

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSTGRADO



Estudio de los Métodos Espectrales en Ecuaciones Diferenciales de una Dimensión y su comparación con el método de Diferencias Finitas

Tesis para optar el grado de Magíster en Matemáticas

AUTOR: Br. David Sáenz López

ASESOR Dr. Rubén Ángel Agapito Ruiz

JURADO Dr. Francisco Pasquel Carbajal Dr. Jorge Richard Chávez Fuentes

> LIMA-PERÚ 2015



Dedicatoria

A mi madre Rosa por su inmenso amor que me dio mientras estuvo a mi lado. A mi padre Wilfredo por su apoyo incondicional ya que gracias a él, he alcanzado esta meta. A mi hermana por su apoyo y cariño.

> A mi querida esposa Karina por su amor y apoyo que me da todos los días. A mi amado hijo David Mateo quien me alegra cada día con sus travesuras y ocurrencias.



Agradecimientos

Agradezco a Dios, por todo lo que tengo y permitirme alcanzar mis metas junto a mi familia.

Quiero expresar mis sinceros agradecimiento al Dr. Rubén Ángel Agapito Ruiz por sus sugerencias, y la oportunidad brindada durante mis estudios de posgrado.

A todas aquellas personas que me apoyaron y me alentaron a terminar lo que había empezado, mis sinceros agradecimientos.

Al Dr. Francisco Pasquel y al Dr. Jorge Richard Chávez, por sus observaciones para esta versión final de la tesis.



Índice General

In	Introducción vi								
Resumen									
Lista de Símbolos									
1	Prel	iminar	res: Aproximación Polinomial	1					
	1.1	Polino	omios Ortogonales	2					
		1.1.1	Existencia y Unicidad	2					
		1.1.2	Sistemas de Polinomios Ortogonales	8					
		1.1.3	Cuadraturas Gaussiana	10					
		1.1.4	Interpolación y Transformadas Discretas	17					
		1.1.5	Diferenciación en el Espacio Físico	20					
		1.1.6	Diferenciación en el Espacio Frecuencia	21					
	1.2	Polino	omios de Chebyshev	22					
		1.2.1	Cálculo de los Coeficientes de Chebyshev	27					
		1.2.2	Diferenciación en el Espacio Frecuencia	28					
		1.2.3	Diferenciación en el Espacio Físico	31					
2	Mét	odos E	spectrales Polinomiales	34					
	2.1	Métoc	lo de Galerkin	36					



	2.2 2.3	Método de Tau3Método de Colocación3						
3	Aplicaciones de los Métodos Espectrales 40							
	3.1	Ecuac	ión del Calor	40				
		3.1.1	Solución Numérica por Diferencias Finitas	40				
		3.1.2	Aproximación Espectral usando Chebyshev Tau	49				
		3.1.3	Aproximación Espectral usando Chebyshev Colocación	52				
	3.2	Ecuac	ión de Burger	55				
		3.2.1	Formulaciones Fuerte y Débil de la Ecuación de Burger	57				
		3.2.2	Aproximación Espectral usando Chebyshev Galerkin	58				
		3.2.3	Aproximación Espectral usando Chebyshev Tau	60				
		3.2.4	Aproximación Espectral usando Chebyshev Colocación	62				
Conclusiones 67								
Ap	oéndi	ce: Cód	ligo Matlab	68				
Bibliografía				82				

v



Introducción

En general, encontrar una solución analítica de una ecuación diferencial parcial no es fácil, y más aún cuando ésta ecuación es no lineal. Debido a esto, surgieron varios métodos numéricos para encontrar una solución aproximada a la deseada. Los métodos numéricos más conocidos son:

- Métodos de Diferencias Finitas que tuvo su gran auge en la década de 1950.
- Métodos de Elementos Finitos que tuvo su gran auge en la década de 1960.
- Métodos Espectrales que tuvo su gran auge en la década de 1970.

Mientras que los métodos de diferencias finitas dan soluciones aproximadas en los puntos de la malla computacional elegida, los métodos de elementos finitos dan aproximaciones polinomiales continuas o continuas por partes en regiones poligonales (generalmente triangulares en dos dimensiones), mientras que los métodos espectrales brindan soluciones aproximadas en la forma de polinomios sobre todo su dominio.

Los métodos espectrales fueron propuestos originalmente en 1944 por Blinova [1], implementado por primera vez en 1954 por Silberman [52], y prácticamente olvidados a mediados de los años 1960. Luego, retomadas por Orzag, Eliason, Machenhauer y Rasmussen que desarrollaron aplicaciones especializadas en la década de 1970, dotados de los primeros fundamentos matemáticos por el trabajo de Gottlieb y Orzag [20].



El origen del término "espectral" [50], no está muy claro pero probablemente surgió del uso de los senos y cosenos de Fourier como funciones base ([3], [20]), especialmente por la relación con el análisis de series de tiempo [49].

El primer esquema de método espectral que se usó para la solución de ecuaciones diferenciales parciales fue el conocido método de Fourier, el cual consiste en representar la solución como una expansión en serie truncada donde las incógnitas son los coeficientes de expansión. Las funciones base que se utilizan para dicho método son apropiadas para problemas periódicos; mientras que para problemas no periódicos, los polinomios de Chebyshev o Legendre se consideran como funciones base, pero otras funciones base (Laguerre, Hermite) se deberían de tener en cuenta según el problema que se esté considerando.

Por mucho años antes de la aparición de las computadoras, los estudios teóricos en física matemática, y especialmente en mecánica de fluidos frecuentemente hacían uso de la expansión en series. Sin embargo, los métodos de expansión en serie demostraron limitaciones debido a la dificultad en el cálculo de las sumas truncadas con un gran número de términos o al tratar problemas no lineales. Estas limitaciones todavía estaban presentes en la etapa más temprana de la dinámica de fluidos computacional cuando los métodos computacionales no eran suficientes para un uso eficaz de expansión en serie. Esto llevó a la decadencia de las técnicas de expansión en serie al beneficio de los métodos numéricos discretos como métodos de diferencias finitas o elementos finitos. Pero, la baja precisión de estos métodos de discretización fue un obstáculo para la representación exacta de flujos como se encontró en muchos problemas de la mecánica de fluidos. Por lo tanto, en la década de 1970, se ha visto un resurgimiento del método de Fourier que se aplicó a la simulación numérica directa de la turbulencia [20].

Los métodos espectrales son una clase de discretización espacial para ecuaciones diferenciales. Las componentes claves para su formulación son las funciones base (llamadas también funciones de aproximación o expansión) y las funciones de prueba. Las funciones base se usan para dar una representación aproximada de la solución. Las funciones de prueba se usan para asegurar que la ecuación diferencial y quizás algunas condiciones de frontera se cumplan tanto como sea posible por la serie truncada de expansión. Esto se consigue minimizando, con respecto a una norma adecuada, el residuo producido por el uso de la expansión truncada en lugar de la solución exacta. Un requerimiento equivalente es que el residuo satisface una condición de ortogonalidad adecuada con respecto a cada una de las funciones de prueba.



La elección de las funciones base es una de las características que distingue a los métodos espectrales de los métodos de elementos finitos y diferencias finitas. Las funciones base son infinitamente diferenciable globalmente y casi ortogonales (es decir, la matriz que contiene los productos internos toman valores muy pequeños, en muchos casos esta matriz es diagonal); mientras que las funciones base en el método de elementos finitos son infinitamente diferenciable localmente y casi ortogonales y en el método de diferencia finita son polinomios superpuestos de orden inferior.

La elección de las funciones de prueba determina los esquemas de método espectral, es decir, los esquemas de Galerkin, Tau y de Colocación o pseudoespectral. En el esquema de Galerkin, las funciones de prueba son las mismas que las funciones base. Por tanto, estas funciones son infinitamente diferenciables y satisfacen todas las condiciones de frontera. La ecuación diferencial se aplica al requerir que el residuo sea cero. El método espectral de Tau es similar al método de Galerkin en el sentido en que la ecuación diferencial se satisface. Sin embargo, ninguna de las funciones base necesita satisfacer las condiciones de frontera. Por lo tanto, se utiliza un conjunto adicional de ecuaciones para hacer cumplir las condiciones de frontera. En el esquema de Colocación, las funciones de prueba son las funciones delta de Dirac centradas en los llamados puntos de colocación. Este enfoque requiere que la ecuación diferencial se cumpla exactamente en los puntos de colocación.

Una característica del esquema de Galerkin es que las funciones base y las funciones de prueba son iguales. La primera aplicación de los métodos espectrales para PDEs, fue hecha por Silberman [52], para el modelo meteorológico, quien usó el método de Galerkin. Sin embargo, los métodos espectrales de Galerkin se convirtieron en algo práctico para los cálculos de alta resolución tales como los problemas no lineales después de Orszag ([37], [38]) y Eliasen, Machenhauer y Rasmussen [12] que desarrollaron métodos para evaluar las sumas de convolución. Para los problemas que contienen términos no lineales más complicadas, los métodos de Galerkin espectrales siguen siendo poco práctico.

El método de Tau es una modificación del método de Galerkin que se puede aplicar a problemas con condiciones de frontera no periódicas. Lanczos [29], desarrolló el método espectral de Tau, y Orszag [39], desarrolló la aplicación del método de Chebyshev Tau para obtener soluciones de gran exactitud a problemas de dinámica de fluidos.

El esquema de colocación parece haber sido utilizado por primera vez por Kantorovic [23] y por Slater [53] en aplicaciones específicas. Frazer, Jones y Skan



[17], desarrollaron un método general para resolver ecuaciones diferenciales ordinarias. Se utilizaron una variedad de funciones base y una distribución arbitraria de puntos de colocación. El trabajo de Lanczos [29], estableció por primera vez que una elección adecuada de las funciones base y distribución de los puntos de colocación es crucial para la exactitud de la solución. Este método fue retomado por Clenshaw [7], Clenshaw y Norton [8] y Wright [58]. Estos estudios incluyeron la aplicación de expansiones polinomiales de Chebyshev a los problemas de valor inicial. Villadsen y Stewart [56], desarrollaron este método para los problemas de valor de frontera.

Los métodos espectrales tienen un amplio uso en diferentes áreas como: teoría cuántica ([31], [36]) basado en la ecuación Schrödinger que proporciona la descripción teórica de numerosos sistemas en química y física; teoría cinética basada en la ecuación de Boltzmann ([27], [32]) o en la ecuación de Fokker-Planck ([5], [45]); problemas en mecánica de fluidos ([4], [20], [42]). También hay importantes aplicaciones en el escape átomos de la atmósfera del planeta ([14], [51]) como la pérdida de carga de partículas de la tierra ([33], [43]) y del sol [11].

El presente trabajo pretende contribuir en sentar los fundamentos sobre métodos espectrales, para que sean aplicados en futuras investigaciones más elaboradas, así como brindar los códigos de implementación (en Matlab), los cuales raramente se encuentran en forma explícita en la literatura.

Este trabajo está organizado de la siguiente manera: el Capítulo 1 abarca las propiedades más importantes de los polinomios ortogonales; en particular, los polinomios de Chebyshev, los cuales son adecuados para representar funciones de dominio finito y sus relaciones de recurrencia asociadas. Además, se presenta un breve repaso de las fórmulas de cuadratura gaussiana. En el Capítulo 2, se presenta en forma detallada los métodos espectrales polinomiales, útiles para problemas con condiciones de frontera no periódicas. Presentamos los métodos de Galerkin, Tau y de Colocación. En el Capítulo 3 se da ejemplos de la implementación numérica de la ecuación del calor usando los métodos de diferencias finitas y los métodos espectrales, usando los polinomios de Chebyshev. Además, se brindan los detalles necesarios para implementar la ecuación de Burger usando los métodos espectrales.



Resumen

Los métodos espectrales se pueden utilizarse para resolver numéricamente ecuaciones diferenciales de la forma:

$$\begin{aligned} &\frac{\partial u}{\partial t}(x,t) = \mathfrak{L}(x,t)u(x,t), \quad x \in D, \ t \ge 0, \\ &B \ u(x,t) = 0, \quad x \in \partial D, \ t > 0, \\ &u(x,0) = g(x), \quad x \in D, \end{aligned}$$

donde *D* es el dominio espacial con frontera ∂D , $\mathfrak{L}(x, t)$ es un operador diferencial (no necesariamente lineal), *B* es el operador de frontera y g(x) es la condición inicial.

Los métodos espectrales son una clase de discretización espacial para ecuaciones diferenciales. Las componentes claves para su formulación son las funciones base (llamadas también funciones de aproximación o expansión) y las funciones de prueba. Las funciones base se usan para dar una representación aproximada de la solución. Las funciones de prueba se usan para asegurar que la ecuación diferencial y quizás algunas condiciones de frontera se cumplan tanto como sea posible por la serie truncada de expansión.

Para encontrar una solución aproximada con métodos espectrales, se utiliza una serie truncada de funciones ortogonales suaves (funciones base) $\Phi_n(x)$, y se





aproxima la solución con la expresión

$$u_N(x,t) = \sum_{n=0}^N \widehat{u}_n(t) \Phi_n(x).$$

Deben encontrarse entonces los coeficientes \hat{u}_n . Existen diversas formas de obtener estos coeficientes, y cada manera de encontrarlos define diferentes esquemas de métodos espectrales. Los más conocidos son: el método de **Galerkin**, el método de **Tau** y el de **Colocación**.

El método de Galerkin es un método donde las funciones base se eligen de forma tal que la solución propuesta satisfaga las condiciones de frontera. Luego se busca una solución $u_N(x, t)$ en forma de serie truncada escrita como combinación lineal de las funciones base, y como ella es una aproximación de la solución de la ecuación diferencial a resolver, al sustituirla en la ecuación, no obtendremos igualdad. A la ecuación diferencial con $u_N(x, t)$ reemplazada la llamaremos residuo y la representaremos con R(x, t). La exactitud de la aproximación dependerá de cuan pequeño, en norma, sea el valor del residuo R(x, t). El método de Tau es una versión modificada del método de Galerkin. La diferencia radica en la forma de manejar las condiciones de frontera, lo cual se ilustrará en este trabajo. El método de colocación parte de la serie truncada propuesta $u_N(x, t)$ y se la evalúa en una partición no uniforme del dominio (llamada puntos de colocación) cuya elección se basa en el método de cuadratura gaussiana, utilizada comúnmente en integración numérica.

Adicionalmente, hacemos una comparación del método de Diferencias Finitas versus el método espectral, llevado a cabo solo para la ecuación del calor. Los principios básicos del método espectral se ilustran con ecuaciones conocidas como la ecuación de Calor y la Ecuación de Burger.

David Sáenz López



Lista de Símbolos

C	_	Conjunto de números complejos
R	-	Conjunto de números reales
Z	1	Conjunto de números enteros
N	—	Conjunto de números enteros no negativos
$C(\Lambda)$	—	espacio de funciones continuas en Λ donde $\Lambda =]a, b[,]a, b], [a, b[, [a, b]$
P_N	-	Conjunto de polinomios de grado $\leq N$
δ_{mn}	<u>-</u>	delta de Kronecker
L_k	_	polinomio de Legendre de grado k
T_k	—	polinomio de Chebyshev de grado k
\mathscr{L}_k	_	polinomio de Laguerre de grado k
H_k	—	polinomio de Hermite de grado k
ω	—	función peso no negativa genérica
$L^p(\Omega)$	—	espacio L^p en Ω con $1 \le p \le \infty$
$L^p_{\omega}(\Omega)$	_	espacio L^p en Ω con $1 \le p \le \infty$ con peso ω
$\langle \cdot, \cdot \rangle$	_	producto interno de $L^2(\Omega)$



- $\langle \cdot, \cdot \rangle_{\omega}$ producto interno de $L^2_{\omega}(\Omega)$
- $\|\cdot\|$ norma de $L^2(\Omega)$
- $\|\cdot\|_{\omega}$ norma de $L^2_{\omega}(\Omega)$
- $\langle \cdot, \cdot \rangle_{N,\omega}$ producto interno discreto asociado a una cuadratura tipo Gauss
- $\langle \cdot, \cdot \rangle_N \quad \quad \langle \cdot, \cdot \rangle_N = \langle \cdot, \cdot \rangle_{N,\omega} \operatorname{con} \omega = 1$
- $\|\cdot\|_{N,\omega}$ norma discreta asociada con $\langle\cdot,\cdot\rangle_{N,\omega}$





1

Preliminares: Aproximación Polinomial

Sabemos que la función exponencial e^x tiene una expansión en serie de Taylor de la forma

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Pero una serie, al considerarse como "suma infinita", no es factible de ser manipulada numéricamente. Así, se prefiere utilizar la aproximación

$$\sum_{k=0}^{N} \frac{x^k}{k!}$$

para *N* grande. Obviamente, se ha cometido un error de aproximación, pero por razones de convergencia éste debe disminuir a medida que *N* se torne más grande.

Sea C[a, b] el espacio vectorial de funciones reales continuas definidas en [a, b]. El teorema de aproximación de Weierstrass nos dice que todo elemento de este espacio se puede aproximar uniformemente por una función polinomial. Ésta a su vez puede escribirse como una combinación lineal de funciones polinomiales ortogonales base { $\phi_k(x)$ }. Así, si formamos la función

$$u_N(x,t) = \sum_{k=0}^N \widehat{u}_k(t)\phi_k(x),$$





para t fijo, es fácil hallar los coeficientes $\hat{u}_k(t)$ realizando un producto interno conveniente, esto es

$$\widehat{u}_k(t) = \frac{\int_a^b u_N(x,t)\phi_k(x)\,dx}{\int_a^b (\phi_k(x))^2 dx}$$

donde $\int_{a}^{b} (\phi_{k}(x))^{2} dx \neq 0$. Utilizaremos como funciones base los polinomios ortogonales de Chebyshev. Así, para resolver una ecuación diferencial se reemplaza la aproximación $u_N(x, t)$ en la ecuación, la cual conlleva a un error, y para minimizar este error, asumimos que es escrito como combinación lineal de funciones polinomiales cuyos coeficientes deberán satisfacer una ecuación diferencial ordinaria, la cual es resuelta con métodos numéricos como Runge-Kutta. Este método de hallar la solución aproximada para t fijo, usando funciones base polinomiales e ir marchando hacia adelante en el tiempo es llamado método espectral.

1.1 **Polinomios Ortogonales**

Mencionaremos las propiedades y definiciones más importantes de polinomios ortogonales en general. Nuestro punto de partida es la generación de los polinomios ortogonales por una relación de recurrencia de tres términos, lo que lleva a algunas fórmulas muy útiles.

Existencia y Unicidad 1.1.1

Dado un intervalo abierto $I :=]a, b[(-\infty \le a < b \le +\infty)]$ y una función peso ω se cumple que

$$\omega(x) > 0$$
, para todo $x \in I$ y $\omega \in L^{1}(I)$ (1.1)

Dos funciones f y g son **ortogonales** con peso $\omega \in L^2_{\omega}(a, b)$ si

$$\langle f,g \rangle_{\omega} := \int_{a}^{b} f(x)g(x)\omega(x) \,\mathrm{d}x = 0.$$

Un polinomio de grado *n* lo denotaremos por

$$p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \quad a_n \neq 0, \tag{1.2}$$

donde a_i son constantes reales, y a_n es el coeficiente principal.



Una sucesión de polinomios $\{p_n\}_{n\geq 0}$ con grad $(p_n) = n$ es **ortogonal** en $L^2_{\omega}(]a, b[)$ si

$$\langle p_n, p_m \rangle_{\omega} := \int_a^b p_n(x) p_m(x) \omega(x) \,\mathrm{d}x = \gamma \delta_{mn},$$
 (1.3)

donde

$$\gamma = \|p_n\|_{\omega}^2 = \int_a^b p_n^2(x)\omega(x) \,\mathrm{d}x \neq 0$$

en el caso m = n, y δ_{mn} es la delta de Kronecker.

A lo largo de este trabajo, $\{p_n\}$ denotará a una sucesión de polinomios ortogonales con respecto a la función peso ω , y p_n es de grado n. Denotaremos por P_n al conjunto generado por todas las combinaciones lineales de polinomios de grado $\leq n$, es decir,

$$P_n := \operatorname{Gen}\{1, x, x^2, \cdots, x^n\}.$$
 (1.4)

Por un simple argumento de dimensión tenemos que

$$P_n := \operatorname{Gen}\{p_0, p_1, p_2, \cdots, p_n\}.$$
 (1.5)

Lema 1. p_{n+m} , $m \ge 1$ es ortogonal a cualquier polinomio $q \in P_n$.

Demostración. Usando (1.5), cualquier $q \in P_n$ se puede escribir como

 $q = \alpha_n p_n + \alpha_{n-1} p_{n-1} + \dots + \alpha_0 p_0.$

Luego, usando (1.3) tenemos que

$$\langle p_{n+m}, q \rangle_{\omega} = 0.$$

De aquí en adelante se denotará al polinomio mónico correspondiente a p_n por

$$\overline{p}_n(x) := \frac{p_n(x)}{a_n} = x^n + b_{n-1}^{(n)} x^{n-1} + \dots + b_0^{(n)},$$
(1.6)

donde $b_i^{(n)} = \frac{a_i}{a_n}$, $i = 0, 1, \dots, n-1$.

Se demuestra a continuación que para cualquier función peso $\omega(x)$ definida en]a, b[, existe una única familia de polinomios ortogonales mónicos generadas por una fórmula de recurrencia de tres términos.



Teorema 1. Para cualquier función peso positiva $\omega \in L^1(I)$, existe una única sucesión de polinomios ortogonales mónicos $\{\overline{p}_n\}$ con grad $(\overline{p}_n) = n$, construida como sigue

$$\overline{p}_0 = 1, \quad \overline{p}_1 = x - \alpha_0,$$

$$\overline{p}_{n+1} = (x - \alpha_n)\overline{p}_n - \beta_n\overline{p}_{n-1}, \quad n \ge 1,$$
(1.7)

donde

$$\alpha_n = \frac{\langle x\overline{p}_n, \overline{p}_n \rangle_{\omega}}{\|\overline{p}_n\|_{\omega}^2}, \quad n \ge 0,$$
(1.8a)

$$\beta_n = \frac{\|\overline{p}_n\|_{\omega}^2}{\|\overline{p}_{n-1}\|_{\omega}^2}, \quad n \ge 1.$$
(1.8b)

Demostración. Se verifica fácilmente que los dos primeros polinomios son

$$\overline{p}_0 \equiv 1$$
, $\overline{p}_1 = x - \alpha_0$

Para determinar α_0 , dado que $\langle \overline{p}_0, \overline{p}_1 \rangle_{\omega} = 0$ se deduce que

$$\alpha_{0} = \frac{\int_{a}^{b} \omega(x) x \, \mathrm{d}x}{\int_{a}^{b} \omega(x) \, \mathrm{d}x} = \frac{\langle x \overline{p}_{0}, \overline{p}_{0} \rangle_{\omega}}{\|\overline{p}_{0}\|_{\omega}^{2}},$$

donde el denominador es positivo porque $\omega(x) > 0$.

