

Introducción al Análisis Numérico de Ecuaciones en Derivadas Parciales de Evolución

January 12, 2003

Enrique Zuazua
enrique.zuazua@uam.es

1 Introducción y motivación

Estas notas constituyen una breve guía de lo que consideramos puede y debe ser un último capítulo de un curso introductorio al Cálculo Numérico de Ecuaciones Diferenciales. En efecto, tras haber estudiado los elementos básicos del Cálculo Numérico para Ecuaciones Diferenciales Ordinarias (EDO) y los aspectos fundamentales de la aproximación numérica de la ecuación de Laplace es coherente y natural combinar y sintetizar estos conocimientos para introducirse en el mundo de las Ecuaciones en Derivadas Parciales (EDP) de evolución.

La forma en la que las EDP se presentan habitualmente en la modelización de fenómenos de la Ciencia y Tecnología es precisamente la de modelos de evolución en los que se describe la dinámica a lo largo del tiempo de determinada cantidad o variable (también a veces denominada *estado*) que puede representar objetos de lo más diversos que van desde la posición de un satélite en el espacio hasta la dinámica de un átomo, pasando por los índices bursátiles o el grado en que una enfermedad afecta a la población. En otras palabras, los modelos dinámicos o de evolución son los más naturales en la medida que reproducen nuestra propia concepción del mundo: un espacio tri-dimensional que evoluciona y cambia en el tiempo¹.

Cuando el estado o variable de un modelo o sistema de evolución es finito-dimensional, el modelo más natural es un sistema de EDO, cuya dimensión coincide precisamente con el del número de parámetros necesarios para describir dicho estado. Así, por ejemplo, para posicionar una partícula en el espacio necesitamos de tres variables dependientes del tiempo y para describir su dinámica un sistema de tres ecuaciones diferenciales. Pero en muchas ocasiones, como es el caso sistemáticamente en el contexto de la Mecánica de Medios Continuos, la variable de estado es infinito-dimensional. Esto ocurre por ejemplo cuando se pretende describir la deformación de cuerpos elásticos o la temperatura de un cuerpo sólido en los que la deformación o temperatura

¹Si bien la Teoría de la Relatividad establece que es mejor considerar a las cuatro del mismo modo, nuestra percepción $3 + 1$ está condicionada por razones puramente fisiológicas y culturales.

de cada uno de los puntos de ese medio continuo constituye una variable o incógnita del sistema. Los modelos matemáticos naturales en este caso son las EDP.

En la teoría clásica de EDP éstas se clasifican en tres grandes grupos: *elípticas*, *parabólicas* e *hiperbólicas*.

El modelo elíptico por excelencia involucra el *operador de Laplace*

$$\Delta = \sum_{i=1}^N \partial^2 / \partial x_i^2 \quad (1.1)$$

y ha sido objeto de estudio en el capítulo anterior. La variable tiempo está ausente en este modelo. Es por eso que sólo permite describir estados estacionarios o de equilibrio.

Las ecuaciones parabólicas y las hiperbólicas, representadas respectivamente por la *ecuación del calor* y la *de ondas*, son los modelos clásicos en el contexto de las EDP de evolución. Sus características matemáticas son bien distintas. Mientras que la ecuación del calor permite describir fenómenos altamente irreversibles en tiempo en los que la información se propaga a velocidad infinita, la ecuación de ondas es el prototipo de modelo de propagación a velocidad finita y completamente reversible en tiempo.

El operador del calor es

$$\partial_t - \Delta, \quad (1.2)$$

de modo que al actuar sobre una función $u = u(x, t)$ que depende de la variable espacio-tiempo $(x, t) \in \mathbb{R}^N \times (0, \infty)$ tiene como resultado

$$[\partial_t - \Delta] u = \frac{\partial u}{\partial t} - \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}. \quad (1.3)$$

Sin embargo, el operador de ondas o de D'Alembert es de la forma

$$\square = \partial_t^2 - \Delta \quad (1.4)$$

y da lugar a

$$\square u = [\partial_t^2 - \Delta] u = \frac{\partial^2 u}{\partial t^2} - \Delta u. \quad (1.5)$$

La irreversibilidad temporal de (1.3) es evidente. Si hacemos el cambio de variable $t \rightarrow \tilde{t} = -t$, el operador (1.3) cambia y da lugar al operador del calor retrógrado $\partial_{\tilde{t}} + \Delta$ mientras que el operador de ondas permanece invariante.

El operador del calor y de ondas se distinguen también por sus ámbitos de aplicación. Mientras que el primero es habitual en la dinámica de fluidos (a través de una versión más sofisticada, el operador de Stokes) o en fenómenos de difusión (del calor, de contaminantes, . . .), el operador de ondas y sus variantes intervienen de forma sistemática en elasticidad (frecuentemente a través de sistemas más sofisticados, como el de Lamé, por ejemplo) o en la propagación de ondas acústicas o electromagnéticas (ecuaciones de Maxwell).

