

**ESTIMACIÓN NUMÉRICA DEL TIEMPO DE
EXPLOSIÓN PARA UN PROBLEMA PARABÓLICO
DEBIDO A CONDICIONES DE FRONTERA NO
LINEALES Y TÉRMINO HOMOGÉNEO**

LUIS ANDRÉS ROSSO CERÓN

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2010**

**ESTIMACIÓN NUMÉRICA DEL TIEMPO DE
EXPLOSIÓN PARA UN PROBLEMA PARABÓLICO
DEBIDO A CONDICIONES DE FRONTERA NO
LINEALES Y TÉRMINO HOMOGÉNEO**

LUIS ANDRÉS ROSSO CERÓN

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

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

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2010**

Dedicatoria

A la memoria de mi madre, por su incondicional apoyo y sincero amor. Donde quiera que estés, mil gracias mamá.

A mi familia toda, porque de una u otra manera han estado conmigo en muchos momentos difíciles.

Al amor de mi vida, por su compañía y aliento de siempre, gracias.

A mis amigos: Sergio, Carlos, Michael, por su ayuda y amistad, y a todas aquellas personas sensatas que durante estos años conocí.

Al profesor Bernardo Mayorga, por ser un verdadero maestro y de quien mucho aprendí. Al doctor Julio César Carrillo, por su impecable dirección en este proyecto de grado, y su ingente labor en la elaboración del mismo.

Tabla de Contenidos

Dedicatoria	4
Introducción	2
1. Preliminares	3
1.1. Formulación de problemas variacionales simétricos de problemas de valor en la frontera unidimensionales	3
1.1.1. Definiciones y teoremas básicos	4
1.1.2. Espacios de funciones	6
1.1.3. Formulación variacional de un problemas de valores en la frontera	9
1.1.4. Métodos de solución de problemas variacionales	11
1.2. Métodos de Euler para problemas de valor inicial	16
1.2.1. El método de Euler	16
1.2.2. El método de diferencias finitas	23
1.2.3. Métodos de Runge-Kutta	27
2. Estimación numérica del tiempo de explosión para un problema parabólico	34
2.1. Problema de conducción de calor en una barra unidimensional	34
2.1.1. Flujo y transferencia de energía térmica	34
2.1.2. Deducción de la ecuación conducción de calor	38
2.1.3. Condiciones iniciales y de frontera	49
2.2. Fenómeno de b -explosión para un problema parabólico	52
2.2.1. Fenómeno de b -explosión	52
2.2.2. Formulación variacional	54
2.2.3. Formulación semidiscreta generalizada	55
2.2.4. Puntos de b -explosión	57
A. Un ejemplo de b-explosión para un problema no lineal	63

B. Algunos teoremas básicos del Cálculo diferencial	66
Bibliografía	71

RESUMEN

TÍTULO: ESTIMACIÓN NUMÉRICA DEL TIEMPO DE EXPLOSIÓN PARA UN PROBLEMA PARABÓLICO DEBIDO A CONDICIONES DE FRONTERA NO LINEALES Y TÉRMINO HOMOGÉNEO*

AUTOR: LUIS ANDRÉS ROSSO CERÓN**

PALABRAS CLAVES: Ecuaciones diferenciales parciales con condiciones iniciales y de frontera, ecuación de calor, Ley de Fourier, fenómeno de *b-explosión*.

DESCRIPCIÓN: El fenómeno de *b-explosión* es un fenómeno físico que aparece con frecuencia en varios experimentos particulares. Por ejemplo, cuando se trata de controlar la temperatura de un reactor nuclear deben tenerse en cuenta condiciones iniciales que, en la mayoría de los casos, constituye la temperatura inicial del mismo. Si en determinado caso la temperatura, considerada como una función de dos variables (el tiempo y la posición), aumenta sin cota superior, podría producirse una explosión en un tiempo finito t . Resulta entonces conveniente determinar el valor de dicho tiempo t para el cual ocurre este fenómeno. La ecuación en derivadas parciales que modela este fenómeno es una ecuación diferencial parabólica, y revisar los métodos numéricos convenientes para determinar el valor de t con un buen grado de aproximación constituye el propósito de esta monografía, por tanto, el interés principal consiste en localizar numéricamente el tiempo de explosión de la solución del problema parabólico,

$$\begin{aligned}u_t - u_{xx} &= 0, & (x, t) &\in (0, 1) \times [0, T) \\u_x(1, t) &= f(u(1, t)), & t &\in [0, T) \\u_x(0, t) &= 0, & t &\in [0, T) \\u(x, 0) &= \phi(x), & x &\in [0, 1]\end{aligned}$$

en donde T es una constante positiva y f y ϕ son funciones positivas y suficientemente regulares para garantizar que el problema tiene una única solución.

*Dr. Julio César Carrillo Escobar, Director del Trabajo de Grado.

**Programa de Licenciatura en Matemáticas, Escuela de Matemáticas, Facultad de Ciencias, Universidad Industrial de Santander.

ABSTRACT

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

AUTHORLUIS ANDRÉS ROSSO CERÓN**

KEYWORDS: Partial differential equations with initial conditions and boundary-value, heat equation, Fourier law, nonlinear boundary conditions, b-blowup phenomenon.

DESCRIPTION: The phenomenon of b-explosion is a physical phenomenon that occurs frequently in several individual experiments. For example, when it comes to controlling the temperature of a nuclear reactor to be taken into account initial conditions, in most cases, the initial temperature is the same. If at any case temperature, considered as a function of two variables (time and position) increases without upper bound, could cause an explosion in a finite time t . It is therefore desirable to determine the value of the time t for which this phenomenon occurs. The partial differential equation that models this phenomenon is a parabolic differential equation, and review numerical methods suitable for determining the value of t with a good degree of approximation is the purpose of this paper therefore the main interest is to find numerically the time of explosion of the solution of parabolic problem,

$$\begin{aligned}u_t - u_{xx} &= 0, & (x, t) &\in (0, 1) \times [0, T) \\u_x(1, t) &= f(u(1, t)), & t &\in [0, T) \\u_x(0, t) &= 0, & t &\in [0, T) \\u(x, 0) &= \phi(x), & x &\in [0, 1]\end{aligned}$$

where T is a positive constant and f and ϕ are positive functions and regular enough to ensure that the problem has a unique solution.

*Dr. Julio César Carrillo Escobar, Undergraduate Dissertation Director.

**Undergraduate Program of Licentiate in Mathematics, School of Mathematics, Faculty of Science, Universidad Industrial de Santander.

Introducción

Las ecuaciones diferenciales parciales constituyen, sin lugar a duda, un instrumento matemático poderoso en lo que concierne a la descripción analítica de una larga lista de fenómenos físicos, por ejemplo, en el modelado de fluidos en la mecánica de los medios continuos. Estudiar los métodos numéricos y teóricos que permitan resolver estos modelos, facilitan el desarrollo tecnológico y propician una acertada comprensión de los procesos en el mundo real.

Para muchas ecuaciones o sistemas de estas ecuaciones diferenciales, su solución se puede volver no acotada en un tiempo finito, fenómeno que es conocido como de explosión de la solución del (fenómeno asociado al) problema en tiempo finito (blow-up time). Ejemplos típicos en donde este fenómeno se observa son los problemas que involucran términos de reacción en la ecuación, o cuando aparecen términos no lineales en las condiciones de frontera asociadas con la ecuación diferencial. Específicamente, el interés de este trabajo consiste en localizar numéricamente el tiempo de explosión de la solución del problema parabólico,

$$\begin{aligned}u_t - u_{xx} &= 0, & (x, t) &\in (0, 1) \times [0, T) \\u_x(1, t) &= f(u(1, t)), & t &\in [0, T) \\u_x(0, t) &= 0, & t &\in [0, T) \\u(x, 0) &= \phi(x), & x &\in [0, 1]\end{aligned}$$

en donde T es una constante positiva y f y ϕ son funciones positivas y suficientemente regulares para garantizar que el problema tiene una única solución.

Capítulo 1

Preliminares

1.1. Formulación de problemas variacionales simétricos de problemas de valor en la frontera unidimensionales

Mediante la teoría usual de ecuaciones diferenciales ordinarias se pueden resolver ciertas clases de ecuaciones diferenciales ordinarias con unas condiciones dadas, llamados problemas de valor inicial. No obstante, muchas de las aplicaciones en la Física y la Ingeniería requieren de resolver este mismo tipo de ecuaciones diferenciales con unas condiciones en la frontera del intervalo, llamados problemas de valor en la frontera, para los cuales las técnicas de solución de problemas de valor inicial ya no funcionan. Una segunda dificultad asociada con la solución de problemas consiste en que los coeficientes o el término no homogéneo de la ecuación diferencial son funciones acotadas que tienen discontinuidades. Tales funciones no son diferenciables en todo el dominio del problema de la ecuación diferencial, pero si son integrables.

Una técnica de solución de problemas en la frontera, motivada en parte por la técnica de solución de problemas de valor inicial mediante la transformada de Laplace, consiste en obtener una formulación integral del problema llamada formulación variacional o débil del problema de valor en la frontera.

Como motivación, consideremos el siguiente problema de valor en la frontera unidimensional. Dada una función f continua en el intervalo $[a, b]$, hallar una función u tal que

$$\begin{aligned} -u'' &= f & a < x < b, \\ u(a) &= u(b) = 0. \end{aligned}$$

En este caso, la solución u del problema, también llamada **solución clásica** o

solución fuerte del problema, es una función con segunda derivadas continuas en $[a, b]$ que satisface el problema en el sentido usual.

Al multiplicar la ecuación diferencial del problema por una función cualquiera v con primera derivada continua en $[a, b]$ tal que $v(a) = v(b) = 0$, e integrar por partes, tenemos

$$\int_a^b u'(x)v'(x) dx = \int_a^b f(x)v(x) dx.$$

La principal ventaja que tiene esta nueva formulación del problema de valor en la frontera dado, es que únicamente requiere que u tenga primera derivada continua. Las demás condiciones son dadas a fin de garantizar la existencia de las integrales. Como en muchos problemas de aplicación, la función f es continua a trozos, entonces es suficiente requerir que la derivada u' sea integrable y que exista en un sentido a precisar.

En términos muy generales, este método de resolver problemas de valores en la frontera, llamado **enfoque variacional**, requiere de precisar la noción de solución débil del problema, para lo cual se requiere de los espacios de Sóbolev. La existencia y unicidad de la solución débil se garantiza mediante el teorema de Lax-Milgram.

1.1.1. Definiciones y teoremas básicos

Definición 1.1. Sea V un espacio vectorial real. Un funcional o forma lineal $F(\cdot)$ en V es un función $F : V \rightarrow \mathbb{R}$ tal que para cada $\alpha, \beta \in \mathbb{R}$ y cada $v, v' \in V$ se tiene que cumple con las siguiente propiedad:

$$F(\alpha v + \beta v') = \alpha F(v) + \beta F(v').$$

Definición 1.2. Una forma $B(\cdot, \cdot)$ en V es una función $B : V \times V \rightarrow \mathbb{R}$. Asociadas con una forma tenemos las siguientes definiciones.

1. $B(\cdot, \cdot)$ es **bilineal** si cada una de los funcionales $v \mapsto B(v, v')$ y $v' \mapsto B(v, v')$ es lineal en V .
2. $B(\cdot, \cdot)$ es **simétrica** si $B(v, v') = B(v', v)$ para cada $v, v' \in V$.
3. $B(\cdot, \cdot)$ es **definido positiva** si $B(v, v) \geq 0$ para cada $v \in V$ y $B(v, v) = 0$ si $v = 0$; o bien, si $B(v, v) > 0$ para cada $v \neq 0$.

Definición 1.3. Un **producto interno** (real), denotado (\cdot, \cdot) , en el espacio vectorial V es una forma bilineal, simétrica y definida positiva. Un espacio vectorial junto con un producto interno (\cdot, \cdot) es llamado un **espacio con producto interno** y es usualmente denotado como $(V, (\cdot, \cdot))$.

Teorema 1.1 (Desigualdad de Cauchy-Schwarz). *Si V es un espacio con producto interno $(V, (\cdot, \cdot))$, entonces*

$$|(u, v)| \leq (u, u)^{1/2}(v, v)^{1/2}$$

para cada $u, v \in V$. La igualdad se cumple si y solo si u y v son linealmente independientes.

Definición 1.4. *Sea V un espacio vectorial. Una norma $\|\cdot\|$ en V es una función $\|\cdot\| : V \rightarrow \mathbb{R}$ que cumple con las siguientes propiedades:*

1. $\|v\| \geq 0$ para cada $v \in V$.
 $\|v\| = 0$ si y solo si $v = 0$.
2. $\|kv\| = |k| \|v\|$ para cada $k \in \mathbb{R}$ y cada $v \in V$.
3. $\|v + v'\| \leq \|v\| + \|v'\|$ para cada $v, v' \in V$ (**desigualdad triangular**).

Un espacio vectorial junto con un norma $\|\cdot\|$ es llamado un **espacio normado** y es usualmente denotado como $(V, \|\cdot\|)$.

La desigualdad triangular nos permite garantizar que una norma no es un funcional lineal.

Una norma $\|\cdot\|$ en un espacio vectorial V induce de manera natural en el espacio vectorial la noción de distancia, o métrica,

$$d(v, v') = \|v - v'\|$$

para cada $v, v' \in V$. Un espacio vectorial dotado con la topología inducida por esta métrica es llamado un **espacio vectorial normado**.

Al igual que en espacios euclidianos, un espacio métrico V es completo si toda sucesión de Cauchy (v_n) de elementos de V tiene límite v en V . Para un espacio normado V , una sucesión (v_n) es de Cauchy si $\|v_m - v_n\| \rightarrow 0$ cuando $m, n \rightarrow \infty$, y completez significa que $\|v_n - v\| \rightarrow 0$ cuando $n \rightarrow \infty$.

Definición 1.5. *Un espacio vectorial normado $(V, \|\cdot\|)$ es llamado un **espacio de Banach** si V es completo en la métrica inducida por la norma $\|\cdot\|$.*

También, un producto interno (\cdot, \cdot) en un espacio vectorial V induce una norma en el espacio vectorial,

$$\|v\| = (v, v)^{1/2}$$

para cada $v \in V$. La desigualdad triangular es fácil de establecer a partir de la desigualdad de Cauchy-Schwarz. En consecuencia, todo espacio con producto interno es un espacio normado.

Definición 1.6. Sea $(V, (\cdot, \cdot))$ un espacio con producto interno. Si el espacio normado asociado $(V, \|\cdot\|)$ es completo entonces $(V, (\cdot, \cdot))$ es llamado un **espacio de Hilbert**.

Definición 1.7. Se dice que un funcional lineal $F(\cdot)$ en un espacio normado $(V, \|\cdot\|)$ es **acotado** (o **continuo**) si existe una constante positiva c tal que

$$|F(v)| \leq c \|v\|, \quad v \in V.$$

Definición 1.8. Se dice que una forma bilineal $a(\cdot, \cdot)$ en un espacio normado $(V, \|\cdot\|)$ es **acotada** (o **continua**) si existe una constante positiva α tal que

$$|a(u, v)| \leq \alpha \|u\| \|v\|, \quad u, v \in V,$$

y **coerciva** en V si existe una constante positiva β tal que

$$a(v, v) \geq \beta \|v\|^2, \quad v \in V.$$

1.1.2. Espacios de funciones

Espacios de funciones continuas

Sea (a, b) un intervalo abierto. Para cada entero no negativo k , $C^k(a, b)$ es el conjunto de las funciones definidas en (a, b) que tienen derivadas continuas hasta de orden k en (a, b) . Este conjunto de funciones es un espacio vectorial bajo las operaciones de suma de funciones y de multiplicación de un escalar por una función.

Una función u definida en (a, b) tiene **soporte compacto** si existe un compacto $K \subset (a, b)$ tal que $u(x) = 0$ para todo $x \in (a, b) \setminus K$. Se designa por $C_0^k(a, b)$ al conjunto de funciones en $C^k(a, b)$ que tienen soporte compacto.

Una función u es continua a trozos en el intervalo $[a, b]$ si el intervalo se puede dividir en un número finito de subintervalos en cada uno de los cuales u es continua y tiene límites laterales. Es obvio de esta definición que una función continua a trozos no es necesariamente continua.

El espacio de Lebesgue L^2

Denotamos como $L^2(a, b)$ al conjunto de todas las funciones u definidas de (a, b) con la propiedad de que

$$\int_a^b |u(x)|^2 dx < \infty,$$

donde la integral se considera en el sentido de Lebesgue más que en el sentido de Riemann (ver Apóstol [1] y Bartle [4] para una discusión más detallada). Es bueno recordar que si una función es Riemann integrable, también es Lebesgue integrable, y que ambas integrales tienen el mismo valor. Así, es entonces claro que toda función continua a trozos y con soporte compacto en (a, b) son funciones que están en $L^2(a, b)$. Existen funciones en $L^2(a, b)$, continuas a trozos, que difieren en los puntos en las cuales ellas son discontinuas. Por ejemplo, en $L^2(-1, 1)$, las funciones

$$f(x) = \begin{cases} x^2 & -1 < x < 0, \\ x + 1 & 0 \leq x < 1, \end{cases} \quad \text{y} \quad g(x) = \begin{cases} x^2 & -1 \leq x < 0, \\ x + 1 & 0 < x < 1, \end{cases}$$

difieren en $x_0 = 0$, y no obstante, $\|f - g\|_2 = 0$. A fin de evitar este tipo de diferencia entre funciones de $L^2(a, b)$, decimos que $f = g$ en c.t.p. (casi todas partes) de (a, b) si

$$\|f - g\|_2 = 0.$$

Por esta razón, el espacio $L^2(a, b)$ puede ser interpretado como el conjunto de clases de equivalencia de funciones que son iguales en casi todas partes de (a, b) (para más detalles, ver Bartle [4]).

El conjunto de funciones $L^2(a, b)$ es un espacio vectorial bajo las operaciones usuales de suma de funciones y de multiplicación de un escalar por una función. También, si u y v son funciones arbitrarias en $L^2(a, b)$ entonces el producto uv es Lebesgue integrable. Las funciones

$$(u, v) = \int_a^b u(x)v(x) dx, \quad u, v \in L^2(a, b),$$

y

$$\|u\|_2 = (u, u)^{1/2}, \quad u \in L^2(a, b)$$

definen en $L^2(a, b)$ un producto interno y una norma, respectivamente, llamados el producto interno y la norma de Lebesgue de orden 2, o simplemente, L^2 .

La demostración de que la norma $\|\cdot\|_2$ cumple con la desigualdad triangular se fundamenta en la desigualdad de Cauchy-Schwarz,

$$|(u, v)| \leq \|u\|_2 \|v\|_2, \quad \text{para cada } u, v \in L^2(a, b).$$

Una sucesión (u_n) en $L^2(a, b)$ se dice que *converge en la media* a $u \in L^2(a, b)$ si

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

Esta propiedad no es lo mismo que la de *convergencia puntual*,

$$\lim_{n \rightarrow \infty} u_n(x) = u(x), \quad a \leq x \leq b.$$

Una sucesión (u_n) en $L^2(a, b)$ se dice que es una *sucesión de Cauchy* si

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

Toda sucesión en $L^2(a, b)$ que converge en la media a una función en $L^2(a, b)$ es una sucesión de Cauchy. Una propiedad fundamental de $L^2(a, b)$ es que el recíproco también es cierto (ver Brézis [5, p. 57]). En virtud de esta propiedad, $L^2(a, b)$ es un espacio de Banach. Como la norma de $L^2(a, b)$ es inducida por un producto interno, tenemos además que $L^2(a, b)$ es un espacio de Hilbert.

El espacio de Sóbolev H^1

Sea (a, b) un intervalo abierto, acotado o no. Se define el espacio de Sóbolev $W^{1,2}(a, b)$ como el conjunto de todas las funciones $u \in L^2(a, b)$ para las cuales existe una función $g \in L^2(a, b)$ tal que

$$\int_a^b u(x)v'(x)dx = - \int_a^b g(x)v(x)dx, \quad v \in C_0^1(a, b).$$

La función g se llama la **derivada débil** de u en (a, b) . Al igual que el espacio de Lebesgue, este conjunto es un espacio vectorial y usualmente se denota este como $H^1(a, b)$.