Procedemos por inducción. Asumiendo que se ha obtenido una sucesión de polinomios mónicos ortogonales $\{\overline{p}_k\}_{k=0}^n$ buscamos \overline{p}_{n+1} de la forma

$$\overline{p}_{n+1} = x\overline{p}_n - \alpha_n\overline{p}_n - \beta_n\overline{p}_{n-1} - \sum_{k=0}^{n-2}\gamma_k^{(n)}\overline{p}_k, \qquad (1.9)$$

donde α_n y β_n deben estar dados por las fórmulas (1.8) y los γ_k están por determinar, es más, probaremos que los γ_k son cero. Se requiere que

$$\langle \overline{p}_{n+1}, \overline{p}_k \rangle_{\omega} = 0, \quad 0 \le k \le n.$$
 (1.10)



Tomando entonces el producto interno con \overline{p}_k en ambos lados de (1.9), y usando la ortogonalidad de $\{\overline{p}_k\}_{k=0}^n$, encontramos que (1.10) se cumple sí y sólo sí

$$\langle \overline{p}_{n+1}, \overline{p}_n \rangle_{\omega} = \langle x \overline{p}_n, \overline{p}_n \rangle_{\omega} - \alpha_n \langle \overline{p}_n, \overline{p}_n \rangle_{\omega} = 0, \langle \overline{p}_{n+1}, \overline{p}_{n-1} \rangle_{\omega} = \langle x \overline{p}_n, \overline{p}_{n-1} \rangle_{\omega} - \beta_n \langle \overline{p}_{n-1}, \overline{p}_{n-1} \rangle_{\omega} = 0, \langle \overline{p}_{n+1}, \overline{p}_j \rangle_{\omega} = \langle x \overline{p}_n, \overline{p}_j \rangle_{\omega} - \sum_{k=0}^{n-2} \gamma_k^{(n)} \langle \overline{p}_k, \overline{p}_j \rangle_{\omega}$$

$$\stackrel{(1.11)}{=} \langle x \overline{p}_n, \overline{p}_j \rangle_{\omega} - \gamma_j^{(n)} || \overline{p}_j ||_{\omega}^2 = 0, \quad 0 \le j \le n-2.$$

Por lo tanto, la primera igualdad implica (1.8a), y por la segunda,

$$\beta_n = \frac{\langle x\overline{p}_n, \overline{p}_{n-1} \rangle_{\omega}}{\|\overline{p}_{n-1}\|_{\omega}^2} = \frac{\langle \overline{p}_n, x\overline{p}_{n-1} \rangle_{\omega}}{\|\overline{p}_{n-1}\|_{\omega}^2} = \frac{\|\overline{p}_n\|_{\omega}^2}{\|\overline{p}_{n-1}\|_{\omega}^2}$$

donde se ha usado el hecho que

$$x\overline{p}_{n-1} = \overline{p}_n + \sum_{k=0}^{n-1} \delta_k^{(n)} \overline{p}_k,$$

y la ortogonalidad de $\{\overline{p}_k\}_{k=0}^n$ implica la última igualdad en la fórmula para β_n . Resta por demostrar que los coeficientes $\{\gamma_k^{(n)}\}_{k=0}^{n-2}$ en (1.9) son todos ceros. En efecto, obtenemos del Lema 1 que

$$\langle x\overline{p}_n, \overline{p}_j \rangle_{\omega} = \langle \overline{p}_n, x\overline{p}_j \rangle_{\omega} = 0, \quad j = 0, 1, \cdots, n-2,$$

el cual, junto con la última ecuación de (1.11), implica $\gamma_j^{(n)} \equiv 0$ para $k = 0, 1, \dots, n-2$, en (1.9). Esto completa la inducción.

A continuación probaremos que la sucesión de polinomios generada por (1.7)-(1.8) es única. Para esto, asumamos que $\{\overline{q}_n\}_{n=0}^n$ es otra sucesión de polinomios mónicos ortogonales. Como \overline{p}_n , dado por (1.7), también es mónico, tenemos grad $(\overline{p}_{n+1} - \overline{q}_{n+1}) \le n$. Por el Lema 1,

$$\langle \overline{p}_{n+1}, \overline{p}_{n+1} - \overline{q}_{n+1} \rangle_{\omega} = 0, \quad \langle \overline{q}_{n+1}, \overline{p}_{n+1} - \overline{q}_{n+1} \rangle_{\omega} = 0,$$

lo cual implica

$$\langle \overline{p}_{n+1} - \overline{q}_{n+1}, \overline{p}_{n+1} - \overline{q}_{n+1} \rangle_{\omega} = 0.$$

Así, $\overline{p}_{n+1}(x) - \overline{q}_{n+1}(x) \equiv 0$. Esto prueba la unicidad.



El teorema anterior nos da una relación de recurrencia de tres términos para la construcción de polinomios ortogonales donde las constantes α_n y β_n pueden ser evaluadas de forma explícita. La relación de recurrencia de tres términos (1.7) es esencial para obtener otras propiedades de los polinomios ortogonales. Por conveniencia, primero la extendemos a polinomios ortogonales no necesariamente mónicos.

Corolario 1. Sea $\{p_n\}$ una sucesión de polinomios ortogonales con coeficiente principal $k_n \neq 0$. Entonces se tiene

$$p_{n+1} = (a_n x - b_n) p_n - c_n p_{n-1}, \quad n \ge 0,$$
(1.12)

donde convenimos en definir $p_{-1} := 0$, $p_0 = k_0$, y

$$a_n = \frac{k_{n+1}}{k_n},\tag{1.13a}$$

$$b_n = \frac{k_{n+1}}{k_n} \frac{\langle x p_n, p_n \rangle_{\omega}}{\|p_n\|_{\omega}^2},$$
(1.13b)

$$c_n = \frac{k_{n-1}k_{n+1}}{k_n^2} \frac{\|p_n\|_{\omega}^2}{\|p_{n-1}\|_{\omega}^2}$$
(1.13c)

Demostración. Este resultado se sigue directamente del Teorema 1 al insertar $\overline{p}_l = \frac{p_l}{k_l} \operatorname{con} l = n - 1, n, n + 1 \operatorname{en} (1.7) \operatorname{y} (1.8).$

Los polinomios ortogonales $\{p_n\}$ con coeficientes principales $\{k_n\}$ son únicamente determinados por (1.12)-(1.13). Una consecuencia importante de la fórmula de recurrencia de tres términos (1.12)-(1.13) es la conocida **fórmula de Christoffel - Darboux.**

Corolario 2. Sea $\{p_n\}$ una sucesión de polinomios ortogonales con grad $(p_n) = n$. Entonces

$$\frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x - y} = \frac{k_{n+1}}{k_n} \|p_n\|_{\omega}^2 \sum_{j=0}^n \frac{p_j(x)p_j(y)}{\|p_j\|_{\omega}^2},$$
(1.14)

y

$$p_{n+1}'(x)p_n(x) - p_n'(x)p_{n+1}(x) = \frac{k_{n+1}}{k_n} \|p_n\|_{\omega}^2 \sum_{j=0}^n \frac{p_j^2(x)}{\|p_j\|_{\omega}^2}.$$
 (1.15)



Demostración. Probaremos (1.14) usando el Corolario 1.

$$p_{j+1}(x)p_j(y) - p_j(x)p_{j+1}(y) = [(a_jx - b_j)p_j(y) - c_jp_{j-1}(x)]p_j(y) - p_j(x)[(a_jy - b_j)p_j(y) - c_jp_{j-1}(y)] = a_j(x - y)p_j(x)p_j(y) + c_j[p_j(x)p_{j-1}(y) - p_{j-1}(x)p_j(y)].$$

Por tanto, usando (1.13) se tiene

$$\frac{k_j}{k_{j+1} \|p_j\|_{\omega}^2} \frac{p_{j+1}(x)p_j(y) - p_j(x)p_{j+1}(y)}{x - y} - \frac{k_{j-1}}{k_j \|p_{j-1}\|_{\omega}^2} \frac{p_j(x)p_{j-1}(y) - p_{j-1}(x)p_j(y)}{x - y}$$
$$= \frac{1}{\|p_j\|_{\omega}^2} p_j(x)p_j(y).$$

Definiendo $p_{-1} := 0$, la relación también se cumple para j = 0. Luego, sumando término a término para $0 \le j \le n$ se obtiene (1.14).

Ahora probaremos (1.15). Se observa que

$$\frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x - y} = \frac{p_{n+1}(x) - p_{n+1}(y)}{x - y}p_n(y) - \frac{p_n(x) - p_n(y)}{x - y}p_{n+1}(y).$$

En consecuencia, tomando límite cuando $y \rightarrow x$ a (1.14), se obtiene (1.15) usando la definición de derivada.

Los ceros de los polinomios ortogonales desempeñan un papel fundamental en la implementación de los métodos espectrales. Por ejemplo, ellos eligen como son los nodos de la cuadratura gaussiana, y como usarlos para generar mallas computacionales para los métodos espectrales. Por lo que es útil para obtener sus propiedades esenciales.

Sea una sucesión de polinomios ortogonales $\{p_n\}$ (con grad $(p_n) = n$) con respecto a la función peso $\omega(x)$ en (a, b). El principal resultado acerca de los ceros de los polinomios ortogonales es el siguiente:

Teorema 2. Los ceros de p_{n+1} son todos reales, simples, y están en el intervalo]a; b[.

Demostración. Probaremos que los ceros son reales. En efecto, sean $\alpha \pm i\beta$ dos raíces complejas de p_{n+1} , entonces

$$\frac{p_{n+1}}{(x-\alpha)^2+\beta^2} \in P_{n-1},$$



usando el Lema 1 se tiene

$$\int_{a}^{b} p_{n+1} \frac{p_{n+1}}{(x-\alpha)^{2} + \beta^{2}} \omega \, \mathrm{d}x = \int_{a}^{b} ((x-\alpha)^{2} + \beta^{2}) \left| \frac{p_{n+1}}{(x-\alpha)^{2} + \beta^{2}} \right|^{2} \omega \, \mathrm{d}x = 0,$$

lo que implica que $\beta = 0$. Por tanto, todos los ceros de p_{n+1} son reales.

Probaremos que los n + 1 ceros de p_{n+1} son simples y que pertenecen al intervalo]a, b[. En efecto, por la ortogonalidad se tiene que

$$\int_{a}^{b} p_{n+1}(x)\omega(x)\,\mathrm{d}x=0,\quad\forall n\geq 0,$$

entonces existe al menos un cero de p_{n+1} en]a, b[. Es decir, p_{n+1} puede cambiar de signo en]a, b[. Sean $x_0, x_1, \dots, x_k \in$]a, b[dichos puntos en el cual p_{n+1} cambia de signo.

- Si k = n se cumple ya que x_0, x_1, \dots, x_n son los n + 1 ceros reales simple de p_{n+1} .
- Si k < n, consideremos el polinomio

$$q(x) = (x - x_0)(x - x_1) \cdots (x - x_k).$$

Como Grad(q) = k + 1 < n + 1, por ortogonalidad se obtiene que

$$\langle p_{n+1}, q \rangle_{\omega} = 0.$$

Además, $p_{n+1}(x)q(x)$ no cambia de signo en el intervalo]a, b[, como cada cambio de signo en p_{n+1} se cancela por un cambio de signo en q(x) se sigue que

 $\langle p_{n+1}, q \rangle_{\omega} \neq 0,$

lo cual nos lleva a una contradicción.

1.1.2 Sistemas de Polinomios Ortogonales

Consideraremos aquí desde un punto de vista general, el problema de la expansión de una función en términos de una sucesión de polinomios ortogonales. Recordemos que P_N es el espacio de todos los polinomios de grado $\leq N$. Supongamos que $\{p_n\}$ es un sistema o sucesión de polinomios (con grad $(p_n) = n$) que





son ortogonales dos a dos en el intervalo]-1,1[con respecto a una función peso positiva ω :

$$\int_{-1}^{1} p_k(x) p_m(x) \omega(x) \, dx = 0 \quad \text{si } m \neq k.$$
 (1.16)

Un resultado fundamental que usaremos es el teorema de aproximación de Weierstrass que establece que cualquier función continua en un intervalo finito se puede aproximar mediante un polinomio.

Teorema 3 (Teorema de Aproximación de Weierstrass (1885)). Sea $f \in C([a, b])$. Entonces existe una sucesión de polinomios $p_n(x)$ tal que

$$\lim_{n \to \infty} p_n(x) = f(x)$$

uniformemente en [a, b].

Ver [30] para una demostración.

Este clásico teorema implica que la sucesión $\{p_n\}_{n=0,1,\cdots}$ forma una base en el espacio $L^2_{\omega}(-1,1)$. Este es el espacio de funciones medibles v con norma

$$\|v\|_{\omega} = \left(\int_{-1}^{1} |v(x)|^2 \omega(x) \mathrm{d}x\right)^{1/2} < \infty.$$

Esta norma proviene del producto interno

$$\langle u, v \rangle_{\omega} = \left(\int_{-1}^{1} u(x)v(x)\omega(x) \,\mathrm{d}x \right)^{1/2}.$$

Cuando $\omega \equiv 1$ (peso de Legendre), a menudo usamos la notación más simple $L^2(-1,1)$ en lugar de $L^2_{\omega}(-1,1)$. La serie formal de una función $u \in L^2_{\omega}(-1,1)$ en términos de la sucesión $\{p_n\}$ es

$$u(x) = \sum_{n=0}^{\infty} \widehat{u}_n p_n,$$

donde los coeficientes de expansión \widehat{u}_n se definen como

$$\widehat{u}_n = \frac{1}{\|p_n\|_{\omega}^2} \int_{-1}^1 u(x) p_n(x) \omega(x) \,\mathrm{d}x.$$



La ecuación anterior representa la **transformada polinomial** de u. Para un entero N > 0, la serie truncada de u de orden N es el polinomio

$$u_N = \sum_{n=0}^N \widehat{u}_n p_n.$$

1.1.3 Cuadraturas Gaussiana

Ahora analizaremos la relación que existe entre los polinomios ortogonales y las fórmulas de integración gaussiana. La idea principal detrás de la cuadratura gaussiana es elegir adecuadamente los nodos de interpolación en el intervalo [a, b] que maximice el grado de exactitud de la regla de cuadratura. A la familia de cua-

draturas numéricas pertenecen las integrales $\int_{a}^{b} f(x)\omega(x) dx$ que se calculan por medio de la suma

$$\int_{a}^{b} f(x)\omega(x) \,\mathrm{d}x = \sum_{j=0}^{N} f(x_{j})\omega_{j} + E_{N}[f], \qquad (1.17)$$

donde $\{x_j, \omega_j\}_{j=0}^N$ son los nodos (ceros de p_{N+1}) y pesos de cuadratura, y $E_N[f]$ es el error de cuadratura. Si $E_N[f] \equiv 0$, decimos que la fórmula de cuadratura (1.17) es exacta para f.

Ahora necesitamos conocer los pesos de cuadratura. Supongamos que los nodos $\{x_j\}_{j=0}^N$ son distintos. Si $f(x) \in C^{N+1}[a, b]$, se tiene (ver [10])

$$E_N[f] := \frac{1}{(N+1)!} \int_a^b f^{(N+1)}(\xi(x)) \prod_{i=0}^N (x-x_i) \, \mathrm{d}x, \qquad (1.18)$$

donde $\xi(x) \in [a, b]$. La base polinomial de Lagrange con $\{x_j\}_{j=0}^N$ está dada por

$$h_{j}(x) = \prod_{\substack{i=0\\i \neq j}}^{N} \frac{x - x_{i}}{x_{j} - x_{i}}, \quad 0 \le j \le N,$$
(1.19)

así que tomando $f(x) = h_j(x)$ en (1.17) y usando (1.18), se encuentran los pesos de cuadratura

$$\omega_j = \int_a^b h_j(x)\omega(x)\,\mathrm{d}x, \quad 0 \le j \le N. \tag{1.20}$$



Se dice que la regla de cuadratura mostrada en (1.17)-(1.20) tiene un grado de exactitud m, si cumple

$$E_N[p] = 0, \forall p \in P_m \text{ pero existe } q \in P_{m+1} \text{ tal que } E_N[q] \neq 0.$$
 (1.21)

En general, para N + 1 nodos distintos $\{x_j\}_{j=0}^N \subseteq]a, b[$, el grado de exactitud de (1.17)-(1.20) está entre N y 2N + 1. Además, si los nodos $\{x_k\}_{k=0}^N$ se eligen como ceros del polinomio ortogonal p_{N+1} con respecto ω , esta regla disfruta del máximo grado de exactitud 2N + 1.

Teorema 4. La aproximación (1.17) es exacta si f es un polinomio de grado $\leq 2N + 1$.

Demostración. Supongamos que *f* es un polinomio de grado $\leq 2N+1$. Entonces $f = p_{N+1}q + r$ donde *q* y *r* son polinomios de grado $\leq N$. Como p_{N+1} es ortogonal para todo los polinomios de grado $\leq N$,

$$\int_{a}^{b} f(x)\omega(x) dx = \int_{a}^{b} p_{N+1}(x)q(x)\omega(x) dx + \int_{a}^{b} r(x)\omega(x) dx = \int_{a}^{b} r(x)\omega(x) dx.$$

Además, $r(x) = \sum_{j=0}^{N} r(x_j)h_j(x)$ y $f(x_j) = p_{N+1}(x_j)q(x_j) + r(x_j) = r(x_j)$ y así
$$\int_{a}^{b} r(x)\omega(x) dx = \sum_{j=0}^{N} f(x_j)\int_{a}^{b} h_j(x)\omega(x) dx = \sum_{j=0}^{N} \omega_j f(x_j).$$

Dicha regla se conoce como la cuadratura de Gauss.

Teorema 5 (**Cuadratura de Gauss**). Sea $\{x_j\}_{j=0}^N$ el conjunto de ceros del polinomio ortogonal p_{N+1} . Entonces existe un único conjunto de pesos de cuadratura $\{\omega_j\}_{j=0}^N$ definido por (1.20) tal que

$$\int_{a}^{b} p(x)\omega(x) \, dx = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N+1},$$
(1.22)

donde los pesos de cuadratura son todos positivos y están dados por

$$\omega_j = \frac{k_{N+1}}{k_N} \frac{\|p_N\|_{\omega}^2}{p_N(x_j)p_{N+1}'(x_j)}, \quad 0 \le j \le N,$$
(1.23)

donde k_i es el coeficiente principal del polinomio p_i .



Demostración.Se
a $\{h_j: 0 \le j \le N\}$ la base polinomial de Lagrange definida en (1.19). Se sabe que

$$p(x) = \sum_{j=0}^{N} p(x_j) h_j(x), \quad \forall p \in P_N.$$

Luego,

$$\int_{a}^{b} p(x)\omega(x) \,\mathrm{d}x = \sum_{j=0}^{N} p(x_j) \int_{a}^{b} h_j(x)\omega \,\mathrm{d}x \stackrel{(1.20)}{=} \sum_{j=0}^{N} p(x_j)\omega_j, \quad (1.24)$$

lo que implica que (1.22) es exacta para cualquier $p \in P_N$, es decir, el grado de precisión de esta regla no es menor a N.

Ahora para cualquier $p \in P_{2N+1}$ se puede escribir como $p = r p_{N+1} + s$ donde $r, s \in P_N$; pero como $p_{N+1}(x_j) = 0$, se tiene $p(x_j) = s(x_j)$ para $0 \le j \le N$. Por el Lema 1, se sabe que p_{N+1} es ortogonal a $r \in P_N$ entonces

$$\int_{a}^{b} p(x)\omega(x) dx = \int_{a}^{b} [r(x)p_{N+1}(x) + s(x)]\omega(x) dx$$
$$= \int_{a}^{b} s(x)\omega(x) dx$$
$$\stackrel{(1.32)}{=} \sum_{j=0}^{N} s(x_{j})\omega_{j}$$
$$= \sum_{j=0}^{N} p(x_{j})\omega_{j}, \quad \forall p \in P_{2N+1}, \quad (1.25)$$

con lo cual se ha probado (1.22).

Probaremos que $\omega_j > 0$ para $0 \le j \le N$. Tomando $p(x) = h_j^2(x) \in P_{2N}$ en (1.25) se obtiene

$$0 < \int_a^b h_j^2(x)\omega(x) \,\mathrm{d}x = \sum_{k=0}^N h_j^2(x_k)\omega_k = \omega_j, \quad 0 \le j \le N.$$

Falta probar (1.23). Como $p_{N+1}(x_j) = 0$, tomamos $y = x_j$ y n = N en la fórmula de Christoff-Darboux (1.14) se obtiene

$$p_N(x_j)\frac{p_{N+1}(x)}{x-x_j} = \frac{k_{N+1}}{k_N} \sum_{i=0}^N \frac{\|p_N\|_{\omega}^2}{\|p_i\|_{\omega}^2} p_i(x_j) p_i(x).$$



Multiplicando por $\omega(x)$ e integrando sobre]a, b[y usando la ortogonalidad $\langle p_i, 1 \rangle = 0$, para $i \ge 1$ se tiene

$$p_N(x_j) \int_a^b \frac{p_{N+1}(x)}{x - x_j} \omega(x) \, \mathrm{d}x = \frac{k_{N+1}}{k_N} \|p_N\|_{\omega}^2 \frac{\langle p_0(x_j) p_0 \rangle_{\omega}}{\|p_0\|_{\omega}^2} = \frac{k_{N+1}}{k_N} \|p_N\|_{\omega}^2.$$
(1.26)

Además, la base polinomial de Lagrange se puede expresar $h_j(x)$ en (1.19) se puede expresar como

$$h_j(x) = \frac{p_{N+1}(x)}{p'_{N+1}(x_j)(x-x_j)}.$$
(1.27)

Reemplazando en el lado izquierdo de (1.26) se obtiene

$$p_{N}(x_{j})\int_{a}^{b} \frac{p_{N+1}(x)}{x-x_{j}}\omega(x) dx = p_{N}(x_{j})p_{N+1}'(x_{j})\int_{a}^{b} h_{j}(x)\omega(x) dx$$
$$\frac{k_{N+1}}{k_{N}} \|p_{N}\|_{\omega}^{2} = p_{N}(x_{j})p_{N+1}'(x_{j})\omega_{j}$$

con lo cual se ha probado (1.23).

El anterior teorema es fundamental ya que muestra que los nodos óptimos de la fórmula de cuadratura de Gauss son precisamente los ceros del polinomio ortogonal para el mismo intervalo y función peso. La cuadratura de Gauss es óptima debido a que se ajusta a todos los polinomios hasta de grado 2N + 1 exactamente, y es imposible encontrar cualquier conjunto de $\{x_j, \omega_j\}_{j=0}^N$ tal que (1.22) se cumpla para todo $p \in P_{2N+2}$.

A continuación, consideraremos las cuadraturas de Gauss-Radau o de Gauss-Lobatto que incluyen uno o ambos extremos como un nodo.

Comenzamos con la cuadratura de Gauss-Radau. Suponiendo que nos gustaría incluir el extremo izquierdo x = a en la cuadratura, definimos

$$q_N(x) = \frac{p_{N+1}(x) + \alpha_N p_N(x)}{x - a} \operatorname{con} \alpha_N = -\frac{p_{N+1}(a)}{p_N(a)}.$$
 (1.28)

Es obvio que $q_N \in P_N$ y para cualquier $r_{N-1} \in P_{N-1}$, obtenemos del Lema 1 que

$$\int_{a}^{b} q_{N}(x)r_{N-1}(x)\omega(x)(x-a)\,\mathrm{d}x = \int_{a}^{b} (p_{N+1}(x) + \alpha_{N}p_{N}(x))r_{N-1}(x)\omega(x)\,\mathrm{d}x = 0.$$
(1.29)

Por lo tanto, { $q_N : N \ge 0$ } define una sucesión de polinomios ortogonales con respecto a la función peso $\tilde{\omega}(x) := \omega(x)(x-a)$, y el coeficiente principal de q_N es k_{N+1} .



