{A.28})$$

Por ser sucesión de Cauchy en H^1 , tenemos que $\|u_n - u_m\|_{H^1} \rightarrow 0$ cuando $n, m \rightarrow \infty$. Entonces la desigualdad (A.28) implica que

$$\|u_n - u_m\|_{\infty} \rightarrow 0, n, m \rightarrow \infty. \quad (\text{A.29})$$

Por lo tanto (u_n) es una sucesión de Cauchy en el espacio completo $(C^o(I), \|\cdot\|_\infty)$, entonces

$$u \rightarrow \tilde{u} \in C^o(\bar{I}). \quad (\text{A.30})$$

Entonces, ahora mostraremos que \tilde{u} es un representante de u .

Como $(u_n) \rightarrow u$ bajo la norma L^2 , esto implica que existe una subsucesión u_{n_k} tal que $u_{n_k}(x) \rightarrow u(x)$ para casi todo $x \in I$. Entonces usando (A.30) tenemos que

$$u(x) = \tilde{u}(x), \text{ para casi todo } x \in I, \quad (\text{A.31})$$

lo que demuestra que \tilde{u} es un representante de u .

Finalmente, tenemos que

$$\|u_n\|_\infty \leq \sqrt{2}\|u_n\|_{H^1}. \quad (\text{A.32})$$

Puesto que $u_n \rightarrow \tilde{u}$ en $C^o(\bar{I})$, y $u_n \rightarrow u$ en $H^1(I)$, cuando $n \rightarrow \infty$, y debido a la continuidad de la norma, esto implica que

$$\|\tilde{u}\|_\infty \leq \sqrt{2}\|u\|_{H^1}, \quad (\text{A.33})$$

lo que concluye la demostración del Lema.

□

A.0.6. Demostración al Lema (2.3.1)

Lema: A.0.2 Sea $g \in C^2([0, 1])$ una función de x que es dos veces continuamente diferenciable y sea $0 = x_0 < x_1 < x_2 < \dots < x_n = 1$, una partición de $[0, 1]$. Entonces

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \frac{1}{2} [g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (\text{A.34})$$

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) \left(\frac{x_{j+1} - x}{x_{j+1} - x_j} \right)^2 dx = \frac{1}{12} [3g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (\text{A.35})$$

$$\frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx = \frac{1}{12} [3g(x_j) + g(x_{j-1})] + O(\rho^2), \quad (\text{A.36})$$

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) \frac{x_{j+1} - x}{x_{j+1} - x_j} \frac{x - x_j}{x_{j+1} - x_j} dx = \frac{1}{12} [g(x_j) + g(x_{j+1})] + O(\rho^2), \quad (\text{A.37})$$

donde $\rho_j := x_{j+1} - x_j$ y $\rho = \max_{j=0,1,\dots,n} \rho_j$.

Demostración.

- Tomemos la integral dada por (A.34),

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \int_0^1 g(x_j + \rho_j h) dh, \quad (\text{A.38})$$

donde se hace el cambio de variable $x = x_j + \rho_j h$.

De la misma forma tomamos la misma integral pero ahora con el cambio de variable $x = x_{j+1} - \rho_j w$, de donde obtenemos

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \int_0^1 g(x_{j+1} - \rho_j w) dw, \quad (\text{A.39})$$

por lo tanto al promediar las integrales (A.38) y (A.39) tenemos

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \frac{1}{2} \int_0^1 [g(x_j + \rho_j h) + g(x_{j+1} - \rho_j h)] dh. \quad (\text{A.40})$$

Tomamos el integrando y desarrollamos en serie de Taylor

$$g(x_j + \rho_j h) = g(x_j) + g'(x_j)\rho_j h + O(\rho_j^2), \quad (\text{A.41})$$

y por otro lado

$$\begin{aligned} g(x_{j+1} - \rho_j h) &= g(x_{j+1}) - g'(x_{j+1})\rho_j h + O(\rho_j^2) \\ &= g(x_{j+1}) - [g'(x_j) + g''(x_j)\rho_j]\rho_j h + O(\rho_j^2), \end{aligned} \quad (\text{A.42})$$

reagrupando los órdenes de h correspondientes tenemos

$$g(x_{j+1} - \rho_j h) = g(x_{j+1}) - g'(x_j)\rho_j h + O(\rho_j^2). \quad (\text{A.43})$$