Las funciones

$$(u, v)_1 = (u, v) + (u', v') = \int_a^b (u(x)v(x) + u'(x)v'(x)) dx, \quad u, v \in H^1(a, b),$$

y

$$\|u\|_1 = (u, u)_1^{1/2} = \left(\|u\|_2^2 + \|u'\|_2^2 \right)^{1/2} = \left(\int_a^b (|u(x)|^2 + |u'(x)|^2) dx \right)^{1/2}, \quad u \in H^1(a, b),$$

definen en $H^1(a, b)$ un producto interno y una norma, respectivamente, llamados el producto interno y la norma de Sóbolev de orden 1.

La ventaja de esta norma y este producto interno, es que ellos hacen del espacio de Sóbolev $H^1(a, b)$ un espacio de Hilbert (ver Brézis [5, p. 150]).

También consideramos el espacio de Sóbolev $H_0^1(a, b)$ como el espacio de las funciones en $H^1(a, b)$ que tienen soporte compacto. También $H^1(a, b)$ es un espacio de Hilbert con respecto al producto interno y la norma de Sóbolev de orden 1; además,

$$H_0^1(a, b) = \{v \in H^1(a, b) \mid v(a) = v(b) = 0\}. \quad (1.1)$$

1.1.3. Formulación variacional de un problemas de valores en la frontera

Consideremos el siguiente problema de valores en la frontera. Dada una función $f \in C[0, 1]$, o en $L^2(0, 1)$, encontrar una función $u(x)$ tal que

$$\begin{cases} -u''(x) + u(x) = f(x), & 0 < x < 1, \\ u(0) = u(1) = 0. \end{cases} \quad (1.2)$$

Este problema modela, por ejemplo, los desplazamientos de una barra de longitud uno, fijada en los extremos y sometida a una carga tangencial o bien la temperatura de un modelo de conducción térmica con una fuente de calor de intensidad $f(x)$.

Definición 1.9. Una solución *clásica* o *fuerte* de (1.2) es una función $u \in C^2[a, b]$ que verifica (1.2), en el sentido usual.

Si u es una solución clásica de (1.2) y v es una función suficientemente regular tal que $v(0) = v(1) = 0$, mediante integración por partes, se tiene que

$$\int_0^1 u''(x)v(x) dx = - \int_0^1 u'(x)v'(x) dx.$$

Por lo tanto, si u es una solución clásica de (1.2) y multiplicamos este problema por este tipo de funciones e integramos por partes, obtenemos

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

Consideremos sobre este espacio,

$$\begin{aligned} a(u, v) &:= (u', v') + (u, v) = \int_0^1 (u'(x)v'(x) + u(x)v(x)) dx \\ F(v) &:= \int_0^1 f(x)v(x) dx. \end{aligned}$$

donde (\cdot, \cdot) denota el producto interno de $L^2(0, 1)$.

Definición 1.10. Una solución *débil* al problema con valores en la frontera (1.2) es una función $u \in H^1(0, 1)$ tal que

$$a(u, v) = (f, v) \quad \text{para todo } v \in H_0^1(0, 1). \quad (1.3)$$

La formulación variacional del problema (1.2) consiste en encontrar una función del problema (1.3).

De la fórmula de integración por partes se sigue que toda solución clásica de (1.2) es una solución débil. A continuación se establece que esta solución débil existe y es única.

Es evidente que la forma $a(\cdot, \cdot)$ es bilineal, simétrica en $H_0^1(0, 1)$. Dado que

$$a(v, v) = \int_0^1 ((v')^2 + v^2) dx = (v', v') + (v, v) = \|v\|_1^2, \quad v \in H_0^1(0, 1),$$

tenemos entonces que la forma $a(\cdot, \cdot)$ es coerciva. De otro lado, como $a(\cdot, \cdot)$ es un producto interno en $H_0^1(0, 1)$, entonces se cumple en $H_0^1(0, 1)$ la desigualdad de Cauchy-Schwarz:

$$|a(u, v)| \leq a(u, u)^{1/2} a(v, v)^{1/2} \quad \text{para todo } u, v \in H_0^1(0, 1).$$

Como

$$a(v, v) = \|v\|_1^2 \quad \text{para todo } v \in H_0^1(0, 1),$$

se sigue que $a(\cdot, \cdot)$ es continua en $H_0^1(0, 1)$:

$$|a(u, v)| \leq \|u\|_1 \|v\|_1, \quad \text{para todo } u, v \in H_0^1(0, 1).$$

De otro lado, si $v \in H_0^1$ existe un compacto $[a, b]$ en $(0, 1)$ tal que

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

donde $c = \max\{\sup_{x \in [a, b]} |f(x)|, 1\}$. Esto demuestra que el funcional lineal F es acotado en $H_0^1(0, 1)$.

Teorema 1.2 (Existencia y unicidad de solución débil). *Para cada $f \in L^2(0, 1)$ existe una única solución u de (1.3) en $H_0^1(0, 1)$. Además,*

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

donde I es el funcional lineal

$$I(v) = \frac{1}{2} a(v, v) - F(v), \tag{1.4}$$

el cual es llamado el **principio de Dirichlet**.

Demostración. Como la forma $a(\cdot, \cdot)$ es bilineal, coerciva y acotada y el funcional $F(\cdot)$ es lineal, el resultado se sigue del teorema de Lax-Milgram (ver Brézis [5, p. 84]). La simetría de $a(\cdot, \cdot)$ garantiza que u es el mínimo del funcional $I(\cdot)$ en $H_0^1(0, 1)$ (ver de nuevo Brézis [5, p. 84]). \square

1.1.4. Métodos de solución de problemas variacionales

Como se mencionó en la introducción de la sección anterior, basta obtener la formulación variacional del problema de valores en la frontera (1.2) en el espacio vectorial

$$C_0^1(0, 1) = \{v \in C^1(0, 1) \mid v(0) = v(1) = 0\},$$

cuando $f \in C(0, 1)$. El problema es que este espacio no es completo en la norma $L^2(0, 1)$, $\|\cdot\|_2$, y por tanto, tampoco lo es en la norma de Sóbolev de orden 1, $\|\cdot\|_1$. Puede demostrarse que al completar el espacio $C_0^1(0, 1)$ en la norma $\|\cdot\|_1$ se obtiene el espacio de Sóbolev $H_0^1(0, 1)$ (ver Brézis [5]). En este espacio, como ya se demostró, existe una única solución débil al problema (1.2), aún en el caso que $f \in L^2(0, 1)$.

Los métodos de solución numérica del problema variacional (1.3), llamados métodos variacionales, utilizan este hecho para obtener una sucesión de soluciones aproximadas, la cual es una sucesión de Cauchy. La completitud de $H_0^1(0, 1)$ garantiza que el límite de esta sucesión es la solución débil del problema variacional (1.3), cosa que no se puede garantizar en el caso en que se resuelva el problema variacional en $C_0^1(0, 1)$.

Método de Ritz

El método para encontrar una solución aproximada del problema variacional asociado con un problema de valor en la frontera, requiere de alguna clase de discretización.

Definición 1.11. *Dado un espacio vectorial de dimensión finita $H_n \subset H_0^1$, una solución de **Ritz** al problema con valores en la frontera (1.2) es una función $u_n \in H_n$ tal que*

$$a(u_n, v) = (f, v) \quad \text{para todo } v \in H_n. \quad (1.5)$$

El problema de aproximación de Ritz del problema (1.2) consiste en encontrar una solución del problema (1.5).

En el caso del método de Ritz, se aplica el teorema de Lax-Milgram para obtener una solución del problema variacional en un espacio H_n de dimensión finita de $H_0^1(0, 1)$. Como H_n es también un espacio de Hilbert, el teorema de Lax-Milgram garantiza la existencia de una única solución del problema variacional en H_n .

Teorema 1.3 (Existencia y unicidad de la solución de Ritz). *Para cada $f \in L^2(0,1)$ existe una única solución u_n de (1.5) en H_1 . Además,*

$$I(u_n) = \min_{v \in H_n} I(v),$$

donde I es el funcional lineal dado en (1.4).

Observación 1.1. *Cuando el problema (1.2) incluye condiciones de frontera de la forma $u(0) = u(1) = 0$, la forma bilineal $a(\cdot, \cdot)$ no es simétrica. En tal caso, el Teorema 1.2 garantiza únicamente la existencia de solución a través del problema variacional (1.3), pero no del funcional I . En este caso, el método variacional se llama de Galerkin. Por las mismas razones, el problema variacional (1.3) se llama simétrico y el segundo no simétrico.*

El Teorema 1.3 no es constructivo en el sentido que no especifica un procedimiento para encontrar a u_n . No obstante, como H_n es un espacio vectorial de dimensión finita, consideremos que este espacio tiene una base de funciones $\{\phi_n\}$. Como todo elemento de H_n tiene una representación única en términos de los elementos de esta base, entonces existen unos únicos escalares $\alpha_1, \dots, \alpha_n$ tales que

$$u_n(x) = \sum_{i=1}^n \alpha_i \phi_i(x). \quad (1.6)$$

Reemplazando en (1.4) y usando la propiedad de bilinealidad de $a(\cdot, \cdot)$ y la propiedad de linealidad de $F(\cdot)$, tenemos que

$$\begin{aligned} I(u_n) &= \frac{1}{2} a \left(\sum_{i=1}^n \alpha_i \phi_i, \sum_{i=1}^n \alpha_i \phi_i \right) - F \left(\sum_{i=1}^n \alpha_i \phi_i \right) \\ &= \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j a(\phi_i, \phi_j) - \sum_{i=1}^n \alpha_i F(\phi_i) \\ &= \frac{1}{2} \boldsymbol{\alpha}^T M \boldsymbol{\alpha} - \boldsymbol{\alpha}^T F \\ &= I(\boldsymbol{\alpha}), \end{aligned}$$

donde

$$\begin{aligned} M &= \begin{bmatrix} a(\phi_1, \phi_1) & a(\phi_1, \phi_2) & \cdots & a(\phi_1, \phi_n) \\ a(\phi_2, \phi_1) & a(\phi_2, \phi_2) & \cdots & a(\phi_2, \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ a(\phi_n, \phi_1) & a(\phi_n, \phi_2) & \cdots & a(\phi_n, \phi_n) \end{bmatrix}, \\ \boldsymbol{\alpha} &= \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad F = \begin{bmatrix} F(\alpha_1) \\ \vdots \\ F(\alpha_n) \end{bmatrix}, \end{aligned} \quad (1.7)$$

M se conoce como la matriz de rigidez, con elementos $m_{ij} = a(\phi_i, \phi_j) = (\phi'_i, \phi'_j)$ y F es la matriz de los términos forzantes $F_i = F(\phi_i)$. La simetría de $a(\cdot, \cdot)$ implica la simetría de la matriz K . Además, de la propiedad de coercividad de $a(\cdot, \cdot)$ tenemos que

$$\boldsymbol{\alpha}^T M \boldsymbol{\alpha} = a(u_n, u_n) = \|u_n\|_1^2 \geq 0.$$

Como $\boldsymbol{\alpha}^T M \boldsymbol{\alpha} = 0$ si y solo si $\boldsymbol{\alpha} = 0$, tenemos demostrado el siguiente teorema.

Teorema 1.4. *Como la forma $a(\cdot, \cdot)$ es bilineal y coerciva entonces la matriz K en (1.7) es definida positiva.*

Como $I(\boldsymbol{\alpha})$ es un campo escalar en \mathbb{R}^n con matriz hessiana definida positiva, entonces $I(\boldsymbol{\alpha})$ es minimizado por el $\boldsymbol{\alpha}$ que cumple

$$M \boldsymbol{\alpha} = F.$$

El mínimo u_n de I sobre H_n es precisamente $\boldsymbol{\alpha}$, en virtud de (1.6).

Método de elementos finitos

En el método de Ritz, las funciones base de H_n tienen soporte global, en el sentido que están definidas en todo el intervalo $[0, 1]$. El método de elementos finitos se entiende como una implementación del método de Ritz con funciones base de elementos finitos. Este tipo de funciones base son polinomios continuos a trozos y que tienen soporte local, en el sentido que las funciones se anulan fuera de un pequeño subintervalo de $[a, b]$. En el caso unidimensional, como es nuestro caso, tales funciones son definidas con respecto a alguna partición del intervalo $[0, 1]$. Así, el método de Ritz se puede entender como una solución aproximada del problema variacional (1.3) mediante funciones base con soporte global y el método de elementos finitos, como una solución del mismo problema con funciones de soporte local, en el sentido definido anteriormente.

Definición de funciones de base global. La construcción de las funciones base global para el método de elementos finitos se fundamenta en la siguiente idea. Sea $v \in C^0([0, 1])$ una función que cumple con las condiciones de frontera homogéneas $v(0) = v(1) = 0$. Consideramos una partición $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ de $[0, 1]$. Entonces la poligonal u_n que une los puntos $(x_0, u(x_0)), (x_1, u(x_1)), \dots, (x_n, u(x_n))$, la cual es un polinomio lineal continuo a trozos, es una aproximación en $H_0^1(0, 1)$ de u .

Definición 1.12 (Funciones base global). *Sea $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ de $[0, 1]$ y sea H_n el espacio vectorial que satisfacen lo siguiente:*

1. $v \in C^0[0, 1]$,
2. v restringido al intervalo $[0, 1]$, $i = 1, 2, \dots, n$ es un polinomio lineal,
3. $v(0) = v(1) = 0$.

Para la misma partición \mathcal{P} e $i = 1, 2, \dots, n$ definimos a ϕ_i como los polinomios lineales continuos a trozos en $H_0^1(0, 1)$ que cumplen la condición

$$\phi_i(x_j) = \delta_{ij},$$

la delta de Kronecker.

Lema 1.1. *El conjunto de funciones $\{\phi_i \mid i = 1, 2, \dots, n\}$ es linealmente independiente.*

Demostración. Evidentemente, como cada ϕ_i ,

$$\sum_{i=1}^n c_i \phi_i(x_j) = 0$$

implica que $c_i = 0$, para cada $i = 1, 2, \dots, n$, y por tanto las funciones ϕ_i son linealmente independientes. \square

Para ver que el conjunto $\{\phi_i \mid i = 1, 2, \dots, n\}$ genera a H_n , consideramos la siguiente definición.

Definición 1.13. *Dada una función $v \in C^0[0, 1]$, la función interpolación $v_I \in H_n$ de v es la poligonal*

$$v_I(x) = \sum_{i=1}^n v(x_i) \phi_i(x), \quad \text{para cada } x \in [0, 1].$$

De esta definición concluimos que v_I está en el espacio generado por el conjunto de funciones $\{\phi_i \mid i = 1, 2, \dots, n\}$.

Lema 1.2. *Si $v \in H_n$ entonces $v = v_I$.*

Demostración. La función $v - v_I$ es lineal en cada subintervalo $[x_{i-1}, x_i]$ y además

$$\begin{aligned} v(x_j) - v_I(x_j) &= v(x_j) - \sum_{i=1}^n v(x_i)\phi_i(x_j) \\ &= v(x_j) - \sum_{i=1}^n v(x_i)\delta_{i,j} \\ &= v(x_j) - v(x_j) \\ &= 0, \end{aligned}$$

por lo cual $v - v_I$ debe ser idénticamente cero. \square

Teorema 1.5. *El conjunto de funciones $\{\phi_i \mid i = 1, 2, \dots, n\}$ forma una base de H_n .*

Demostración. El resultado se sigue de los lemas anteriores. \square

Construcción de las funciones de base local y global. Sea $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ una partición de $[0, 1]$. En cada subintervalo $e_i = [x_{i-1}, x_i]$, llamado elemento finito, e $i = 1, \dots, n$, consideramos las funciones base local $\phi_1^{(i)}(x)$ y $\phi_2^{(i)}(x)$ en e_i , definidas como

$$\phi_1^{(i)}(x) = \frac{x_i - x}{x_i - x_{i-1}}, \quad \phi_2^{(i)}(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}}, \quad \text{si } x_{i-1} \leq x \leq x_i.$$

Para $i = 1, \dots, n-1$, definimos la función global $\phi_i(x)$ en el intervalo $[x_{i-1}, x_{i+1}] = e_i \cup e_{i+1}$ como

$$\begin{aligned} \phi_i(x) &= \begin{cases} \phi_2^{(i)}(x), & \text{si } x \in e_i, \\ \phi_1^{(i+1)}(x), & \text{si } x \in e_{i+1}, \\ 0, & \text{en caso contrario.} \end{cases} \\ &= \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & \text{si } x_{i-1} \leq x \leq x_i, \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & \text{si } x_i \leq x \leq x_{i+1}, \\ 0, & \text{en caso contrario.} \end{cases} \end{aligned} \tag{1.8}$$

$$\phi_n(x) = \begin{cases} \frac{x - x_{n-1}}{x_n - x_{n-1}}, & \text{si } x_{n-1} \leq x \leq x_n, \\ 0, & \text{en caso contrario.} \end{cases} \tag{1.9}$$

Es fácil verificar que estas funciones cumplen con las condiciones de la Definición 1.12.

1.2. Métodos de Euler para problemas de valor inicial

A continuación se discutirán los métodos de Euler, de diferencias finitas y de Runge-Kutta. El primero de ellos tiene como característica fundamental el hecho de que la variable espacial x se mantiene fija o constante, mientras que se permite la variación de la variable temporal t , pudiendo tener condiciones iniciales (sobre la función o sus derivadas) o de frontera en distintos valores de t , para un valor fijo de x .

De otro lado el método de diferencias finitas hace uso del teorema de Taylor para aproximar una función f mediante polinomios en torno a un punto x_0 del dominio de f . Esta idea es extendida para aproximar la primera y segunda derivada de f , lo que origina métodos particulares de aproximación (diferencias retrógradas, centradas, progresivas), y que se generaliza en los métodos de Runge-Kutta, ver [11].

1.2.1. El método de Euler

Definición 1.14. *Se dice que el problema de valor inicial*

$$\frac{dy}{dt} = f(t, y), \quad a \leq t \leq b, \quad y(a) = \alpha$$

es un problema bien planteado si:

1. *existe una única solución, $y(t)$, para este problema;*
2. *existen constantes positivas ε y k , con la propiedad de que existe una única solución $z(t)$, al problema*

$$\frac{dz}{dt} = f(t, z) + \delta(t), \quad a \leq t \leq b, \quad z(a) = \alpha + \varepsilon_0,$$

con

$$|z(t) - y(t)| < k\varepsilon, \quad \text{para toda } a \leq t \leq b$$

siempre que $|\varepsilon_0| < \varepsilon$ y $|\delta(t)| < \varepsilon$.

El objetivo del método de Euler es obtener una aproximación al problema de valor inicial bien planteado

$$\frac{dy}{dt} = f(t, y), \quad (1.10)$$

donde $a \leq t \leq b$ y $y(a) = \alpha$. Con este método no se obtiene una aproximación continua de la solución $y(t)$, sino que se generan aproximaciones de y en varios puntos, llamados **puntos de red**, en el intervalo $[a, b]$. Una vez que se obtenga la solución aproximada en estos puntos, la solución aproximada en otros puntos del intervalo se puede encontrar utilizando interpolación.

Se hace la suposición de que los puntos de red están distribuidos uniformemente sobre el intervalo $[a, b]$. Tal condición se puede garantizar escogiendo un entero positivo N y seleccionando los puntos de red $\{t_0, t_1, t_2, \dots, t_N\}$ donde

$$t_i = a + ih \quad \text{para cada } i = 1, 2, \dots, N.$$

La distancia entre dos puntos consecutivos de la red, t_i y t_{i+1} está dada por $h = \frac{b-a}{N}$ y es llamada **tamaño de paso**.

Se usará el Teorema de Taylor (B.3) para derivar el método de Euler.