Teorema 6 (**Cuadratura de Gauss-Radau**). Sea $x_0 = a y \{x_j\}_{j=1}^N$ los ceros de q_N definidos en (1.28). Entonces existe un único conjunto de pesos de cuadratura $\{\omega_j\}_{j=0}^N$, definidos por (1.20), tal que

$$\int_{a}^{b} p(x)\omega(x) dx = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N}.$$
(1.30)

Además, los pesos de cuadratura son positivos y se pueden expresar como

$$\omega_0 = \frac{1}{q_N(a)} \int_a^b q_N(x) \omega(x) \, dx, \qquad (1.31a)$$

$$\omega_{j} = \frac{1}{x_{j} - a} \frac{k_{N+1}}{k_{N}} \frac{\|q_{N-1}\|_{\tilde{\omega}}^{2}}{q_{N-1}(x_{j})q_{N}'(x_{j})}, \quad 1 \le j \le N.$$
(1.31b)

Demostración. Sea $\{h_j : 0 \le j \le N\}$ la base polinomial de Lagrange definida en (1.19). Se sabe que

$$p(x) = \sum_{j=0}^{N} p(x_j) h_j(x), \quad \forall p \in P_N$$

Luego,

$$\int_{a}^{b} p(x)\omega(x) \,\mathrm{d}x = \sum_{j=0}^{N} p(x_j) \int_{a}^{b} h_j(x)\omega \,\mathrm{d}x \stackrel{(1.20)}{=} \sum_{j=0}^{N} p(x_j)\omega_j, \tag{1.32}$$

lo que implica que (1.30) es exacta para cualquier $p \in P_N$, es decir, el grado de precisión de esta regla no es menor a N.

Para cualquier $p \in P_{2N}$, se puede expresar como

 $p=(x-a)rq_N+s,\quad r\in P_{N-1},\ s\in P_N.$

Como $(x-a)q_N(x)\Big|_{x=x_j} = 0$, se tiene $p(x_j) = s(x_j)$ para $0 \le j \le N$. Usando (1.29) se llega a

$$\int_{a}^{b} p(x)\omega(x) dx = \int_{a}^{b} [(x-a)q_{N}(x)r(x) + s(x)]\omega(x) dx$$
$$= \int_{a}^{b} s(x)\omega(x) dx$$
$$= \sum_{j=0}^{N} s(x_{j})\omega_{j} = \sum_{j=0}^{N} p(x_{j})\omega_{j}, \quad \forall p \in P_{2N}.$$





Probaremos que $\omega_j > 0$ para $0 \le j \le N$. Tomando $p(x) = h_k^2(x) \in P_{2N}$ en la igualdad anterior se tiene que

$$0 < \int_a^b h_k^2(x)\omega(x) \,\mathrm{d}x = \sum_{j=0}^N h_k^2(x_j)\omega_j = \omega_k, \quad 0 \le k \le N.$$

Además, la base polinomial de Lagrange se puede expresar $h_j(x)$ en (1.19) se puede expresar como

$$h_{j}(x) = \frac{(x-a)q_{N}(x)}{((x-a)q_{N}(x))'\Big|_{x=x_{j}}(x-x_{j})}$$
$$= \frac{(x-a)q_{N}(x)}{(q_{N}(x_{j})+(x-a)q'_{N}(x_{j}))(x-x_{j})}, \quad 0 \le j \le N$$

Tomando j = 0 se obtiene (1.31a) de la definición de ω_0 , y para $1 \le j \le N$

$$\omega_j = \int_a^b h_j(x)\omega(x) \,\mathrm{d}x = \frac{1}{(x_j - a)} \int_a^b \frac{q_N(x)}{q_N'(x_j)(x - x_j)} \widetilde{\omega}(x) \,\mathrm{d}x$$

Como $\{q_n\}$ son ortogonales con respecto a $\tilde{\omega}$ por lo que la parte integral resulta ser los pesos de la cuadratura de Gauss asociado con *N* nodos que son los ceros de $q_n(x)$. Así (1.31b) se obtiene de (1.23).

Observación 1. Del mismo modo, una segunda cuadratura de Gauss-Radau se puede construir si queremos incluir el extremo derecho x = b en lugar del extremo izquierdo x = a.

Pasamos ahora a la cuadratura de Gauss-Lobatto, cuyos nodos incluyen los dos extremos x = a, b. En este caso, elegimos α_N y β_N tal que

$$p_{N+1}(x) + \alpha_N p_N(x) + \beta_N p_{N-1}(x) = 0 \text{ para } x = a, b.$$
(1.33)

y sea

$$z_{N-1}(x) = \frac{p_{N+1}(x) + \alpha_N p_N(x) + \beta_N p_{N-1}(x)}{(x-a)(b-x)}.$$
(1.34)

Es obvio que $z_{N-1} \in P_{N-1}$ y para cualquier $r_{N-2} \in P_{N-2}$, se obtiene del Lema 1 que

$$\int_{a}^{b} z_{N-1} r_{N-2}(x-a)(b-x)\omega \,\mathrm{d}x = \int_{a}^{b} (p_{N+1} + \alpha_N p_N + \beta_N p_{N-1}) r_{N-2}\omega \,\mathrm{d}x = 0.$$
(1.35)



Por lo tanto, { z_{N-1} : $N \ge 1$ } define una sucesión de polinomios ortogonales con respecto a la función peso $\widehat{\omega}(x) := (x - a)(b - x)\omega$, y el coeficiente principal de z_{N-1} es $-k_{N+1}$.

Teorema 7 (**Cuadratura de Gauss-Lobatto**). Sea $x_0 = a$, $x_N = b$ y $\{x_j\}_{j=1}^{N-1}$ los ceros de z_{N-1} definidos en (1.33)-(1.34). Entonces existe un único conjunto de pesos de cuadratura $\{\omega_j\}_{j=0}^N$, definidos por (1.20), tal que

$$\int_{a}^{b} p(x)\omega(x) \, dx = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N-1},$$
(1.36)

donde los pesos de cuadratura se expresan como

$$\omega_0 = \frac{1}{(b-a)z_{N-1}(a)} \int_a^b (b-x)z_{N-1}(x)\omega(x) \, dx, \qquad (1.37a)$$

$$\omega_j = \frac{1}{(x_j - a)(b - x_j)} \frac{k_{N+1}}{k_N} \frac{\|z_{N-2}\|_{\widehat{\omega}}^2}{z_{N-2}(x_j) z'_{N-1}(x_j)}, \quad 1 \le j \le N - 1,$$
(1.37b)

$$\omega_N = \frac{1}{(b-a)z_{N-1}(b)} \int_a^b (x-a)z_{N-1}(x)\omega(x) \, dx.$$
(1.37c)

Además, se tiene $\omega_j > 0$ para $1 \le j \le N - 1$.

Demostración. Sea $\{h_j\}_{j=0}^N$ la base polinomial de Lagrange definido en (1.19). Se sabe que

$$p(x) = \sum_{j=0}^{N} p(x_j) h_j(x), \quad \forall p \in P_N.$$

En consecuencia,

$$\int_{a}^{b} p(x)\omega(x) \, \mathrm{d}x = \sum_{j=0}^{N} p(x_j) \int_{a}^{b} h_j(x)\omega(x) \, \mathrm{d}x \stackrel{(1.20)}{=} \sum_{j=0}^{N} p(x_j)\omega_j$$

el cual implica que (1.36) es exacta para cualquier $p \in P_N$. Esto implica que el grado de precisión de esta regla no es menor o igual a N.

Luego, para cualquier $p \in P_{2N-1}$, se puede escribir

$$p = (x-a)(b-x)rz_{N-1} + s, r \in P_{N-2}, s \in P_N.$$



Como $(x-a)(b-x)z_{N-1}(x)\Big|_{x=x_j} = 0$ se tiene $p(x_j) = s(x_j)$ para $0 \le j \le N$. Por lo tanto, se deduce de (1.35) que

$$\int_{a}^{b} p(x)\omega(x) dx = \int_{a}^{b} [(x-a)(b-x)r(x)z_{N-1}(x) + s(x)]\omega(x) dx$$
$$= \int_{a}^{b} s(x)\omega(x) dx$$
$$= \sum_{j=0}^{N} s(x_{j})\omega_{j}$$
$$= \sum_{j=0}^{N} p(x_{j})\omega_{j}, \quad \forall p \in P_{2N-1}.$$

Falta justificar que $\omega_j > 0$ para $1 \le j \le N - 1$. Como los $\{z_{N-1}\}$ son ortogonales con respecto a la función peso $\widehat{\omega}$, y $z_{N-1}(x_j) = 0$ para $1 \le j \le N - 1$, se obtiene de la fórmula de Christoff-Darboux (1.15) que

$$\frac{k_N}{k_{N+1}} z_{N-2}(x_j) z'_{N-1}(x_j) = \sum_{j=0}^{N-2} \frac{\|z_{N-2}\|_{\widehat{\omega}}^2}{\|z_j\|_{\widehat{\omega}}^2} z_j^2(x_j) > 0, \quad 1 \le j \le N-1.$$

Usando esto en (1.37b) se llega a $\omega_j > 0$ para $1 \le j \le N - 1$.

Las fórmulas de cuadratura gaussiana proporcionan potentes herramientas para la evaluación de las integrales y productos internos en un método espectral. También juegan un papel importante en la diferenciación espectral, lo cual mostraremos más adelante.

1.1.4 Interpolación y Transformadas Discretas

Sea $\{x_j, \omega_j\}_{j=0}^N$ un conjunto de nodos y pesos de cuadratura de Gauss, Gauss-Radau o Gauss-Lobatto. Se define el correspondiente producto interno y norma como

$$\langle u, v \rangle_{N,\omega} \coloneqq \sum_{j=0}^{N} u(x_j) v(x_j) \omega_j, \qquad \|u\|_{N,\omega} \coloneqq \sqrt{\langle u, u \rangle_{N,\omega}}.$$
(1.38)

Observe que $\langle \cdot, \cdot \rangle_{N,\omega}$ es una aproximación del producto interno continuo $(\cdot, \cdot)_{\omega}$, y la exactitud de las fórmulas de cuadratura tipo Gauss implican

$$\langle u, v \rangle_{N,\omega} := \langle u, v \rangle_{\omega}, \quad \forall u, v \in P_{2N+\delta},$$
(1.39)



donde $\delta = 1$, 0 y -1 para la cuadratura de Gauss, Gauss-Radau y Gauss-Lobato, respectivamente.

Los nodos de las fórmulas tipo Gauss juegan un papel importante en las aproximaciones de colocación, estos son los puntos de colocación en el cual la ecuación diferencial se satisfaga. Asumiremos que la función peso ω está dada junto con la sucesión de polinomios ortogonales correspondiente { p_k }, $k = 0, 1, 2, \cdots$.

En un método de colocación, la representación fundamental de una función suave u en]-1; 1[está en términos de sus valores en los puntos discretos tipo Gauss. Las derivadas de la función se aproximan a través de derivadas analíticas del operador de interpolación.

Definición. Para cualquier $u \in C(\Lambda)$, se define el **operador de interpolación** \mathfrak{I}_n : $C(\Lambda) \rightarrow P_N$ tal que

$$(\mathfrak{I}_n u)(x_j) = u(x_j), \quad 0 \le j \le N, \tag{1.40}$$

donde $\Lambda =]a, b[, [a, b], [a, b]$ para la cuadratura de Gauss, Gauss-Radau y Gauss-Lobato, respectivamente.

 \mathfrak{I}_n está bien definido debido a que los x_j son distintos. Además, la condición de interpolación (1.40) implica que $\mathfrak{I}_n p = p$ para todo $p \in P_n$. Por otro lado, como $\mathfrak{I}_n u \in P_n$, se puede escribir

$$(\mathfrak{I}_n u)(x) = \sum_{n=0}^N \widetilde{u}_n p_n(x), \qquad (1.41)$$

lo que implica que $u(x_j) = \sum_{n=0}^{N} \tilde{u}_n p_n(x_j)$. Los \tilde{u}_n se conocen como los coeficientes polinomiales discretos de u.

Teorema 8. Usando el producto interno discreto de (1.41) con $\{p_k\}_{k=0}^N$, se pueden determinar los coeficientes $\{\tilde{u}_n\}$ como sigue

$$\widetilde{u}_n = \frac{1}{\gamma_n} \sum_{j=0}^N u(x_j) p_n(x_j) \omega_j, \quad 0 \le n \le N,$$
(1.42)

donde $\gamma_n = ||p_n||_{\omega}^2 para \ 0 \le n \le N-1, y$

$$\gamma_N = \begin{cases} \|p_N\|_{\omega}^2, & \text{para Gauss y Gauss-Radau,} \\ \langle p_N, p_N \rangle_{N,\omega}, & \text{para Gauss-Lobatto} \end{cases}$$
(1.43)



o en forma matricial

$$\begin{bmatrix} \widetilde{u}_0\\ \widetilde{u}_1\\ \vdots\\ \widetilde{u}_N \end{bmatrix} = \begin{bmatrix} \gamma_0^{-1} p_0(x_0)\omega_0 & \gamma_0^{-1} p_0(x_1)\omega_1 & \cdots & \gamma_0^{-1} p_0(x_N)\omega_N\\ \gamma_1^{-1} p_1(x_0)\omega_0 & \gamma_1^{-1} p_1(x_1)\omega_1 & \cdots & \gamma_1^{-1} p_1(x_N)\omega_N\\ \vdots & \vdots & \ddots & \vdots\\ \gamma_N^{-1} p_N(x_0)\omega_0 & \gamma_N^{-1} p_N(x_1)\omega_1 & \cdots & \gamma_N^{-1} p_N(x_N)\omega_N \end{bmatrix} \begin{bmatrix} u(x_0)\\ u(x_1)\\ \vdots\\ u(x_N) \end{bmatrix}$$

denotando por $\tilde{\mathbf{u}} = (\tilde{u}_n)_{0 \le n \le N}$, $\mathbf{p} = (\gamma_n^{-1} p_n(x_j) \omega_j)_{0 \le n, j \le N}$ y $\mathbf{u} = (u(x_n))_{0 \le n \le N}$, la forma matricial queda definida como

ũ = pu

Las fórmulas (1.42)-(1.43) definen la **transformada polinomial discreta directa**, la cual transforma los valores físicos $\{u(x_j)\}_{j=0}^N$ a los coeficientes de expansión $\{\tilde{u}_n\}_{n=0}^N$. Recíprocamente, la **transformada polinomial discreta inversa** se formula por

$$u(x_j) = (I_n u)(x_j) = \sum_{n=0}^N \tilde{u}_n p_n(x_j), \quad 0 \le j \le N,$$
(1.44)

o en forma matricial

$$\begin{bmatrix} u(x_{0})\\ u(x_{1})\\ \vdots\\ u(x_{N}) \end{bmatrix} = \begin{bmatrix} \gamma_{0}^{-1}p_{0}(x_{0})\omega_{0} & \gamma_{0}^{-1}p_{0}(x_{1})\omega_{1} & \cdots & \gamma_{0}^{-1}p_{0}(x_{N})\omega_{N}\\ \gamma_{1}^{-1}p_{1}(x_{0})\omega_{0} & \gamma_{1}^{-1}p_{1}(x_{1})\omega_{1} & \cdots & \gamma_{1}^{-1}p_{1}(x_{N})\omega_{N}\\ \vdots & \vdots & \ddots & \vdots\\ \gamma_{N}^{-1}p_{N}(x_{0})\omega_{0} & \gamma_{N}^{-1}p_{N}(x_{1})\omega_{1} & \cdots & \gamma_{N}^{-1}p_{N}(x_{N})\omega_{N} \end{bmatrix} \begin{bmatrix} \widetilde{u}_{0}\\ \widetilde{u}_{1}\\ \vdots\\ \widetilde{u}_{N} \end{bmatrix}$$

la cual queda definida $\mathbf{u} = \mathbf{p}\tilde{\mathbf{u}}$, esto significa que lleva los coeficientes de expansión $\{\tilde{u}_n\}_{n=0}^N$ a los valores físicos $\{u(x_j)\}_{j=0}^N$.

Vemos que si las matrices $(p_n(x_j))_{0 \le n, j \le N}$ o $(\gamma_n^{-1}p_n(x_j)\omega_j)_{0 \le n, j \le N}$ son precalculadas, entonces las transformadas discretas (1.42) y (1.44) se pueden manipular directamente por una multiplicación matriz-vector usual de orden $O(N^2)$. Como las transformadas discretas se usan con frecuencia en los códigos espectrales, es deseable reducir la complejidad computacional, especialmente para casos multidimensionales. En particular, la transformada rápida de Fourier (FFT) (Cooley y Tukey (1965)) y la transformada discreta de Chebyshev (tratada como una transformada de cosenos de Fourier) se puede lograr por $O(N \log_2 N)$ operaciones. Sin embargo, con la ventaja de computadoras más potentes, este aspecto no debe ser una gran preocupación para problemas de escala moderada.



1.1.5 Diferenciación en el Espacio Físico

Ahora estamos listos para hacer frente a un problema importante, las técnicas de diferenciación espectral basada en polinomios. Comencemos con la implementación en el espacio físico.

Supongamos que $u \in P_N$ es una aproximación de la solución desconocida U. Sea $\{h_j\}_{j=0}^N$ la base polinomial de Lagrange asociada con un conjunto de puntos tipo Gauss $\{x_j\}_{j=0}^N$. Es claro que,

$$u(x) = \sum_{j=0}^{N} u(x_j) h_j(x).$$
(1.45)

Luego, derivando *m* veces se tiene

$$u^{(m)}(x_k) = \sum_{j=0}^{N} h_j^{(m)}(x_k) u(x_j), \quad 0 \le k \le N.$$
(1.46)

Denotemos

$$\mathbf{u}^{(m)} := (u^{(m)}(x_0), u^{(m)}(x_1), \cdots, u^{(m)}(x_N))^T, \quad \mathbf{u} := \mathbf{u}^{(0)};$$

$$D^{(m)} := (d_{k\,i}^{(m)})_{0 \le k, \, j \le N}, \quad D := D^{(1)}.$$
(1.47)

Teorema 9. La matriz de diferenciación de orden superior se puede calcular mediante el producto de primer orden como sigue

$$D^{(m)} = D \cdot D \cdots D = D^m, \quad m \ge 1, \tag{1.48}$$

y

$$\mathbf{u}^{(m)} = D^m \mathbf{u}, \quad m \ge 1. \tag{1.49}$$

Ver [49] para una demostración.

Gracias al teorema 9, basta calcular la matriz de diferenciación de primer orden *D*. Presentamos las fórmulas explícitas para las entradas de *D*.

Teorema 10. Las entradas de D se determinan por

$$d_{kj} = h'_{j}(x_{k}) = \begin{cases} \frac{Q'(x_{k})}{Q'(x_{j})} \frac{1}{x_{k} - x_{j}}, & si \ k \neq j \\ \frac{Q''(x_{k})}{2Q'(x_{k})}, & si \ k = j \end{cases}$$
(1.50)



donde

$$Q(x) = p_{N+1}(x), \quad (x-a)q_N(x), \quad (x-a)(b-x)z_{N-1}(x)$$
(1.51)

son los polinomios de cuadratura (1.28)-(1.34) de la cuadratura de Gauss, Gauss-Radau y Gauss-Lobatto, respectivamente.

Ver [49] para una demostración.

Por lo tanto, después de haber precalculado la matriz de diferenciación de primer orden, la diferenciación en el espacio físico puede llevarse a cabo a través de la multiplicación matriz-matriz y matriz-vector.

1.1.6 Diferenciación en el Espacio Frecuencia

La diferenciación en el espacio de frecuencia radica en expresar los coeficientes de expansión de las derivadas de una función en términos de los coeficientes de expansión de la misma función. Más precisamente, dada $u \in \mathbb{P}_N$, en lugar de utilizar la base polinomial de Lagrange, se expande u en términos de los polinomios ortogonales:

$$u(x) = \sum_{j=0}^{N} \hat{u}_n p_n(x).$$
 (1.52)

Usando la propiedad de ortogonalidad, se pueden determinar los coeficientes por

$$\widehat{u}_n = \frac{1}{\|p_n\|_{\omega}^2} \int_a^b u(x) p_n(x) \omega(x) \, dx, \quad 0 \le n \le N.$$
(1.53)

Como $u' \in \mathbb{P}_{N-1}$, se tiene

$$u'(x) = \sum_{n=0}^{N} \widehat{u}_n^{(1)} p_n(x) \qquad \text{con } \widehat{u}_N^{(1)} = 0.$$
(1.54)

Con el fin de expresar $\{\widehat{u}_n^{(1)}\}_{n=0}^N$ en términos $\{\widehat{u}_n\}_{n=0}^N$, se asume que los $\{p'_n\}$ también son ortogonales. En efecto, esta propiedad se cumple para los polinomios ortogonales clásicos tales como los polinomios de Legendre, Chebyshev, Jacobi y Hermite. En otras palabras, $\{p'_n\}$ satisface la relación de recurrencia de tres términos debido al Corolario 1:

$$p_{n+1}'(x) = (a_n^{(1)}x - b_n^{(1)})p_n'(x) - c_n^{(1)}p_{n-1}'(x).$$
(1.55)



Derivando la regla de recurrencia de tres términos (1.12) y usando (1.55), se obtiene

$$p_n(x) = \tilde{a}_n p'_{n-1}(x) + \tilde{b}_n p'_n(x) + \tilde{c}_n p'_{n+1}(x).$$
(1.56)

Teorema 11. Los coeficientes $\{\hat{u}_n^{(1)}\}$ en (1.54) pueden calcularse mediante la fórmula de recurrencia inversa (1.57) como sigue

$$\widehat{u}_{n-1}^{(1)} = \frac{1}{\widetilde{c}_{n-1}} [\widehat{u}_n - \widetilde{b}_n \widehat{u}_n^{(1)} - \widetilde{a}_{n+1} \widehat{u}_{n+1}^{(1)}], \quad n = N - 1, \cdots, 1,
\widehat{u}_N^{(1)} = 0, \quad \widehat{u}_{N-1}^{(1)} \frac{1}{\widetilde{c}_{N-1}} \widehat{u}_N.$$
(1.57)

Ver [49] para una demostración.

Las derivadas de orden superior en el espacio de la frecuencia se pueden llevar a cabo mediante el uso de la fórmula (1.57) repetidas veces. Es importante señalar que la **diferenciación espectral** junto con las **transformadas discretas** forman los ingredientes básicos para la denominada **"técnica pseudo-espectral"** (particularmente útil para problemas no lineales); las **derivadas** son manipuladas en el espacio frecuencia, los productos internos se calculan en el espacio fisico, y los dos espacios se comunican a través de las transformadas discretas.

1.2 Polinomios de Chebyshev

Los polinomios de Chebyshev de primer tipo, $T_k(x)$, $k = 0, 1, \dots$, son las autofunciones del problema singular de Sturm-Liouville

$$-\left(\sqrt{1-x^2}T'_k(x)\right)' + \frac{n^2}{\sqrt{1-x^2}}T_k(x) = 0, \qquad (1.58)$$

o equivalentemente,

$$(1-x^2)T_k''(x) - xT_k'(x) + n^2T_k(x) = 0.$$
(1.59)

Sean P_N el espacio de polinomios de grado no mayor a $N \in \mathbb{N}, \, N > 0, \, \mathbf{y}$ la función peso

$$\omega(x) := \frac{1}{\sqrt{1 - x^2}}.$$

Sea I = [-1; 1], definimos el espacio $L^2_{\omega}(I)$ por

 $L^{2}_{\omega}(I) := \{ v : I \to \mathbb{R} / v \text{ medible Lebesgue y } \|v\|_{\omega} < \infty \},$



donde la norma

$$\|v\|_{\omega} := \left(\int_{-1}^{1} |v(x)|^2 \omega(x) \, dx\right)^{1/2},$$

está inducida por el producto escalar (interno) ponderado

$$\langle u, v \rangle_{\omega} := \left(\int_{-1}^{1} u(x)v(x)\omega(x) \, dx \right). \tag{1.60}$$

Definición. Los polinomios $T_n(x)$, $n \in \mathbb{N}$, definidos por

 $T_n(x) := \cos(n\theta), \ \theta = \arccos(x) \quad x \in [-1, 1],$

son llamados los polinomios de Chebyshev de primer tipo.