Finalmente la integral (A.40), al implementar las aproximaciones (A.41) y (A.42) resulta

$$\frac{1}{\rho_j} \int_{x_j}^{x_{j+1}} g(x) dx = \frac{1}{2} [g(x_j) + g(x_{j+1})] + O(\rho_j^2). \quad (\text{A.44})$$

- Ahora para resolver la integral (A.36) tenemos que

$$\frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx = \int_0^1 g(x_{j-1} + h\rho_{j-1}) h^2 dh, \quad (\text{A.45})$$

donde hemos usado el cambio de variable $x = x_{j-1} + h\rho_{j-1}$.

Usando la misma integral, pero con el cambio de variable $x = x_j - \rho_{j-1}w$ tenemos

$$\frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx = \int_0^1 g(x_j - w\rho_{j-1})(1-w)^2 dw. \quad (\text{A.46})$$

Entonces la integral (A.36) se puede escribir como la suma de (A.45) por un factor $\alpha \in \mathbb{R}$ con (A.46) por un factor $\beta \in \mathbb{R}$, tal que $\alpha + \beta = 1$.

Por lo tanto

$$\begin{aligned} & \frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx \\ &= \int_0^1 [\alpha g(x_{j-1} + h\rho_{j-1})h^2 + \beta g(x_j - h\rho_{j-1})(1-h)^2] dh, \end{aligned} \quad (\text{A.47})$$

donde tenemos que desarrollar en serie de Taylor la parte derecha.

Entonces tenemos

$$g(x_{j-1}+h\rho_{j-1}) = g(x_{j-1})+g'(x_{j-1})h\rho_{j-1}+O(\rho^2) = g(x_{j-1})+g'(x_j)h\rho_{j-1}+O(\rho^2). \quad (\text{A.48})$$

Y el otro término

$$g(x_j - h\rho_{j-1}) = g(x_j) - g'(x_j)h\rho_{j-1} + O(\rho^2). \quad (\text{A.49})$$

Al implementar las aproximaciones en (A.47) tenemos

$$\begin{aligned} & \frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx \\ &= \int_0^1 [\alpha h^2 g(x_{j-1}) + \alpha g'(x_j) h^3 \rho_{j-1} + \beta(1-h)^2 g(x_j) - \beta g'(x_j) h(1-h)^2 \rho_{j-1} + O(\rho^2)] dh, \end{aligned} \quad (\text{A.50})$$

donde podemos apreciar que para eliminar término de orden $O(\rho)$, hay que elegir α y β de forma adecuada, tomando en cuenta que las integrales que acompañan a dichos factores son

$$\int_0^1 h^3 dh = \frac{1}{4}, \quad \int_0^1 h(1-h)^2 dh = \frac{1}{12}. \quad (\text{A.51})$$

Al sustituir estas integrales resulta

$$\frac{\alpha}{4} - \frac{\beta}{12} = 0, \quad (\text{A.52})$$

además de la ecuación

$$\alpha + \beta = 1. \quad (\text{A.53})$$

Al resolver para α y β tenemos que $\alpha = 1/4$ y $\beta = 3/4$. Podemos entonces calcular la aproximación para (A.50) de donde obtenemos

$$\begin{aligned} & \frac{1}{\rho_{j-1}} \int_{x_{j-1}}^{x_j} g(x) \left(\frac{x - x_{j-1}}{x_j - x_{j-1}} \right)^2 dx \\ &= \frac{1}{12} g(x_{j-1}) + \frac{1}{4} g(x_j) + O(\rho^2). \end{aligned} \quad (\text{A.54})$$

- La demostración de (A.35) es similar a la de (A.36).
- Por último demostremos (A.37), para esto tomemos la integral

$$\int_{x_j}^{x_{j+1}} g(x) \frac{x_{j+1} - x}{x_{j+1} - x_j} \frac{x - x_j}{x_{j+1} - x_j} dx = \frac{1}{\rho_j^2} \int_{x_j}^{x_{j+1}} g(x) (x_{j+1} - x)(x - x_j) dx. \quad (\text{A.55})$$