Supongamos que $y(t)$, la única solución de (1.10), tiene dos derivadas continuas en $[a, b]$, de tal manera que para cada $i = 0, 1, 2, \dots, N-1$, $y(t_{i+1})$ puede escribirse como

$$y(t_{i+1}) = y(t_i) + (t_{i+1} - t_i)y'(t_i) + \frac{(t_{i+1} - t_i)^2}{2}y''(\xi_i)$$

para algún ξ_i , donde $t_i < \xi_i < t_{i+1}$. Usando la notación $h = t_{i+1} - t_i$, se tiene

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(\xi_i)$$

y, como $y(t)$ satisface la ecuación diferencial (1.10), entonces

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}y''(\xi_i).$$

El método de Euler construye $w_i \approx y(t_i)$ para cada $i = 1, 2, \dots, N$, donde

$$\begin{aligned} w_0 &= \alpha \\ w_{i+1} &= w_i + hf(t_i, w_i). \end{aligned}$$

La anterior ecuación se llama la **ecuación de diferencia** asociada con el método de Euler.

Ejemplo 1.1. Encuentre aproximaciones al problema de valor inicial

$$y' = -y + t + 1, \text{ si } 0 \leq t \leq 1, y(0) = 1$$

En primer lugar, la solución característica de esta ecuación diferencial es $y_c = c_1 e^{-t}$; $c_1 \in \mathbb{R}$, y la solución particular propuesta en este caso es de la forma $y_p = At + B$. Por lo tanto $A + At + B = t + 1$, de donde se deduce que $A = 1$ y $B = 0$, lo cual implica que la solución general a esta ecuación diferencial es

$$y(t) = Ce^{-t} + t,$$

y haciendo uso de la condición inicial $y(0) = 1$, se obtiene que $C = 1$, es decir

$$y(t) = t + e^{-t}$$

es la solución al problema.

Supongamos que $N = 10$, como $0 \leq t \leq 1$, entonces $h = 0,1$ y $t_i = 0,1i$. Usando el hecho de que $f(t, y) = -y + t + 1$, se tiene

$$\begin{aligned} w_0 &= 1 \\ w_i &= w_{i-1} + h(-w_{i-1} + t_{i-1} + 1) \\ &= w_{i-1} + 0,1(-w_{i-1} + 0,1(i-1) + 1) \\ &= 0,9w_{i-1} + 0,01(i-1) + 0,1 \end{aligned}$$

para $i = 1, 2, \dots, 10$. La siguiente tabla compara los valores aproximados en t_i y los valores exactos.

t_i	w_i	y_i	Error = $ w_i - y_i $
0,0	1,000000	1,000000	0,0
0,1	1,000000	1,004837	0,004837
0,2	1,010000	1,018731	0,008731
0,3	1,029000	1,040818	0,011818
0,4	1,056100	1,070320	0,014220
0,5	1,090490	1,106531	0,016041
0,6	1,131441	1,148812	0,017371
0,7	1,178297	1,196585	0,018288
0,8	1,230467	1,249329	0,018862
0,9	1,287420	1,306570	0,019150
1,0	1,348678	1,367879	0,019201

Nótese que el error crece ligeramente conforme el valor de t_i aumenta. El crecimiento controlado del error es una consecuencia de la estabilidad del método de

Euler, por lo que se espera que los errores ocasionados por el redondeo crezcan cuando mucho en forma lineal.

Para obtener una cota del error para el caso general del método de Euler necesitamos los siguientes lemas

Lema 1.3. *Para toda $x \geq -1$ y para cualquier m positiva,*

$$0 \leq (1+x)^m \leq e^{mx}.$$

Demostración. Aplicando el Teorema de Taylor (B.3) con $f(x) = e^x$, $x_0 = 0$ y $n = 1$, obtenemos

$$e^x = 1 + x + \frac{x^2}{2}e^\xi,$$

con $\xi \in (x, 0)$. Entonces

$$0 \leq 1 + x \leq 1 + x + \frac{x^2}{2}e^\xi = e^x$$

y, como $1 + x \geq 0$,

$$0 \leq (1+x)^m \leq (e^x)^m = e^{mx}. \quad \square$$

Lema 1.4. *Si s y t son números reales positivos, $\{a_i\}_{i=0}^k$ es una sucesión que satisface $a_0 \geq -\frac{t}{s}$, y*

$$a_{i+1} \leq (1+s)a_i + t, \text{ para cada } i = 0, 1, 2, \dots, k,$$

entonces

$$a_{i+1} \leq e^{(i+1)s} \left(\frac{t}{s} + a_0 \right) - \frac{t}{s}.$$

Demostración. Para un entero fijo i , la desigualdad de la hipótesis implica que

$$\begin{aligned} a_{i+1} &\leq (1+s)a_i + t \\ &\leq (1+s)[(1+s)a_{i-1} + t] + t \\ &\leq (1+s)(1+s)[(1+s)a_{i-2} + t] + t + t \\ &\vdots \\ &\leq (1+s)^{i+1}a_0 + [1 + (1+s) + (1+s)^2 + \dots + (1+s)^i]t. \end{aligned}$$

Ahora bien,

$$1 + (1 + s) + (1 + s)^2 + \cdots + (1 + s)^i = \sum_{j=0}^i (1 + s)^j$$

es una serie geométrica con razón $(1 + s)$ y, por tanto, su suma es

$$\frac{1 - (1 + s)^{(i+1)}}{1 - (1 + s)} = \frac{1}{s}[(1 + s)^{i+1} - 1].$$

Así que,

$$a_{i+1} \leq (1 + s)^{i+1} a_0 + \frac{(1 + s)^{i+1} - 1}{s} t = (1 + s)^{i+1} \left(\frac{t}{s} + a_0 \right) - \frac{t}{s},$$

y en virtud del lema (1.3),

$$a_{i+1} \leq e^{(i+1)s} \left(\frac{t}{s} + a_0 \right) - \frac{t}{s}. \quad \square$$

Teorema 1.6. *Sea $y(t)$ la única solución al problema de valor inicial*

$$y' = f(y, t); \quad a \leq t \leq b, \quad y(a) = \alpha$$

y sean w_0, w_1, \dots, w_N las aproximaciones generadas por el método de Euler para algún entero positivo N .

Si f satisface una condición de Lipschitz con constante L en

$$D = \{(t, y) : a \leq t \leq b, -\infty < y < \infty\},$$

y existe una constante M con la propiedad de que

$$|y''(t)| \leq M, \quad \text{para toda } t \in [a, b],$$

entonces

$$|y(t_i) - w_i| \leq \frac{hM}{2L} [e^{L(t_i-a)} - 1]$$

para cada $i = 0, 1, 2, \dots, N$.

Demostración. Cuando $i = 0$ el resultado es claramente cierto, ya que $y(t_0) = w_0 = \alpha$.

Del método de Euler se tiene que para cada $i = 0, 1, 2, \dots, N - 1$,

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(\xi_i);$$

y también, de la ecuación de diferencia

$$w_{i+1} = w_i + hf(t_i, w_i).$$

Por consiguiente, usando la notación $y_i = y(t_i)$ e $y_{i+1} = y(t_{i+1})$, resulta

$$y_{i+1} - w_{i+1} = y_i - w_i + h[f(t_i, y_i) - f(t_i - w_i)] + \frac{h^2}{2}y''(\xi_i) \quad \text{y}$$

$$|y_{i+1} - w_{i+1}| \leq |y_i - w_i| + h|f(t_i, y_i) - f(t_i - w_i)| + \frac{h^2}{2}|y''(\xi_i)|.$$

Las suposiciones de que f satisface una condición de Lipschitz para la segunda variable con constante L y que $y''(t) \leq M$ implican que

$$|y_{i+1} - w_{i+1}| \leq |y_i - w_i|(1 + hL) + \frac{h^2M}{2}.$$

Usando el lema (1.4) y tomando $a_j = |y_j - w_j|$ para cada $j = 0, 1, \dots, N$, mientras que $s = hL$ y $t = \frac{h^2M}{2}$, vemos que

$$|y_{i+1} - w_{i+1}| \leq e^{hL(i+1)} \left(|y_0 - w_0| + \frac{h^2M}{2hL} \right) - \frac{h^2M}{2hL}.$$

Como $|y_0 - w_0| = 0$ y $(i+1)h = t_{i+1} - t_0 = t_{i+1} - a$, tenemos el resultado deseado:

$$|y_{i+1} - w_{i+1}| \leq [e^{L(t_{i+1}-a)} - 1]$$

para cada $i = 0, 1, 2, \dots, N - 1$. □

Ejemplo 1.2. *Regresando al problema de valor inicial*

$$y' = -y + t + 1, \quad \text{si } 0 \leq t \leq 1, \quad y(0) = 1$$

Se observa que con $f(t, y) = -y + t + 1$, $\frac{\partial f}{\partial y} = -1$, así que el teorema (1.6) implica que f satisface una condición de Lipschitz con $L = 1$.

Además, como en este caso se conoce la solución exacta que es

$$y(t) = t + e^{-t},$$

se tiene que $y''(t) = e^{-t}$, y

$$|y''(t)| \leq e^{-0} = 1 \text{ para toda } t \in [0, 1].$$

Usando la desigualdad

$$|y(t_i) - w_i| \leq \frac{hM}{2L} [e^{L(t_i-a)} - 1]$$

con $h = 0,1$ y $M = L = 1$ obtenemos la cota de error

$$|y_i - w_i| \leq 0,005(e^{t_i} - 1).$$

La siguiente tabla muestra el error real encontrado en el ejemplo (1.1) junto con esta cota de error.

Nótese que, aun cuando se usó la cota real para la segunda derivada de la solución, la cota de error es considerablemente mayor que el error real.

t_i	0,1	0,2	0,3	0,4	0,5
Error real	0,00484	0,00873	0,01182	0,01422	0,01604
Cota de error	0,00526	0,01107	0,01749	0,02459	0,03244
t_i	0,6	0,7	0,8	0,9	1,0
Error real	0,01737	0,01829	0,01886	0,01915	0,01920
Cota de error	0,04111	0,05069	0,06128	0,07298	0,08591

Lo más importante de la fórmula de la cota de error dada por el Teorema (1.6) es el hecho de que la cota depende linealmente del tamaño de paso h . Esto significa que si se disminuye el tamaño de paso las aproximaciones serán más precisas. En el resultado del Teorema (1.6) no se considera el efecto que el error de redondeo tiene en la elección del tamaño de paso. Es evidente que si se escoge h menor, se necesitan llevar a cabo más cálculos y que cuando se usa aritmética de dígitos finitos, que es lo más usual se espera que el error de redondeo sea mayor. Por ello, en la práctica no se usa la forma de la ecuación de diferencia

$$\begin{aligned} w_0 &= \alpha, \\ w_i &= w_{i-1} + h f(t_{i-1}, w_{i-1}) \end{aligned}$$

para cada $i = 1, 2, \dots, N$, para calcular la aproximación a la solución y_i en el punto de red t_i , en su lugar, se usa una ecuación de la forma

$$\begin{aligned} u_0 &= \hat{\alpha}, \\ u_i &= u_{i-1} + h f(t_{i-1}, u_{i-1}) + \delta_i \end{aligned}$$

para cada $i = 1, 2, \dots, N$, donde $\hat{\alpha}$ es la aproximación al valor inicial α y δ_i denota el error de redondeo asociado con los cálculos de $u_{i-1} + h f(t_{i-1}, u_{i-1})$.

1.2.2. El método de diferencias finitas

Aproximaciones polinómicas

La idea fundamental para los cálculos numéricos en la técnica de diferencias finitas se basa en las aproximaciones de una función $f(x)$ mediante polinomios de Taylor cerca de $x = x_0$. Sea $x = x_0 + \Delta x$, esto es, $\Delta x = x - x_0$. Si se aproximara a $f(x)$ por una constante c en cercanías de $x = x_0$, se elegiría como aproximación $f(x_0) = c$, o sea, un polinomio de grado cero. Una mejor aproximación a $f(x)$ cerca de x_0 se obtiene usando la recta tangente de f en el punto $x = x_0$, de la siguiente forma,

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0) = f(x_0) + \Delta x f'(x_0),$$

es decir, una aproximación lineal. Puede también considerarse una aproximación cuadrática a $f(x)$ en cercanías de x_0 ,

$$f(x) \approx f(x_0) + \Delta x f'(x_0) + (\Delta x)^2 \frac{f''(x_0)}{2!},$$

de tal forma que los valores de la función, y de su primera y segunda derivada en $x = x_0$, coincidan con los de $f(x)$. Se puede pensar que la aproximación a $f(x)$ cuando $\Delta x \rightarrow 0$, es mejor a medida en que el grado del polinomio aumenta. De hecho, el Teorema de Taylor (B.3) así lo demuestra.

Error de truncamiento

En virtud de Teorema de Taylor (B.3), se tiene que

$$f(x) = f(x_0) + \Delta x f'(x_0) + (\Delta x)^2 \frac{f''(x_0)}{2!} + \cdots + (\Delta x)^n \frac{f^{(n)}(x_0)}{n!} + R_n,$$

donde

$$R_n = (\Delta x)^{n+1} \frac{f^{(n+1)}(\xi_{n+1})}{(n+1)!}, \text{ con } x_0 < \xi_{n+1} < x = x_0 + \Delta x.$$

La expresión anterior se llama *resto de la serie de Taylor*, y tiene una forma parecida al último término en la fórmula de (B.3), pero está evaluado en un punto intermedio, casi siempre desconocido.

Ejemplo 1.3. Cuando se calcula f en $x = x_0 + \Delta x$, mediante un polinomio lineal, se obtiene

$$f(x_0 + \Delta x) = f(x_0) + \Delta x f'(x_0) + (\Delta x)^2 \frac{f''(\xi_2)}{2!}, \quad x_0 < \xi_2 < x, \quad (1.11)$$

si $\Delta x \rightarrow 0$, entonces ξ_2 está contenido en un intervalo pequeño y, suponiendo que $\frac{d^2 f}{dx^2}$ es continua, el error de truncamiento es aproximadamente

$$R_1 \approx (\Delta x)^2 \frac{f''(x_0)}{2!}.$$

Si además se tiene que $\frac{d^2 f}{dx^2}$ está acotada, o sea, $\left| \frac{d^2 f}{dx^2} \right| < M$, el error de truncamiento es del orden de $(\Delta x)^2$, es decir,

$$|R_1| \leq K(\Delta x)^2, \text{ con } K = \frac{M}{2},$$

y se denota por $R_1 = O((\Delta x)^2)$.

Aproximaciones de la primera derivada

Usando la serie de Taylor, es posible aproximar las derivadas de varias maneras. Por ejemplo, a partir de (1.11)

$$\frac{df}{dx}(x_0) = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - (\Delta x)^2 \frac{f''(\xi_2)}{2!}. \quad (1.12)$$

Introducimos entonces una aproximación por diferencias finitas, la aproximación por **diferencias progresivas** a $\frac{df}{dx}$:

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \quad (1.13)$$

Esta expresión es muy parecida a la definición de derivada. Aquí utilizamos una diferencia progresiva, pero no se toma el límite cuando $\Delta x \rightarrow 0$. Como (1.12) es válida para todo Δx , puede sustituirse Δx por $-\Delta x$ y obtener así la aproximación por **diferencias regresivas** a $\frac{df}{dx}$,

$$\frac{df}{dx}(x_0) = \frac{f(x_0 - \Delta x) - f(x_0)}{-\Delta x} + (\Delta x)^2 \frac{f''(\bar{\xi}_2)}{2!}, \quad (1.14)$$

de donde:

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0 - \Delta x) - f(x_0)}{-\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \quad (1.15)$$

Comparando (1.12) con (1.13) y (1.14) con (1.15), se observa que el error de truncamiento es $O(\Delta x)$, y casi idéntico para las aproximaciones por diferencias progresivas y retrógradas de la primera derivada. Para obtener una aproximación más precisa de $\frac{df}{dx}(x_0)$, podemos calcular la media de las aproximaciones progresivas y retrógradas. Sumando (1.12) y (1.14), se tiene:

$$2\frac{df}{dx}(x_0) = \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{\Delta x} + \frac{\Delta x}{2} \left[\frac{d^2 f}{dx^2}(\bar{\xi}_2) - \frac{d^2 f}{dx^2}(\xi_2) \right]. \quad (1.16)$$

Como $\bar{\xi}_2$ está cerca de ξ_2 , se espera que el error casi se cancele y sea mucho menor que $O(\Delta x)$. Para reducir el error en esta aproximación, volvemos a la serie de Taylor para $f(x_0 - \Delta x)$ y $f(x_0 + \Delta x)$:

$$f(x_0 + \Delta x) = f(x_0) + \Delta x f'(x_0) + \frac{(\Delta x)^2}{2!} f''(x_0) + \frac{(\Delta x)^3}{3!} f'''(x_0) + \dots \quad (1.17)$$

$$f(x_0 - \Delta x) = f(x_0) - \Delta x f'(x_0) + \frac{(\Delta x)^2}{2!} f''(x_0) - \frac{(\Delta x)^3}{3!} f'''(x_0) + \dots \quad (1.18)$$

Restando (1.18) de (1.17), se obtiene:

$$f(x_0 + \Delta x) - f(x_0 - \Delta x) = 2\Delta x f'(x_0) + \frac{2}{3!} (\Delta x)^3 f'''(x_0) + \dots$$

Así que se espera que

$$f'(x_0) = \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x} - \frac{(\Delta x)^2}{6} f'''(\xi_3). \quad (1.19)$$

Esto nos lleva a la aproximación por **diferencias centrales** de $\frac{df}{dx}(x_0)$:

$$f'(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x}. \quad (1.20)$$

Ejemplo 1.4. Para $f(x) = \cos x$ y utilizando $\Delta x = 0,1$, considérese la aproximación numérica de $\left(\frac{df}{dx}\right)\left(\frac{\pi}{2}\right)$. A diferencia de los problemas prácticos aquí se conoce la respuesta exacta, $f'\left(\frac{\pi}{2}\right) = -\text{sen}\left(\frac{\pi}{2}\right) = -1$. Utilizando una calculadora, y haciendo $x_0 = \frac{\pi}{2}$, se obtiene

$$f\left(\frac{\pi}{2} + 0,1\right) = f\left(\frac{5\pi + 1}{10}\right) = \cos\left(\frac{5\pi + 1}{10}\right) \approx -0,099833$$

y

$$f\left(\frac{\pi}{2} - 0,1\right) = f\left(\frac{5\pi - 1}{10}\right) = \cos\left(\frac{5\pi - 1}{10}\right) \approx 0,099833,$$

luego

$$f'\left(\frac{\pi}{2}\right) \approx \frac{f\left(\frac{5\pi+1}{10}\right) - f\left(\frac{5\pi-1}{10}\right)}{2(0,1)} = \frac{-0,099833 - 0,099833}{2(0,1)} = -0,990815.$$

Usando el resto en la fórmula de Taylor para la función f , se calcula el error esperado. Para las diferencias progresivas y retrógradas se tiene

$$E \approx \left| \frac{\Delta x}{2} \frac{d^2 f}{dx^2} \left(\frac{\pi}{2}\right) \right| = \left| -\frac{0,1}{2} \cos\left(\frac{\pi}{2}\right) \right| = 0,$$

mientras que para diferencias centradas, es

$$E \approx \frac{(\Delta x)^2}{6} \frac{d^3 f}{dx^3} \left(\frac{\pi}{2}\right) = \frac{(0,1)^2}{6} \operatorname{sen}\left(\frac{\pi}{2}\right) = \frac{1}{600} = 0,001\bar{6}$$

se observa, en este caso, que las diferencias centradas dejan un error no nulo en la aproximación del valor $f'\left(\frac{\pi}{2}\right)$, y por tanto es el método apropiado para tal cálculo.

Segundas derivadas

Sumando (1.17) y (1.18), obtenemos

$$f(x_0 + \Delta x) + f(x_0 - \Delta x) = 2f(x_0) + (\Delta x)^2 f''(x_0) + \frac{2(\Delta x)^4}{4!} f^{(iv)}(x_0) + \dots$$

Por lo que se espera que

$$f''(x_0) = \frac{f(x_0 + \Delta x) - 2f(x_0) + f(x_0 - \Delta x)}{(\Delta x)^2} - \frac{(\Delta x)^2}{12} f^{(iv)}(\xi). \quad (1.21)$$

Esto nos da una aproximación por diferencias finitas para la segunda derivada con error de truncamiento $O(\Delta x)^2$:

$$\frac{d^2 f}{dx^2}(x_0) \approx \frac{f(x_0 + \Delta x) - 2f(x_0) + f(x_0 - \Delta x)}{(\Delta x)^2}. \quad (1.22)$$

La anterior ecuación se llama aproximación por **diferencias centradas** para la segunda derivada, ya que también se puede obtener aplicando varias veces las fórmulas de diferencias centradas para las primeras derivadas.