Los polinomios de Chebyshev son funciones coseno después de un cambio en la variable independiente. Esta propiedad es el origen de su gran popularidad en la aproximación numérica de problemas de valores en la frontera no periódicas. La transformación $x = \cos \theta$ permite muchas relaciones matemáticas como resultados teóricos relativos al sistema de Fourier que se adaptan fácilmente al sistema de Chebyshev.

Observación 2. Para establecer una relación entre polinomios algebraicos y trigonométricos recurriremos a la identidad trigonométrica

$$\cos(n\theta) + i \sin(n\theta) = (\cos\theta + i \sin\theta)^n$$
$$= \cos^n \theta + i \left(\frac{n}{1}\right) \cos^{n-1} \theta \sin\theta + i^2 \left(\frac{n}{2}\right) \cos^{n-2} \theta \sin^2 \theta + \cdots$$

Los términos del lado derecho que usan potencias pares de sen θ son reales mientras que los de potencia impar de sen θ son imaginarios. Además, se sabe que sen^{2m} $\theta = (1 - \cos^2 \theta)^m$, $m \in \mathbb{N}$. En consecuencia, para cualquier *n* natural se puede escribir

$$T_n(\cos\theta) = \cos(n\theta), \qquad n \ge 0$$

donde $T_n(x) := \cos(n \arccos(x)) = \alpha_0^{(n)} + \alpha_1^{(n)}x + \dots + \alpha_n^{(n)}x^n$ es el polinomio de Chebyshev de orden (grado) *n* el cual es un polinomio algebraico de grado *n* con


coeficientes reales. Se sabe que,

$$T_{0}(x) = 1,$$

$$T_{1}(x) = \cos(\theta) = x,$$

$$T_{2}(x) = \cos(2\theta) = 2x^{2} - 1,$$

$$T_{3}(x) = \cos(3\theta) = 4x^{3} - 3x,$$

$$T_{4}(x) = \cos(4\theta) = 8x^{4} - 8x^{2} - 1, \cdots$$

(1.61)

En general, usando la fórmula de Moivre, tenemos

$$\cos(n\theta) = \operatorname{Re}\{(\cos(\theta) + i \operatorname{sen}(\theta))^n\}$$

y luego aplicando la fórmula binomial, podemos expresar T_n como

$$T_n(x) = \frac{n}{2} \sum_{k=0}^{\lfloor \frac{k}{2} \rfloor} (-1)^n \frac{(n-k-1)!}{k!(n-2k)!} (2x)^{n-2k}$$
(1.62)

donde [w] denota la parte entera de w.

De la identidad trigonométrica

$$\cos(k+1)\theta + \cos(k-1)\theta = 2\cos\theta\cos(k\theta)$$

se deduce la relación de recurrencia

$$T_{k+1}(x) = 2x T_k(x) - T_{k-1}(x), \quad k > 1, \quad T_0(x) = 1, \quad T_1(x) = x.$$
 (1.63)

la cual nos permite, en particular, deducir la expresión de los polinomios T_k , $k \ge 2$, a partir de T_0 y T_1 . La expresión (1.62) puede ser útil en circunstancias especiales pero la representación $T_n(x) = \cos(n\theta)$ es generalmente usada en computación así como en estudios teóricos.

Proposición 1. [Ortogonalidad] Los polinomios $T_n(x)$ son ortogonales en [-1;1] con peso $w(x) = (1-x^2)^{-1/2}$, es decir,

$$\langle T_n, T_m \rangle_{\omega} = \int_{-1}^{1} T_n(x) T_m(x) \frac{1}{\sqrt{1-x^2}} \, dx = \frac{\pi}{2} c_n \delta_{n,m}, \quad n, m \in \mathbb{N}$$

donde el símbolo $\delta_{n,m}$ representar la delta de Kronecker y a lo largo de este trabajo, los coeficientes c_n se definen por

$$c_n := \begin{cases} 2, & n = 0\\ 1, & n > 0. \end{cases}$$
(1.64)



Demostración. En base al cambio de variable $z = \arccos(x)$ tenemos los siguientes casos

 $n \neq k$ $\int_{0}^{\pi} \cos nz \cos kz \, dz = \frac{1}{2} \int_{0}^{\pi} [\cos(n+k)z + \cos(n-k)z] \, dz$ $= \frac{1}{2} \Big[\frac{\sin(n+k)z}{n+k} + \frac{\sin(n-k)z}{n-k} \Big]_{0}^{\pi} = 0$ $n = k \neq 0$ $\int_{0}^{\pi} \cos^{2} nz \, dz = \frac{1}{2} \int_{0}^{\pi} [1 + \cos 2nz] \, dz$ $= \frac{1}{2} \Big[z + \frac{\sin 2nz}{2n} \Big]_{0}^{\pi} = \frac{\pi}{2}$ n = k = 0 $\int_{0}^{\pi} dz = \pi$

que es lo que se quería probar.

A continuación se indican algunas propiedades útiles para la aplicación de los polinomios de Chebyshev a la solución de ecuaciones diferenciales ordinarias y parciales.

$$T_k(-x) = (-1)^k T_k(x)$$
(1.65a)

$$T_k(\pm 1) = (\pm 1)^k,$$
 (1.65b)

$$T'_{k}(\pm 1) = (\pm 1)^{k+1} k^{2}, \quad |T_{k}(x)| \le 1,$$
 (1.65c)

$$|T'_k(x)| \le k^2, \quad -1 \le x \le 1$$
 (1.65d)

$$(1-x^2)T'_k(x) = \frac{k}{2}T_{k-1}(x) - \frac{k}{2}T_{k+1}(x), \qquad (1.65e)$$

$$2T_k(x)T_n(x) = T_{|k+n|}(x) + T_{|k-n|}(x), \qquad (1.65f)$$

$$\int_{-1}^{1} T'_{n}(x)T'_{m}(x)\sqrt{1-x^{2}}\,\mathrm{d}x = \frac{n^{2}c_{n}\pi}{2}\delta_{n,m}, \quad n,m \in \mathbb{N}$$
(1.65g)



El conocimiento de estas propiedades es útil cuando se usan condiciones de frontera. Es importante observar que

$$T_k(-x) = (-1)^k T_k(x) \tag{1.66}$$

lo cual lleva a una propiedad de paridad del polinomio de grado k, es decir, $T_k(x)$ es par si k es par, e impar si k es impar. El polinomio $T_k(x)$ se anula en los puntos x_i (puntos de Gauss) definidos por

$$x_i = \cos\left(i + \frac{1}{2}\right)\frac{\pi}{k}, \quad i = 0, \cdots, k-1.$$
 (1.67)

y alcanza sus valores extremos ± 1 en los puntos \overline{x}_i (puntos de Gauss-Lobatto) definidos por

$$\overline{x}_i = \cos\frac{\pi i}{k}, \quad i = 0, \cdots, k.$$
(1.68)

Una relación de recurrencia de la derivada de $T_k(x)$ es

$$\frac{T'_{k+1}(x)}{k+1} - \frac{T'_{k-1}(x)}{k-1} = 2T_k(x)$$
(1.69)

válida para k > 1. Una fórmula similar para la p-ésima derivada se obtiene por derivación sucesiva de (1.69).

De (1.69) se obtiene

$$T_k'(x) = 2k \sum_{\substack{n=0\\n+k \text{ impar}}}^{k-1} \frac{1}{c_n} T_n(x), \qquad (1.70a)$$

$$T_k''(x) = \sum_{\substack{n=0\\n+k \text{ par}}}^{k-2} \frac{1}{c_n} k(k^2 - n^2) T_n(x).$$
(1.70b)

Observación 3. Por el teorema de aproximación de Weierstrass, el conjunto de polinomios ortogonales $\{T_k(x)\}_{k\in\mathbb{N}}$ también es completo en el espacio $L^2_{\omega}(I)$, y en consecuencia, toda función $u \in L^2_{\omega}(-1, 1)$ se puede expresar en una serie de Chebyshev como sigue

$$u(x) = \sum_{k=0}^{\infty} \widehat{u}_k T_k(x), \qquad (1.71)$$

donde los coeficientes \hat{u}_k son

$$\widehat{u}_k = \frac{\langle u, T_k \rangle_{\infty}}{\|T_k\|_{\omega}^2} = \frac{2}{\pi c_k} \langle u, T_k \rangle_{\omega} = \frac{2}{\pi c_k} \int_{-1}^1 u(x) T_k(x) w(x) \, dx.$$



Otra consecuencia de los polinomios de Chebyshev es que los nodos y pesos de cuadratura se pueden obtener explícitamente.

Teorema 12. Sea $\{x_j, \omega_j\}_{j=0}^N$ un conjunto de nodos y pesos de cuadratura tipo *Chebyshev-Gauss.*

• Para cuadratura Chebyshev Gauss (CG).

$$x_j = \cos\frac{(2j+1)\pi}{2N+2}, \quad w_j = \frac{\pi}{N+1}, \quad j = 0, 1, \cdots, N.$$
 (1.72)

• Para cuadratura Chebyshev Gauss-Radau (CGR).

$$x_{j} = \cos \frac{2\pi j}{2N+1}, \quad w_{j} = \begin{cases} \frac{\pi}{2N+1}, & j = 0\\ \frac{2\pi}{2N+1}, & j = 1, \cdots, N. \end{cases}$$
(1.73)

• Para cuadratura Chebyshev Gauss-Lobatto (CGL).

$$x_{j} = \cos \frac{\pi j}{N}, \quad w_{j} = \begin{cases} \frac{\pi}{2N}, & j = 0, N \\ \frac{\pi}{N}, & j = 1, \cdots, N-1. \end{cases}$$
 (1.74)

Se debe tener en cuenta que los nodos de cuadratura de Chebyshev como se acaban de definir se ordenan de derecha a izquierda. Esto viola la convención general de que los nodos de cuadratura estén ordenados de izquierda a derecha. Prácticamente la totalidad de la literatura clásica sobre los métodos espectrales de Chebyshev utiliza este orden inverso. Por lo tanto, en el caso especial de los nodos de cuadratura de Chebyshev vamos a tomar la convención de orden que se utiliza ampliamente. Nos damos cuenta que la solución de este dilema es solo de invertir mentalmente el orden de los nodos de Chebyshev cada vez que se utilicen las fórmulas generales para los polinomios ortogonales.

1.2.1 Cálculo de los Coeficientes de Chebyshev

Consideremos la aproximación de Chebyshev de la función u(x) definida para $x \in [-1; 1]$:

$$u_N(x) = \sum_{k=0}^{N} \hat{u}_k T_k(x).$$
(1.75)



Los coeficientes de expansión \hat{u}_k , $k = 0, \dots, N$ se determinan mediante la técnica de Galerkin, la cual consiste en que el residuo $R_N(x) = u(x) - u_N(x)$ se anula en el sentido del promedio ponderado

$$\langle R_N(x), T_l(x) \rangle_{\omega} = 0, \qquad l = 0, \cdots, N,$$
 (1.76)

es decir,

$$\int_{-1}^{1} \left(u(x)T_l(x)\omega(x) - \sum_{k=0}^{N} \widehat{u}_k T_k(x)T_l(x)\omega(x) \right) dx = 0, \qquad l = 0, \cdots, N.$$

Luego, tomando en cuenta la condición de ortogonalidad (Proposición 1), se obtiene la expresión para los coeficientes de expansión de Chebyshev

$$\widehat{u}_{k} = \frac{2}{\pi c_{k}} \int_{-1}^{1} u(x) T_{k}(x) \omega(x) \, dx.$$
(1.77)

Además, se puede expresar (1.75) por medio de la representación $T_k = \cos kz$ con $x = \cos z$ como sigue

$$u_N(x) = \sum_{k=0}^{N} \widehat{u}_k \cos kz, \qquad (1.78)$$

probando que la expansión (1.75) con respecto a x es equivalente a una serie de Fourier de cosenos en z.

1.2.2 Diferenciación en el Espacio Frecuencia

La obtención de la derivada en la base de Chebyshev es más complicada que en la de Fourier. De hecho, la expresión de la derivada de $T_k(x)$ implica todos los polinomios de paridad opuesta y de grado inferior, mientras que la derivada de e^{ikx} es simplemente ike^{ikx} . Esto hace que los aspectos computacionales de las dos aproximaciones sean diferentes: las matrices de diferenciación de Chebyshev en los espacios espectral y físico son no esparcidas, mientras que las matrices de Fourier análogas son no esparcidas sólo en el espacio físico.

De la relación de recurrencia (1.69), se obtiene

$$T'_{k}(x) = 2k \sum_{n=0}^{M} \frac{1}{c_{k-1-2n}} T_{k-1-2n}(x),$$



donde $M = \left\lfloor \frac{k-1}{2} \right\rfloor$. Por lo tanto, considerando la primera derivada de la función, se tiene que

$$u_N'(x) = \sum_{k=0}^N \widehat{u}_k T_k'(x) = \sum_{k=1}^N \widehat{u}_k^{(1)} T_k(x).$$
(1.79)

Usando (1.69) en la ecuación anterior, llegamos a la siguiente expresión

$$2k\,\widehat{u}_{k} = c_{k-1}\,\widehat{u}_{k-1}^{(1)} - \widehat{u}_{k+1}^{(1)}, \quad 1 \le k \le N.$$
(1.80)

La relación anterior nos va a permitir encontrar los coeficientes \hat{u}_k en función de los $\hat{u}_k^{(1)}$. La ecuación (1.80) la obtenemos a partir de integrar (1.79)

$$u_N(x) = \int u'_N(x) \, \mathrm{d}x = \int (\widehat{u}_0^{(1)} T_0 + \widehat{u}_1^{(1)} T_1 + \dots + \widehat{u}_N^{(1)} T_N) \, \mathrm{d}x$$

considerando

$$T_0(x) = T'_1(x)$$

 $T_1(x) = \frac{1}{4}T'_2(x)$

y usando la ecuación (1.69)

$$u_N(x) = \widehat{u}_0^{(1)} T_1 + \frac{\widehat{u}_1^{(1)}}{2} \frac{T_2}{2} + \frac{\widehat{u}_2^{(1)}}{2} \left(\frac{T_3}{3} - T_1\right) + \frac{\widehat{u}_3^{(1)}}{2} \left(\frac{T_4}{4} - \frac{T_2}{2}\right) + \dots + \\ = \frac{\widehat{u}_{k-1}^{(1)}}{2} \left(\frac{T_k}{k} - \frac{T_{k-2}}{k-2}\right) + \frac{\widehat{u}_k^{(1)}}{2} \left(\frac{T_{k+1}}{k+1} - \frac{T_{k-1}}{k-1}\right) + \frac{\widehat{u}_{k+1}^{(1)}}{2} \left(\frac{T_{k+2}}{k+2} - \frac{T_k}{k}\right) \\ = \dots$$

Si agrupamos los coeficientes del polinomio T_1 , se obtiene \hat{u}_1 , ahora si agrupamos los coeficientes del polinomio T_2 , se obtiene \hat{u}_2 , si agrupamos los coeficien-



tes del polinomio T_n , se obtiene \hat{u}_n , es decir

$$\begin{array}{lll} T_{1} \rightarrow & \widehat{u}_{1} = \frac{1}{2} (2 \, \widehat{u}_{0}^{(1)} - \widehat{u}_{2}^{(1)}) & \Rightarrow & 2 \, \widehat{u}_{1} = 2 \, \widehat{u}_{0}^{(1)} - \widehat{u}_{2}^{(1)} \\ T_{2} \rightarrow & \widehat{u}_{2} = \frac{1}{2} \left(\frac{\widehat{u}_{1}^{(1)}}{2} - \frac{\widehat{u}_{3}^{(1)}}{2} \right) & \Rightarrow & 4 \, \widehat{u}_{2} = \widehat{u}_{1}^{(1)} - \widehat{u}_{3}^{(1)} \\ \vdots & & \\ T_{N-2} \rightarrow & \widehat{u}_{N-2} = \frac{1}{2} \left(\frac{\widehat{u}_{N-3}^{(1)}}{N-2} - \frac{\widehat{u}_{N-1}^{(1)}}{N-2} \right) & \Rightarrow & 2(N-2) \, \widehat{u}_{N-2} = \widehat{u}_{N-3}^{(1)} - \widehat{u}_{N-1}^{(1)} \\ T_{N-1} \rightarrow & \widehat{u}_{N-1} = \frac{1}{2} \left(\frac{\widehat{u}_{N-2}^{(1)}}{N-1} \right) & \Rightarrow & 2(N-1) \, \widehat{u}_{N-1} = \widehat{u}_{N-2}^{(1)} \\ T_{N} \rightarrow & \widehat{u}_{N} = \frac{1}{2} \left(\frac{\widehat{u}_{N-1}^{(1)}}{N} \right) & \Rightarrow & 2N \, \widehat{u}_{N} = \widehat{u}_{N-1}^{(1)} \end{array}$$

donde $\hat{u}_N^{(1)} = \hat{u}_{N+1}^{(1)} = 0$. La relación \hat{u}_0 no se obtiene de la misma manera, ya que aparece en la constante de integración.

Usando las ecuaciones (1.69) y (1.80), se deducen los coeficientes $\widehat{u}_k^{(1)}$:

$$\widehat{u}_{k}^{(1)} = \frac{2}{c_{k}} \sum_{p=k+1}^{N} p \,\widehat{u}_{p}, \quad k = 0, \cdots, N-1$$
(1.81)

con (p+k) impar y $\widehat{u}_N^{(1)} = 0$. Esto se puede escribir en forma matricial como $\widehat{U}^{(1)} = \widehat{\mathscr{D}}\widehat{U}$,

donde $\widehat{U} = (\widehat{u}_0, \dots, \widehat{u}_N)^T$, $\widehat{U}^{(1)} = (\widehat{u}_0^{(1)}, \dots, \widehat{u}_N^{(1)})^T$ y $\widehat{\mathcal{D}}$ es una matriz triangular superior y sus entradas se deducen de (1.81).

La expansión de la segunda derivada de la función u_N es

$$u_N''(x) = \sum_{k=0}^N \widehat{u}_k^{(2)} T_k(x).$$

con

$$\widehat{u}_{k}^{(2)} = \frac{1}{c_{k}} \sum_{p=k+2}^{N} p(p^{2} - k^{2}) \widehat{u}_{p}, \quad k = 0, \cdots, N-2$$
(1.82)

 $\mathrm{con}\,(p+k)\,\mathrm{par}\,\mathrm{y}\,\widehat{u}_{N-1}^{(2)}=\widehat{u}_{N}^{(2)}=0.$ Expres ando $\widehat{u}_{k}^{(2)}\,\mathrm{en}$ forma matricial

$$\widehat{U}^{(2)} = \widehat{\mathscr{D}}^2 \widehat{U},$$



donde $\widehat{U}^{(2)} = (\widehat{u}_0^{(2)}, \cdots, \widehat{u}_N^{(2)})^T$.

Las expresiones analíticas (1.81) y (1.82) son de interés cada vez que los coeficientes de expansión de las derivadas estén implicados en los cálculos. Por otro lado, si sólo se necesitan los valores numéricos de los coeficientes, se pueden calcular del producto matricial, o a partir de fórmulas de recurrencia deducidas a partir de (1.69). En general, la expresión para $T_k(x)$ dada por (1.69) se reemplaza en $T'_k(x)$. Luego, mediante la identificación de la derivada T'_k con el mismo índice, se obtiene la fórmula de recurrencia para la primera derivada. En general, la fórmula de recurrencia para los coeficientes $\hat{u}_k^{(p)}$ para la p—ésima derivada se obtiene por diferenciación sucesiva,

$$c_{k-1} \widehat{u}_{k-1}^{(p)} = \widehat{u}_{k+1}^{(p)} + 2k \, \widehat{u}_{k}^{(p-1)}, \quad 1 \leq k \leq N-p+1,$$

donde $\hat{u}_{N-p+2}^{(p)} = \hat{u}_{N-p+3}^{(p)} = 0$. Esta relación de recurrencia nos permitirá obtener los coeficientes $\hat{u}_k^{(p)}$ de los coeficientes $\hat{u}_k^{(p-1)}$. Para la primera derivada

$$\widehat{u}_N^{(1)} = 0, \qquad \widehat{u}_{N-1}^{(1)} = 2N\,\widehat{u}_N,$$

y para la segunda derivada

$$\widehat{u}_{N-1}^{(2)} = \widehat{u}_N^{(2)} = 0, \qquad \widehat{u}_{N-2}^{(2)} = 2(N-1)\widehat{u}_{N-1}^{(1)} = 2N(N-1)\widehat{u}_N.$$

1.2.3 Diferenciación en el Espacio Físico

La aplicación de métodos de colocación para la solución de ecuaciones diferenciales puede considerar como incógnitas a los coeficientes de expansión, así como los valores de la cuadrícula. El primer método rara vez se emplea en el caso de Chebyshev ([34]). Por otro lado, el segundo enfoque (también conocido como "colocación ortogonal", [15]) es de uso actual en la mecánica de fluidos computacional como en las obras de Wengle ([57]) para la ecuación de advección-difusión y de Orszag y Patera ([40]) o Ouazzani y Peyret ([41]) para las ecuaciones de Navier-Stokes.

Por lo tanto, en el contexto del método de colocación en donde las incógnitas son los valores de la cuadrícula, es necesario expresar las derivadas en cualquier punto de colocación en términos de los valores de la cuadrícula de la función, es decir, para la p-ésima derivada $\hat{u}_N^{(p)}$:

$$\widehat{u}_N^{(p)}(x_i) = \sum_{j=0}^N d_{ij}^{(p)} u_N(x_j), \quad i = 0, \cdots, N.$$



De acuerdo a la Subsección 1.1.5, se tiene

$$\mathbf{u}^{(p)} = D^p \mathbf{u}, \quad p \ge 1,$$

donde

$$D = (d_{ij} = h'_i(x_i))_{0 \le i, j \le N}, \quad \mathbf{u}^{(p)} = (u^{(p)}(x_0), \cdots, u^{(p)}(x_N)), \quad \mathbf{u} = \mathbf{u}^{(0)}.$$

Las entradas de la matriz de diferenciación de primer orden *D* se pueden determinar por las fórmulas explícitas siguientes:

• Para el caso de Chebyshev-Gauss-Lobatto ($x_0 = -1$ y $x_N = 1$), la matriz de la primera derivada es:

$$d_{ij} = \begin{cases} \frac{2N^2 + 1}{6}, & i = j = 0, \\ \frac{\tilde{c}_i}{\tilde{c}_j} \frac{(-1)^{i+j}}{x_i - x_j}, & i \neq j, \ 0 \le i, j \le N, \\ -\frac{x_j}{2(1 - x_j^2)}, & 1 \le i = j \le N - 1, \\ -\frac{2N^2 + 1}{6}, & i = j = N, \end{cases}$$

y la matriz de la segunda derivada es

$$d_{ij}^{(2)} = \begin{cases} \frac{(-1)^{i+j}}{\widetilde{c}_j} \frac{x_i^2 + x_i x_j - 2}{(1 - x_i^2)(x_i - x_j)^2}, & 1 \le i \le N - 1, \ 0 \le j \le N, \ i \ne j, \\ \frac{(N^2 - 1)(1 - x_i^2) + 3}{3(1 - x_i^2)^2}, & 1 \le i = j \le N - 1, \\ \frac{2}{3} \frac{(-1)^j}{\widetilde{c}_j} \frac{(2N^2 + 1)(1 - x_j) - 6}{(1 - x_i)^2}, & i = 0, \ 1 \le j \le N, \\ \frac{2}{3} \frac{(-1)^{j+N}}{\widetilde{c}_j} \frac{(2N^2 + 1)(1 + x_j) - 6}{(1 + x_j)^2}, & i = N, \ 0 \le j \le N - 1, \\ \frac{N^4 - 1}{15}, & i = j = 0, \ i = j = N. \end{cases}$$

donde $\tilde{c}_0 = \tilde{c}_N = 2$ y $\tilde{c}_j = 1$ para $1 \le j \le N - 1$. Para otros puntos de colocación las matrices cambian.