Definimos a $\varphi \equiv \frac{x_j + x_{j+1}}{2}$ y

$$x = \varphi + \frac{h\rho_j}{2}, \quad -1 \leq h \leq 1. \quad (\text{A.56})$$

La integral anterior se reescribe como

$$\begin{aligned} & \int_{x_j}^{x_{j+1}} g(x) \frac{x_{j+1} - x}{x_{j+1} - x_j} \frac{x - x_j}{x_{j+1} - x_j} dx \\ &= \frac{1}{\rho_j^2} \int_{-1}^1 g(\varphi + h\rho_j/2) (x_{j+1} - \varphi - \frac{h\rho_j}{2}) (\varphi + \frac{h\rho_j}{2} - x_j) \frac{\rho_j}{2} dh \\ &= \frac{1}{2\rho_j} \int_{-1}^1 g(\varphi + h\rho_j/2) \frac{\rho_j^2}{4} (1-h)(1+h) dh. \end{aligned} \quad (\text{A.57})$$

Expandemos en esta integral $g(\varphi + h\rho_j/2) = g(\varphi) + h\rho_j g'(\varphi)/2 + O(\rho_j^2)$, también hacemos el producto notable $(1-h^2) = (1+h)(1-h)$, y encontramos que la integral se aproxima de la siguiente forma

$$\begin{aligned} & \int_{x_j}^{x_{j+1}} g(x) \frac{x_{j+1} - x}{x_{j+1} - x_j} \frac{x - x_j}{x_{j+1} - x_j} dx \\ &= \frac{\rho_j}{8} g(\varphi) \int_{-1}^1 (1-h^2) dh + O(\rho^3) = \frac{\rho_j}{6} g(\varphi) + O(\rho^3). \end{aligned} \quad (\text{A.58})$$

Por último usamos que

$$g(\varphi) = g\left(\frac{x_j + x_{j+1}}{2}\right) = \frac{1}{2} [g(x_j) + g(x_{j+1})] + O(\rho^2). \quad (\text{A.59})$$

Usando esta aproximación en la última integral, obtenemos la afirmación.

□

Bibliografía

- [1] T. Regge and J. Wheeler. Stability of a Schwarzschild singularity. *Phys. Rev.*, 108:1063–1069, 1957.
- [2] F. Zerilli. Effective potential for even-parity Regge-Wheeler gravitational perturbation equations. *Phys. Rev. Lett.*, 24:737–738, 1970.
- [3] S.A. Teukolsky. Rotating black holes: Separable wave equations for gravitational and electromagnetic perturbations. *Phys. Rev. Lett.*, 29:1114–1118, 1972.
- [4] S. E. Gralla and R. M. Wald. A rigorous derivation of gravitational self-force. *Class. Quantum Grav.*, 25:205009, 2008.
- [5] C. F. Sopuerta and P. Laguna. Finite element computation of the gravitational radiation emitted by a pointlike object orbiting a nonrotating black hole. *Phys. Rev. D*, 73:044028, 2006.
- [6] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley and Sons, 1989.
- [7] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Dover, Mineola, New York, 2009.
- [8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes*. Cambridge University Press, Cambridge, England, 1986.
- [9] J. D. Hoffman. *Numerical Methods for Engineers and Scientists*. McGraw-Hill, Inc., New-York, 1992.
- [10] H. B. Keller E. Isaacson. *Analysis of Numerical Methods*. Dover Publications, Inc., New York, 1994.
- [11] Y. Saad. *Iterative Methods for Sparse Linear Systems, 2nd edition*. SIAM, Philadelphia, PA, 2003.
- [12] H. Beyer. *Beyond partial differential equations. On linear and quasi-linear abstract hyperbolic evolution equations*. Lecture Notes in Mathematics 1898, Springer, 2007.

- [13] L.C. Evans. *Partial differential equations*. American Mathematical Society, Providence, Rhode Island, 1998.
- [14] J. W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Springer, 1995.