Derivadas parciales

Al resolver ecuaciones en derivadas parciales, se analizan funciones de dos o más variables, por ejemplo $u(x, y)$, $u(x, t)$ y $u(x, y, t)$. Los métodos numéricos utilizan a menudo aproximaciones por diferencias finitas. Algunas derivadas parciales (aunque no todas) se pueden obtener usando los métodos anteriores para funciones de una variable. Por ejemplo, para una función $u(x, y)$, la derivada parcial $\frac{\partial u}{\partial x}$ es, en realidad, una derivada ordinaria $\frac{du}{dx}$ si mantenemos a y fijo. Por tanto, es posible utilizar las fórmulas de diferencias progresivas, retrógradas o centradas. Usando la fórmula de diferencias centradas, por ejemplo, se obtiene

$$\frac{\partial u}{\partial x}(x_0, y_0) \approx \frac{u(x_0 + \Delta x, y_0) - u(x_0 - \Delta x, y_0)}{2\Delta x}.$$

Para la derivada $\frac{\partial u}{\partial y}$, si se mantiene a x fijo, se obtiene

$$\frac{\partial u}{\partial y}(x_0, y_0) \approx \frac{u(x_0, y_0 + \Delta y) - u(x_0, y_0 - \Delta y)}{2\Delta y},$$

en uso de las fórmulas de diferencias centradas.

1.2.3. Métodos de Runge-Kutta

Los métodos de Runge-Kutta son usados en la práctica para reemplazar a los métodos de Taylor. Si bien es cierto que estos últimos tienen la propiedad de tener un error de truncamiento local grande, tienen la desventaja de requerir el cálculo y evaluación de las derivadas de la función que se está trabajando. Esto puede resultar un procedimiento muy complicado y tedioso, más aun si se tiene una gran cantidad de problemas. En cambio, los métodos de Runge-Kutta eliminan el cálculo y evaluación de las derivadas de la función, conservando un error de truncamiento local grande.

Los métodos de Runge-Kutta son aquellos que se pueden escribir en la forma

$$\begin{cases} k_i = f\left(x_n + c_i h, y_n + h \sum_{j=1}^s a_{ij} k_j\right), & i = 1, 2, \dots, s \\ y_{n+1} = y_n + h \sum_{i=1}^s b_i k_i. \end{cases}$$

Cada una de las evaluaciones de la función k_i es una *etapa*.

Para derivar el método se requiere del siguiente teorema, cuya demostración puede encontrarse, por ejemplo, en [10].

Teorema 1.7. *Supongamos que $f(t, y)$ y todas sus derivadas parciales de orden menor o igual a $n + 1$ son continuas en $D = \{(t, y) : a \leq t \leq b, c \leq y \leq d\}$. Sea $(t_0, y_0) \in D$. Para toda $(t, y) \in D$, existe ξ entre t y t_0 y η entre y y y_0 con*

$$f(t, y) = P_n(t, y) + R_n(t, y),$$

donde

$$\begin{aligned} P_n(t, y) = & f(t_0, y_0) + \left[(t - t_0) \frac{\partial f}{\partial t}(t_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(t_0, y_0) \right] \\ & + \left[\frac{(t - t_0)^2}{2} \frac{\partial^2 f}{\partial t^2}(t_0, y_0) + (t - t_0)(y - y_0) \frac{\partial^2 f}{\partial t \partial y}(t_0, y_0) + \frac{(y - y_0)^2}{2} \frac{\partial^2 f}{\partial y^2}(t_0, y_0) \right] \\ & + \cdots + \left[\frac{1}{n!} \sum_{j=0}^n \binom{n}{j} (t - t_0)^{n-j} (y - y_0)^j \frac{\partial^n f}{\partial t^{n-j} \partial y^j}(t_0, y_0) \right] \end{aligned}$$

y

$$R_n(t, y) = \frac{1}{(n+1)!} \sum_{j=0}^{n+1} \binom{n+1}{j} (t - t_0)^{n+1-j} (y - y_0)^j \frac{\partial^{n+1} f}{\partial t^{n+1-j} \partial y^j}(\xi, \eta).$$

$P_n(t, y)$ se llama el *polinomio de Taylor de grado n en dos variables* para la función f alrededor de (t_0, y_0) , y $R_n(t, y)$ es el término de residuo asociado con $P_n(t, y)$.

Ejemplo 1.5. *El polinomio de Taylor de grado tres para $f(t, y) = \sin(ty)$ alrededor de $(0, \pi)$ se encuentra a partir de*

$$\begin{aligned} P_3(t, y) = & f(t_0, y_0) + (t - t_0) \frac{\partial f}{\partial t}(t_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(t_0, y_0) \\ & + \left[\frac{(t - t_0)^2}{2} \frac{\partial^2 f}{\partial t^2}(t_0, y_0) + (t - t_0)(y - y_0) \frac{\partial^2 f}{\partial t \partial y}(t_0, y_0) + \frac{(y - y_0)^2}{2} \frac{\partial^2 f}{\partial y^2}(t_0, y_0) \right] \\ & + \left[\frac{(t - t_0)^3}{6} \frac{\partial^3 f}{\partial t^3}(t_0, y_0) + \frac{(t - t_0)^2 (y - y_0)}{2} \frac{\partial^3 f}{\partial t^2 \partial y}(t_0, y_0) + \right. \\ & \left. + \frac{(t - t_0)(y - y_0)^2}{2} \frac{\partial^3 f}{\partial t \partial y^2}(t_0, y_0) + \frac{(y - y_0)^3}{6} \frac{\partial^3 f}{\partial y^3}(t_0, y_0) \right]. \end{aligned}$$

Evaluando cada una de estas derivadas parciales en $(t_0, y_0) = (0, \pi)$, el polinomio anterior se reduce a

$$P_3(t, y) = \pi t + t(y - \pi) - \frac{\pi^3}{6} t^3.$$

Este polinomio dará una buena aproximación de $\text{sen}(ty)$ siempre y cuando t esté cerca de cero e y esté cerca de π . Por ejemplo,

$$P_3(0, 01; \pi + 0, 01) \approx 0, 03151076,$$

mientras que

$$\text{sen}(0, 01(\pi + 0, 01)) \approx 0, 03151071$$

El primer paso para derivar un método de Runge-Kutta consiste en determinar valores para a_1, α_1 y β_1 con la propiedad de que $a_1 f(t + \alpha_1, y + \beta_1)$ se aproxime a

$$T^{(2)}(t, y) = f(t, y) + \frac{h}{2} f'(t, y)$$

con un error no mayor que $O(h^2)$, que es el error de truncamiento local para el método de Taylor de orden dos.

Como

$$f'(t, y) = \frac{df}{dt}(t, y) = \frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) \cdot y'(t) \quad \text{y} \quad y'(t) = f(t, y),$$

esto implica que

$$T^{(2)}(t, y) = f(t, y) + \frac{h}{2} \frac{\partial f}{\partial t}(t, y) + \frac{h}{2} \frac{\partial f}{\partial y}(t, y) \cdot f(t, y). \quad (1.23)$$

Expandiendo a $f(t + \alpha_1, y + \beta_1)$ en su polinomio de Taylor de grado uno alrededor de (t, y) implica que:

$$a_1 f(t + \alpha_1, y + \beta_1) = a_1 f(t, y) + a_1 \alpha_1 \frac{\partial f}{\partial t}(t, y) + a_1 \beta_1 \frac{\partial f}{\partial y}(t, y) + a_1 R_1(t + \alpha_1, y + \beta_1), \quad (1.24)$$

donde

$$R_1(t + \alpha_1, y + \beta_1) = \frac{\alpha_1^2}{2} \frac{\partial^2 f}{\partial t^2}(\xi, \eta) + \alpha_1 \beta_1 \frac{\partial^2 f}{\partial t \partial y}(\xi, \eta) + \frac{\beta_1^2}{2} \frac{\partial^2 f}{\partial y^2}(\xi, \eta) \quad (1.25)$$

para algunas ξ entre t y $t + \alpha_1$, y η entre y y $y + \beta_1$. Igualando los coeficientes en las ecuaciones (1.23) y (1.24) se obtienen las tres ecuaciones:

$$\begin{aligned} f(t, y) : a_1 &= 1, \\ \frac{\partial f}{\partial t}(t, y) : a_1 \alpha_1 &= \frac{h}{2}, \\ \frac{\partial f}{\partial y}(t, y) : a_1 \beta_1 &= \frac{h}{2} f(t, y). \end{aligned}$$

Los parámetros a_1 , α_1 y β_1 están determinados de manera única por

$$a_1 = 1, \quad \alpha_1 = \frac{h}{2} \quad \text{y} \quad \beta_1 = \frac{h}{2} f(t, y);$$

así,

$$T^{(2)}(t, y) = f\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) - R_1\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right)$$

y de la ecuación (1.25),

$$\begin{aligned} R_1\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right) &= \frac{h^2}{8} \frac{\partial^2 f}{\partial t^2}(\xi, \eta) + \frac{h^2}{4} f(t, y) \frac{\partial^2 f}{\partial t \partial y}(\xi, \eta) + \\ &+ \frac{h^2}{8} (f(t, y))^2 \frac{\partial^2 f}{\partial y^2}(\xi, \eta). \end{aligned}$$

Si todas las segundas derivadas parciales de f están acotadas, el error

$$R_1 = \left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right)$$

será de $O(h^2)$, el orden del error de truncamiento local del método de Taylor de orden dos.

El método de diferencia que resulta al reemplazar $T^{(2)}(t, y)$ en el método de Taylor de orden dos por $f(t + (h/2), y + (h/2)f(t, y))$ es un método de Runge-Kutta específico, conocido como el **método del punto medio**.

Método del punto medio

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + hf\left(t_i + \frac{h}{2}, w_i + \frac{h}{2} f(t, y)\right) \end{cases}$$

para cada $i = 0, 1, \dots, N - 1$. Ya que en $a_1 f(t + \alpha_1, y + \beta_1)$, solamente están presentes tres parámetros y todos son necesarios para la comparación con $T^{(2)}$, se necesitará una forma más compleja para satisfacer las condiciones requeridas para cualquier método de Taylor de orden mayor.

La forma de cuatro parámetros más apropiada para aproximar

$$T^{(3)}(t, y) = f(t, y) + \frac{h}{2} f'(t, y) + \frac{h^2}{6} f''(t, y),$$

es

$$a_1 f(t, y) + a_2 f(t + \alpha_2, y + \delta_2 f(t, y)); \quad (1.26)$$

y aún con esto, no hay suficiente flexibilidad para hacer coincidir el término

$$\frac{h^2}{6} \left[\frac{\partial f}{\partial y}(t, y) \right] f(t, y)$$

que resulta de la expansión de $\frac{h^2}{6} f''(t, y)$. Por lo tanto, lo mejor que se puede obtener usando (1.26) son métodos con una cota de error de $O(h^2)$. Sin embargo, el hecho de que (1.26) tenga cuatro parámetros permite cierta flexibilidad en su elección de tal manera que se puedan derivar una serie de métodos de $O(h^2)$. Los dos más importantes son, el **método de Euler modificado**, que corresponde a escoger $a_1 = a_2 = \frac{1}{2}$ y $\alpha_2 = \delta_2 = h$ y tiene la forma de ecuación de diferencia:

Método de Euler modificado

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + \frac{h}{2}[f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))] \end{cases}$$

para cada $i = 0, 1, 2, \dots, N - 1$, y el **método de Heun**, que corresponde a $a_1 = \frac{1}{4}$, $a_2 = \frac{3}{4}$, y $\alpha_2 = \delta_2 = \frac{2}{3}h$ y que tiene la forma de ecuación de diferencia

Método de Heun

$$\begin{cases} w_0 = \alpha \\ w_{i+1} = w_i + \frac{h}{4}[f(t_i, w_i) + 3f(t_{i+1} + \frac{2}{3}h, w_i + \frac{2}{3}hf(t_i, w_i))] \end{cases}$$

para cada $i = 0, 1, 2, \dots, N - 1$.

Ambos métodos están clasificados como métodos de Runge-Kutta de orden dos, que es el orden de su error de truncamiento local.

Ejemplo 1.6. *Aplicando cualquiera de los métodos de Runge-Kutta de orden 2 al Ejemplo (1.1), se obtiene la ecuación de diferencia*

$$\begin{aligned} w_0 &= 1, \\ w_{i+1} &= 0,905w_i + 0,0095i + 0,1. \end{aligned}$$

Para comparar los diferentes resultados de estos métodos usaremos en su lugar, el problema de valor inicial

$$y' = -y + t^2 + 1, \quad 0 \leq t \leq 1, \quad y(0) = 1.$$

cuya solución exacta es $y(t) = -2e^{-t} + t^2 - 2t + 3$. En la siguiente tabla aparecen los resultados de estos cálculos.

t_i	Valor exacto	Método del punto medio		Método de Euler modificado		Método de Heun	
		Valor	Error	Valor	Error	Valor	Error
0,0	1,0000000	1,0000000	0	1,0000000	0	1,0000000	0
0,1	1,0003252	1,0002500	$7,52 \cdot 10^{-5}$	1,0005000	$1,75 \cdot 10^{-4}$	1,0003333	$8,10 \cdot 10^{-6}$
0,2	1,0025385	1,0024263	$1,12 \cdot 10^{-4}$	1,0029025	$3,64 \cdot 10^{-4}$	1,0025850	$4,65 \cdot 10^{-5}$
0,3	1,0083636	1,0082458	$1,18 \cdot 10^{-4}$	1,0089268	$5,63 \cdot 10^{-4}$	1,0084728	$1,09 \cdot 10^{-4}$
0,4	1,0193599	1,0192624	$9,75 \cdot 10^{-5}$	1,0201288	$7,69 \cdot 10^{-4}$	1,0195512	$1,91 \cdot 10^{-4}$
0,5	1,0369387	1,0368825	$5,62 \cdot 10^{-5}$	1,0379166	$9,78 \cdot 10^{-3}$	1,0372272	$2,88 \cdot 10^{-4}$
0,6	1,0623767	1,0623787	$2,00 \cdot 10^{-6}$	1,0635645	$1,19 \cdot 10^{-3}$	1,0627739	$3,97 \cdot 10^{-4}$
0,7	1,0968294	1,0969027	$7,33 \cdot 10^{-5}$	1,0982259	$1,40 \cdot 10^{-3}$	1,0973437	$5,14 \cdot 10^{-4}$
0,8	1,1413421	1,1414969	$1,55 \cdot 10^{-4}$	1,1429444	$1,60 \cdot 10^{-3}$	1,1419794	$6,37 \cdot 10^{-4}$
0,9	1,1968607	1,1971047	$2,44 \cdot 10^{-4}$	1,1986647	$1,80 \cdot 10^{-3}$	1,1976247	$7,64 \cdot 10^{-4}$
1,0	1,2642411	1,2645798	$3,39 \cdot 10^{-4}$	1,2662416	$2,00 \cdot 10^{-3}$	1,2651337	$8,93 \cdot 10^{-4}$

Aun cuando $T^3(t, y)$ se puede aproximar con un error de $O(h^3)$ mediante una expresión de la forma

$$f(t + \alpha_1, y + \delta_1 f(t + \alpha_2, y + \alpha_2 f(t, y))),$$

que involucra cuatro parámetros, el álgebra necesaria para determinar a α_1 , δ_1 , α_2 , y δ_2 es bastante complicada y no entraremos en detalles aquí. De hecho, el método de Runge-Kutta de orden tres que resulta de esta expresión no se usa generalmente en la práctica. El método de Runge-Kutta que se usa más a menudo es de orden cuatro y, en forma de diferencia, está dado por:

$$\begin{aligned} w_0 &= \alpha, \\ k_1 &= hf(t_i, w_i), \\ k_2 &= hf\left(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_1\right), \\ k_3 &= hf\left(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_2\right), \\ k_4 &= hf(t_{i+1}, w_i + k_3), \\ w_{i+1} &= w_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \end{aligned}$$

para cada $i = 0, 1, \dots, N - 1$. Este método tiene un error de truncamiento de $O(h^4)$, siempre y cuando la solución $y(t)$ tenga cinco derivadas continuas.

Ejemplo 1.7. Usando el método de Runge-Kutta de cuarto orden para determinar aproximaciones a la solución del problema de valor inicial

$$y' = -y + t + 1, \quad 0 \leq t \leq 1, \quad y(0) = 1,$$

con $h = 0,1$, $N = 10$, y $t_i = 0,1i$ se obtienen los resultados y errores que se muestran en la siguiente tabla.

t_i	Valores exactos	Valores de Runge-Kutta de orden cuatro	Error
0,0	1,0000000000	1,0000000000	0
0,1	1,0048374180	1,0048375000	$8,200 \cdot 10^{-8}$
0,2	1,0187307531	1,0187309014	$1,483 \cdot 10^{-7}$
0,3	1,0408182207	1,0408184220	$2,013 \cdot 10^{-7}$
0,4	1,0703200460	1,0703202889	$2,429 \cdot 10^{-7}$
0,5	1,1065306597	1,1065309344	$2,747 \cdot 10^{-7}$
0,6	1,1488116360	1,1488119344	$2,984 \cdot 10^{-7}$
0,7	1,1965853038	1,1965856187	$3,149 \cdot 10^{-7}$
0,8	1,2493289641	1,2493292897	$3,256 \cdot 10^{-7}$
0,9	1,3065696597	1,3065699912	$3,315 \cdot 10^{-7}$
1,0	1,3678794412	1,3678797744	$3,332 \cdot 10^{-7}$

Ejemplo 1.8. Para el problema

$$y' = -y + 1, 0 \leq t \leq 1, y(0) = 0,$$

compararemos el método de Euler con $h = 0,025$, el método de Euler modificado con $h = 0,05$ y el método de Runge-Kutta de cuarto orden con $h = 0,1$ en los puntos de red $t_i = 0,1; 0,2; 0,3; 0,4$, y $0,5$. Los resultados se dan en la siguiente tabla.

t	Método de Euler	Método de Euler modificado	Método de Runge-Kutta de cuarto orden	Valor real
	$h = 0,025$	$h = 0,05$	$h = 0,1$	
0,1	0,096312	0,095123	0,09516250	0,095162582
0,2	0,183348	0,181198	0,18126910	0,181269247
0,3	0,262001	0,259085	0,25918158	0,259181779
0,4	0,333079	0,329563	0,32967971	0,329679954
0,5	0,397312	0,393337	0,39346906	0,393469340

Para este ejemplo, el método de cuarto orden es claramente superior.

Capítulo 2

Estimación numérica del tiempo de explosión para un problema parabólico

2.1. Problema de conducción de calor en una barra unidimensional

Para comenzar revisaremos los conceptos físicos que sustentan la formulación de la ecuación de calor, tales como *el primer principio de la termodinámica, los modos de transferencia de calor, la ley de Fourier* y el muy importante *principio de la conservación de la energía térmica*. Con base en lo anterior se planteará la ecuación de calor con condiciones iniciales y de frontera que serán ajustadas al problema que nos atañe, explicando detenidamente lo que cada una de ellas significa. Si se desea conocer más detalle acerca de la deducción de esta ecuación pueden consultarse las referencias bibliográficas [2] y [10].