Es útil recordar que

$$d_{ij}^{(2)} = \sum_{k=0}^{N} d_{ik}^{(1)} d_{kj}^{(1)}.$$

En forma vectorial, las derivadas se pueden expresar como

$$U^{(1)} = \mathscr{D}U, \qquad U^{(2)} = \mathscr{D}^2U$$

donde

$$U = (u_N(x_0), u_N(x_1), \cdots, u_N(x_N))^T, \qquad U^{(p)} = (u_N^{(p)}(x_0), u_N^{(p)}(x_1), \cdots, u_N^{(p)}(x_N))^T$$

con p=1,2. Por último, la matriz de diferenciación ${\mathcal D}$ se define por





2

Métodos Espectrales Polinomiales

En este capítulo, se presenta la formulación de métodos espectrales polinomiales para resolver ecuaciones diferenciales parciales, utilizando los métodos de Galerkin, Tau y de Colocación. La presencia de condiciones de frontera no periódicas nos da la motivación para el uso de métodos polinomiales. Solo trataremos problemas que necesiten soluciones suaves.

Los métodos espectrales se usan para resolver numéricamente ecuaciones diferenciales de la forma

$$\frac{\partial u}{\partial t}(x,t) = \mathfrak{L}u(x,t), \quad x \in D, \ t \ge 0$$

$$\mathfrak{B}u(x,t) = 0, \quad x \in \partial D, \ t > 0$$

$$u(x,0) = f(x), \quad x \in D,$$
(2.1)

donde *D* es el dominio espacial y ∂D su frontera, \mathfrak{L} es un operador diferencial (que puede ser no lineal), \mathscr{B} representa las condiciones de frontera y f(x) es la condición inicial.

La idea básica de los métodos espectrales es la búsqueda de una solución aproximada al problema, es decir, usa una serie truncada de la forma

$$u_N(x,t) = \sum_{k=0}^{N} a_k(t) \Phi_k(x).$$
 (2.2)



donde $\Phi_k(x)$ son funciones ortogonales suaves conocidas. Luego, debemos encontrar los coeficientes a_k , con los cuales tendremos la solución, y no se tendría ninguna discretización espacial, esto es, la solución se tendría en **todo el dominio espacial**. Esta característica contrasta con el método de diferencias finitas, donde la solución solo se da en los puntos de la malla. Existen varias formas de obtener los coeficientes, y cada manera de encontrarlos define los diferentes esquemas de métodos espectrales. Existen muchas formas de elegir el conjunto ortogonal usado en la expansión (2.2), las cuales dependen mucho de las condiciones de frontera.

- Para las funciones de frontera periódica, la elección natural de la base de funciones es $\Phi_k(x) = e^{ikx}$ (método espectral de Fourier).
- Para condiciones de frontera no periódica, se usan los polinomios de Chebyshev $\Phi_k(x) = T_k(x)$ (método espectral de Chebyshev) o los polinomios de Legendre $\Phi_k(x) = L_k(x)$ (método espectral de Legendre).
- Para aproximar en intervalos semi-infinitos, x ∈ [0; +∞[, se usan los polinomios de Laguerre Φ_k(x) = ℒ_k(x) (método espectral de Laguerre).
- Para aproximar en la recta real, x ∈]−∞;+∞[, se usan los polinomios de Hermite Φ_k(x) = ℋ_k(x) (método espectral de Hermite).

La elección de las funciones de prueba distingue las formulaciones siguientes:

- Galerkin. Las funciones de prueba son las mismas que las funciones base, es decir, $\phi_k = \psi_k$, ψ_k satisface las condiciones de frontera sean periódicas o no.
- Tau. $\phi_k = \psi_k$, pero ψ_k no satisface las condiciones de frontera.
- **Colocación.** Las funciones de prueba ψ_k son los polinomios base de Lagrange tal que $\psi_k(x_j) = \delta_{jk}$, donde $\{x_j\}$ son los puntos de colocación preasignados. Se requiere que el residuo sea cero en $\{x_i\}$.

En situaciones más generales, los polinomios de Chebyshev son una mejor opción. Los polinomios de Chebyshev, se amoldan muy bien a condiciones de frontera arbitrarias, que en diferencias finitas sería difícil de implementar.



2.1 Método de Galerkin

El método de Galerkin es un método donde las funciones base se eligen de forma tal que la solución propuesta satisfaga las condiciones de frontera. Cuando desarrollamos una función continua y derivable en su dominio en términos de una serie truncada de un conjunto ortogonal { Φ_n }, los coeficientes de la expansión decrecen muy rápidamente. En este método se busca soluciones, $u_N(x, t) \in B_N$ de la forma

$$u_N(x,t) = \sum_{k=0}^N a_k(t)\Phi_k(x),$$

donde $\Phi_k(x)$ es un polinomio tomado del espacio

$$B_N = \operatorname{Gen}\{\Phi_k(x) \in \operatorname{Gen}\{x^n, n = 0, \dots, k\} / \mathscr{B}\Phi_k = 0, k = 0, \dots, n\}.$$

Al reemplazar la expansión $u_N(x,t) = \sum_{k=0}^N a_k(t)\Phi_k(x)$ en la ecuación, se tiene el residuo R(x,t) que evidentemente no se anula

$$R(x,t) = \frac{\partial u_N}{\partial t}(x,t) - \mathcal{L} u_N(x,t) \neq 0.$$
(2.3)

Usando el producto interno

$$r_n = \langle R, \Phi_n(x) \rangle_{\omega} = \int R(x, t) \Phi_n(x) \omega(x) \, dx, \qquad (2.4)$$

se pide que el residuo sea pequeño; es decir, para todo $n = 0, 1, \dots, N$,

$$\langle \Phi_n(x), R(x,t) \rangle_{\omega} = 0 \Leftrightarrow \left\langle \Phi_n, \frac{\partial u_N}{\partial t}(x,t) - \mathcal{L} u_N(x,t) \right\rangle_{\omega} = 0$$

$$\Leftrightarrow \left\langle \Phi_n(x), \frac{\partial}{\partial t} \sum_{k=0}^N a_k(t) \Phi_k(x) \right\rangle_{\omega} - \left\langle \Phi_n(x), \mathcal{L} \sum_{k=0}^N a_k(t) \Phi_k(x) \right\rangle_{\omega} = 0$$

$$\Leftrightarrow \sum_{k=0}^N \frac{\mathrm{d}}{\mathrm{d}t} a_k(t) \langle \Phi_n, \Phi_k(x) \rangle_{\omega} - \sum_{k=0}^N a_k(t) \langle \Phi_n, \mathcal{L} \Phi_k(x) \rangle_{\omega} = 0$$

$$\Leftrightarrow \sum_{k=0}^N M_{kn} \frac{\mathrm{d}a_k}{\mathrm{d}t} = \sum_{k=0}^N S_{kn} a_k(t), \ \forall n = 0, \cdots, N,$$

$$(2.5)$$



donde $M_{kn} := \langle \Phi_k, \Phi_n(x) \rangle_{\omega}$ son las entradas de la matriz masa M y $S_{kn} := \langle \Phi_n, \mathcal{L} \Phi_k(x) \rangle_{\omega}$ son las entradas de la matriz de rigidez S. Observe que la matriz masa es simétrica y definida positiva. La simetría es evidente, mientras que ser definida positiva puede comprobarse considerando un vector **u** distinto de cero con entradas u_j y usando la definición de M:

$$\mathbf{u}^{T} M \mathbf{u} = \sum_{i,j=0}^{N} u_{i} u_{j} M_{ij}$$

=
$$\sum_{i,j=0}^{N} u_{i} u_{j} \int_{-1}^{1} \Phi_{i}(x) \Phi_{j}(x) \omega(x) dx$$

=
$$\int_{-1}^{1} \left(\sum_{i=0}^{N} u_{i} \Phi_{i}(x) \right)^{2} \omega(x) dx > 0,$$

ya que $\|\cdot\|_{L^2_{\omega}[-1,1]}$ es una norma.

La base $\Phi_k(x)$ se puede construir a partir de una combinación lineal de polinomios de Chebyshev $T_k(x)$ para garantizar que las condiciones de frontera se cumplan para todos los $\phi_k(x)$. Por ejemplo, si la condición de frontera es u(1, t) = 0, podemos elegir la base { $T_k(x) - T_k(1)$ }, la cual satisface las condiciones de frontera. Ya que estamos usando los polinomios $T_k(x)$ como base, es natural elegir la función peso, $\omega(x) = (1 - x^2)^{-1/2}$, tal que $\langle T_k(x), T_n(x) \rangle_{\omega} = \frac{\pi}{2} c_k \delta_{kn}$.

Los métodos de Galerkin utilizan polinomios de Chebyshev con el fin de construir las bases de espacios de dimensión finita. Las dificultades aparecen siempre que se apliquen a las funciones base, condiciones de frontera muy complicadas

• En los trabajos de D. Gottlieb (ver [20]) y C. Canuto (ver [6]), se usa la base

$$\Phi_k := \begin{cases} T_k - T_0, & k \text{ par} \\ T_k - T_1, & k \text{ impar} \end{cases}$$
(2.6)

y en consecuencia $B_N = \text{Gen}\{\Phi_1, \Phi_2, \cdots, \Phi_N\}.$

• Jie Shen (ver [46] y [47]) introdujo funciones base diferentes usando polinomios de Legendre y Chebyshev. En el último caso, definió

$$\Phi_k(x) := T_k(x) - T_{k+2}(x), \quad k = 0, 1, 2, \dots, N.$$
(2.7)



2.2 Método de Tau

El método de Tau es una modificación del método de Galerkin, cuya diferencia entre ellos es que $\psi_n = \Phi_n$ para todo $n = 1, \dots, N$, pero Φ_n no satisface las condiciones de frontera $\mathscr{B}\Phi_n(x) \neq 0$. Se toma una base ortogonal de funciones $\Phi_n(x)$, y se considera una solución aproximada de la forma

$$u_N(x,t) = \sum_{n=0}^{N} a_n(t)\Phi_n(x)$$
 (2.8)

donde *N* es el número de términos de la expansión. En los últimos *K* términos se incluyen las condiciones de frontera, donde *K* es el número de condiciones de frontera $\mathcal{B}u(x, t) = 0$ del problema.

Con el método de Tau se realiza los mismos pasos hechos con el método de Galerkin, es decir,

$$\langle \Phi_k, R \rangle_{\omega} = 0 \Leftrightarrow \sum_{k=0}^N M_{kn} \frac{\mathrm{d}a_k}{\mathrm{d}t} = \sum_{k=0}^N S_{kn} a_k(t), \ \forall n = 0, \cdots, N-K,$$

Las condiciones de frontera nos conducen a ecuaciones algebraicas para los *K* coeficientes restantes.

Como ejemplo tomemos los dos tipos de frontera más conocidos, tipo Dirichlet y tipo de Neumann, para un problema de una dimensión. Usando la base de Chebyshev y las identidades

$$T_n(\pm 1) = (\pm 1)^n$$

$$T'_n(\pm 1) = (\pm 1)^{n+1} n^2$$

tendremos que

$$u(\pm 1, t) = \sum_{n=0}^{N} (\pm 1)^{n} a_{n}(t) \text{ (Dirichlet)}$$
$$u'(\pm 1, t) = \sum_{n=0}^{N} n^{2} (\pm 1)^{n+1} a_{n}(t) \text{ (Neumann)}$$

Con las condiciones de frontera Dirichlet, Neumann o una combinación de estas, obtenemos las K ecuaciones faltantes para obtener los coeficientes y poder resolver así la ecuación diferencial. El número de condiciones de frontera depende de las dimensiones espaciales del problema, es decir, para una dimensión será K = 2, para dos dimensiones K = 4 y en general será $K = 2\beta$ donde β es la dimensión del problema.



2.3 Método de Colocación

En este método se asume que la solución aproximada u_N satisface la ecuación diferencial parcial en algunos puntos $x_0, x_1, ..., x_N$, llamados **puntos de colocación** del dominio respectivo. Luego se pide *colocar* la función en esos puntos, es decir, se requiere que la ecuación diferencial se satisfaga en los puntos de colocación, esto es, $R(x_i) = 0$ donde R es el residuo.

Se puede escribir el método de colocación de la misma forma que los métodos de Galerkin y de Tau, es decir, minimizar el residuo a través de (2.4), pero en el método de colocación se realiza un producto interno del residuo con $\delta(x-x_j)$, siendo los x_j los puntos de colocación, esto es, se fuerza a que la ecuación se cumpla en los puntos de colocación

$$\langle R, \delta(x-x_j) \rangle = \int_a^b R(x) \,\delta(x-x_j) \,dx = \mathbf{R}(x_j) = 0 \tag{2.9}$$

para j = 0, 1, 2, ..., N. Los puntos j = 0 y j = N corresponden a las condiciones de frontera.





3

Aplicaciones de los Métodos Espectrales

En este capítulo se dan ejemplos del método de diferencias finitas y el método espectral y de sus esquemas presentados en el capítulo anterior. La ecuación que analizamos con ambos métodos es la ecuación del calor. Posteriormente se resuelve la ecuación de Burger utilizando solo el método espectral. Con estos ejemplos se pretende brindar una base suficiente para aplicar los métodos espectrales para cualquier ecuación diferencial parcial en una dimensión.

3.1 Ecuación del Calor

3.1.1 Solución Numérica por Diferencias Finitas

La ecuación del calor describe la difusión del calor en el tiempo, en una región espacial. Esta se describe como

$$\frac{\partial u}{\partial t} = \alpha \nabla^2 u,$$

donde α es el coeficiente de difusividad térmica.

Tomemos la ecuación de calor descrita de la forma anterior, pero en una dimensión, esto es,

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}, \quad a < x < b, \ t > 0$$
(3.1)



La solución de la ecuación (3.1) requiere que se especifiquen las condiciones de frontera en x = a y x = b, y la condición inicial en t = 0. Las condiciones de frontera e inicial son

$$u(a, t) = u(b, t) = 0,$$
 $u(x, 0) = f(x).$ (3.2)

Otras condiciones de frontera, como las condiciones de gradiente (Neumann) o las mixtas, se pueden especificar. En este trabajo solo se considerarán las condiciones dadas en (3.2).

El método de diferencias finitas es una de las técnicas para la obtención de soluciones numéricas de la ecuación (3.1). En todas las soluciones numéricas, la ecuación diferencial parcial (EDP) se reemplaza con una aproximación discreta. La palabra "discreta" significa que la solución numérica es conocida sólo en un número finito de puntos en el dominio físico. El número de esos puntos puede ser seleccionado por el usuario. En general, aumentar el número de los puntos no sólo aumenta la resolución, sino también la exactitud de la solución numérica.

De la aproximación discreta resulta un conjunto de ecuaciones algebraicas que se resuelven para los valores de las incógnitas discretas. La figura 3.1 es una representación esquemática de la solución numérica.



Figura 3.1: Relación entre los problemas de la forma continua y discreta

La malla es el conjunto de puntos donde se calcula la solución discreta, estos puntos se denominan nodos. Dos parámetros importantes de la malla son Δx , que es la distancia entre puntos adyacentes en el espacio, y Δt , es la la distancia entre los pasos de tiempo adyacentes. Para los ejemplos considerados en la tesis, Δx y Δt son constantes en toda la malla.

La idea central del método de diferencias finitas es reemplazar derivadas con fórmulas en diferencias, que utilizan sólo valores discretos asociados con posiciones en la malla. Aplicar el método de diferencias finitas a una ecuación diferencial implica reemplazar todas las derivadas con las fórmulas en diferencias. En la ecuación del calor existen derivadas con respecto al tiempo, y derivadas con respecto al espacio. El uso de diferentes combinaciones de puntos de malla en



las fórmulas en diferencia resulta en diferentes esquemas. Cuando el espaciado de la malla ($\Delta x \ y \ \Delta t$) tiende a cero, la solución numérica obtenida con cualquier esquema se aproxima a la verdadera solución de la ecuación diferencial inicial. Esta aproximación varía con el esquema.

Discretización de Dominio

Particionamos el intervalo [a, b] como

$$x_i = a + (i-1)\Delta x, \quad i = 1, 2, ..., N$$

donde N es el número total de nodos en el espacio incluyendo la frontera, y

$$\Delta x = \frac{b-a}{N-1}$$

es el espacio entre los x_i . De forma similar, particionamos el intervalo [0, T] como

$$t_m = (m-1)\Delta t, \quad m = 1, 2, \dots, M$$

donde M es el número de pasos en el tiempo y

$$\Delta t = \frac{T}{M-1}$$

es el tamaño de paso temporal.

Esquemas para la Ecuación del Calor

Las derivadas respecto al tiempo y al espacio son reemplazadas por diferencias finitas. Para ello se requiere especificar las posiciones de los valores de u tanto en el tiempo como en el espacio en las fórmulas de diferencias finitas. Introducimos el supraíndice m para designar el paso de tiempo en la solución discreta.

Tiempo Hacia Adelante, Espacio Centrado (FTCS)

Este esquema es conocido como FTCS (Forward in Time Central in Space Scheme) por sus siglas en inglés. En este esquema se aproxima la derivada respecto al tiempo en la ecuación (3.1) con **diferencia finita hacia adelante**

$$\left. \frac{\partial u}{\partial t} \right|_{t_{m+1}, x_i} = \frac{u_i^{m+1} - u_i^m}{\Delta t} + \mathcal{O}(\Delta t).$$
(3.3)



Observamos que los términos del lado derecho involucran u en $x = x_i$.

Usando la aproximación de diferencia central para $\left(\frac{\partial^2 u}{\partial x^2}\right)_{x_i}$, y evaluando todos los términos en el tiempo *m* obtenemos

$$\frac{\partial^2 u}{\partial x^2}\Big|_{x_i} = \frac{u_{i-1}^m - 2u_i^m + u_{i+1}^m}{\Delta x^2} + \mathcal{O}(\Delta x^2).$$
(3.4)

Reemplazando (3.3) y (3.4) en la ecuación (3.1), y agrupando los términos del error de truncación se llega

$$\frac{u_i^{m+1} - u_i^m}{\Delta t} = \alpha \frac{u_{i-1}^m - 2u_i^m + u_{i+1}^m}{\Delta x^2} + \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x^2).$$
(3.5)

Los errores temporales y errores espaciales tienen diferentes órdenes. Además, observemos que podemos resolver en forma explícita u_i^{m+1} en términos de los otros valores de u. En efecto, eliminando los términos de los errores de truncación de la ecuación (3.5) y resolviendo para u_i^{m+1} se tiene

$$u_i^{m+1} = u_i^m + \frac{\alpha \Delta t}{\Delta x^2} (u_{i+1}^m - 2u_i^m + u_{i-1}^m).$$
(3.6)

con condición inicial

$$u(x_i, 0) = f(x_i), \quad i = 1, ..., N$$

y las condiciones de frontera son

$$u(a, t_m) = u(b, t_m) = 0, \quad m = 2, ..., M.$$

La ecuación (3.6) se le llama aproximación en **tiempo hacia adelante, espacio centrado** o aproximación FTCS de la ecuación del calor. Una ligera mejora en la eficiencia computacional se puede obtener con un pequeño reordenamiento de la ecuación (3.6)

$$u_i^{m+1} = r u_{i+1}^m + (1-2r)u_i^m + r u_{i-1}^m$$
(3.7)

para cada $i = 2, \dots, N-1$ y $m = 1, \dots, M$, donde $r = \frac{\alpha \Delta r}{\Delta x^2}$.

Esto genera el siguiente sistema

$$\begin{aligned} u_2^{m+1} &= r u_1^m + (1-2r) u_2^m + r u_3^m \\ u_3^{m+1} &= r u_2^m + (1-2r) u_3^m + r u_4^m \\ \vdots &\vdots \\ u_{N-2}^{m+1} &= r u_{N-3}^m + (1-2r) u_{N-2}^m + r u_{N-1}^m \\ u_{N-1}^{m+1} &= r u_{N-2}^m + (1-2r) u_{N-1}^m + r u_N^m \end{aligned}$$



Se observa que este sistema se puede expresar como una multiplicación matricial

$$u^{(m+1)} = A u^{(m)}$$

donde A es la matriz tridiagonal

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ r & 1 - 2r & r & \cdots & 0 & 0 & 0 \\ 0 & r & 1 - 2r & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \cdots & r & 1 - 2r & r \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

 $u^{(m+1)}$ es el vector de valores de u en el paso de tiempo m + 1, y $u^{(m)}$ es el vector de valores u en el paso de tiempo m. La primera y última fila de A se ajustan de manera que los valores frontera de u no cambian cuando se calcula el producto matricial.

El esquema FTCS es fácil de implementar ya que los valores de u_i^{m+1} se pueden encontrar independientemente uno del otro. La solución completa está contenida en dos bucles: un bucle exterior sobre todos los pasos de tiempo, y un bucle interior sobre todos los nodos interiores. Las soluciones de la ecuación (3.1) sujetas a las condiciones de frontera e inicial en la ecuación (3.2) son todas funciones acotadas y decrecientes. La aproximación FTCS puede dar soluciones inestables (esto quiere decir, que cuando Δt es demasiado grande, las soluciones se alejan entre sí). Las soluciones estables (es decir, para condiciones iniciales cercanas las soluciones se mantienen cercanas entre sí) con el esquema FTCS se obtienen si

$$r = \frac{\alpha \Delta t}{\Delta x^2} < \frac{1}{2}.$$
(3.8)

Tiempo Hacia Atrás, Espacio Centrado (BTCS)

Este esquema es conocido como BTCS (Backward in Time Central in Space Scheme) por sus siglas en inglés. En la obtención de (3.6), la diferencia hacia adelante fue usado para aproximar la derivada temporal en el lado izquierdo de la ecuación (3.1). Ahora, se elige la diferencia hacia atrás,

$$\left. \frac{\partial u}{\partial t} \right|_{t_{m-1}, x_i} = \frac{u_i^m - u_i^{m-1}}{\Delta t} + \mathcal{O}(\Delta t).$$
(3.9)



Reemplazamos la ecuación (3.9) y la ecuación (3.4) en la ecuación (3.1); y agrupando los términos del error de truncación se llega

$$\frac{u_i^m - u_i^{m-1}}{\Delta t} = \alpha \frac{u_{i-1}^m - 2u_i^m + u_{i+1}^m}{\Delta x^2} + \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x^2).$$
(3.10)

Los errores de truncación en esta aproximación tienen el mismo orden de magnitud que los errores de truncación de la ecuación (3.5). A diferencia de la ecuación (3.5), la ecuación (3.10) no se puede ordenar para obtener una simple fórmula algebraica para calcular u_i^m en términos de sus vecinos u_{i+1}^m , u_{i-1}^m , $y u_i^{m-1}$. Desarrollando una ecuación similar para u_{i+1}^m se demuestra que u_{i+1}^m depende de u_{i+2}^m y u_i^m . Por tanto, la ecuación (3.10) es una ecuación en un sistema de ecuaciones para los valores de u en los nodos internos de la malla espacial ($i = 2, 3, \dots, N-1$).