La Mecánica de Medios Continuos está repleta también de otras ecuaciones, operadores y modelos, pero en todos ellos, de una u otra manera, encontraremos siempre el operador del calor, de ondas o una variante muy próxima de los mismos.

Frecuentemente los modelos son más sofisticados que una “simple” ecuación aislada. Se trata a menudo de sistemas acoplados de EDP en los que es habitual encontrar tanto componentes parabólicos como hiperbólicos. Es el caso por ejemplo de las ecuaciones de la *termoelasticidad*. En estos casos, si bien un buen conocimiento de los aspectos más relevantes de la ecuación del calor y de ondas aisladamente puede no ser suficiente a causa de las interacciones de los diferentes componentes, sí que resulta indispensable.

Por todo ello es natural e importante entender todos los aspectos matemáticos fundamentales de estas dos piezas clave: la ecuación del calor y la de ondas. Evidentemente esto es también cierto desde el punto de vista del Análisis y del Cálculo Numérico.

Hasta ahora nos hemos referido sólo a las ecuaciones del calor y de ondas en su expresión más sencilla: con coeficientes constantes. Estas ecuaciones, cuando modelizan fenómenos en medios heterogéneos (compuestos por materiales de diversa naturaleza) adoptan formas más complejas y se presentan con coeficientes variables, dependientes de la variable espacial x , de la variable temporal t o de ambas.

Por limitaciones de tiempo nos centraremos esencialmente en el estudio de estas ecuaciones en el caso más sencillo de los coeficientes constantes y lo haremos, sobre todo, en una variable espacial. A pesar de ello, creemos que quien asimile bien los conceptos que aquí expondremos y entienda las técnicas y los resultados principales que presentaremos estará en condiciones de abordar con éxito situaciones más complejas, incluyendo EDP con coeficientes variables y en varias dimensiones espaciales.

En esta introducción no hemos mencionado para nada otras palabras clave en la modelización de fenómenos complejos como son los términos “no-lineal” y “no-determinista”. La aproximación numérica de modelos de EDP que involucran estos fenómenos queda fuera de los objetivos de este curso pero, nuevamente, se puede asegurar que los elementos que aquí expondremos serán sin duda de gran utilidad, si no indispensables, a la hora de adentrarse en otros modelos más complejos que involucren términos no-lineales y estocásticos.

Habiendo ya motivado la necesidad de proceder al desarrollo de métodos numéricos para la resolución de la ecuación del calor y de la ecuación de ondas, veamos cual es la forma o, más bien, cuáles son las formas más naturales de proceder. Hay al menos tres

- a) Discretizamos simultáneamente las variable de espacio y de tiempo. De este modo pasamos directamente de la EDP de evolución a un sistema puramente discreto. Es lo que se denomina una *discretización completa*.
- b) Mantenemos la variable temporal continua y discretizamos la variable espacial. En este caso se trata de una *semi-discretización espacial* y el problema se reduce a un sistema de ecuaciones diferenciales de dimensión igual al de nodos espaciales que tenga el mallado utilizado en la discretización espacial. Estos métodos se conocen también como *métodos de líneas*.

- c) Mantenemos la variable espacial continua y discretizamos el tiempo. Se trata en este caso de una *semi-discretización temporal*. El sistema se reduce a la resolución iterada, discretamente en tiempo, de ecuaciones de Laplace.

De entre todas estas vías la última es la menos habitual (si bien se trata de un método frecuente a la hora de probar resultados analíticos de existencia de soluciones, idea que inspira, por ejemplo, la teoría de semigrupos no-lineales) y actualmente es objeto de estudio intensivo de cara, en particular, a desarrollar algoritmos paralelizables.