2.1.1. Flujo y transferencia de energía térmica

Flujo de energía térmica

El proceso de transferencia de calor nos es familiar a todos. En días fríos nos ponemos más ropa a fin de reducir la transferencia de calor de nuestro cuerpo tibio al frío ambiente. Para preparar una taza de café podemos usar una cafetera dentro de la cual se transfiere calor de una resistencia eléctrica al agua, calentándola hasta su ebullición. La disciplina de ingeniería conocida como **transferencia de calor** trata los métodos para calcular **velocidades** de transferencia de energía térmica. Los ingenieros utilizan dichos métodos en el diseño de componentes y sistemas en

los que tiene lugar una transferencia de calor. Este tema es importante para casi todas las ramas de la tecnología. Sin embargo, tradicionalmente la disciplina que más se ha ocupado de la transferencia de calor ha sido la ingeniería mecánica, en virtud de la importancia del proceso de transferencia en los sistemas de conversión de energía, desde centrales eléctricas de carbón hasta calentadores de agua que utilizan energía solar.

Muchos problemas de *diseño térmico* requieren la disminución de las tasas de transferencia de calor por medio de un *aislante* adecuado. El aislamiento térmico de los edificios en climas extremos es un ejemplo muy conocido, pero hay muchos más. El transbordador espacial está equipado con losas térmicas que aíslan al vehículo de las altas temperaturas producidas por la onda de choque durante la reentrada a la atmósfera. Los criostatos, que mantienen las temperaturas criogénicas necesarias para el uso de superconductores, deben aislarse eficazmente con objeto de reducir la carga de enfriamiento del sistema de refrigeración. Con frecuencia la única manera de garantizar la protección contra el calentamiento intenso consiste en utilizar un fluido como *sumidero* de calor. Para enfriar las toberas de los motores de combustible líquido de los cohetes, se hace pasar el combustible frío a través de conductos dispuestos alrededor de las toberas antes de inyectarlo en la cámara de combustión. Uno de los componentes críticos de un reactor de fusión es la *primera pared* del recipiente, que debe ser capaz de soportar el intenso calor del plasma candente. Tales paredes pueden enfriarse por medio de un flujo de helio gaseoso o de litio líquido. La transferencia de calor desde un fluido a otro es un problema común en diseño térmico. Los dispositivos que se usan para tal efecto se llaman *intercambiadores de calor*. Un ejemplo cotidiano es el radiador de un automóvil, en el cual el calor del refrigerante caliente se transmite al aire frío que pasa a través del armazón del radiador.

En la producción de energía y en las industrias procesadoras se necesitan muchos tipos distintos de intercambiadores de calor. Una central eléctrica, ya sea de combustible fósil o nuclear, tiene una *caldera* en la que se produce vapor de agua para impulsar las turbinas, y un *condensador* en el que el vapor se condensa a fin de reducir la sobrepresión en las turbinas y recuperar el agua. El condensador que patentó James Watt en 1769 aumentaba a más del doble el rendimiento de las máquinas de vapor de la época; esto puso en marcha la Revolución Industrial. Los sistemas comunes de refrigeración o de aire acondicionado tienen un *evaporador* que absorbe calor a baja temperatura y un *condensador* que cede energía calorífica a una temperatura superior. El condensador de un refrigerador doméstico consiste, por lo general, en un serpentín con *aletas* de enfriamiento que aumentan el intercambio de calor con el medio ambiente. Una refinería de petróleo posee una gran variedad de equipos para la transferencia de calor, que incluye columnas de rectificación y *crackers* o alambiques de descomposición térmica. Se

emplean muchos intercambiadores para transferir calor entre diferentes estadios del proceso con objeto de reducir el consumo total de energía de la refinería.

A menudo el problema de diseño es un problema de *control térmico*, es decir, se pretende que la temperatura de funcionamiento de los componentes sensibles al calor se mantenga dentro de un margen específico. El enfriamiento de toda clase de elementos electrónicos es un ejemplo de control térmico. El desarrollo de computadores más rápidos está hoy seriamente limitado por la dificultad de controlar la temperatura de componentes muy pequeños que disipan grandes cantidades de calor. Un problema muy complicado es el control térmico de dispositivos sensibles a la temperatura en satélites de comunicaciones en órbita alrededor de la Tierra. Es preciso que los transistores y los diodos no se sobrecalienten, que las baterías no se congelen, que los componentes ópticos de los telescopios no se desajusten debido a la dilatación térmica, y que las fotografías se procesen a la temperatura adecuada para garantizar una alta resolución. El control térmico de las estaciones espaciales del futuro presentará obstáculos aún mayores, pues deberán incluir los sistemas necesarios para hacerlas habitables.

De los ejemplos anteriores se desprende que la transferencia de calor abarca una amplia gama de fenómenos físicos y de sistemas de ingeniería. Es preciso comprender y cuantificar estos fenómenos antes de desarrollar una metodología para el diseño térmico de sistemas.

La transferencia de calor y su relación con la termodinámica

Cuando colocamos un objeto caliente en un ambiente frío, se enfría: el objeto pierde energía interna, mientras que el ambiente la gana. Habitualmente describimos esta interacción como una *transferencia de calor* del objeto a la región que lo rodea. Puesto que la teoría del calórico cayó en descrédito hace mucho tiempo, no creemos que una *sustancia calórica* fluya del objeto a su entorno. Más bien decimos que se ha transferido energía interna por medio de complicadas interacciones en la escala atómica o subatómica. A pesar de todo, sigue siendo común describir estos procesos como transferencia, transporte, o flujo, de calor. La transferencia de calor en ingeniería se ocupa del cálculo de la velocidad a la que el calor fluye en un medio dado, a través de una interfaz o entre dos superficies, así como de la determinación de las temperaturas asociadas.

Es importante entender la diferencia esencial entre transferencia de calor en ingeniería y lo que usualmente se llama termodinámica. La termodinámica clásica aborda sistemas en equilibrio. Su metodología puede usarse para calcular la energía necesaria para llevar a un sistema de un estado de equilibrio a otro, pero no permite determinar la velocidad a la que ocurre el cambio. Por ejemplo, si la temperatura de un lingote de hierro de 1 kg se hace descender bruscamente de

1000°C A 100°C en un baño de aceite, la termodinámica nos dice que el lingote ha perdido una cantidad de energía igual a su masa (1kg) multiplicada por su calor específico (450J/kgK) y por la variación de temperatura (900K), es decir, aproximadamente 405kJ. Sin embargo, la termodinámica no puede decirnos cuánto tiempo tendremos que esperar para que la temperatura descienda a 100°C. El tiempo depende de la temperatura del baño de aceite, de las propiedades físicas del aceite, de su movimiento, y de otros factores. Un análisis adecuado de la transferencia de calor considerará todos estos elementos.

El estudio de procesos de transferencia de calor requiere el uso de algunos conceptos de termodinámica. En particular, se aplica la **primera ley de la termodinámica**, generalmente en formas muy sencillas, puesto que los efectos de trabajo a menudo pueden despreciarse. La primera ley es una manera de enunciar el *principio de conservación de la energía*, que es una de las leyes fundamentales de la física. Este principio puede formularse de muchas maneras si se excluyen las formas de energía que no atañen al problema en cuestión, o sencillamente redefiniendo qué se entiende por energía. En transferencia de calor, con frecuencia nos referimos a la primera ley como *principio de conservación de la energía*, o simplemente como *balance de calor* o de *energía* cuando no se realiza trabajo. Sin embargo, como ocurre en la termodinámica, es esencial usar la forma correcta de la primera ley. A continuación se presentan algunas formas sencillas del principio de conservación de la energía.

Supongamos que un sistema cerrado consistente en una masa sólida fija tiene un volumen $V[m^3]$, y el sólido una densidad $\rho[kg/m^3]$. Se transfiere calor al sistema a una velocidad $Q[J/s \text{ o } W]$; dentro del sólido puede generarse calor, por ejemplo mediante fisión nuclear o por una corriente eléctrica, a la velocidad $Q_v[W]$. Puede suponerse que los sólidos son incompresibles, de modo que no hay trabajo realizado por el sistema o sobre él. El principio de conservación de la energía exige que en un intervalo de tiempo $\Delta t[s]$,

$$\boxed{\text{Variación de energía interna dentro del sistema}} = \boxed{\text{Calor transferido hacia el interior del sistema}} + \boxed{\text{Calor generado dentro del sistema}},$$

es decir,

$$\Delta U = Q\Delta t + Q_v\Delta t$$

Modos de transferencia de calor

En termodinámica el *calor* se define como la energía que se transfiere debido a gradientes o diferencias de temperatura. De manera consistente, con este punto de vista, la termodinámica sólo reconoce dos modos de transferencia de calor: *conducción* y *radiación*. Por ejemplo, la transferencia de calor a través de la pared

de una tubería de acero se realiza por conducción, mientras que la transferencia del Sol a la Tierra o a una nave espacial se efectúa por radiación térmica. Estas modalidades de transferencia de calor se producen a escala molecular o subatómica.

En aire a presión normal la conducción se realiza por medio de moléculas que se desplazan distancias muy cortas ($0,65\mu m$) antes de chocar e intercambiar energía con otra molécula. Por otro lado, la radiación se efectúa por medio de fotones que viajan casi sin encontrar obstáculos a través del aire, de una superficie a otra. Así, una diferencia importante entre la conducción y la radiación es que los portadores de energía en la conducción tienen un *camino libre medio* corto, mientras que en la radiación ocurre lo contrario. Sin embargo, en el aire y a presiones muy bajas, típicas en dispositivos de alto vacío, el camino libre medio de las moléculas puede ser mucho mayor que las dimensiones del dispositivo, de modo que las moléculas pueden desplazarse sin obstáculos de una superficie a otra. En este caso la transferencia de calor por medio de moléculas se rige por leyes análogas a las de radiación.

Ahora estudiaremos la teoría física que sustenta la formulación de la ecuación de calor en el caso unidimensional, específicamente sobre una barra rígida orientada en la dirección del eje x ; así como las condiciones iniciales y de frontera que sobre esta se imponen.

2.1.2. Deducción de la ecuación conducción de calor

Densidad de energía térmica

Partimos considerando una barra de área transversal constante A orientada en la dirección del eje x , cuyos puntos inicial y final se encuentran en las coordenadas $x = 0$ y $x = 1$, como se observa en la Figura 2.1. Se introduce ahora la cantidad de energía térmica por unidad de volumen de dicha barra como una cantidad desconocida, dependiendo de la posición y del tiempo.

$$\varrho(x, t) \equiv \text{densidad de energía térmica.}$$

No hay que confundir la cantidad anteriormente definida con la densidad del material, que en nuestro caso es una barra de longitud uno orientada a lo largo del eje x . Si la barra no es uniforme, la densidad depende de la posición sobre la barra, es decir,

$$\rho(x) = \frac{M}{V}(x) \equiv \text{densidad de la barra.}$$

A las cantidades físicas que dependen del tiempo se les llama *no estacionarias*. Consideremos ahora un *volumen de control* V^c sobre la barra, que no es otra cosa

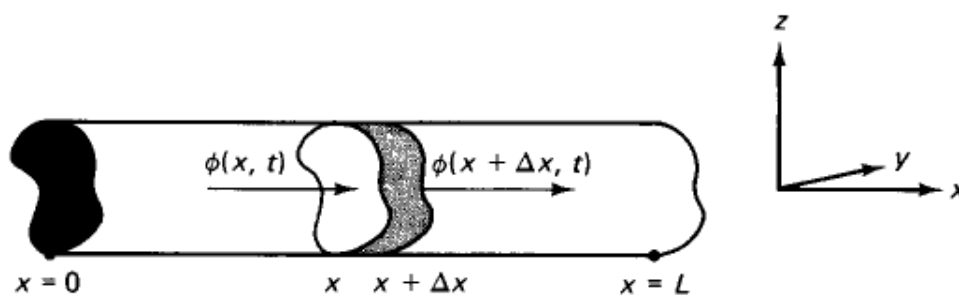


Figura 2.1: Flujo de energía en un subvolumen de control.

que una pequeña sección fina de esta, es decir, si $[a, b] \subset (0, 1)$, entonces

$$V^c \equiv [a, b] \times A,$$

si $a = x$ y $b = x + \Delta x$, con $x \in (0, 1)$ y $\Delta x > 0$, entonces afirmar que el volumen de control es una sección fina de la barra equivale a decir que $\Delta x \rightarrow 0$. Daremos por cierto que todas las cantidades térmicas se mantienen constantes a través del volumen de control considerado. La manera más sencilla en que esto se hace es aislando perfectamente la superficie lateral. En este caso la dependencia de x y t corresponde a una situación en la que la barra no es calentada uniformemente (de ahí que aparezca la variable tiempo), y además que la densidad de energía térmica varía de una sección transversal a otra (por eso la presencia de la variable x o de posición).

La energía térmica

El volumen de control está comprendido entre x y $x + \Delta x$, si la densidad de energía térmica es constante en todo el volumen, la energía total en el volumen de control está dada por el producto de la densidad de energía térmica y del volumen. En términos generales esta magnitud no es constante; pero si es que Δx es *considerablemente pequeño*, es decir, si el volumen de control es *muy* fino, entonces $\varrho(x, t)$ puede aproximarse como una constante en este volumen y estaría dada por la fórmula

$$\text{energía térmica} = \varrho(x, t)A \Delta x,$$

siendo $A \Delta x$ el volumen del volumen de control considerado.

Principio de la conservación de la energía térmica

La energía térmica entre x y $x + \Delta x$ varía con el tiempo debido únicamente a la energía que fluye a través de los bordes (x y $x + \Delta x$) y por aquella generada en el interior. No hay cambios en la energía térmica que fluyan a través de la superficie lateral, ya que hemos supuesto que la superficie lateral está aislada. El proceso fundamental de flujo de calor está descrito en palabras por la siguiente ecuación :

$$\boxed{\text{Tasa de variación de la energía térmica en el tiempo}} = \boxed{\text{energía térmica que fluye a través de las fronteras por unidad de tiempo}} + \boxed{\text{energía térmica generada en el interior por unidad de tiempo}}$$

A esta relación se le llama *principio de la conservación de la energía térmica*. Para un volumen de control dado, la tasa de cambio de la energía térmica está dada por

$$\frac{\partial}{\partial t}[\varrho(x, t)A\Delta x],$$

donde la derivada parcial $\frac{\partial}{\partial t}$ depende sólo del tiempo debido a que x se mantiene fijo.

Flujo de calor

Como consideramos el caso unidimensional de una barra, la energía térmica puede fluir solamente en dos sentidos: de izquierda a derecha o viceversa. Definimos ahora el **flujo de calor**, denotado por ϕ , de la siguiente forma:

$$\phi(x, t) = \text{flujo de calor},$$

y representa la *cantidad de calor por unidad de tiempo que fluye a través de las caras laterales de la barra*. Si $\phi(x, t) < 0$, se asume que el calor está fluyendo hacia la derecha.

Consideremos una partición del intervalo $[a, b]$, o lo que es igual una partición longitudinal del volumen de control en la dirección del eje x , con

$$a = x_0 < x_1 < \cdots < x_n = b,$$

cada *subrebanada* originada por esta partición será denominada subvolumen de control, así, el i -ésimo subvolumen de control dentro del volumen de control limitado por los puntos $a = x$ y $b = x + \Delta x$ está dado por

$$V_i^c = [x_{i-1}, x_i] \times A,$$



Figura 2.2: Flujo de energía en un volumen de control.

con $a < x_{i-1} < x_i < b$.

En cada subvolumen de control asumiremos que el flujo de calor es uniforme (constante), no estacionario (es decir independiente del tiempo) y que el material de la barra es uniforme (en el sentido de que la densidad de la barra ρ , en cada subvolumen de control se mantiene constante).

Vamos a probar que la energía térmica que fluye por unidad de tiempo a través de los bordes de la barra, entre las posiciones x y $x + \Delta x$, es decir en el volumen de control, es

$$\phi(x, t)A - \phi(x + \Delta x, t)A,$$

donde el flujo de calor, en un mismo instante de tiempo t , debe ser multiplicado por el área superficial A de la barra.

En efecto, según la definición del flujo de calor

$$\phi = \frac{Q}{At},$$

donde Q es la energía térmica generada por unidad de volumen en la unidad de tiempo, cantidad que será definida formalmente más adelante. De esta ecuación se deduce que

$$\frac{Q}{t} = \phi A,$$

luego la variación de la energía térmica que fluye a través de las caras del i -ésimo subvolumen de control, en el instante de tiempo t , está dada por

$$\frac{\Delta Q_i}{t} = \frac{\phi(x_{i-1}, t)tA - \phi(x_i, t)tA}{t} = A[\phi(x_{i-1}, t) - \phi(x_i, t)].$$

Para encontrar la variación de la energía térmica que fluye por unidad de tiempo a través del volumen de control, es decir, entre las posiciones $a = x$ y $b = x + \Delta x$,

basta sumar las contribuciones de cada uno de los subvolúmenes de control dentro del volumen de control, es decir

$$\begin{aligned}\frac{\Delta Q}{t} &= \sum_{i=0}^n A[\phi(x_{i-1}, t) - \phi(x_i, t)] \\ &= A \sum_{i=0}^n [\phi(x_{i-1}, t) - \phi(x_i, t)] = A[\phi(x_0, t) - \phi(x_n, t)] \\ &= A[\phi(a, t) - \phi(b, t)] = \phi(x, t)A - \phi(x + \Delta x, t)A,\end{aligned}$$

porque $x_0 = a = x$ y $x_n = b = x + \Delta x$.

Si $\phi(x, t) > 0$ y $\phi(x + \Delta x, t) > 0$, como se ilustra en la Figura 2.2, entonces la energía térmica que fluye por unidad de tiempo en x contribuye a un aumento de la energía térmica en el corte, mientras que el flujo de calor en $x + \Delta x$ disminuye la energía térmica.

Fuentes de calor

Como se había dicho anteriormente, la energía térmica considerada en el volumen de control de la barra comprendida entre x y $x + \Delta x$ se debe a aquella que fluye por los bordes en la unidad de tiempo y aquella que se genera en el interior. Ahora bien, respecto a las fuentes internas de energía térmica, es claro que han de tener dependencia con la posición sobre la barra así como con el instante de tiempo en el que se han efectuado las mediciones; se define entonces

$$q(x, t) = \boxed{\begin{array}{l} \text{energía térmica generado} \\ \text{dentro del volumen de} \\ \text{control, por unidad de vol-} \\ \text{umen por unidad de tiem-} \\ \text{po} \end{array}} = \frac{\Delta Q}{A \Delta x t},$$

esta cantidad $q(x, t)$ es, como se ha dicho antes, aproximadamente constante en un volumen de control fino de la barra, por lo que la energía térmica total generada por unidad de tiempo en este volumen de control es de aproximadamente $q(x, t)A \Delta x$. Reiterando lo expuesto anteriormente puede decirse que, en un volumen de control considerablemente pequeño de la barra, la tasa de cambio con que varía la energía térmica es aproximadamente igual al cambio de la energía térmica que fluye por los bordes por unidad de tiempo, más la energía térmica del interior por unidad de tiempo.