Para ver el sistema de ecuaciones más claro, eliminamos los términos de error de truncación de la ecuación (3.10) y reordenando la ecuación resultante se obtiene

$$-\frac{\alpha}{\Delta x^2}u_{i-1}^m + \left(\frac{1}{\Delta t} + \frac{2\alpha}{\Delta x^2}\right)u_i^m - \frac{\alpha}{\Delta x^2}u_{i+1}^m = \frac{1}{\Delta t}u_i^{m-1}.$$
 (3.11)

para para cada $i = 2, \dots, N-1$ y $m = 1, \dots, M$ con condición inicial

$$u(x_i, 0) = f(x_i), \quad i = 2, ..., N-1$$

y las condiciones de frontera son

$$u(a, t_m) = u(b, t_m) = 0, \quad m = 2, ..., M.$$

El sistema de ecuaciones se puede representar en forma matricial como

$$\begin{bmatrix} b_{1} & c_{1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_{2} & b_{2} & c_{2} & 0 & \cdots & 0 & 0 & 0 \\ 0 & a_{3} & b_{3} & c_{3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_{N-1} & b_{N-1} & c_{N-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_{N} & b_{N} \end{bmatrix} \begin{bmatrix} u_{1}^{m} \\ u_{2}^{m} \\ u_{3}^{m} \\ \vdots \\ u_{N-1}^{m} \\ u_{N}^{m} \end{bmatrix} = \begin{bmatrix} d_{1} \\ d_{2} \\ d_{3} \\ \vdots \\ d_{N-1} \\ d_{N} \end{bmatrix}$$
(3.12)



donde los coeficientes de los nodos interiores son

$$a_{i} = -\frac{\alpha}{\Delta x^{2}}, \quad i = 2, 3, \dots, N-1$$

$$b_{i} = \frac{1}{\Delta t} + \frac{2\alpha}{\Delta x^{2}},$$

$$c_{i} = -\frac{\alpha}{\Delta x^{2}},$$

$$d_{i} = \frac{1}{\Delta t} u_{i}^{m-1}.$$

Para las condiciones de frontera de Dirichlet en la ecuación (3.2)

$$b_1 = 1$$
, $c_1 = 0$, $d_1 = u(a, t)$
 $a_N = 0$, $b_N = 1$, $d_N = u(b, t)$.

La implementación del esquema BTCS requiere resolver un sistema de ecuaciones en cada paso de tiempo. Además de la dificultad para desarrollar el código, el esfuerzo computacional por paso de tiempo para el esquema BTCS es mayor que el esfuerzo computacional por paso de tiempo del esquema FTCS. El esquema BTCS tiene una gran ventaja sobre el esquema FTCS: es incondicionalmente estable (para soluciones a la ecuación del calor). El esquema BTCS es tan exacto como el esquema FTCS. Por lo tanto, con un poco de esfuerzo extra, el esquema BTCS produce un modelo computacional que es robusto al elegir Δt y Δx . Esta ventaja no siempre es abrumadora, sin embargo, el esquema FTCS sigue siendo útil para algunos problemas.

Crank-Nicolson

Los esquemas FTCs y BTCS tienen un error de truncamiento temporal de $\mathcal{O}(\Delta t)$. Cuando las soluciones en tiempo exacto son importantes, el esquema de Crank-Nicolson tiene ventajas significativas. El esquema de Crank-Nicolson no es más difícil de implementar que el esquema BTCS, y tiene un error de truncamiento temporal que es $\mathcal{O}(\Delta t^2)$. El esquema de Crank-Nicolson es implícito, como el de BTCS, además es incondicionalmente estable.

El lado izquierdo de la ecuación del calor se aproxima con la diferencia regresiva en el tiempo usado en el esquema BTCS, es decir, la ecuación (3.9). El lado derecho de la ecuación del calor se aproxima con el promedio del esquema de diferencia central evaluado en el paso de tiempo actual y anterior. Por lo tanto, la



ecuación (3.1) se aproxima con

$$\frac{u_i^m - u_i^{m-1}}{\Delta t} = \frac{\alpha}{2} \left[\frac{u_{i-1}^m - 2u_i^m + u_{i+1}^m}{\Delta x^2} + \frac{u_{i-1}^{m-1} - 2u_i^{m-1} + u_{i+1}^{m-1}}{\Delta x^2} \right].$$
(3.13)

Observamos que los valores de u a partir de los pasos de tiempo m y m-1 aparecen en el lado derecho. La ecuación (3.13) se usa para predecir los valores de u en el tiempo m, así que todos los valores de u en el tiempo m-1 se asumen que se conocen. Reordenando la ecuación (3.13), los valores de u en el tiempo m se dejan en el lado izquierdo, y los valores de u en el tiempo m-1 se dejan en el lado derecho.

$$-\frac{\alpha}{2\Delta x^2}u_{i-1}^m + \left(\frac{1}{\Delta t} + \frac{\alpha}{\Delta x^2}\right)u_i^m - \frac{\alpha}{2\Delta x^2}u_{i+1}^m = \frac{\alpha}{2\Delta x^2}u_{i-1}^{m-1} + \left(\frac{1}{\Delta t} - \frac{\alpha}{\Delta x^2}\right)u_i^{m-1} + \frac{\alpha}{2\Delta x^2}u_{i+1}^{m-1} \quad (3.14)$$

para para cada $i = 2, \dots, N-1$. El esquema de Crank-Nicolson es implícito, y resulta un sistema de ecuaciones para u que se puede resolver en cada paso de tiempo. El sistema de ecuaciones es idéntico al de la ecuación (3.12) pero con diferentes coeficientes:

$$\begin{aligned} a_i &= -\frac{\alpha}{2\Delta x^2}, & i = 2, 3, \cdots, N-1 \\ b_i &= \frac{1}{\Delta t} + \frac{\alpha}{\Delta x^2}, \\ c_i &= -\frac{\alpha}{2\Delta x^2}, \\ d_i &= a_i u_{i-1}^{m-1} + \left(\frac{1}{\Delta t} + a_i + c_i\right) u_i^{m-1} + c_i u_{i+1}^{m-1} \end{aligned}$$

Para las condiciones de frontera de Dirichlet en la ecuación (3.2)

$$b_1 = 1, \quad c_1 = 0, \quad d_1 = u(a, t)$$

 $a_N = 0, \quad b_N = 1, \quad d_N = u(b, t)$

Observe que a_i , b_i y c_i para el esquema de Crank-Nicolson difiere de sus contrapartes en el esquema BTCS por un factor de dos. Aparte de esta pequeña diferencia, la única diferencia significativa de implementación entre el esquema BTCS y Crank-Nicolson es la aparición de los términos extras u^{m-1} en d_i .



Los esquemas BTCS y Crank-Nicholson son muy similares en forma algorítmica. El esquema de Crank-Nicolson tiene un error de truncamiento de $\mathcal{O}(\Delta t^2) + \mathcal{O}(\Delta x^2)$, es decir, el error de truncamiento temporal es significativamente menor que el error de truncamiento temporal del esquema BTCS.

Implementación

Considere la ecuación de calor

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad |x| < 1$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0,$$

y la condición inicial

 $u(x,0) = \operatorname{sen}(\pi x) = f(x).$

La solución exacta de la ecuación del calor con la condición inicial y las condiciones de frontera dadas es

$$u(x,t) = e^{-\pi^2 t} \operatorname{sen} \pi x.$$
 (3.15)

El esquema Crank-Nicholson es implementado aquí usando Matlab. El sistema de ecuaciones para el método de Crank-Nicholson se resuelve con la función sistema_tridiagonal. Para calcular el error se usa la norma del máximo.

Ejecutando cranknicholson con los parámetros asignados:

>>[x,t,U,err]=cranknicholson(@funcion1,0,0,-1,1,0.2,1,20,100)

El error es 0.279368006727355 en t = 0, 2 y el ploteo se muestra en la figura 3.2





Figura 3.2: Solución de la ecuación del calor

La figura 3.3 muestra las gráficas de los errores máximos de la aproximación respecto a N para los tiempos t = 0.2, t = 0.3 y t = 0.4, se puede observar que el error no crece con el tiempo. Para obtener la gráfica se usó el programa errorcrank

3.1.2 Aproximación Espectral usando Chebyshev Tau

Considere la ecuación de calor

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}, \quad |x| \le 1$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0$$

y la condición inicial $u(x, 0) = sen(\pi x) = f(x)$.

En el método de Chebyshev Tau buscamos soluciones $u_N(x, t)$ de la forma

$$u_N(x,t) = \sum_{n=0}^{N} a_n(t) T_n(x)$$

que al reemplazar en la ecuación del calor se obtiene el residuo

$$R_N(x,t) = \sum_{n=0}^{N-2} \left(\frac{\mathrm{d}a_n}{\mathrm{d}t}(t) - \alpha a_n^{(2)}(t) \right) T_n(t).$$





Figura 3.3: La gráfica nos muestra el error máximo de la función con respecto a N en los tiempos t = 0.2, t = 0.3, t = 0.4.

Para minimizar el residuo $R_N(x, t)$ se requiere que sea ortogonal al espacio expandido por $\{T_k(x)\}_{k=0}^{N-2}$, es decir,

$$\frac{2}{\pi c_k} \int_{-1}^{1} R_N(x,t) T_k(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = 0, \qquad k = 0, 1, \cdots, N-2$$

Debido a la ortogonalidad, la integral anterior se convierte en

$$\frac{\mathrm{d}a_n}{\mathrm{d}t} = \alpha a_n^{(2)}, \qquad n = 0, 1, \cdots, N-2, \tag{3.16}$$

donde

$$a_n^{(2)} = \frac{1}{c_n} \sum_{\substack{p=n+2\\p+n \text{ par}}} p(p^2 - n^2) a_p(t)$$

con

$$c_n = \begin{cases} 2, & \text{si } n = 0, \\ 1, & \text{si } n \ge 1 \end{cases}$$

Se obtiene (N-1) ecuaciones diferenciales ordinarias para los coeficientes

$$\frac{\mathrm{d}a_n}{\mathrm{d}t} = \frac{1}{c_n} \sum_{\substack{p=n+2\\p+n \text{ par}}} p(p^2 - n^2) a_p(t), \quad 0 \le n \le N - 2.$$



Faltan dos ecuaciones para resolver el sistema de ecuaciones y estas las obtenemos de las condiciones de frontera, donde al hacer la expansión con los polinomios de Chebyshev y usando la identidad (1.65a) se obtiene

$$\sum_{n=0}^{N} (-1)^{N} a_{n} = 0, \qquad \sum_{n=0}^{N} a_{n} = 0,$$

que al considerar N como un número par se puede reescribir como

$$a_N = -\sum_{\substack{n=0\\n \text{ par}}}^{N-2} a_n,$$
 (3.17)

$$a_{N-1} = -\sum_{\substack{n=1\\n \text{ impar}}}^{N-3} a_n.$$
 (3.18)

Con las ecuaciones (3.16)-(3.18) se tienen los coeficientes para cualquier tiempo y para evolucionar la ecuación en el tiempo, necesitamos los coeficientes en tiempo inicial, por lo que usamos la condición

$$u_N(x,0) = \sum_{n=0}^{N} a_n(0)T_n(x) = f(x) = \operatorname{sen}(\pi x),$$
(3.19)

donde los coeficientes son

$$a_n(0) = \frac{2}{\pi c_k} \int_{-1}^{1} \frac{f(x)T_k(x)}{\sqrt{1-x^2}} \, \mathrm{d}x = \frac{2}{\pi c_k} \int_{-1}^{1} \frac{T_k(x)\mathrm{sen}(\pi x)}{\sqrt{1-x^2}} \, \mathrm{d}x.$$
(3.20)

Existen varios métodos con los cuales podemos evolucionar los coeficientes en el tiempo pero el que usamos en esta tesis es el método de Runge-Kutta de orden 4.

La solución exacta de la ecuación del calor con la condición inicial $f(x) = sen(\pi x) y$ las condiciones de frontera que se usó para comparar la solución aproximada es

$$u(x,t) = u(x,t) = e^{-\pi^2 t} \operatorname{sen} \pi x.$$
 (3.21)

Ejecutando calor_tau.m con los parámetros asignados:

>>[x,t,a,err] = calor_tau(20,[1 0 0;1 0 0])





Figura 3.4: Solución de la ecuación del calor

El error es 0.1384366811961 en t = 0, 2 y el ploteo se muestra en la figura **??** La figura 3.4 muestra las gráficas de los errores máximos de la aproximación respecto a N para los tiempos t = 0.2, t = 0.3 y t = 0.4, se puede observar que el error no crece con el tiempo. Para obtener la gráfica se usó el programa errortau

3.1.3 Aproximación Espectral usando Chebyshev Colocación

Considere la ecuación de calor

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad |x| < 1$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0,$$

y la condición inicial

$$u(x,0) = \operatorname{sen}(\pi x) = f(x).$$

Elegimos las funciones base

$$\Phi_k(x) = T_k(x), \quad k = 0, 1, \cdots, N.$$





Figura 3.5: La gráfica nos muestra el error máximo de la función con respecto a N en los tiempos t = 0.2, t = 0.3, t = 0.4.

La solución aproximada que se busca es

$$u_N(x,t) = \sum_{k=0}^N a_k(t) \Phi_k(x).$$

En el método de colocación se requiere que

$$R_N(x,t) = \frac{\partial u_N}{\partial t} - \frac{\partial^2 u_N}{\partial x^2}$$

se anule en los puntos interiores, proporcionándonos las (N-1) ecuaciones

$$\frac{\mathrm{d}u_N}{\mathrm{d}t}(x_j,t) = \sum_{k=0}^N D_{jk}^{(2)} u_N(x_k,t), \quad \forall j = 1, \cdots, N-1.$$

Aquí $D_{jk}^{(2)}$ es la matriz de diferenciación de segundo orden basado en los puntos de cuadratura Chebyshev-Gauss-Lobatto.

Cerramos el sistema mediante el uso de las condiciones de frontera $u(x_0, t) = u(x_N, t) = 0$, esto nos da el sistema lineal

$$\sum_{n=0}^{N} (-1)^{n} a_{n}(t) = \sum_{n=0}^{N} a_{n}(t) = 0$$





y la condición inicial

$$u_N(x_k, 0) = u(x_k, 0), \quad k = 0, 1, \cdots, N.$$

Las ecuaciones anteriores están basados en la formulación fuerte de la ecuación diferencial, ya que la solución aproximada requiere satisfacer la ecuación diferencial exactamente en un conjunto de puntos discretos, en este caso llamado los puntos de colocación. Se puede obtener formalmente las mismas ecuaciones a partir de una formulación débil del problema tomando como funciones de prueba a las funciones delta de Dirac (distribuciones, ver [16]).

La solución exacta de la ecuación del calor con la condición inicial y las condiciones de frontera que se usó para comparar la solución aproximada es

$$u(x,t) = e^{-\pi^2 t} \operatorname{sen} \pi x.$$
 (3.22)

Ejecutando calor_colocacion con los parámetros asignados:

>>[x,t,u,err] = calor_colocacion(20,0.2,[1 0 0;1 0 0])

El error es 8.516338265845969e-08 en t=0.2 y el ploteo se muestra en la figura 3.6



Figura 3.6: Solución de la ecuación del calor

54



La figura 3.7 muestra las gráficas de los errores máximos de la aproximación respecto a N para los tiempos t = 0.2, t = 0.3 y t = 0.4 con esto se puede observar que la exactitud de la función solución no crece a cada paso evolutivo, por lo que nos da la seguridad de que nuestra aproximación tendrá estabilidad temporal, es decir, el error no crece con el tiempo. Para obtener la gráfica se usó el programa errorcolocacion



Figura 3.7: La gráfica nos muestra el error máximo de la función con respecto a N en los tiempos t = 0.2, t = 0.3, t = 0.4.

3.2 Ecuación de Burger

La ecuación de Burgers fue propuesta por el holandés Johannes M. Burgers. Esta ecuación (forma fuerte) se escribe como

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} - v \frac{\partial^2 u}{\partial x^2} = 0, \quad x \in \Omega, \quad \forall t > 0,$$
(3.23)

donde Ω es el dominio espacial, v es una constante positiva que podemos considerarla como la viscosidad del fluido y u como la velocidad de este, que nos da una idea para problemas más complejos de dinámica de fluidos tales como las ecuaciones de Navier Stokes. La ecuación (3.23) se puede escribir como

$$\frac{\partial u}{\partial t} + \frac{\partial \mathscr{F} u}{\partial x} = 0, \quad x \in \Omega, \quad \forall t > 0, \tag{3.24}$$



donde ${\mathscr F}$ está dada por

$$\mathscr{F}(u) = \frac{1}{2}u^2 - v\frac{\partial u}{\partial x}.$$
(3.25)

Además, la ecuación (3.23) se complementa con una condición inicial

$$u(x,0) = u_0(x), \quad x \in \Omega,$$
 (3.26)

y condiciones de frontera de Dirichlet

$$u(-1, t) = u_L(t), \quad u(1, t) = u_R(t).$$
 (3.27)

La ecuación de Burgers es uno de las pocas ecuaciones diferenciales parciales no lineales para las que las soluciones exactas se conocen en términos de los valores iniciales (Cole [9], Hopf [22]). Hopf (1950) y Cole (1951) mostraron que la transformación

$$u = -2v \frac{\varphi_x}{\phi} \tag{3.28}$$

reduce la ecuación de Burgers (3.23) para u a la ecuación del calor para ϕ :

$$\frac{\partial \phi}{\partial t} - v \frac{\partial^2 \phi}{\partial x^2} = 0 \tag{3.29}$$

Observe que si $u_b(x, t)$ es una solución de (3.23) y c y t_0 son constantes, entonces

$$u(x, t) = c + u_b(x - ct, t + t_0)$$
(3.30)

también es solución. Para los ejemplos numéricos, utilizamos dos soluciones para u basadas en soluciones para u_b obtenidas de la transformación de Hopf-Cole. Para problemas no periódicos se usa

$$\phi_b(x,t) = \frac{x}{t} \frac{\sqrt{\frac{a}{t}} e^{\frac{-x^2}{4\nu t}}}{1 + \sqrt{\frac{a}{t}} e^{\frac{-x^2}{4\nu t}}},$$
(3.31)

donde *a* es una constante. La expresión ϕ_b corresponde a u_b en (3.30). Para problemas periódicos es una suma infinita de soluciones

$$\phi_b(x,t) = \frac{1}{\sqrt{4\pi v t}} \sum_{n=-\infty}^{\infty} e^{\frac{-(x-2\pi n)^2}{4v t}}.$$
(3.32)



3.2.1 Formulaciones Fuerte y Débil de la Ecuación de Burger

Tanto (3.23) y (3.24) están en la **forma fuerte**, es decir, la ecuación diferencial parcial se cumpla en cada punto de su dominio y para cada tiempo. Una **forma débil** de la ecuación diferencial parcial se obtiene al exigir que la integral de la EDP cerca de todas las funciones en un espacio apropiado X de funciones de prueba se cumpla; precisamente, multiplicamos ambos lados de (3.23) por cada función de prueba e integrar en el espacio (para cada tiempo), para obtener

$$\int_{a}^{b} \frac{\partial u}{\partial t} \psi \, dx + \int_{a}^{b} u \frac{\partial u}{\partial x} \psi \, dx - \int_{a}^{b} v \frac{\partial^{2} u}{\partial x^{2}} \psi \, dx = 0, \quad \forall v \in X, \quad \forall t > 0, \quad (3.33)$$

donde $\Omega = (a, b)$. Esto a menudo se refiere como una **forma intergral** de la EDP.

Otra forma débil de la EDP, que a menudo se usa, se obtiene al aplicar una integración por partes a (3.33), se tiene

$$\begin{split} \int_{a}^{b} \frac{\partial u}{\partial t} \psi \, dx - \frac{1}{2} \int_{a}^{b} u^{2} \frac{\partial \psi}{\partial x} \, dx + \nu \int_{a}^{b} \frac{\partial u}{\partial x} \frac{\partial \psi}{\partial x} \, dx + \\ \left(\frac{1}{2} u^{2} \psi - \nu \frac{\partial u}{\partial x} \psi \right) \Big|_{x=b} - \left(\frac{1}{2} u^{2} \psi - \nu \frac{\partial u}{\partial x} \psi \right) \Big|_{x=a} = 0, \quad \forall v \in X, \quad \forall t > 0. \end{split}$$

$$(3.34)$$

La ecuación (3.33) es significativa si *u* es dos veces diferenciable, mientras que las funciones de prueba no tienen que ser diferenciables. Cualquier método de discretización que se considere se obtiene de la formulación fuerte o débil del problema. Los métodos espectrales de colocación utilizan la forma fuerte de la EDP, al igual que los métodos de diferencias finitas. Para los métodos espectrales de Galerkin y de tau es preferible utilizar la EDP en su forma débil. Mientras que todas las formulaciones del problema diferencial son equivalentes (ya que la solución es suficientemente suave), este no es el caso en general para las diversas formulaciones discretas obtenidas de formulaciones alternativas de la EDP; esto quiere decir, la solución discreta que se basa en el método de Galerkin no tiene por qué coincidir con la solución discreta basada en un método de colocación. Por otra parte, los métodos de Galerkin que se basan en (3.33) no son necesariamente equivalentes a los métodos de Galerkin basados en (3.34)

La forma fuerte (3.23) se puede escribir en forma compacta como

$$u_t + \mathscr{G}(u) + \mathscr{L} u = 0 \quad \text{en } \Omega, \quad \forall t > 0, \tag{3.35}$$

TESIS PUCP



Teoría Espectral

donde el operador no lineal \mathscr{G} se define por $\mathscr{G}(u) = u \frac{\partial u}{\partial x}$, y el operador lineal \mathscr{L} es $-v \frac{\partial^2}{\partial x^2}$. La versión compacta correspondiente de la forma débil (3.33) es

 $(u_t + \mathscr{G}(u) + \mathscr{L}u, \psi) = 0 \quad \forall \psi \in X, \quad \forall t > 0, \tag{3.36}$

donde (u, v) denota el producto interno en X.

En esta sección se ilustran los procesos de discretización para diversas aproximaciones espectrales de la ecuación de Burgers. Consideramos aquí diferentes formas de tratar los términos no lineales y lineales, así como diferentes formas de tratar las condiciones de frontera. Cada discretización que presentamos se obtiene de una de las tres formulaciones de la EDP.

3.2.2 Aproximación Espectral usando Chebyshev Galerkin

Considere la ecuación de Burger

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2},$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0,$$

y la condición inicial u(x, 0) = f(x).