Desde un punto de vista estrictamente computacional sólo la primera es válida y realmente programable en el ordenador. Pero hay varias razones para no descartar la segunda. En primer lugar, cuando realizamos la discretización espacial estamos sustituyendo la dinámica infinito-dimensional de la EDP por una dinámica en dimensión finita. El estudio de la legitimidad de esta sustitución, en sí, es ya un objetivo interesante no sólo desde un punto de vista práctico sino también en el plano conceptual. Por supuesto, en este empeño lo estudiado sobre la ecuación de Laplace nos resultará de suma utilidad pues la semi-discretización consiste precisamente en discretizar el laplaciano en la variable espacial dejando intacta la variable temporal. Una vez justificada la idoneidad de esta primera discretización espacial, lo cual pasa obviamente por un análisis de la convergencia a medida que el paso del mallado espacial tiende a cero, nos encontramos pues frente a un sistema de EDO. Algunos de los programas comerciales (Matlab, por ejemplo) están equipados de rutinas de resolución de EDO. Esto hace que se puedan obtener con facilidad aproximaciones numéricas y visualizaciones gráficas de las soluciones de dicho sistema de EDO y, por consiguiente de la EDP, lo cual supone sin duda una razón importante para proceder de este modo. Pero no debemos olvidar que la teoría desarrollada en la primera parte de este curso está precisamente orientada a la discretización temporal de sistemas de EDO, con su consiguiente análisis de convergencia. Al final de este doble proceso de aproximación nos encontraremos por tanto con un sistema completamente discreto, igual que si hubiésemos procedido directamente por la primera vía, pero esta vez lo haremos habiendo utilizado las dos teorías de convergencia previamente desarrolladas. Ni que decir tiene que, si realizamos los dos procesos de aproximación (el del laplaciano y el de la EDO) con cuidado, obtendremos no sólo los resultados de convergencia de las soluciones del problema discreto al continuo sino estimaciones del error. Las estimaciones de error junto con el coste computacional del método numérico es lo que al final establece su bondad.

Es por esto que, en cada uno de los ejemplos (ecuación del calor y de ondas) analizaremos las dos primeras vías: discretización completa y semi-discretización temporal. Por supuesto lo que aquí presentaremos no serán más que algunos aspectos, conceptos y resultados básicos y fundamentales. El lector interesado en un análisis más detallado encontrará en los textos de la Bibliografía que incluimos al final de estas notas un excelente material para profundizar en el estudio de este campo, además de diversas y útiles referencias complementarias.

Tanto en la sección dedicada a la ecuación del calor como a la de ondas comenzaremos recordando algunas de sus propiedades analíticas más importantes, para después abordar los aspectos numéricos. No conviene olvidar que la eficacia de un método numérico depende en gran medida de la fidelidad con que consigue reproducir a nivel discreto las propiedades analíticas

del modelo continuo.

Antes de concluir esta introducción conviene hacer una observación sobre la notación que usaremos a lo largo de las notas:

1. • El índice j será utilizado para denotar la componente j -ésima de la solución numérica, que será de hecho una aproximación de la solución de la EDP en el punto nodal $x_j = jh$, siendo $h = \Delta x$ el paso del mallado espacial.
2. • El exponente k se utiliza para denotar la solución numérica en el paso temporal k -ésimo, aproximación de la solución de la EDP en el instante de tiempo $t = k\Delta t$.
3. • El superíndice $\hat{\cdot}$ se utiliza para denotar las componentes de Fourier de las soluciones tanto en el marco continuo como en el discreto.

2 La ecuación del calor

Como hemos mencionado anteriormente, la ecuación del calor es el prototipo de ecuación de evolución de tipo parabólico cuyas variantes están presentes de manera sistemática en todos los modelos matemáticos de la difusión y de la Mecánica de Fluidos.

Como hemos dicho antes, la ecuación del calor es un modelo fuertemente irreversible en tiempo en el que la información se propaga a velocidad infinita. Estas propiedades quedarán claramente de manifiesto en la siguiente sección en la que recordamos sus principales propiedades analíticas.

2.1 Propiedades básicas de la ecuación del calor

Consideremos en primer lugar el problema de Cauchy

$$\begin{cases} u_t - \Delta u = 0 & \text{en } \mathbb{R}^N \times (0, \infty) \\ u(x, 0) = \varphi(x) & \text{en } \mathbb{R}^N. \end{cases} \quad (2.1)$$

Se trata de un problema característico en el sentido de Cauchy-Kowaleski (ver F. John [J]). Precisamente por serlo cabe esperar que (2.1) esté bien planteado a pesar de que no damos dos datos de Cauchy como es habitual en una ecuación de orden dos, sino sólo una.

La solución fundamental de (2.1) se puede calcular explícitamente. Obtenemos así el núcleo de Gauss:

$$G(x, t) = (4\pi t)^{-N/2} \exp(-|x|^2 / 4t). \quad (2.2)$$

No es difícil comprobar que G es efectivamente la solución de (2.1) con $\varphi = \delta_0$, la delta de Dirac en $x = 0^2$.

Por consiguiente, para “cualquier” φ ,

$$u = G * \varphi \quad (2.3)$$

²Recordemos que δ_0 es la medida tal que $\langle \delta_0, \phi \rangle = \phi(0)$ para toda función continua ϕ .

representa la única solución de (2.1). (Hemos entrecomillado el cuantificador “cualquier” puesto que se requieren algunas condiciones mínimas sobre φ y la propia solución para que ésta pueda escribirse de manera única como en (2.3). Basta por ejemplo con tomar $\varphi \in L^2(\mathbb{R}^N)$ o $\varphi \in L^\infty(\mathbb{R}^N)$ y buscar soluciones u tales que, para cualquier $t > 0$, sean funciones acotadas (véase F. John [J])).