La conservación de energía térmica (rebanada fina)

El enunciado: *la tasa de cambio de la energía térmica se debe a la energía térmica que fluye a través de las fronteras y a las fuentes internas*, se halla expresado

matemáticamente por la ecuación

$$\frac{\partial}{\partial t}[\varrho(x, t)A\Delta x] \approx \phi(x, t)A - \phi(x + \Delta x, t)A + q(x, t)A\Delta x. \quad (2.1)$$

La ecuación (2.1) no es del todo exacta, puesto que diversas cantidades se suponían aproximadamente constantes en el volumen de control, sin embargo (2.1) se vuelve cada vez más precisa a medida que $\Delta x \rightarrow 0$. Antes de dar un cuidadoso (y matemáticamente riguroso) estudio a la ecuación anterior, nos detendremos un poco en tratar de explicar las ideas básicas del proceso límite $\Delta x \rightarrow 0$. En el límite cuando $\Delta x \rightarrow 0$, la ecuación (2.1) no da información muy interesante, porque se obtiene simplemente la identidad $0 = 0$. Sin embargo, si en primer lugar se divide por Δx y luego se toma el límite cuando $\Delta x \rightarrow 0$, obtenemos

$$\frac{\partial \varrho}{\partial t} = \lim_{\Delta x \rightarrow 0} \frac{\phi(x, t) - \phi(x + \Delta x, t)}{\Delta x} + q(x, t), \quad (2.2)$$

donde el área de la sección transversal A ha sido cancelada. Este resultado es ahora exacto, y por lo tanto se reemplaza \approx en (2.1) por $=$ en (2.2). En este proceso de límite, $\Delta x \rightarrow 0$, t se mantiene fijo (o lo que es igual, en cada volumen de control la densidad de energía térmica es una cantidad estacionaria). En consecuencia, según la definición de derivada parcial,

$$\frac{\partial \varrho}{\partial t} = -\frac{\partial \phi}{\partial x} + q. \quad (2.3)$$

La conservación de energía térmica (exacta)

Una formulación alternativa del principio de conservación de la energía de calor tiene la ventaja de que no nos limita a pequeños trozos, para ello hacemos uso del *principio de superposición*. El cálculo aproximado que resulta del proceso de límite $\Delta x \rightarrow 0$ se evita. Consideramos cualquier volumen de control finito (desde $x = a$ hasta $x = b$) de la barra original de una dimensión (ver fig. 1.2.2). Analizaremos la conservación de la energía térmica en esta región. La energía térmica total está dada por $\int_a^b \varrho(x, t) dx$, y es la suma de las contribuciones de cortes infinitesimales (principio de superposición). Una vez más, los cambios en la energía térmica total sólo se deben a la energía térmica que fluye por los bordes laterales ($x = a$ hasta $x = b$) y a la energía térmica generada dentro de la región, y así (después de simplificar la constante A)

$$\frac{d}{dt} \int_a^b \varrho dx = \phi(a, t) - \phi(b, t) + \int_a^b q dx \quad (2.4)$$

La derivada ordinaria $\frac{d}{dt}$ aparece en (2.4) ya que $\int_a^b \varrho dx$ sólo depende de t , no de x . No obstante

$$\frac{d}{dt} \int_a^b \varrho dx = \int_a^b \frac{\partial \varrho}{\partial t} dx,$$

si a y b son constantes (y si e es continua). Cada término en (2.4) es ahora una integral ordinaria, al cumplirse la igualdad

$$\phi(a, t) - \phi(b, t) = - \int_a^b \frac{\partial \phi}{\partial x} dx,$$

que no es más que el enunciado de que el cambio en el flujo de calor en el corte comprendido entre a y b es la suma de las contribuciones de los flujos de calor en los cortes infinitesimales comprendidos en este mismo segmento, (lo que es válido si ϕ es continuamente diferenciable). Se sigue que,

$$\int_a^b \left(\frac{\partial \varrho}{\partial t} + \frac{\partial \phi}{\partial x} - q \right) dx = 0.$$

Esta integral debe ser cero para *valores arbitrarios* de a y b , lo cual es posible sólo si el propio integrando es idénticamente igual a cero.

Lema 2.1. *Sea $I = [a, b] \subseteq \mathbb{R}$ y $f : I \rightarrow \mathbb{R}$ una función continua en I , tal que $\int_a^b f(x) dx = 0$. Entonces $f(x) = 0$, para todo $x \in I$.*

Demostración. Supongamos que la conclusión no es cierta, esto es, existe por lo menos un punto $c \in I$ tal que $f(c) \neq 0$. Puesto que f es continua en I y $c \in I$, se tiene que

$$\lim_{x \rightarrow c} f(x) = f(c).$$

Esto es, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que si $x \in I$ y $|x - c| < \delta$ entonces $|f(x) - f(c)| < \varepsilon$.

Como $f(c) \neq 0$, entonces $f(c) > 0$ ó $f(c) < 0$. Si $f(c) > 0$, entonces para $\varepsilon = \frac{f(c)}{2}$, existe un número $\delta > 0$, tal que si $x \in I \cap (c - \delta, c + \delta)$, entonces

$$|f(x) - f(c)| < \frac{f(c)}{2}.$$

De esto se deduce que

$$\frac{f(c)}{2} < f(x) < \frac{3f(c)}{2}.$$

Esto muestra que si $x \in I \cap (c - \delta, c + \delta)$, entonces $f(x) > 0$.

Ahora bien, como f es continua en $[a, b]$, f es Riemann integrable en $[a, b]$, y

$$\int_a^b f(x) dx \equiv \sup \sum_{k=1}^n m_k(x_k - x_{k-1}),$$

donde $m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}$ y $P = \{a = x_0 < x_1 < x_2 < \dots < x_n = b\}$ es cualquier *partición* de $I = [a, b]$.

Sea \widehat{P} una partición de I tal que

$$\widehat{P} = \{y_0 = a \leq y_1 = c - \delta < y_2 < y_3 < \dots < y_{n-1} = c + \delta \leq y_n = b\},$$

es decir, \widehat{P} contiene puntos del intervalo $(c - \delta, c + \delta)$, donde f es positiva, entonces

$$m_k = \inf\{f(y) : y \in [y_{k-1}, y_k]\} > 0, \quad \forall k = 1, 2, \dots, n.$$

De aquí se sigue que $m_k(y_k - y_{k-1}) > 0$, para todo $k = 1, 2, \dots, n$; y por tanto

$$\sum_{k=1}^n m_k(y_k - y_{k-1}) > 0, \text{ así que}$$

$$\sup \left[\sum_{k=1}^n m_k(x_k - x_{k-1}) \right] > 0,$$

esto es

$$\int_a^b f(x) dx > 0,$$

lo cual es contradictorio con la hipótesis. De esta forma se prueba que no existe $c \in I$ tal que $f(c) > 0$. Análogamente, si $f(c) < 0$, basta elegir $\varepsilon = -f(c)$, para encontrar una partición R del intervalo $[a, b]$ tal que $\int_a^b f(x) dx < 0$, lo que contradice nuevamente la hipótesis. Por tanto, $f(x) = 0$ para todo $x \in I$, como se afirmó. \square

Así, obtenemos nuevamente la expresión

$$\frac{\partial e}{\partial t} = -\frac{\partial \phi}{\partial x} + q. \quad (2.5)$$

La ecuación (2.4), **ley de conservación integral**, es más relevante que la forma diferencial (2.5). La ecuación (2.5) es válida en el caso en que las variables físicas sean continuas.

Una explicación detallada del signo menos en la ecuación (2.5) se hace enseguida. Si $\frac{\partial \phi}{\partial x} > 0$, para $a \leq x \leq b$, entonces el flujo de calor ϕ es una función creciente de x . El calor está fluyendo más hacia la derecha en $x = b$ que en $x = a$ (asumiendo por supuesto que $b > a$). Así entonces (si se desprecian los efectos de las fuentes internas de calor, es decir, si q es despreciable), la energía térmica debe disminuir entre $x = a$ y $x = b$, resultando así el signo menos en (2.5).

Temperatura y calor específico

En la práctica los materiales se describen por su **temperatura**,

$$u(x, t) \equiv \text{temperatura},$$

y no por su densidad térmica. Distinguir entre los conceptos de temperatura y energía térmica no fue una tarea fácil. Solamente a mediados del siglo XVIII los físicos pudieron utilizar aparatos que les permitieron realizar experimentos precisos para determinar diferentes cantidades de energía térmica en dos materiales diferentes a distintas temperaturas. Esto obligó a la introducción del concepto de **calor específico** (o capacidad calorífica):

$$c = \text{calor específico},$$

y es la *energía térmica que requiere la una unidad de masa de una sustancia para elevar su temperatura una unidad*. En general, a partir de experimentos (y de la definición) el calor específico c de un material depende de la temperatura u . Por ejemplo, la energía térmica necesaria para elevar una unidad de masa de 0°C a 1°C podría ser diferente de la necesaria para elevar esta misma cantidad de materia de 85°C a 86°C para la misma sustancia. De otro lado, resolver problemas de flujo de calor donde se involucra el calor específico en función de la temperatura son matemáticamente complicados de resolver y no entraremos en detalles aquí. A menudo, para intervalos restringidos de temperatura, el calor específico es independiente de esta. Sin embargo, los experimentos sugieren que materiales diferentes requieren cantidades diferentes de energía térmica para calentarse. Puesto que sería satisfactorio formular la ecuación correcta que involucre las cantidades físicas expuestas anteriormente en situaciones en que el calor específico varía sobre la barra con la posición, este dependerá de x , esto es, $c = c(x)$. De otro lado, en muchos problemas se considera que la barra de un material es uniforme, en cuyo caso el calor específico c será una constante.

Nuevamente la energía térmica

La energía térmica en una rebanada fina está definida, como se dijo antes, por $q(x, t)A dx$. Sin embargo, también se define como la energía que se necesita para

elevar la temperatura desde una de referencia a 0° a su temperatura actual $u(x, t)$. Como el calor específico es independiente de la temperatura, para intervalos restringidos de esta, la energía térmica *por unidad de masa* es $c(x)u(x, t)$. Se define ahora la **densidad de masa** $\rho(x)$.

$\rho(x)$ = densidad de masa (masa por unidad de volumen),

lo que le permite variar esta cantidad con x , dejando libre así la posibilidad que la barra esté compuesta de un material no uniforme. Como la masa de cualquier rebanada fina de la barra está dada por $\rho A \Delta x$, entonces la energía térmica en cualquier rebanada fina es $c(x)u(x, t) \cdot \rho A \Delta x$, así que

$$\varrho(x, t)A \Delta x = c(x)u(x, t) \cdot \rho A \Delta x.$$

De esta manera se ha mostrado la relación básica entre la energía térmica y la temperatura:

$$\varrho(x, t) = c(x)\rho(x)u(x, t). \quad (2.6)$$

Cuando la densidad de energía térmica se simplifica utilizando (2.6), la conservación de la energía térmica, (2.3) o (2.5), se convierte en

$$c(x)\rho(x)\frac{\partial u}{\partial t} = -\frac{\partial \phi}{\partial x} + q. \quad (2.7)$$

La ley de Fourier

Por lo general, (2.7) es considerada como una ecuación con dos incógnitas, la temperatura $u(x, t)$ y el flujo de calor (flujo por unidad de área en la unidad de tiempo) $\phi(x, t)$. Vamos ahora a resumir algunas propiedades cualitativas del flujo de calor muy conocidas:

1. Si la temperatura es constante en una región, no hay flujo de energía térmica.
2. Si existen diferencias de temperatura, la energía térmica fluye de la región más caliente a la región más fría.
3. Cuanto mayor sea la diferencia de temperatura (para el mismo material), mayor será el flujo de energía calorífica.
4. El flujo de energía térmica variará para diferentes materiales, incluso con las mismas diferencias de temperatura.

Jean-Baptiste Joseph Fourier (1768-1830) reconoció estas cuatro propiedades, y tras numerosos experimentos las resumió en la fórmula

$$\phi = -K_0 \frac{\partial u}{\partial x}, \quad (2.8)$$

ecuación conocida como la **Ley de Fourier de conducción del calor**. Aquí $\frac{\partial u}{\partial x}$ es la derivada de la temperatura (en función de x para t fijo), y representa las diferencias de temperatura (por unidad de longitud). La ecuación (2.8) establece que el flujo de calor es proporcional a la diferencia de temperatura (por unidad de longitud). Si u aumenta al aumentar x (es decir, si la temperatura es más caliente a la derecha, $\frac{\partial u}{\partial x} > 0$), entonces, en virtud de la propiedad 2, el calor fluye hacia la izquierda. Esto explica el signo menos en (2.8).

Se ha designado por K_0 al coeficiente de proporcionalidad entre las cantidades ϕ y $\frac{\partial u}{\partial x}$, este mide la capacidad del material para conducir el calor y se llama la **conductividad térmica**. Los experimentos indican que los distintos materiales conducen el calor de manera diferente; por lo que K_0 depende del material en particular. Cuanto mayor sea K_0 , mayor será el flujo de energía térmica con las mismas diferencias de temperatura. Un material con un bajo valor de K_0 sería un mal conductor de la energía térmica (y por tanto un perfecto aislante térmico). Para una barra compuesta de diferentes materiales, K_0 será una función de x . Por otra parte, los experimentos muestran que la capacidad de conducir el calor para la mayoría de los materiales es diferente a diferentes temperaturas, esto es, K_0 es función de u y de x ; $K_0(x, u)$. Sin embargo, vamos a suponer que la conductividad térmica K_0 depende solamente de x , $K_0(x)$.

La ecuación de calor

Si la ley de Fourier, (2.8), se sustituye en la ecuación de conservación de la energía térmica, (2.7), resulta una ecuación en derivadas parciales con una sola incógnita:

$$c\rho \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(K_0 \frac{\partial u}{\partial x} \right) + q. \quad (2.9)$$

Ahora bien, en esta ecuación se considera que las fuentes de energía calórica Q son cantidades conocidas, siendo la única incógnita la temperatura $u(x, t)$. Los coeficientes térmicos (c, ρ, K_0), dependen todos del material y por lo tanto pueden ser funciones de x . En el caso especial de una barra uniforme, en la que (c, ρ, K_0) son constantes, la ecuación diferencial (2.9) se convierte en

$$c\rho \frac{\partial u}{\partial t} = K_0 \frac{\partial^2 u}{\partial x^2} + q$$

Si además no hay fuentes internas de calor, es decir, si $q = 0$; entonces, después de dividir por la constante $c\rho$, la ecuación diferencial anterior se convierte en

$$\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}, \quad (2.10)$$

donde la constante,

$$k = \frac{K_0}{c\rho},$$

se llama la **difusividad térmica**, y es la conductividad térmica del material dividida por el producto del calor específico y la densidad de masa del mismo. La ecuación (2.10) es llamada algunas veces la **ecuación de calor** cuando las fuentes internas de energía térmica son despreciables. Si la energía térmica se concentró inicialmente en un sólo lugar, la ecuación (2.10) describirá el proceso físico conocido como **difusión**. Por ejemplo, la concentración de $u(x, t)$ de productos químicos (tales como perfumes y contaminantes) satisface la ecuación de difusión (2.8) en determinadas situaciones en una dimensión.

Nótese que si $k = 1$, entonces (2.10) toma la forma

$$u_t(x, t) - u_{xx}(x, t) = 0,$$

que es la ecuación diferencial parcial planteada en nuestro problema.

2.1.3. Condiciones iniciales y de frontera

Condiciones iniciales

Las ecuaciones en derivadas parciales que describen el flujo de energía térmica, (2.9) o (2.10), tienen una derivada respecto al tiempo. En el caso de una ecuación diferencial ordinaria con una sola derivada, el problema de valor inicial consiste en resolver la ecuación diferencial considerando dicha condición inicial. Haciendo uso de la segunda ley de Newton de la dinámica se puede determinar la función de posición $x(t)$ de una partícula cuando se conoce la función aceleración de la misma, planteándose para ello una ecuación diferencial ordinaria de segundo orden, $m \frac{d^2 x}{dt^2} = F$; siendo m la masa inercial de la partícula y F la fuerza resultante de todas aquellas fuerzas que actúan *sobre* la misma. En este caso el problema de valores iniciales consiste en resolver la ecuación diferencial junto con dos condiciones conocidas, la posición inicial $x(0)$ y la velocidad inicial $\frac{dx}{dt}(0)$. A partir de esta información (incluyendo el conocimiento de las fuerzas actuantes),

y resolviendo la ecuación diferencial con las condiciones iniciales, es posible predecir la posición (y por tanto la trayectoria) de la partícula en la dirección del eje x . Aunque el propósito aquí no es el de predecir la temperatura en un instante de tiempo determinado, sí es importante considerar las condiciones iniciales que sobre la ecuación de calor se tienen, por cuanto estas son determinantes en el tipo de solución(es) obtenida(s), para lo cual se tendrá que estudiar si hay o no *blow up*.

Dado que la ecuación de calor tiene una derivada respecto al tiempo, hay que darse una **condición general inicial** (usualmente en $t = 0$), esto es, la temperatura inicial. Es posible que la temperatura inicial no sea constante, sino que dependa de x , por lo que debe tenerse la distribución de la temperatura inicial, $u(x, 0) = \phi(x)$. ¿Es esta información suficiente para predecir la temperatura en cualquier instante de tiempo? Se sabe que la distribución de la temperatura inicial y la temperatura cambian de acuerdo a la ecuación (2.9) ó (2.10). Sin embargo, debe conocerse lo que sucede en los dos extremos $x = 0$ y $x = 1$. Sin conocer esta información, no es posible predecir la temperatura, estas son precisamente condiciones de contorno o frontera y serán tratadas a continuación.

Al resolver la ecuación de calor, ya sea (2.9) ó (2.10), una **condición de contorno** o **frontera** es necesaria en cada extremo de la barra. La condición adecuada depende del mecanismo físico que actúa en cada extremo. A menudo la condición en la frontera depende tanto del material interior como exterior de la barra. Para evitar un problema matemático más complejo, se supondrá que el ambiente exterior es conocido, y que no es alterado significativamente por la barra.

Temperatura prescrita

En ciertas situaciones, la temperatura del extremo de la varilla, por ejemplo en $x = 0$, se puede aproximar por una **temperatura prescrita**,

$$u(0, t) = u_B(t) \quad (2.11)$$

donde $u_B(t)$ es la temperatura de algún líquido con el que la barra esté en contacto.

Frontera aislada

En otras situaciones es posible prescribir el flujo de calor en vez de la temperatura,

$$-K_0(0) \frac{\partial u}{\partial x}(0, t) = \phi(t), \quad (2.12)$$

donde $\phi(t)$ es conocido. Esto es equivalente a dar una condición para la primera derivada, $\frac{\partial u}{\partial x}$, en $x = 0$. La pendiente está dada en $x = 0$. La ecuación (2.12) no

se puede integrar con respecto a x , porque la derivada se conoce sólo para un sólo valor de x . El ejemplo más sencillo de la condición de contorno de flujo térmico establecido es cuando un extremo está *perfectamente aislado*. En este caso no hay flujo de calor en la frontera. Si en $x = 0$ hay aislamiento, entonces

$$\frac{\partial u}{\partial x}(0, t) = 0 \quad (2.13)$$

La ley de enfriamiento de Newton

Cuando una barra unidimensional está en contacto en la frontera con un fluido en movimiento (por ejemplo aire) entonces, ni la temperatura prescrita ni el flujo térmico establecido son cantidades apropiadas para describir la temperatura en esta posición. Por ejemplo, imaginemos una barra muy caliente en contacto con aire frío en movimiento. El calor va a salir de la barra, calentando el aire y este se llevará el calor. Este proceso de transferencia de calor se llama **convección**. Naturalmente, el aire estará más caliente cerca de la barra. Nuevamente este es un problema complicado, en realidad la temperatura del aire varía con la distancia de la barra. Los experimentos muestran que, con una buena aproximación, el flujo de calor que sale de la barra es proporcional a la diferencia de temperatura entre la barra y la temperatura exterior prescrita. Esta condición de frontera se conoce como **ley de enfriamiento de Newton**. En el punto $x = 0$ esta relación implica

$$-K_0 \frac{\partial u}{\partial x}(0, t) = -H[u(0, t) - u_B(t)], \quad (2.14)$$

donde la constante de proporcionalidad H se llama el **coeficiente de transferencia de calor** (o **coeficiente de convección**). Se observa que esta condición de frontera involucra una combinación lineal de u y $\frac{\partial u}{\partial x}$. Ahora bien, debe tenerse mucho cuidado con el signo de proporcionalidad de H ; así por ejemplo, si la barra está más caliente que el líquido con el que está en contacto, es decir, si $u(0, t) > u_B(t)$, entonces el calor fluye hacia afuera de la barra en $x = 0$, dirigido este hacia la izquierda, lo que implica que el flujo de calor sea negativo. Por esta razón se ha puesto el signo menos en (2.14) (si $H > 0$). Según esto la temperatura es mayor a la derecha de $x = 0$ y debería esperarse que esta siga en aumento a la derecha. Por lo tanto $\frac{\partial u}{\partial x}$ debe ser positiva en $x = 0$.