Para usar el método de Chebyshev Galerkin elegimos la base

$$\Phi_k(x) = T_{k+1}(x) - T_{k-1}(x),$$

(por tanto $\phi_k(\pm 1) = 0$), y buscamos soluciones $u_N(x, t)$ de la forma

$$u_N(x,t) = \sum_{k=0}^{N} a_k(t) \Phi_k(x) = \sum_{k=0}^{N} a_k(t) (T_{k+1}(x) - T_{k-1}(x)).$$

Como es usual, se requiere que el residuo

$$R_N(x,t) = \frac{\partial u_N}{\partial t} + u_N \frac{\partial u_N}{\partial x} - \nu \frac{\partial^2 u_N}{\partial x^2}$$

sea ortogonal a $\Phi_n(x)$:

$$\frac{2}{\pi} \int_{-1}^{1} R_N(x,t) \Phi_n(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = 0, \qquad n = 1, 2, \cdots, N-1,$$



produciendo el esquema de Chebyshev Galerkin

$$\sum_{k=1}^{N-1} L_{nk} \frac{\mathrm{d}a_k}{\mathrm{d}t} = \sum_{k=1}^{N-1} S_{nk}(\mathbf{a}(t))a_k(t), \qquad n = 1, 2, \cdots, N-1,$$

donde $\mathbf{a}(t)$ es el vector de elementos $a_k(t)$. La matriz masa tridiagonal está dada por

$$L_{nk} = \frac{2}{\pi} \int_{-1}^{1} \Phi_n(x) \Phi_k(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = 2\delta_{nk} - \delta_{n,k+2} - \delta_{n,k-2}.$$

La no linealidad complica la expresión de los elementos de la matriz de rigidez, ya que las entradas de S ahora consisten de las contribuciones de la parte no lineal, S^h , y la parte lineal, S^p :

$$S_{nk}(\mathbf{a}(t)) = -S_{nk}^{h}(\mathbf{a}(t)) + \nu S_{nk}^{p}$$

= $-\frac{2}{\pi} \int_{-1}^{1} \Phi_{n}(x) \Phi_{k}(x) \sum_{l=0}^{N-1} a_{l}(t) \frac{d\Phi_{l}(x)}{dx} \frac{1}{\sqrt{1-x^{2}}} dx + \frac{2}{\pi} \int_{-1}^{1} \Phi_{n}(x) \frac{d^{2}\Phi_{k}(x)}{dx^{2}} \frac{1}{\sqrt{1-x^{2}}} dx.$

Para simplificar la parte no lineal, podemos usar la definición de $\Phi_k(x)$ y la identidad $2T_k(x)T_n(x) = T_{|k+n|}(x) + T_{|k-n|}(x)$. (1.65f) Sin embargo, el cálculo de las integrales que participan en la matriz de rigidez sigue siendo más complicado.

Una vez que se calculan los componentes de la matriz de rigidez, tenemos (N-1) ecuaciones diferenciales ordinarias para los (N-1) incógnitas, $\mathbf{a} = (a_1, \dots, a_{N-1})^T$,

$$\frac{\mathrm{d}\mathbf{a}}{\mathrm{d}t} = M^{-1}(-S^h\mathbf{a} + \nu S^p)\mathbf{a}(t),$$

con condición inicial

$$a_k(0) = \frac{2}{\pi} \int_{-1}^{1} f(x)(T_{k+1}(x) - T_{k-1}(x)) \frac{1}{\sqrt{1 - x^2}} \, dx, \quad k = 1, 2, \cdots, N - 1$$

Los métodos polinomiales de Galerkin son de considerable complejidad, tanto analítica como computacional. El requisito es que cada polinomio de la base elegida debe satisfacer las condiciones de frontera de forma individual lo que a veces causa que la formulación de Galerkin llegue a ser complicado, en particular cuando las condiciones de frontera son dependientes del tiempo. Además, aunque la


matriz de masa depende sólo de la base que se utiliza, el cálculo de las entradas de la matriz de rigidez tiene que ser completado para cada problema individual y no es en absoluto simple, incluso para problemas lineales con coeficientes constantes. La presencia de términos no lineales complica aún más el cálculo de la matriz de rigidez, lo que hace al método de Galerkin extremadamente complejo. Una simplificación del procedimiento de Galerkin se conoce como el método de tau, y se presenta en la siguiente sección.

3.2.3 Aproximación Espectral usando Chebyshev Tau

Considere la ecuación de Burger

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2},$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0,$$

y la condición inicial u(x, 0) = f(x).

Para resolver la ecuación de Burger usando el método de Chebyshev Tau buscamos soluciones $u_N(x, t)$ de la forma

$$u_N(x,t) = \sum_{n=0}^N a_n(t)T_n(x)$$

con las condiciones

$$\sum_{n=0}^{N} (-1)^{n} a_{n}(t) = \sum_{n=0}^{N} a_{n}(t) = 0.$$

Se requiere que el residuo

$$R_N(x,t) = \frac{\partial u_N}{\partial t} + u_N \frac{\partial u_N}{\partial x} - \nu \frac{\partial^2 u_N}{\partial x^2}$$

sea ortogonal al espacio expandido por $\{T_k(x)\}_{k=0}^{N-2}$, es decir

$$\frac{2}{\pi c_k} \int_{-1}^{1} R_N(x,t) T_k(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = 0, \qquad k = 0, 1, \cdots, N-2,$$



originando el sistema de EDOs

$$\frac{\mathrm{d}a_k}{\mathrm{d}t} + \frac{2}{\pi c_k} \int_{-1}^1 u_N \frac{\partial u_N}{\partial x} T_k(x) \frac{1}{\sqrt{1-x^2}} \,\mathrm{d}x = v a_k^{(2)}(t), \qquad k = 0, 1, \cdots, N-2.$$

Recuerde que

$$a_k^{(2)} = \frac{1}{c_k} \sum_{\substack{p=k+2\\p+k \text{ par}}} p(p^2 - k^2) a_p(t).$$

El cálculo del término no lineal requiere un poco más de trabajo. Introduciendo la expansión $u_N(x, t)$ y su derivada se obtiene

$$\frac{2}{\pi c_k} \int_{-1}^{1} u_N \frac{\partial u_N}{\partial x} T_k(x) \frac{1}{\sqrt{1-x^2}} dx = \frac{2}{\pi c_k} \int_{-1}^{1} \sum_{n,l=0}^{N} a_n(t) a_l^{(1)}(t) T_n(x) T_l(x) T_k(x) \frac{1}{\sqrt{1-x^2}} dx.$$

Usando la identidad $2T_k(x)T_n(x) = T_{|k+n|}(x) + T_{|k-n|}(x)$ nos da

$$\begin{aligned} \frac{2}{\pi c_k} \int_{-1}^1 u_N \frac{\partial u_N}{\partial x} T_k(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x = \\ &= \frac{1}{2} \frac{2}{\pi c_k} \int_{-1}^1 \sum_{n,l=0}^N a_n(t) a_l^{(1)}(t) (T_{n+l} + (x)T_{|n-l|}(x)) T_k(x) \frac{1}{\sqrt{1-x^2}} \, \mathrm{d}x \\ &= \frac{1}{2} \left(\sum_{\substack{n,l=0\\n+l=k}}^N a_n(t) a_l^{(1)}(t) + \sum_{\substack{n,l=0\\|n-l|=k}}^N a_n(t) a_l^{(1)}(t) \right), \\ &\text{donde } a_l^{(1)} = \frac{2}{c_l} \sum p a_p(t). \end{aligned}$$

 $c_l = c_l = c_{l+1}$ p+l imparEsto establece las ecuaciones para los primeros N-1 coeficientes de expansión, $a_k(t)$. Los dos restantes se encuentran haciendo cumplir las condiciones de

$$\sum_{n=0}^{N} (-1)^{n} a_{n}(t) = \sum_{n=0}^{N} a_{n}(t) = 0.$$

frontera





Como es usual, las condiciones de frontera son

$$a_n(0) = \frac{2}{\pi c_n} + \int_{-1}^{1} f(x) T_n(x) \frac{1}{\sqrt{1 - x^2}} \, \mathrm{d}x, \qquad n = 0, 1, \cdots, N - 2.$$

Aunque el método de tau evita algunos de los problemas de los métodos de Galerkin y permite el desarrollo de métodos bastante simples y eficientes para resolver la ecuación diferencial parcial que se utiliza principalmente para la solución de ecuaciones lineales elípticas. La razón principal es que el método de tau aún requiere obtener ecuaciones para los coeficientes de expansión para cada problema individual. Para coeficientes no lineales o problemas no lineales este proceso puede ser muy complicado y en muchos casos imposible. Sin embargo, para el coeficiente lineal constante o casos especiales de problemas no lineales de coeficiente variables, el método tau resultante proporciona un enfoque eficaz.

3.2.4 Aproximación Espectral usando Chebyshev Colocación

Considere la ecuación de Burger

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2}, \quad x \in [-1, 1] \times [0, T[$$

con las condiciones de frontera

$$u(-1, t) = u(1, t) = 0,$$

y la condición inicial $u(x, 0) = f(x), x \in [-1, 1].$

Para resolver la ecuación de Burger usando el método de Colocación-Chebyshev buscamos soluciones $u_N(x, t)$ de la forma

$$u_N(x,t) = \sum_{k=0}^{N} a_k(t) T_k(x)$$

donde los puntos de colocación están dados por

$$x_j = \cos \frac{\pi j}{N}, \quad w_j = \begin{cases} \frac{\pi}{2N}, & j = 0, N\\ \frac{\pi}{N}, & j = 1, \cdots, N-1. \end{cases}$$

Usando la relación de ortogonalidad discreta

$$\sum_{n=0}^{N} T_i(x_n) T_j(x_n) = \alpha_i \delta_{ij},$$



con

$$\alpha_i = \begin{cases} \frac{N}{2}, & i \neq 0, N \\ N, & i = 0, N \end{cases}$$

e invirtiendo $u_N(x, t)$ se obtiene

$$a_k(t) = \frac{2}{N} \sum_{n=0}^{N} T_k(x_n) u_N(x_n, t).$$

Las derivadas $\frac{\partial u}{\partial x}$ y $\frac{\partial^2 u}{\partial x^2}$ se pueden calcular en los puntos de colocación usando la expansión $u_N(x, t)$ y los coeficientes de expansión $a_k(t)$. La derivada $\frac{\partial u}{\partial x}$ se aproxima como sigue

$$\begin{aligned} \frac{\partial u}{\partial x}(x_i,t) &= \sum_{k=0}^N a_k(t) T'_k(x_i) \\ &= \sum_{n=0}^N \left(\frac{2}{N} \sum_{k=0}^N T'_k(x_i) T_k(x_n) \right) u_N(x_n,t) \\ \frac{\partial u}{\partial x}(x_i,t) &= \sum_{n=0}^N [A_x]_{in} u_N(x_n,t), \end{aligned}$$

donde $[A_x]_{in} = \frac{2}{N} \sum_{k=0}^{N} T'_k(x_i) T_k(x_n), \quad k, n = 0, 1, \dots, N.$ La primera derivada de las funciones de Chebyshev está formado por

$$T'_{k}(x_{i}) = 2k \sum_{\substack{p=0\\k+p \text{ impar}}}^{k-1} c_{p} T_{p}(x_{i}).$$

Similarmente, la segunda derivada $\frac{\partial^2 u}{\partial x^2}$ es aproximado por

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2}(x_i, t) &= \sum_{n=0}^N [A_x]_{in} \frac{\partial u}{\partial x} u_N(x_n, t) \\ &= \sum_{j=0}^N \left(\sum_{n=0}^N [A_x]_{in} [A_x]_{nj} \right) u_N(x_j, t) \end{aligned}$$



Teoría Espectral

$$\frac{\partial^2 u}{\partial x^2}(x_i,t) = \sum_{j=0}^N [B_x]_{ij} u_N(x_j,t),$$

donde $B_x = A_x^2$ y los elementos de la matriz B_x están dados por

$$[B_x]_{kj} = [A_x]_{in} [A_x]_{nj}, \quad k, n = 0, 1, \cdots, N.$$

Usando las condiciones de frontera se tiene

$$\frac{\partial u}{\partial x}(x_i, t) = d_i(t) + \sum_{n=1}^{N-1} [A_x]_{in} u_N(x_n, t), \qquad (3.37)$$
$$\frac{\partial^2 u}{\partial x^2}(x_i, t) = \overline{d}_i(t) + \sum_{n=1}^{N-1} [B_x]_{in} u_N(x_n, t), \qquad (3.38)$$

donde $d_i(t) = [A_x]_{i0} u_N(x_0, t) + [A_x]_{iN} u_N(x_N, t) y \overline{d}_i(t) = [B_x]_{i0} u_N(x_0, t) + [B_x]_{iN} u_N(x_N, t)$. Reemplazando (3.37) y (3.38) en la ecuación de Burger, se obtiene

$$\frac{\partial u_N}{\partial t}(x_i, t) + u(x_i, t) \sum_{n=1}^{N-1} [A_x]_{in} u_N(x_n, t) - \nu \sum_{n=1}^{N-1} [B_x]_{in} u_N(x_n, t) + u(x_i, t) d_i(t) - \nu \overline{d}_i(t) = 0,$$

$$u(x_i, 0) = f(x_i).$$
(3.39)

Luego, el sistema (3.39) se puede escribir en la forma siguiente:

$$\frac{\partial u_N}{\partial t}(t) = F(t, u_N(t))$$

$$u_N(0)u_0.$$
(3.40)

donde

$$u_N(t) = [u_N(x_1, t), u_N(x_2, t), \cdots, u_N(x_{N-1}, t)]$$

$$\frac{\partial u_N}{\partial t}(t) = \left[\frac{\partial u_N}{\partial t}(x_1, t), \frac{\partial u_N}{\partial t}(x_2, t), \cdots, \frac{\partial u_N}{\partial t}(x_{N-1}, t)\right]^T$$

$$F(t, u_N(t)) = [F_1(t, u_N(t)), F_2(t, u_N(t)), \cdots, F_{N-1}(t, u_N(t))]^T$$

у

$$F_{i}(t, u_{N}(t)) = -u(x_{i}, t) \sum_{n=1}^{N-1} [A_{x}]_{in} u_{N}(x_{n}, t) + \nu \sum_{n=1}^{N-1} [B_{x}]_{in} u_{N}(x_{n}, t) - u(x_{i}, t)d_{i}(t) + \nu \overline{d}_{i}(t).$$



Las ecuaciones (3.40) forman un sistema de ecuaciones diferenciales ordinarias (EDO) en el tiempo. Por lo tanto el avance de la solución en el tiempo, usa solucionadores de ODE tales como el método de Runge-Kutta de orden cuatro, ya que es un método explícito que da una buena precisión y se extiende trivialmente a no lineal. El método de Runge-Kutta de cuarto orden está dada por

$$u^{(1)} = u(t_n) + \frac{1}{2}F(t_n, u_N(t_n))$$

$$u^{(2)} = u(t_n) + \frac{1}{2}F(t_{n+\frac{\Delta t}{2}}, u^{(1)})$$

$$u^{(3)} = u(t_n) + F(t_{n+\frac{\Delta t}{2}}, u^{(2)})$$

$$u(t_{n+1}) = u(t_n) + \frac{\Delta t}{6}[F(t_n, u_N(t_n)) + 2F(t_{n+\frac{\Delta t}{2}}, u^{(1)}) + 2F(t_{n+\frac{\Delta t}{2}}, u^{(2)}) + F(t_{n+\Delta t}, u^{(3)})].$$

Ejemplo 1 (Ejemplo Numérico de la Ecuación de Burger No Periódica). La solución exacta no periódica *u* correspondiente a (3.31), (3.30) and (3.28) para v = 0.01, c = 1, a = 16 y $t_0 = 1$ se muestra en la siguiente figura (izquierda) en t = 0 y t = 1. La ecuación de Burger se resuelve en el intervalo]-1; 1[con condiciones de iniciales y de frontera tomadas de la solución exacta.



Figura 3.8: La solución exacta para la ecuación de Burger no periódica (izquierda) y el cálculo de los errores máximos en t = 1 (derecha). Ver [4].



La figura 3.8 (derecha) ilustra los errores de los esquemas numéricos de Chebyshevtau, Chebyshev-colocación y Galerkin con Integración de este problema, integrado en el tiempo con el método Runge Kutta de orden 4. También se incluyen para la comparación, soluciones para diferencias finitas de cuarto y sexto orden.





Conclusiones

En este trabajo, se observa que el método espectral da resultados más precisos cuando se usa un número N de términos en la expansión de la solución aproximada $u_N(x, t)$ comparada con el mismo número de particiones del dominio espacial usado en el método de Diferencias Finitas. Aunque el método espectral es más laborioso de implementar, el trabajo extra se compensa al obtener una precisión muy aceptable para un número N no muy grande.

Se ha mostrado cómo usar el método espectral en una dimensión. En un posterior trabajo de investigación se pretende aplicarlo a dimensiones más altas.

67



Apéndice: Código Matlab

1. funcion1 Programa que define función que se usa como condición inicial

```
function y=funcion1(x)
%y=sin(pi*x);
y=1-cos(pi.*(x+1));
```

2. **cranknicholson**: Programa para resolver la ecuación del calor usando el método de Crank-Nicholson

```
function [x,t,U,err]=cranknicholson(f,c1,c2,a1,a2,b,c,n,m)
% Resuelve el problema de la ecuación del calor
%du/dt=S(x)u''(x,t) usando el método de Crank-Nicholson
%Entrada
%
    f=u(x,0)=@funcion1 como una cadena 'f'
%
    c1=u(a1,t) y c2=u(a2,t)
%
    a1, a2 y b son los extremos de [a1,a2] y [0,b]
%
    c coeficiente de difusividad
   n y m número de puntos de la malla sobre [a1,a2] y [0,b]
%
%Salida
    x discretización espacial [a1,a2]
%
```



```
t discretización temporal [0,b]
%
%
   U matriz solución
%
    err error cometido
%Parámetros iniciales y U
h=(a2-a1)/(n-1);
k=b/(m-1);
x=a1:h:a2;
t=0:k:b;
r=c*k/h^2;
s1=2+2/r;
s2=2/r-2;
U=zeros(n,m);
%Condiciones de frontera
U(1,1:m)=c1;
U(n, 1:m) = c2;
%Generación de la primera fila
U(2:n-1,1)=feval(f,h:h:(n-2)*h)';
%Forma los elementos de la diagonal y fuera de la diagonal de A
%y el vector constante y resuelve el sistema tridiagonal AX=B
Vd(1,1:n)=s1*ones(1,n);
vd(1)=1;
vd(n)=1;
Va=-ones(1,n-1);
Va(n-1)=0;
Vc=-ones(1,n-1);
Vc(1)=0;
Vb(1)=c1;
Vb(n)=c2;
for j=2:m
    for i=2:n-1
        Vb(i)=U(i-1,j-1)+U(i+1,j-1)+s2*U(i,j-1);
    end
    X=sistema_tridiagonal(Va,Vd,Vc,Vb);
    U(1:n,j)=X';
end
%Representación gráfica de la temperatura en diferentes tiempos
%seleccionados
```



```
for l=1:m% u(1,k)=0;% u(n+1,k)=0;
    time(1)=(1-1)*k;
end
subplot(3,1,1); plot(x,U(:,1)),ylabel('u0(x,0)'),
title('Temperatura Inicial')
subplot(3,1,2); plot(x,U(:,end)),ylabel('u(x,tfinal)'),
title('Temperatura final')
subplot(3,1,3); mesh(x,time,U'), xlabel('x'), ylabel('t'),
zlabel('u(x,t)'),
title('Solución del problema del calor usando Método de
Crank-Nicholson')
U=U';
\% --- Compare with exact solution at end of simulation
ue = sin(pi*x).*exp(-time(m)*pi^2);%sin(pi*x).*exp(-time(m)*pi^2)+
sin(3*pi*x).*exp(-9*time(m)*pi^2);
err=max(abs(U(m,:)-ue));
```

3. sistema_tridiagonal: Programa para resolver un sistema tridiagonal

```
function X=sistema_tridiagonal(A,D,C,B)
%Entrada
%
   A es la subdiagonal de la matriz de coeficientes
%
    D es la diagonal de la matriz de coeficientes
%
    C es la superdiagonal de la matriz de coeficientes
%
   B es el vector constante del sistema lineal
%Salida
    X es el vector solución
%
N=length(B);
for k=2:N
    mult=A(k-1)/D(k-1);
    D(k)=D(k)-mult*C(k-1);
    B(k)=B(k)-mult*B(k-1);
end
X(N) = B(N)/D(N);
for k=N-1:-1:1
    X(k) = (B(k) - C(k) * X(k+1)) / D(k);
end
```



Teoría Espectral

4. errorcrank: Programa para obtener el gráfico de los errores máximos

```
%clc, close all
t1 = 0.2;
g=[1 \ 0 \ 0; 1 \ 0 \ 0];
N = 5:5:25;
for k = 1:5
    [x,t,U,err]=cranknicholson(@funcion1,0,0,-1,1,t1,1,5*k,100);
   err1(k) =err;
end
err1
semilogy(N,err1,'-o'), hold on
%-----
t2 = 0.3;
for k = 1:5
   [x,t,U,err]=cranknicholson(@funcion1,0,0,-1,1,t2,1,5*k,100);
   err2(k) = err;
end
err2
semilogy(N,err2,'-s')
%_-----
t3 = 0.4;
for k = 1:5
    [x,t,U,err]=cranknicholson(@funcion1,0,0,-1,1,t3,1,5*k,100);
   err3(k) = err;
end
err3
semilogy(N,err3,'-*'),hold off
```

5. **calor_tau:** Programa para resolver la ecuación del calor usando el método de Chebyshev Tau

```
%chebishev Tau Calor
function [x,t,a,err] = calor_tau(N,tfinal,g);
% Extracción de los coeficientes de las condiciones de frontera
```



```
a1=g(1,1); b1=g(1,2); c1=g(1,3);
   aN=g(2,1); bN=g(2,2); cN=g(2,3);
%Construcción de la matriz
   A=zeros(N-2,N);
   for i=1:N-2
       for j=i+2:2:N
           if i==1
                A(i,j)=(j-1)*((j-1)^2-(i-1)^2)/2;
           else
                A(i,j)=(j-1)*((j-1)^2-(i-1)^2);
           end
       end
   end
   Α;
   B=A(1:N-2,1:N-2);
   for i=1:2
       for j=1:N-2
           M(i,j)=((-1)^(j-1)*(a1-b1*(j-1)^2)*(aN+bN*(N-i)^2)-
           (-1)^(N-i)*(a1-b1*(N-i)^2)*(aN+bN*(j-1)^2))/
            (-(-1)^(N-3+i)*(a1-b1*(N-3+i)^2)*(aN+bN*(N-i)^2)+
            (-1)^{(N-i)*(a1-b1*(N-i)^2)*(aN+bN*(N-3+i)^2))};
       end
   end
   Μ;
   C=A(1:N-2,N-1:N);
   D=C*M;
   G1=(c1*(aN+bN*(N-1)^2)-cN*(a1+b1*(N-1)^2))/(-(-1)^(N-2)*
   (a1-b1*(N-2)^{2})*(aN+bN*(N-1)^{2})+(-1)^{(N-1)}*(a1-b1*(N-1)^{2})*
   (aN+bN*(N-2)^{2});
   G2=(c1*(aN+bN*(N-2)^{2})-cN*(a1+b1*(N-2)^{2}))/(-(-1)^{(N-1)}*
   (a1-b1*(N-1)^2)*(aN+bN*(N-2)^2)+(-1)^(N-2)*(a1-b1*(N-2)^2)*
   (aN+bN*(N-1)^{2})):
   G = [G1; G2];
   F=C*G;
   F2=B+D;
   tfinal=10.; % Tiempo final de integración
%Grafica de la condición inicial
```