El coeficiente H en la ley de enfriamiento de Newton es determinado experimentalmente y depende de las propiedades de la barra, así como de las propiedades de los fluidos con los que ésta esté en contacto (incluyendo la velocidad del fluido). Si el coeficiente es muy pequeño, entonces fluye muy poca energía térmica a través de la frontera. En el límite, cuando $H \rightarrow 0$, la ley de enfriamiento de

Newton se aproxima a la condición de frontera con aislamiento. Puede pensarse en la ley de enfriamiento de Newton, para $H \neq 0$, como el caso en que la barra no está perfectamente aislada del medio que la rodea. Si $H \rightarrow \infty$, la condición de frontera $u(0, t)$ se acerca a la temperatura prescrita, esto es, $u(0, t) = u_B(t)$. Esto se observa fácilmente dividiendo (2.14), entre H

$$\frac{-K_0}{H} \frac{\partial u}{\partial x}(0, t) = -[u(0, t) - u_B(t)]$$

Por tanto, el caso en que $H \rightarrow \infty$ no corresponde a ninguna situación de aislamiento de la barra.

2.2. Fenómeno de b -explosión para un problema parabólico

2.2.1. Fenómeno de b -explosión

Consideremos el problema de valores en la frontera y condición inicial, PVFCI,

$$\begin{aligned} u_t - u_{xx} &= 0, & (x, t) &\in (0, 1) \times [0, T), \\ u_x(1, t) &= f(u(1, t)), & t &\in [0, T), \\ u_x(0, t) &= 0, & t &\in [0, T), \\ u(x, 0) &= \phi(x), & x &\in [0, 1]. \end{aligned} \tag{2.15}$$

Las ecuaciones $u_x(1, t) = f(u(1, t))$ y $u_x(0, t) = 0$ representan *condiciones de frontera* para la función temperatura $u(x, t)$; mientras que $u(x, 0) = \phi(x)$ es una *condición inicial* de dicha función. En este caso, $u_x(1, t) = f(u(1, t))$ indica que el *cambio de temperatura con respecto a la posición* sobre la barra en el extremo $x = 1$, para cualquier instante de tiempo t , viene dado por la función f ; la cual es positiva, suave y depende de la temperatura de la barra en $x = 1$ en un instante t .

De otro lado, la condición de frontera $u_x(0, t) = 0$ muestra que el *cambio de temperatura con respecto a la posición* sobre la barra en el extremo $x = 0$, para cualquier instante de tiempo t , es cero. Lo anterior implica que no hay flujo de calor en esta posición para cualquier instante de tiempo; porque anulándose la derivada parcial con respecto a x en el punto $(0, t)$ entonces la temperatura en esta posición permanece constante y, de acuerdo a la primera ley de Fourier, el flujo de calor en $(0, t)$ es cero.

Por último, la condición inicial $u(x, 0) = \phi(x)$ indica que la temperatura inicial de la barra, esto es, en el instante $t = 0$; en cualquier posición x de la misma, está dada por la función suave y positiva $\phi(x)$.

Problemas del tipo (2.15) se han utilizado para modelar, por ejemplo, reacciones químicas, transferencia de calor y dinámica de poblaciones, ver por ejemplo [8]. Para muchas ecuaciones o sistemas de diferenciales, sus soluciones pueden llegar a ser no acotadas, fenómeno que es conocido como de *b*-explosión.

Definición 2.1 (Fenómeno de *b*-explosión). *Un punto (\tilde{x}, t_b) es llamado un punto de *b*-explosión ('blow-up' en inglés) de $u(x, t)$ si existe una sucesión de puntos (x_k, t_k) en $(0, 1) \times (0, T]$, donde $T > 0$, tal que*

$$u(x_k, t_k) \rightarrow \infty \text{ cuando } (x_k, t_k) \rightarrow (\tilde{x}, t_b^-)$$

Si u y t son interpretados como temperatura y tiempo, entonces t_b es el tiempo de *b*-explosión y \tilde{x} es el punto de *b*-explosión. El tiempo t_b establece cuándo la temperatura *b*-explota y \tilde{x} en qué punto (dónde) de la barra la temperatura *b*-explota.

Definición 2.2. *El conjunto de puntos donde ocurre *b*-explosión puede tener más de un elemento. Si ocurre *b*-explosión en todo el intervalo $[0, 1]$, a esto se llama ***b*-explosión global**; o si ocurre un un subintervalo $[l, 1]$, con $0 < l < 1$, se llama ***b*-explosión regional**. Si ocurre *b*-explosión en un único punto a esto se llama **único punto de *b*-explosión**. Por otro lado, si el fenómeno de *b*-explosión ocurre entonces el tiempo *b*-explosión es único, y puede suceder en tiempo finito o infinito.*

Ejemplo típicos donde este fenómeno ocurre son problemas que involucran términos de reacción en la ecuación en derivadas parciales, ver por ejemplo [8]. De otro lado, cuando el término no lineal aparece en la frontera, como es el caso aquí, Walter demostró en [13] que si f es convexa y satisface la condición

$$\int^{+\infty} \frac{1}{f(s)f'(s)} ds < +\infty, \quad (2.16)$$

entonces una *b*-explosión ocurre. De otro lado, que si f es creciente y satisface

$$\int^{+\infty} \frac{1}{f(s)} ds < +\infty, \quad (2.17)$$

entonces u tiene un único punto de *b*-explosión en $x = 1$. Además, que si f cumple (2.16) pero no (2.17), puede presentarse *b*-explosión global o regional; por ejemplo, la función $f(s) = s \log^p(s)$, cuando $1/2 < p < 1$ ó $p = 1$, respectivamente.

Si la condición inicial $\phi(x)$ es regular (como se supondrá aquí), entonces existe una única solución para el problema y es regular. Sin embargo, no importa cuán

suaves sean ϕ en su dominio, existen casos en los cuales existe un tiempo finito t_b en el cual la solución b -explota.

El propósito es estudiar numéricamente, mediante aproximaciones semidiscretas (que involucran el método de elementos finitos), dónde ocurre b -explosión de la solución del problema (2.15), con fronteras no lineales, según lo establecido por los autores en [6]. Ellos establecen que ocurre b -explosión regional en cercanías de la frontera, ó bien, b -explosión global.

2.2.2. Formulación variacional

Sea

$$H_E^1(0, 1) = \{v \in H^1(0, 1) \mid v(0) = 0\},$$

el cual es también denotado como $H_0^1(0, 1)$. Sea $t \in [0, T)$. Entonces, multiplicando al sistema (2.15) por una función $v \in H_E^1(0, 1)$ e integrando con respecto a x , se obtiene:

$$\int_0^1 u_t(x, t)v(x) dx - \int_0^1 u_{xx}(x, t)v(x) dx = 0. \quad (2.18)$$

Mediante integración por partes, en la segunda integral del primer miembro de la ecuación anterior, y las condiciones de frontera se deduce que

$$\begin{aligned} \int_0^1 u_{xx}(x, t)v(x) dx &= u_x(x, t)v(x) \Big|_{x=0}^1 - \int_0^1 u_x(x, t)v'(x) dx \\ &= v(1)u_x(1, t) - v(0)u_x(0, t) - \int_0^1 u_x(x, t)v'(x) dx \\ &= v(1)f(u(1, t)) - \int_0^1 u_x(x, t)v'(x) dx. \end{aligned}$$

De (2.18) tenemos entonces que

$$\int_0^1 u_t(x, t)v(x) dx - v(1)f(u(1, t)) + \int_0^1 u_x(x, t)v'(x) dx = 0,$$

de lo cual se obtiene la **formulación variacional** del problema (2.15) como sigue: encontrar una función $u = u(\cdot, t) \in H_E^1(0, 1)$, tal que

$$\begin{aligned} (u_t, v) + a(u, v) &= v(1)f(u(1, t)), \quad (x, t) \in (0, 1) \times [0, T), \quad \forall v \in V, \\ u(x, 0) &= \phi(x), \quad x \in [0, 1] \end{aligned} \quad (2.19)$$

donde (\cdot, \cdot) denota el producto interno de $L^2(0, 1)$ y

$$a(u, v) = (u_x, v') = \int_0^1 u_x(x, t)v'(x) dx.$$

2.2.3. Formulación semidiscreta generalizada

La formulación semidiscreta generalizada del problema (2.15), es un método que se obtiene cuando el problema se resuelve en la variable espacial x por el método de elementos finitos. El procedimiento es el siguiente.

Sea $t > 0$, y $H_n \subset H_E^1(0, 1)$ un espacio de elementos finitos con respecto a una partición \mathcal{P}_h de $\Omega = [0, 1]$. De la formulación variacional (2.19) del problema (2.15), la formulación semidiscreta generalizada consiste en el siguiente problema: encontrar una función $u_n(t) = u_n(\cdot, t)$ en H_n para cada t (es decir, $u_n : [0, \infty) \rightarrow H_n$), tal que

$$\begin{aligned} (u_{n,t}, v) + a(u_n, v) &= v(1)f(u_n(1, t)), & (x, t) \in (0, 1) \times [0, T], \quad \forall v \in H_n, \\ u_n(x, 0) &= v_n(x), & x \in [0, 1] \end{aligned} \quad (2.20)$$

donde v_n es una aproximación de ϕ en H_n .

Sea $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$, donde $x_i = i/n$, una partición de $[0, 1]$ en n subintervalos $I_i = [x_{i-1}, x_i]$, de longitud $h = 1/n$, y sea H_n el espacio de funciones que tiene funciones base $\{\phi_i \mid i = 0, 1, 2, \dots, n\}$ de H_n como están definidas en (1.8)-(1.9), en la página 15.

Para un $t \in (0, T)$ y $0 < x < 1$, supongamos que $u_n(x, t)$ es una aproximación de $u(x, t)$ en H_n . Entonces podemos asumir que

$$u_n(x, t) = \sum_{i=1}^n u_i(t)\phi_i(x), \quad (2.21)$$

la cual pertenece a $H^1(0, 1)$ para un valor fijo de t . Al reemplazar este u_n y hacer $v = \phi_j$, $j = 1, \dots, n$, en (2.15) obtenemos el sistema de ecuaciones diferenciales ordinarias

$$\begin{cases} \sum_{i=1}^n u_i'(t)(\phi_i, \phi_j) + \sum_{i=1}^n u_i(t)a(\phi_i, \phi_j) = \delta_{jn}f(u_n(t)), & \text{para } j = 1, 2, \dots, n, \\ u_j(0) = \gamma_j, & \text{para } j = 1, 2, \dots, n, \end{cases}$$

dado que, $\phi_n(1) = \phi_j(x_n) = \delta_{jn}$,

$$\begin{aligned} u_n(1, t) &= \sum_{i=1}^n u_i(t)\phi_i(1) \\ &= \sum_{i=1}^n u_i(t)\delta_{in} \\ &= u_n(t), \end{aligned}$$

y además los γ_j son las componentes dados por la aproximación inicial v_n de $\phi(x)$. Este sistema de ecuaciones diferenciales se puede escribir en forma matricial para $U = (u_0(t), \dots, u_n(t))$ como

$$\begin{aligned} MU' + AU &= f(U), \\ U_0 &= \gamma. \end{aligned} \quad (2.22)$$

donde

$M = [m_{ij}]$ es la matriz de masa con elementos $m_{ij} = (\phi_i, \phi_j)$

$A = [a_{ij}]$ es la matriz de rigidez con elementos $a_{ij} = (\phi'_i, \phi'_j)$

$f(U)$ es el vector de componentes $f_j = \delta_{jn} f(u_n(t))$

Dado que el dominio de la función ϕ_i es el intervalo $[x_{i-1}, x_{i+1}]$ y el de la función ϕ_j es $[x_{j-1}, x_{j+1}]$ entonces se tiene que la matriz M es una matriz tridiagonal de la forma

$$M = \begin{bmatrix} m_{11} & m_{12} & 0 & \cdots & 0 & 0 & 0 \\ m_{21} & m_{22} & m_{23} & \cdots & 0 & 0 & 0 \\ 0 & m_{32} & m_{33} & \cdots & 0 & 0 & 0 \\ & & & \dots & & & \\ 0 & 0 & 0 & \cdots & m_{n-2,n-2} & m_{n-2,n-1} & 0 \\ 0 & 0 & 0 & \cdots & m_{n-1,n-2} & m_{n-1,n-1} & m_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & m_{n,n-1} & m_{n,n} \end{bmatrix}$$

dado que

$$m_{ij} = (\phi_i, \phi_j) = \begin{cases} \int_{[x_{i-1}, x_{i+1}] \cap [x_{j-1}, x_{j+1}]} \phi_i(x) \phi_j(x) dx & \text{si } |j - i| \geq 2 \\ 0 & \text{si } |j - i| < 2, \end{cases}$$

donde la integral es diferente de cero (ver [9]). Por las mismas razones, la matriz A es tridiagonal. Haciendo uso de este hecho, en [9] se muestra aplicando el algoritmo de Thomas a la matriz M se puede resolver el sistema de ecuaciones diferenciales ordinarias (2.22) para U' , obteniendo

$$\begin{aligned} u'_0(t) &= \frac{2}{h^2}(-u_0(t) + u_1(t)), \\ u'_i(t) &= \frac{1}{h^2}(u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)), \\ &\vdots \\ u'_n(t) &= \frac{2}{h^2}(-u_n(t) + u_{n-1}(t)) + \frac{2}{h}f(u_n(t)). \end{aligned} \quad (2.23)$$

Para este sistema se sabe que u_h converge uniformemente a u , dado que u es una función suave. Puede probarse que las soluciones explotan si y solo si f satisface (2.17), por lo que existen funciones no lineales f tales que la solución continua de (2.15) explota mientras que la solución numérica no (por ejemplo para $f(s) = s \ln(s)^p$, con $\frac{1}{2} < p \leq 1$).

En [6] los autores estudian una aproximación numérica de (2.15) en el caso especial en que $f(s) = s^p$ probando que la solución numérica aproxima al tiempo de b -explosión cuando el parámetro h de la malla es pequeño. La diferencia en el comportamiento asintótico entre las soluciones de la ecuación diferencial y su aproximación numérica ha sido estudiado por varios autores en los últimos años, por ejemplo Elliot y Stewart en [7] o Fernández y Rossi en [8], y es un fenómeno a considerar cuando se realizan simulaciones de las aproximaciones numéricas de dichas ecuaciones diferenciales.

2.2.4. Puntos de b -explosión

Para estudiar el comportamiento asintótico de las soluciones al problema parabólico (2.15) se requieren los siguientes resultados.

Lema 2.2 (Principio del máximo). *Sean $h > 0$ fijo y $U = (u_0, \dots, u_n)$ una solución de*

$$\begin{aligned} u'_0 &\leq \frac{2}{h^2}(-u_0 + u_1), \\ u'_i &\leq \frac{1}{h^2}(u_{i+1} - 2u_i + u_{i-1}), \\ u'_n &\leq \frac{2}{h^2}(-u_n + u_{n-1} + hf(U_n)). \end{aligned} \tag{2.24}$$

Entonces

$$\max_{k=0, \dots, n} u_k(t) \leq \max \left\{ \max_{k=0, \dots, n} u_k(0); \sup_{0 < \tau < t} u_n(\tau) \right\}.$$

Demostración. Supongamos primero que $U(u_0, \dots, u_n)$ verifica (2.24) con las desigualdades estrictas. Ahora, si el máximo se alcanza en un nodo interior, digamos $0 < j < n$, sea t_0 el primer tiempo donde esto sucede, entonces tenemos

$$u'_j(t_0) \geq 0 \text{ y } u_j(t_0) \geq u_k(t_0), \quad 0 \leq k \leq n.$$

De otro lado, por nuestra suposición de U tenemos $u'_j(t_0) < 0$, lo que es contradictorio; además puede verse que el máximo no puede ser alcanzado en u_0 . Para completar la demostración, se observa que si $z = h^2(0, \dots, k^2, \dots, n^2)$, entonces $U_\varepsilon \equiv U(t) + \varepsilon z$ verifica (2.24). Como $\varepsilon > 0$ es arbitrario, el lema queda probado. \square

Lema 2.3. *Sea U una solución de (2.23) que explota, entonces existe un tiempo $t_0 < T_h$ tal que*

$$\max_{0 \leq i \leq n} u_i(t) = u_n(t) \quad (2.25)$$

para $t_0 < t < T_h$.

Demostración. En primer lugar, por hipótesis, se tiene que

$$\lim_{s \rightarrow \infty} \frac{f(s)}{s} = +\infty. \quad (2.26)$$

Ahora, por el principio del máximo, se sigue que si $\mathbf{m\acute{a}x}\{u_i(t)\}$ es *grande* entonces coincide con $\sup_{0 < s < t} u_n(s)$. Y, por (2.23) se obtiene que

$$u'_n \geq -\frac{2}{h^2}u_n + \frac{2}{h}f(u_n) \geq \frac{2}{h}f(u_n) \left(1 - \frac{u_n}{hf(u_n)}\right). \quad (2.27)$$

Como $u_n(t)$ explota y $\lim_{s \rightarrow \infty} \frac{f(s)}{s} = +\infty$, $u_n(t)$ es una función creciente de t cerca del tiempo de explosión T_h , y por tanto se concluye (2.25) para valores de t próximos a T_h . \square

Aquí C representará una constante que puede depender de h , pero no de t , y que puede cambiar de una línea a otra. Para la demostración del primer teorema se requiere el siguiente,

Lema 2.4. *La tasa de b -explosión para $u_n(t)$ (y por tanto para $U(t)$) está dada por*

$$\mathbf{m\acute{a}x}_k u_k(t) = u_n(t) \sim F^{-1}(C(T_h - t)),$$

donde $F(z)$ está definida como

$$F(z) = \int_z^{+\infty} \frac{1}{f(y)} dy.$$

Demostración. Integrando (2.27), se obtiene

$$u_n(t) \leq c_1 F^{-1}(C(T_h - t)).$$

De esta desigualdad, se observa que para valores de t cercanos a T_h , por el Lema (2.3), se obtiene

$$u'_n(t) = \frac{2}{h^2}(u_{n-1} - u_n) + \frac{2}{h}f(u_n) \leq \frac{2}{h}f(u_n),$$

por tanto, integrando esta desigualdad resulta

$$u_n(t) \geq F^{-1}(C(T_h - t)),$$

y el lema queda probado. \square

Teorema 2.1. *Asuma que f satisface (2.17) y es creciente. Entonces, el número de nodos donde ocurre b -explosión en las soluciones de (2.23) depende sólo de f y es independiente de h . Además la explosión numérica es global o se localiza en una pequeña vecindad de $\{x = 1\}$ para h suficientemente pequeña. De otro lado, la propagación de b -explosión está dada de la siguiente forma: Sea T_h el tiempo de b -explosión de U y sea*

$$F(z) = \int_z^\infty \frac{1}{f(y)} dy, \quad G_0(t) = F^{-1}(T_h - t), \quad G_{k+1}(t) = \int_0^t G_k(s) ds.$$

Demostración. Ahora, en uso del Lema anterior se busca el comportamiento del nodo, u_{n-1}

$$u'_{n-1}(t) = \frac{u_n - 2u_{n-1}(t) + u_{n-2}(t)}{h^2} \geq \frac{1}{h^2}(u_n(t) - 2u_{n-1}(t)).$$

Así,

$$(e^{\frac{2t}{h^2}} u_{n-1}(t))' \geq \frac{e^{\frac{2t}{h^2}}}{h^2} u_n(t),$$

e integrando se obtiene,

$$u_{n-1}(t) \geq C \int_{t_0}^t u_n(s) ds + C.$$

De otro lado, por (2.25) se tiene

$$u'_{n-1}(t) \leq \frac{2}{h^2}(u_n(t) - u_{n-1}(t)),$$

por lo tanto

$$u_{n-1}(t) \leq C \int_{t_0}^t u_n(s) ds + C.$$

Usando el Lema (2.4) se obtiene

$$u_{n-1}(t) \sim \int_{t_0}^t u_n(s) ds \sim \int_{t_0}^t F^{-1}(C(T_h - s)) ds.$$

Recordando que $G_0(t) = F^{-1}(T_h - t)$, y

$$G_1(t) = \int_{t_0}^t F^{-1}(T_h - t) ds = \int_{T_h-t}^{T_h-t_0} F^{-1}(w) dw.$$

Dado que

$$\int_{t_0}^t F^{-1}(C(T_h - s)) ds = \int_{C(T_h-t)}^{C(T_h-t_0)} F^{-1}(w) dw$$

se tiene que $u_{n-1}(t)$ explota si y sólo si $G_1(t)$ lo hace. Repitiendo este procedimiento se puede comprobar que u_{n-k} explota si y sólo si $G_k(t)$ lo hace, por lo que el número de nodos que explotan depende solo de f y no de h . Esto implica el resultado del Teorema (2.1). \square

Entonces u_{n-k} explota si y solo si G_k lo hace, en este caso la tasa de explosión está dada por

$$u_{n-k}(t) \sim \int_0^t u_{n-k+1}(s) ds.$$

También se prueba la monotonía de explosión en términos de la no linealidad de f .