Teoría Espectral

```
x=-1:2/(N-1):1;
  u0=x.^2;%sin(pi*x);
   subplot(3,1,1); plot(x,u0),ylabel('u0(x,0)'),
  title('Initial data')
   tspan=[0:tfinal/100:tfinal]; % Step size
   a0=zeros(N-2,1);
   for k=1:N-2
       if k==1
           a0(k,1)=chebyint(k-1)/pi;
       else
           a0(k,1)=2*chebyint(k-1)/pi;
       end
   end
   a0;
   options=odeset('RelTol',1e-04,'AbsTol',1e-6);
   [t,a]=ode45('caltau',tspan,a0,options,F2,F); % Se usa
  %Runge Kutta para resolver el sistema de EDOs
  H=a*M';
  a=[a,H];
   [m,n]=size(a);
  ufinal=0;
  for i=1:N
       ufinal=ufinal+a(m,i)*chebypoly(i-1);
   end
  ufinal;
   subplot(3,1,2);
  hg=ezplot(ufinal,[-1,1]);
   subplot(3,1,3); mesh(x,t,a), xlabel('x'), ylabel('t'),
  zlabel('u(x,t)'),
  title('Solution of heat initial-boundary value problem')
  \% --- Compare with exact solution at end of simulation
  ue = sin(pi*x).*exp(-t(m)*pi^2);
  err=max(abs(a(m,:)-ue));
end
```

6. caltau: Programa que calcula el lado derecho de la ecuación del calor



```
function dudt=caltau(t,w,flag,F2,F)
% Función para calcular el lado derecho de la ecuación del calor 1D
dudt=zeros(size(w));
dudt=F2*w+F;
```

7. chebyint: Programa que calcula la integral de los polinomios de Chebyshev

```
function q=chebyint(n) %n es grado 0,1,2...
syms x
u0=x.^2;%sin(pi*x);
q=int(chebypoly(n)*u0./sqrt(1-x.^2),-1,1);
q=double(q);
```

8. chebypoly: Programa que halla los polinomios de Chebyshev

```
function tk = chebypoly(n)
if n==0
   tk = 1;
  tk=poly2sym(tk);
elseif n==1
   tk = [1 0]';
   tk=poly2sym(tk);
else
   tkm2 = zeros(n+1,1);
   tkm2(n+1) = 1;
   tkm1 = zeros(n+1,1);
   tkm1(n) = 1;
   for k=2:n
       tk = zeros(n+1,1);
        for e=n-k+1:2:n
            tk(e) = 2*tkm1(e+1) - tkm2(e);
        end
```



Teoría Espectral

```
if mod(k,2)==0
    tk(n+1) = (-1)^(k/2);
end

if k<n
    tkm2 = tkm1;
    tkm1 = tk;
end

tk=poly2sym(tk);
end</pre>
```

9. errortau: Programa para obtener el gráfico de los errores máximos

```
%clc, close all
t1 = 0.2;
g=[1 \ 0 \ 0; 1 \ 0 \ 0];
N = 5:5:25;
for k = 1:5
   [x,t,u,err] = calor_tau(5*k,t1,g);
   err1(k) =err;
end
err1
semilogy(N,err1,'-o'), hold on
%-----
t2 = 0.3;
for k = 1:5
   [x,t,u,err] = calor_tau(5*k,t2,g);
   err2(k) = err;
end
err2
semilogy(N,err2,'-s')
%-----
t3 = 0.4;
for k = 1:5
```



Teoría Espectral

```
[x,t,u,err] = calor_tau(5*k,t3,g);
err3(k) = err;
end
err3
semilogy(N,err3,'-*'),hold off
```

10. **calor_colocacion**: Programa para resolver la ecuación del calor usando el método de Chebyshev Colocación.

```
function [x,t,u,err] = calor_colocacion(N,tfinal,g);
% Resuelve el problema de la ecuación del calor
% du/dt=S(x)u''(x,t) usando el método espectral de Chebyshev
%Colocación
%Condición Initial u(x,0)=u0
%Condiciones de frontera
               u(-1,t)+u_x(-1,t)=0;
%
%
                u(1,t)+u_x(1,t)=0;
%Entrada
    N es el número de punto de cradratura de
%
% Chebyshev-Gauss-Lobatto
% tfinal tiempo final
   n y m número de puntos de la malla sobre [0,a] y [0,b]
%
%Salida
%
  x discretización espacial [a1,a2]
%
   t discretización temporal [0,b]
%
  U matriz solución
%
    err error cometido
%Parámetros iniciales y U
pi=4.*atan(1.); % Order de aproximación
[x,D2t]=cheb2bc(N,g); % Matriz de diferenciación
%(segundo orden) con las condiciones de frontera dadas
% Se usa la rutina cheb2bc de Weideman & Reddy
u=zeros(size(x));
u0=sin(pi*x);
a=flipud(x);
subplot(3,1,1); plot(x,u0),ylabel('u0(x,0)'),
title('Temperatura Inicial con el Método de Colocación')
```



```
%-----
tspan=[0:tfinal/100:tfinal]; % Step size
options=odeset('RelTol',1e-04,'AbsTol',1e-6);
[t,u]=ode45('sist_edo',tspan,u0,options,D2t); % Se usa
%Runge-Kutta para
%resolver el sistema de EDOs
x=[1;x;-1]; % Malla espacial
[m,n]=size(u);
ufinal=[0,u(m,:),0];
subplot(3,1,2); plot(x,ufinal),ylabel('u(x,tfinal)'),
title('Solución del problema del calor usando Colocación')
%-----
u1=zeros(m,1);
u = [u1, u, u1];
subplot(3,1,3); mesh(x,t,u), xlabel('x'), ylabel('t'),
zlabel('u(x,t)'),
title('Solución del problema del calor usando Colocación')
%-----
% --- Compare with exact solution at end of simulation
ue = sin(pi*x).*exp(-t(m)*pi^2);
err=max(abs(u(m,:)-ue'));
```

11. sist_edo: Función para calcular el lado derecho de la ecuación del calor

```
function dudt=sist_edo(t,w,options,D2t)
% Función para calcular el lado derecho de la ecuación del calor 1D
dudt=zeros(size(w));
dudt=D2t*w;
dudt=dudt;
```

12. **cheb2bc:** Programa para calcular las matrices de primer y segundo orden y las condiciones de frontera evaluadas en los puntos de frontera

```
function [xt,D2t,D1t,phip,phim]=cheb2bc(N,g);
%Program for computing first and second derivative
%matrices and boundary condition functions for 2 point
```



```
%boundary conditions
%
        a_1 u(1) + b_1 u'(1) = c_1
%
        a_N u(-1) + b_N u'(-1) = c_N
%INPUT
%N=number of Chebyshev points in [-1,1]
%g=boundary condition matrix = [a_1 b_1 c_1; a_N b_N c_N]
%OUTPUT
%xt=Chebyshev points corresponding to rows and columns
    of D1t and D2t
%
%D1t=1st derivative matrix incorporating bc
%D2t=2nd derivative matrix incorporating bc
%phip=1st and 2nd derivative of bc function at x=1
%
      array with 2 columns)
%phim=1st and 2nd derivative of bc function at x=-1
% (array with 2 columns)
% S.C. Reddy, J.A.C. Weideman 1998
% Get differentiation matrices
   [x, DM] = chebdif(N, 2);
   DO=eve(N,N);
   D1=DM(:,:,1);
   D2=DM(:,:,2);
% extract boundary condition coefficients
   a1=g(1,1); b1=g(1,2); c1=g(1,3);
   aN=g(2,1); bN=g(2,2); cN=g(2,3);
% Case 0: Invalid boundary condition information
if ((a1==0 & b1==0) | (aN==0 & bN==0)),
   fprintf('Invalid boundary condition information
   (no output) n';
elseif (b1==0 & bN==0)
                                % Dirichlet/Dirichlet
   J=2:N-1;
   K=(2:N-1)';
   D1t=D1(J,K);
   D2t=D2(J,K);
   phip=c1*[D1(K,1) D2(K,1)]/a1;
                                          % phi_+
   phim=cN*[D1(K,N) D2(K,N)]/aN;
                                          % phi_-
   xt=x(K);
                                          % node vector
elseif (b1~=0 & bN==0),
                             % Dirichlet x=-1, Robin x=1
```



```
J=2:N-1;
   K = (1:N-1)';
   xjrow=2*sin((J-1)*pi/2/(N-1)).^2;% 1-x_j, using trig identity
   xkcol=2*sin((K-1)*pi/2/(N-1)).^2;% 1-x_k, using trig identity
   oner=ones(size(xkcol));% column of ones
   fac0 = oner*(1./xjrow);% matrix -1/(1-x_j)
   fac1 = xkcol*(1./xjrow); matrix (1-x_k)/(1-x_j)
   D1t = fac1.*D1(K,J)-fac0.*D0(K,J);
   D2t = fac1.*D2(K,J)-2*fac0.*D1(K,J);
   cfac = D1(1,1)+a1/b1;% compute phi'_1, phi''_1
   fcol1 = -cfac*D0(K,1)+(1+cfac*xkcol).*D1(K,1);
   fcol2 = -2*cfac*D1(K,1)+(1+cfac*xkcol).*D2(K,1);
   D1t = [fcol1 D1t];
   D2t = [fcol2 D2t];
   phim = xkcol.*D1(K,N)/2-D0(K,N)/2;% phi'_-, phi''_-
   phim = cN*[phim xkcol.*D2(K,N)/2-D1(K,N)]/aN;
  phip= -xkcol.*D1(K,1)+D0(K,1);% phi'_+, phi''_+
  phip= c1*[phip -xkcol.*D2(K,1)+2*D1(K,1)]/b1;
  xt = x(K);% node vector
elseif (b1==0 & bN~=0).
% Case 3: Dirichlet at x=1 and Neumann or Robin boundary x=-1.
   J=2:N-1;
  K = (2:N)';
   xjrow=2*cos((J-1)*pi/2/(N-1)).^2;% 1+x_j, using trig identity
   xkcol=2*cos((K-1)*pi/2/(N-1)).^2;% 1+x_k, using trig identity
   oner=ones(size(xkcol));% column of ones
   fac0 = oner*(1./xjrow);% matrix 1/(1+x_j)
   fac1 = xkcol*(1./xjrow);% matrix (1+x_k)/(1+x_j)
   D1t = fac1.*D1(K,J)+fac0.*D0(K,J);
   D2t = fac1.*D2(K,J)+2*fac0.*D1(K,J);
   cfac = D1(N,N)+aN/bN;% compute phi'_N, phi''_N
   lcol1 = -cfac*D0(K,N)+(1-cfac*xkcol).*D1(K,N);
   lcol2 = -2*cfac*D1(K,N)+(1-cfac*xkcol).*D2(K,N);
   D1t = [D1t lcol1];
   D2t = [D2t \ lcol2];
   phip= xkcol.*D1(K,1)/2+D0(K,1);% compute phi'_+,phi''_+
   phip= c1*[phip xkcol.*D2(K,1)/2+D1(K,1)]/a1;
```



```
phim= xkcol.*D1(K,N)+D0(K,N);% compute phi'_-,phi''_-
   phim= cN*[phim xkcol.*D2(K,N)+2*D1(K,N)]/bN;
   xt = x(K);% node vector
elseif (b1~=0 & bN~=0),
% Case 4: Neumann or Robin boundary conditions at both
%endpoints.
J=2:N-1:
K = (1:N)';
xkcol0=sin((K-1)*pi/(N-1)).^2;% 1-x_k^2 using trig identity
xkcol1=-2*x(K);% -2*x_k
xkcol2=-2*ones(size(xkcol0));% -2
xjrow=1./(sin((J-1)*pi/(N-1)).^2);% 1-x_j^2 using trig identity
fac0=xkcol0*xjrow;
fac1=xkcol1*xjrow;
fac2=xkcol2*xjrow;
D1t=fac0.*D1(K,J)+fac1.*D0(K,J);
D2t=fac0.*D2(K,J)+2*fac1.*D1(K,J)+fac2.*D0(K,J);
omx=sin((K-1)*pi/2/(N-1)).^2;% (1-x_k)/2
opx=cos((K-1)*pi/2/(N-1)).^2;% (1+x_k)/2
r0=opx+(0.5+D1(1,1)+a1/b1)*xkcol0/2;% compute phi'_1, phi''_1
r1=0.5-(0.5+D1(1,1)+a1/b1)*x;
r2=-0.5-D1(1,1)-a1/b1;
rcol1=r0.*D1(K,1)+r1.*D0(K,1);
rcol2=r0.*D2(K,1)+2*r1.*D1(K,1)+r2.*D0(K,1);
10=omx+(0.5-D1(N,N)-aN/bN)*xkcol0/2;% compute phi'_N, phi''_N
11=-0.5+(D1(N,N)+aN/bN-0.5)*x;
12=D1(N,N)+aN/bN-0.5;
lcol1=l0.*D1(K,N)+l1.*D0(K,N);
lcol2=l0.*D2(K,N)+2*l1.*D1(K,N)+l2.*D0(K,N);
D1t=[rcol1 D1t lcol1];
D2t=[rcol2 D2t lcol2];
phim1=(xkcol0.*D1(K,N)+xkcol1.*D0(K,N))/2;
phim2=(xkcol0.*D2(K,N)+2*xkcol1.*D1(K,N)+xkcol2.*D0(K,N))/2;
phim=cN*[phim1 phim2]/bN;% compute phi'_-, phi''_-
phip1=(-xkcol0.*D1(K,1)-xkcol1.*D0(K,1))/2;
phip2=(-xkcol0.*D2(K,1)-2*xkcol1.*D1(K,1)-xkcol2.*D0(K,1))/2;
phip=c1*[phip1 phip2]/b1;% compute phi'_+, phi''_+
```



```
xt=x(K);% node vector
end
```

13. errorcolocacion: Programa para obtener el gráfico de los errores máximos

```
%clc, close all
t1 = 0.2;
g=[1 0 0;1 0 0];
N = 5:5:25;
for k = 1:5
   [x,t,u,err] = calor_colocacion(5*k,t1,g);
   err1(k) =err;
end
err1
semilogy(N,err1,'-o'), hold on
%-----
t2 = 0.3;
for k = 1:5
   [x,t,u,err] = calor_colocacion(5*k,t2,g);
   err2(k) = err;
end
err2
semilogy(N,err2,'-s')
%-----
t3 = 0.4;
for k = 1:5
   [x,t,u,err] = calor_colocacion(5*k,t3,g);
   err3(k) = err;
end
err3
semilogy(N,err3,'-*'),hold off
```



Bibliografía

- Blinova, E. N., *Hydrodynamic theory of pressure and temperature waves and center of action of the atmosphere*. Trans. No. 113 (Regional Control Office, Second Weather Region, Patterson Field, OH), 1944
- [2] Boyd, J. P., *Chebyshev and Fourier Spectral Methods*. Second Edition, DOVER Publications, Inc., 2000
- [3] Brown, J., Churchill, R., Fourier Series and Boundary Value Problems. McGraw-Hill, Inc, 1993
- [4] Canuto, C., Hussaini, M.Y., Quarteroni, A., Zang, T.A., Spectral Methods. Fundamentals in Single Domains. Springer Verlag, 2006
- [5] Chandrasekhar, S.: *Brownian motion, dynamical friction, and stellar dynamics.* Rev. Mod. Phys. 21, 383-388 (1949)
- [6] Canuto, C., Hussaini, M.Y., Quarteroni, A., Zang, T.A., *Spectral Methods in Fluid Dynamics*, Springer Verlag, 1987
- [7] Clenshaw, C.W. *The numerical solution of linear differential equations in Chebyshev series*. Proc. Cambridge Philos. Soc.53, 134-149. 1957
- [8] Clenshaw, C.W., H.J. Norton *The solution of nonlinear or-dinary differential* equations in Chebyshev series. Comput. J.6, 88-92.1963



- [9] Cole, J.D. On a quasilinear parabolic equation occurring in aerodynamics. Q. Appl. Math.9, 225-236. 1951
- [10] Davis, P., Rabinowitz P. *Methods of Numerical Integration*, 2nd edn. Computer Science and Applied Mathematics, Academic Press Inc., Orlando, FL
- [11] Echim, M.M., Lemaire, J., Lie-Svendsen, O. A review on solar wind modeling: kinetic and fluid aspects. Surv. Geophys.32, 1-70 (2011)
- [12] Eliasen, E., Machenhauer, B., Rasmussen, E. On a Numerical Method for Integration of the Hydrodynamical Equations with a Spectral Representation of the Horizontal Fields. Rep. No. 2(Institut for Teoretisk Meteorologi, Univ. Copenhagen)
- [13] Epperson, James F. An Introduction to Numerical Methods and Analysis. Second Edition, Mathematical Reviews. 2013
- [14] Fahr, F.J., Shizgal, B., Modern exospheric theories and their observational relevance. Rev. Geophys. Space Phys.21, 75-124 (1983)
- [15] Finlayson, B. A. *The Method of Weighted Residuals and Variational Principles*. Academic Press, New York. 1972
- [16] Folland, Gerald B. Fourier Analysis and its Applications. Pacific Grove, California. 1992
- [17] Frazer, R.A., Jones, W.P., Skan, S.W. Approximation to Func-tions and to the Solution of Differential Equations. Rep. and Mem. 1799 (Aeronautical Research Council, London). 1937
- [18] Gheorghiu, C. I., *Spectral Methods for Differential Problems*. "T. Popoviciu" Institute of Numerical Analysis, Cluj-Napoca, ROMANIA, 2005-2007
- [19] Gottlieb, S., Gottlieb, D., Hesthaven, J., Spectral Methods for Time-Dependent Problems. Cambridge Monographs on Applied and Computational Mathematics, 2007
- [20] Gottlieb, D., Orszag, S. A., *Numerical Analysis of Spectral Methods: Theory and Applications* (SIAM-CBMS, Philadelphia), 1977
- [21] Grandclément, Philippe. *Introduction to Spectral Methods*. Laboratoire Univers et ses Th´eories, Observatoire de Paris, 5 place J. Janssen, 92195 Meudon Cedex, France. 2006



- [22] Hopf E. *The partial differential equation* $u_t + uu_x = vu_{xx}$. Commun. Pure Appl. Math.3, 201-230. 1950
- [23] Kantorovic, L.V. On a new method of approximate solution of partial differential equations. Dokl. Akad. Nauk SSSR4, 532-536 (in Russian). 1934
- [24] Khater, A.H., Temsah, R.S., Hassan, M.M., A Chebyshev spectral collocation method for solving Burgers-type equations. Journal of Computational and Applied Mathematics 222 (2008) 333-350
- [25] Khalifa, A. K., Noor, Khalida I., Noor, Muhammad A. Some numerical methods for solving Burgers equation. International Journal of the Physical Sciences Vol. 6(7), pp. 1702-1710, 4 April, 2011
- [26] Kopriva, David A. Implementing Spectral Methods for Partial Differential Equations. Algorithms for Scientists and Engineers. Springer Sciences, 2009
- [27] Kremer, G.M., An Introduction to the Boltzmann Equation and Transport Processes in Gases. Springer, New York (2010)
- [28] Kutluay, S., Bahadir, A.R., Özdes, A. *Numerical solution of one-dimensional Burgers equation: explicit and exact-explicit finite difference methods.* Journal of Computational and Applied Mathematics 103, 251-261. 1999
- [29] Lanczos, C. Trigonometric interpolation of empirical and analytical functions. J. Math. Phys.17, 123-199. 1938
- [30] Lages Lima, E. Espacios Métricos. Proyecto Euclides, IMPA, 1977.
- [31] Liboff, R.L.: *Introductory Quantum Mechanics*, 4th edn. Addison-Wesley, New York (2002)
- [32] Liboff, R.L.: *Kinetic Theory: Classical, Quantum, and Relativistic Descriptions*, 3rd edn. Springer, New York. 2003
- [33] Lie-Svendsen, O., Rees, M.H., *An improved kinetic model for the polar outflow of a minor ion.* J. Geophys. Res.101, 2415-2433. 1996
- [34] Marion, M. Gay, B. Resolution des equations de Navier-Stokes par methode pseudo-spectrale via une technique de coordination. In: Bristeau, M.-O., Glowinski, R, Hauguel, A., Periaux, J. (eds.): Sixth Int. Symp. Finite Elements Methods in Flow Problems, pp.239-243. 1986



- [35] Mason, J. C. Chebyshev Polynomials. Chapman & Hall / CRC(2003)
- [36] Messiah, A. Quantum Mechanics, vol. I. North Holland, Amsterdam (1961)
- [37] Orszag, S.A. *Numerical methods for the simulation of turbulence*. Phys. Fluids Suppl. II. 12, 250-257. 1969
- [38] Orszag, S.A. Transform method for calculation of vector coupled sums: Application to the spectral form of the vorticity equation. J. Atmosph. Sci.27, 890-895. 1970
- [39] Orszag, S.A. Accurate solution of the Orr-Sommerfeld sta-bility equation. J. Fluid Mech.50, 689-703. 1971
- [40] Orszag, S.A., Patera, A.T. Secondary instability of wall-bounded shear flow. J. Fluid Mech. 128,347-385. 1983
- [41] Ouazzani, J., Peyret, R. A Pseudo-Spectral Solution of Binary Gas Mixture Flows. In: Pandolfi, M., Piva, R (eds.): Fifth GAMM Conference on Numerical Methods in Fluid Mechanics. Vieweg, Braunschweig, pp.275-282. 1984
- [42] Peyret, R. Spectral Methods for Incompressible Viscous Flow. (Springer, Heidelberg). 2002
- [43] Pierrard, V., Lemaire, J. Acollisional model of the polar wind. J. Geophys. Res.103, 11701-11709 (1998)
- [44] Pop, I.S., Gheorghiu, C.I., A Chebyshev-Galerkin Method for Fourth Order Problems, Proceedings of ICAOR, vol. II, pp.217-220, 1997 MR 98g;41002
- [45] Risken, H., Till, F., *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd edn. Springer, Berlin (1996)
- [46] Shen, J. Efficient Spectral-Galerkin Method I. Direct Solvers of Ssecond and Fourth-Oorder Equations using Legendre Polynomials. SIAM J. Sci. Comput. Vol. 15, No. 1, pp. 1489-1505, 1994
- [47] Shen, J. Efficient Spectral-Galerkin Method II. Direct Solvers of Ssecond and Fourth-Oorder Equations using Chebyshev Polynomials. SIAM J. Sci. Comput. Vol. 16, No. 1, pp. 74-87, 1995
- [48] Shen, J., Tang, T., Spectral anf High-Order Methods with Applications. Mathematics Monograph Series. Editorial Sciences Press, 2006



- [49] Shen, J., Tang, T., Wang, L., Spectral Methods. Algorithms, Analysis and Applications. Springer Series in Computational Mathematics. Editorial Board, 2011
- [50] Shizgal, B., Spectral Methods in Chemistry and Physics. Applications to Kinetic Theory and Quantum Mechanics. Springer Series in Computational Mathematics. Editorial Board, 2014
- [51] Shizgal, B.D., Arkos, G.G. *Nonthermal escape of the atmospheres of Venus, Earth, and Mars.* Rev. Geophys. 34, 483-505 (1996)
- [52] Silberman, I., Planetary waves in the atmosphere. J. Meteorol. 11, 27-34, 1954
- [53] Slater, J.C. Electronic energy bands in metal. Phys. Rev. 45, 794-801. 1934
- [54] Stanoyevitch, Alexander. Introduction to Numerical Ordinary and Partial Differential Equations Using MATLAB. 2005
- [55] Trefethen, Lloyd N. *Spectral Methods in Matlab*. Society for Industrial and Applied Mathematics. 2000
- [56] Villadsen, J.V., Stewart, W.E. Solution of boundary value problems by orthogonal collocation. Chem. Eng. Sci.22, 1483-1501. 1967
- [57] Wengle, H. Numerical Solution of Advection-Diffusion Problems by Collocation Methods. In: Hirschel, E. (ed.): Third GAMM Conference on Numerical Methods in Fluid Mechanics. Vieweg, Braunschweig, pp.295-304. 1979
- [58] Wright, K. Chebyshev collocation methods for ordinary dif-ferential equations. Computer J. 6, 358-365. 1964