Teorema 2.2. *Bajo las hipótesis del Teorema (2.1), el número de nodos de explosión para las soluciones de (2.23) es monótono decreciente con el término no lineal f . Esto es, si $f_1(s) > f_2(s)$, entonces $B_h(U_1) \subset B_h(U_2)$.*

Demostración. Observemos que si se define la función

$$F_i(z) = \int_z^{+\infty} \frac{1}{f_i}(y) dy, \quad i = 1, 2,$$

entonces se tiene que

$$f_1(s) \geq f_2(s) \Rightarrow F_1(z) \leq F_2(z),$$

de lo cual se sigue el resultado. \square

Teorema 2.3. *La tasa de explosión para los nodos x_i donde hay b -explosión es estrictamente monótona en el sentido de que*

$$\lim_{t \rightarrow T_h} \frac{u_i(t)}{u_{i+1}(t)} = 0.$$

Demostración. Consideremos el cociente

$$\frac{u_n(t)}{u_{n-1}(t)}.$$

Del comportamiento asintótico de u encontrado en el Teorema (2.1), se tiene que

$$\frac{u_n(t)}{u_{n-1}(t)} \sim \frac{u_n(t)}{\int_{t_0}^t u_n(s) ds}.$$

Ahora, por la regla de L'Hôpital y (2.26), se obtiene

$$\lim_{t \rightarrow T_h} \frac{u_n(t)}{\int_{t_0}^t u_n(s) ds} = \lim_{t \rightarrow T_h} \frac{u_n'(t)}{u_n(t)} = \lim_{t \rightarrow T_h} \frac{Cf(u_n(t))}{u_n(t)} = +\infty$$

y el resultado se sigue por inducción. \square

Veamos ahora algunos ejemplos.

1. $f(s) = s^p$.

En este caso se tienen soluciones numéricas si y solo si $p > 1$. La tasa de b -explosión para los distintos nodos está dada por $u_{n-k}(t) \sim (T_h - t)^{-\frac{1}{(p-1)}+k}$ para $0 \leq k < \frac{1}{(p-1)}$, si $\frac{1}{(p-1)} = k$, es decir un entero, entonces $u_{n-k} \sim -\ln(T_h - t)$. Aquí el conjunto b -explosión está compuesto por K nodos exactamente, siendo $K = \left\lceil \frac{1}{(p-1)} \right\rceil$ ($[x]$ representa la parte entera de x).

En este caso, el b -explosión aproximado numéricamente con condición $p > 1$ es el mismo que en el caso continuo en (2.15), y el conjunto de puntos donde ocurre b -explosión verifica $B(u_h) = [1 - Kh, 1] = B(u) + [-Kh, 0]$. Este es el caso que se considera en [6].

2. $f(s) = s(\ln s)^p$.

Similarmente que en el ejemplo anterior se tienen soluciones numéricas si y solo si $p > 1$. La tasa de b -explosión está dada por

$$\max_i u_i(t) \sim \exp\left(\frac{1}{(T_h - t)^{1/(p-1)}}\right).$$

Como consecuencia del primer ejemplo y del Teorema (2.2), el conjunto donde ocurre b -explosión es todo el intervalo $[0, 1]$. En este caso la condición numérica de b -explosión, $p > 1$, es diferente a la condición en el caso continuo (2.15),

$p > \frac{1}{2}$. Sucede que el conjunto de b -explosión para el problema continuo contiene un solo punto, $B(u) = \{1\}$, si $p > 1$; y es todo el intervalo $[0, 1]$ si $\frac{1}{2} < p < 1$. Se observa que el comportamiento del problema continuo es radicalmente diferente al problema numérico, puesto que no sólo se tienen diferencias en la ocurrencia de b -explosión, sino que también cuando en ambas situaciones las soluciones explotan, el conjunto b -explosión en el caso numérico puede ser global mientras el problema continuo podría tener tan sólo un punto en el conjunto de explosión. Este es el primer ejemplo de estos fenómenos.

3. $f(s) = e^s$.

En este caso con la solución numérica el b -explosión ocurre en solo un punto $B(u_h) = \{1\}$. La tasa de de explosión está dada por $u_n(t) \sim \ln(T_h - t)$. Se prueba que debido a la no linealidad de f el conjunto b -explosión y la tasa de explosión en los dos problemas (el continuo y el discreto) coinciden.

Apéndice A

Un ejemplo de b -explosión para un problema no lineal

Un ejemplo de que en las ecuaciones no lineales puede aparecer este fenómeno debido simplemente a la estructura no lineal del problema la brinda el siguiente problema. Considérese la ecuación ordinaria.

$$\begin{cases} u' = f(u), \\ u(0) = u_0 > 0, \end{cases}$$

donde f es positiva, creciente y regular. En particular, cuando $f(u) = u^p$ con $p > 1$, esta ecuación tiene como única solución a la función

$$u(t) = C_p(T - t)^{-\frac{1}{(p-1)}}$$

donde,

$$T = u_0^{1-p}(p-1)^{-1} \text{ y } C_p = (p-1)^{-\frac{1}{(p-1)}}.$$

En efecto,

$$\begin{aligned}
u'(t) &= C_p \left[\frac{1}{(p-1)} (T-t)^{-\frac{1}{(p-1)}-1} \right] \\
&= C_p \left[\frac{1}{(p-1)} (T-t)^{-\frac{-1-(p-1)}{(p-1)}} \right] \\
&= C_p \left[\frac{1}{(p-1)} (T-t)^{-\frac{p}{(p-1)}} \right] \\
&= (p-1)^{-\frac{1}{(p-1)}} \left[\frac{1}{(p-1)} (T-t)^{-\frac{p}{(p-1)}} \right] \\
&= (p-1)^{-\frac{1}{(p-1)}-1} (T-t)^{-\frac{p}{(p-1)}} \\
&= (p-1)^{-\frac{-1-(p-1)}{(p-1)}} (T-t)^{-\frac{p}{(p-1)}} \\
&= (p-1)^{-\frac{p}{(p-1)}} (T-t)^{-\frac{p}{(p-1)}} \\
&= \left[(p-1)^{-\frac{1}{(p-1)}} (T-t)^{-\frac{1}{(p-1)}} \right]^p \\
&= \left[C_p (T-t)^{-\frac{1}{(p-1)}} \right]^p \\
&= u^p.
\end{aligned}$$

Veamos ahora que $u(t) \rightarrow +\infty$ cuando $t \rightarrow T^-$. Si $g(t) = T-t$, entonces $g(t) \neq 0$ si $t < T$.

Ahora bien, dado $\varepsilon > 0$; al considerar $\delta = \varepsilon$ tenemos que si $T - \delta < t < T$, entonces $|T - t| < \varepsilon$; es decir,

$$\lim_{t \rightarrow T^-} g(t) = 0.$$

Puesto que $p > 1$, también tenemos que

$$\lim_{t \rightarrow T^-} [g(t)]^{\frac{1}{(p-1)}} = 0.$$

En consecuencia podemos garantizar que dado $\varepsilon > 0$, por ejemplo $\varepsilon = \frac{1}{M}$, con $M > 0$, existe un número positivo $\delta' = \delta'(\varepsilon)$, tal que $T - \delta' < t < T$, entonces

$$\left| [g(t)]^{\frac{1}{(p-1)}} \right| < \varepsilon = \frac{1}{M};$$

luego

$$M < \frac{1}{[g(t)]^{\frac{1}{(p-1)}}} = \frac{1}{(T-t)^{\frac{1}{(p-1)}}} = (T-t)^{-\frac{1}{(p-1)}}.$$

De aquí se deduce que

$$(T - t)^{-\frac{1}{(p-1)}} \rightarrow +\infty, \text{ cuando } t \rightarrow T^-;$$

y puesto que $C_p = (p - 1)^{-\frac{1}{(p-1)}} > 0$, entonces $u(t) \rightarrow +\infty$ cuando $t \rightarrow T^-$. En este sencillo ejemplo puede verse que la solución $u(t)$ es regular para todo $t < T$ y que $u(t) \rightarrow +\infty$ cuando $t \rightarrow T^-$. En este caso $u(t)$ b -explota.

Apéndice B

Algunos teoremas básicos del Cálculo diferencial

Teorema B.1 (Teorema de Rolle). Sean $[a, b] \subset \mathbb{R}$ y $f : [a, b] \rightarrow \mathbb{R}$ una función continua en $[a, b]$. Si f es derivable en (a, b) y $f(a) = f(b)$, entonces existe $c \in (a, b)$ tales que $f'(c) = 0$.

Demostración. Como $f : [a, b] \rightarrow \mathbb{R}$ es continua, entonces tiene máximo y mínimo en $[a, b]$.

Si el máximo y mínimo ocurre en los extremos (a y b); como $f(a) = f(b)$, entonces f es constante y por tanto $f'(c) = 0$, para todo $c \in (a, b)$.

Si el máximo y el mínimo no ocurre en los extremos, entonces existe $c \in (a, b)$ tales que f tiene máximo o mínimo en $x = c$, y puesto que existe $f'(c)$, se tiene que $f'(c) = 0$. En cualquier caso la proposición es cierta. \square

Teorema B.2 (Teorema del valor medio). Sea $f : [a, b] \rightarrow \mathbb{R}$ una función continua en $[a, b]$, derivable en (a, b) . Entonces existe $c \in (a, b)$ tales que

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Demostración. Sea $g(x)$ la recta que pasa por los puntos $(a, f(a))$ y $(b, f(b))$ y sea $h(x) = f(x) - g(x)$.

Entonces,

$$h(x) = f(x) - \left(\frac{x(f(b) - f(a)) + bf(a) - af(b)}{b - a} \right),$$

y por tanto $h(a) = 0 = h(b)$. Como h es continua en $[a, b]$ y es derivable en (a, b) ; el teorema de Rolle implica que $h'(c) = 0$, para algún $c \in (a, b)$; es decir,

$$f'(c) = g'(c) = \frac{f(b) - f(a)}{b - a}. \quad \square$$

Teorema B.3 (Teorema de Taylor). *Sea $n \in \mathbb{N}$, sea $I = [a, b]$ y sea $f : I \rightarrow \mathbb{R}$ tal que f y sus derivadas $f', f'', \dots, f^{(n)}$ son continuas en I y que $f^{(n+1)}$ existe en (a, b) . Si $x_0 \in I$, entonces para cualquier $x \in I$ existe un punto c entre x y x_0 tal que*

$$f(x) = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(c) + \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

Demostración. Sean x_0 y x dadas y sea que J denote el intervalo cerrado con puntos terminales x_0 y x . Se define la función F en J por

$$F(t) = f(x) - \sum_{k=0}^n \frac{(x-t)^k}{k!} f^{(k)}(t)$$

para $t \in J$.

Entonces,

$$\begin{aligned} F'(t) &= -f'(t) - [-f'(t) + (x-t)f''(t)] \\ &- \left[-(x-t)f''(t) + \frac{(x-t)^2}{2!} f'''(t) \right] \\ &- \left[-\frac{(x-t)^2}{2!} f'''(t) + \frac{(x-t)^3}{3!} f^{(4)}(t) \right] \\ &- \dots - \left[-\frac{(x-t)^{n-2}}{(n-2)!} f^{(n-1)}(t) + \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) \right] \\ &- \left[-\frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) + \frac{(x-t)^n}{n!} f^{(n+1)}(t) \right] \\ &= -\frac{(x-t)^n}{n!} f^{(n+1)}(t). \end{aligned}$$

Si se define ahora G en J por

$$G(t) = F(t) - \left(\frac{x-t}{x-x_0} \right)^{n+1} F(x_0)$$

para $t \in J$, entonces $G(x_0) = F(x_0) - F(x_0) = 0$ y $G(x) = F(x) - F(x) = 0$. En virtud del teorema de Rolle, existe un punto $c \in (x_0, x)$ tal que

$$G'(c) = 0 = F'(c) + (n+1) \frac{(x-c)^n}{(x-x_0)^{n+1}} F(x_0).$$

Se deduce por tanto que

$$\begin{aligned} F(x_0) &= -\frac{1}{n+1} \frac{(x-x_0)^{n+1}}{(x-c)^n} F'(c) \\ &= \frac{1}{n+1} \frac{(x-x_0)^{n+1}}{(x-c)^n} \frac{(x-c)^n}{n!} f^{(n+1)}(c) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-x_0)^{n+1}, \end{aligned}$$

lo cual implica que

$$f(x) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(c) + \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x-x_0)^k,$$

como se afirmó. □

Definición B.1. Sea $A \subseteq \mathbb{R}$ y sea $f : A \rightarrow \mathbb{R}$. Si existe una constante $K > 0$ tal que

$$|f(x) - f(u)| \leq K|x - u|$$

para toda $x, u \in A$, entonces se dice que f es una **función de Lipschitz** (o que satisface una condición de **Lipschitz**) en A .

La condición de que la función $f : A \rightarrow \mathbb{R}$ en un intervalo I es una función de Lipschitz se puede interpretar geoméricamente de la siguiente manera. Si se escribe la condición como

$$\left| \frac{f(x) - f(u)}{x - u} \right| \leq K,$$

con $x, u \in I; x \neq u$, entonces la cantidad que está entre los signos de valor absoluto es la pendiente del segmento de recta que une los puntos $(x, f(x))$ y $(u, f(u))$. Por tanto, una función f satisface una condición de Lipschitz si y sólo si las pendientes de todos los segmentos de recta que unen los puntos de la gráfica de $y = f(x)$ en I están acotadas por algún número K .

Definición B.2. Sea $A \subseteq \mathbb{R}$ y $x_0 \in \mathbb{R}$. Se dice que x_0 es un **punto de acumulación** del conjunto A , si para cualquier número $r > 0$ se tiene que $B(x_0; r) - \{x_0\} \cap A \neq \emptyset$, es decir, si el conjunto $(x_0 - r, x_0 + r)$ contiene al menos un punto de A distinto de x_0 .

Ejemplo B.1. Sea $A = \{\frac{1}{n} : n \in \mathbb{N}\}$. Entonces 0 es un punto de acumulación de A .

En efecto, por la propiedad arquimediana de los números reales se tiene que para cualquier número $\varepsilon > 0$; existe un número natural N tal que

$$1 < N\varepsilon, \text{ es decir, } \frac{1}{N} < \varepsilon.$$

Ahora, si $n \in \mathbb{N}$ y $N \leq n$; entonces

$$\frac{1}{n} \leq \frac{1}{N} < \varepsilon,$$

y por tanto, si $N \leq n$, se tiene que

$$\frac{1}{n} \in (-\varepsilon, \varepsilon) - \{0\}$$

lo que prueba la afirmación.

Definición B.3. Sean $A \subseteq \mathbb{R}$, $f : A \rightarrow \mathbb{R}$ y c un punto de acumulación de A . Se dice que el número real L es un **límite de f en c** , y se escribe $\lim_{x \rightarrow c} f(x) = L$, si: para cualquier número $\varepsilon > 0$; existe un número $\delta > 0$ tales que si

$$x \in A \text{ y } 0 < |x - c| < \delta, \text{ entonces } |f(x) - L| < \varepsilon$$

Definición B.4. Sean $A \subseteq \mathbb{R}$, $f : A \rightarrow \mathbb{R}$ y $c \in A$. Se dice que f es **continua en c** , y se denota $\lim_{x \rightarrow c} f(x) = f(c)$, si para cualquier número $\varepsilon > 0$, existe un número $\delta > 0$ tales que si

$$x \in A \text{ y } |x - c| < \delta, \text{ entonces } |f(x) - f(c)| < \varepsilon$$

Ejemplo B.2. Sea

$$\begin{aligned} f : \mathbb{R}^+ &\longrightarrow \mathbb{R}^+ \\ x &\longmapsto f(x) = \frac{1}{x}. \end{aligned}$$

Pruebe que f es continua en \mathbb{R}^+ , es decir, para cualquier $c \in \mathbb{R}^+$ demuestre que f es continua en c .

Demostración. Dado $c \in \mathbb{R}^+$ y $\varepsilon > 0$, sea $\delta = \frac{\varepsilon c^2}{1 + \varepsilon c} > 0$, tales que si

$$|x - c| < \frac{\varepsilon c^2}{1 + \varepsilon c},$$

entonces

$$c - \frac{\varepsilon c^2}{1 + \varepsilon c} < x < \frac{\varepsilon c^2}{1 + \varepsilon c} + c,$$

luego,

$$0 < \frac{c}{1 + \varepsilon c} < x < \frac{2\varepsilon c^2 + c}{1 + \varepsilon c},$$

y por tanto

$$\frac{1 + \varepsilon c}{c(2\varepsilon c + 1)} < \frac{1}{x} < \frac{1}{c} + \varepsilon.$$

Restando $\frac{1}{c}$ a la anterior desigualdad se deduce que

$$-\frac{\varepsilon}{2\varepsilon c + 1} < \frac{1}{x} - \frac{1}{c} < \varepsilon.$$

Ahora bien,

$$-\varepsilon < -\frac{\varepsilon}{2\varepsilon c + 1},$$

porque

$$1 + 2\varepsilon c > 1,$$

puesto que $\varepsilon c > 0$, y por tanto

$$-\varepsilon < -\frac{\varepsilon}{2\varepsilon c + 1} < \frac{1}{x} - \frac{1}{c} < \varepsilon,$$

es decir,

$$\left| \frac{1}{x} - \frac{1}{c} \right| < \varepsilon.$$

□

Bibliografía

- [1] Tom M. Apostol, *Mathematical Analysis*. Second Edition. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1974.
- [2] Richard Haberman, *Elementary Applied Partial Differential Equations with Fourier Series and Boundary Values Problems*. Prentice Hall, Englewood Cliffs, New Jersey, Second Edition, 1987.
- [3] Robert G. Bartle, Donald R. Sherbert, *Introducción al análisis matemático de una variable*. Editorial Limusa Wiley, México, Segunda edición, 2002.
- [4] Robert G. Bartle, *The Elements of Real Analysis*. Second Edition. John Wiley & Sons, New York-London-Sydney, 1976.
- [5] Häim Brézis, *Análisis funcional*, Alianza editorial, S.A., Madrid, 1984.
- [6] Gabriel Acosta, Julian Fernández Bonder, Pablo Groisman, *Numerical Approximation of the Heat Equation with Nonlinear Boundary Conditions in Several Space Dimensions*. Estados Unidos, 2002.
- [7] C.M. Elliott, A. Stuart, *Global Dynamics of Discrete Semilinear Parabolic Equations*. IMA, preprint series No. 973, 1992.
- [8] Julian Fernández Bonder, Julio D. Rossi, *Blow-up vs. Spurious Steady Solutions*. American Mathematical Society. Artículo electrónico, 2000.
- [9] Julio C. Carrillo E., *The Finite Element Method for Parabolic Problems*. Technical Report No. SP05-1, University of Louisiana at Lafayette, Lafayette, LA, USA, 2005.
- [10] Watson Fulks, *Advance Calculus: An Introduction to Analysis*. Wiley & Sons, Third edition, New York, 1978.
- [11] Richard L. Burden, J. Douglas Faires, *Análisis numérico*. Grupo Editorial Iberoamericana, México D.F. 1991.

- [12] Pablo Groisman, Julio D. Rossi, *Aproximando soluciones que explotan*. Boletín de la Sociedad Española de Matemática Aplicada, 2003.
- [13] Walter W. *On Existence and Nonexistence in the Large of Solutions Parabolic Differential Equations with a Nonlinear Boundary Conditions*, SIAM J. Math. Anal 6(1), 85-90, (1975).