En (2.3) $*$ representa la convolución espacial de modo que

$$u(x, t) = (4\pi t)^{-N/2} \int_{\mathbb{R}^N} \exp(-|x - y|^2 / 4t) \varphi(y) dy. \quad (2.4)$$

En esta expresión se observa inmediatamente la velocidad infinita de propagación. En efecto, todos los valores de φ , en cualquier punto y de \mathbb{R}^n , intervienen a la hora de calcular u en cualquier punto espacio-temporal (x, t) .

En (2.4) es también fácil comprobar el enorme efecto regularizante de la ecuación del calor. En efecto, basta que $\varphi \in L^1(\mathbb{R}^N)$ o que $\varphi \in L^\infty(\mathbb{R}^N)$ para que la solución $u(\cdot, t)$ ³ en cada instante $t > 0$ sea una función de $C^\infty(\mathbb{R}^N)$. Este efecto regularizante implica también la irreversibilidad temporal.⁴

De la fórmula (2.4) se deducen otras propiedades de la solución de la ecuación del calor:

- *Principio del máximo:* Si $\varphi \geq 0$ entonces $u \geq 0$ y en realidad $u > 0$ en $\mathbb{R}^N \times (0, \infty)$ salvo que $\varphi \equiv 0$.
- *Conservación de la masa:*

$$\int_{\mathbb{R}^N} u(x, t) dx = \int_{\mathbb{R}^N} \varphi(x) dx, \quad \forall t > 0. \quad (2.5)$$

- *Decaimiento:*

$$\|u(t)\|_{L^\infty(\mathbb{R}^N)} \leq C t^{-N/2} \|\varphi\|_{L^1(\mathbb{R}^N)}, \quad \forall t > 0. \quad (2.6)$$

Todas ellas admiten claras interpretaciones físicas y obedecen, efectivamente, al comportamiento habitual en un proceso de difusión.

Consideramos ahora el problema de la difusión del calor en un dominio acotado Ω de \mathbb{R}^N . En esta ocasión, con el objeto de que el sistema de ecuaciones sea completo tenemos también que imponer condiciones de contorno que determinen la interacción del medio Ω con el medio circundante. Desde un punto de vista matemático las condiciones más simples son las de Dirichlet. Obtenemos así el sistema

$$\begin{cases} u_t - \Delta u = 0 & \text{en } \Omega \times (0, \infty) \\ u = 0 & \text{en } \partial\Omega \times (0, \infty) \\ u(x, 0) = \varphi(x) & \text{en } \Omega. \end{cases} \quad (2.7)$$

³Interpretamos la función $u = u(x, t)$ como una función del tiempo t que, a cada instante t , tiene como imagen una función de x que varía en el tiempo.

⁴En efecto, si la ecuación del calor estuviese bien puesta en el sentido retrógrado del tiempo, como la solución es regular para $t > 0$, volviendo hacia atrás en el tiempo, obtendríamos en el instante inicial $t = 0$ una función $C^\infty(\mathbb{R}^N)$. De este modo acabaríamos probando que toda función de $L^1(\mathbb{R}^N)$ o $L^\infty(\mathbb{R}^N)$ está en $C^\infty(\mathbb{R}^N)$, cosa falsa evidentemente.

Las condiciones de contorno $u = 0$ en $\partial\Omega$ indican que las paredes del dominio Ω se mantienen a temperatura constante $u = 0$. En la práctica, frecuentemente, se utilizan otras condiciones de contorno no tanto sobre la variable u que en la ecuación del calor representa la temperatura, sino sobre el flujo de calor a través de la frontera. Así, por ejemplo, en el caso en que queramos representar que el dominio Ω está completamente aislado de su entorno impondremos condiciones de flujo nulo, i.e.

$$\frac{\partial u}{\partial n} = 0 \text{ en } \partial\Omega \times (0, \infty).$$

Aquí $\partial/\partial n$ denota el operador derivada normal y n es el vector normal exterior unitario a $\partial\Omega$ que varía en función de la geometría del dominio al variar el punto $x \in \partial\Omega$. Se trata de una derivada direccional, de modo que

$$\frac{\partial}{\partial n} = \nabla \cdot n,$$

donde ∇ denota el operador gradiente $\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_N} \right)$ y \cdot el producto escalar euclideo en \mathbb{R}^N .

Pero, con el objeto de simplificar y no hacer demasiado larga la presentación, en estas notas nos limitaremos a considerar las condiciones de contorno de Dirichlet como en (2.7).

En este caso la solución no es tan fácil de obtener explícitamente como lo fue para el problema de Cauchy en \mathbb{R}^N . Son diversos los métodos disponibles para su resolución: Galerkin, semigrupos, series de Fourier, ... El lector interesado en el estudio de estos métodos puede consultar el texto de L. Evans [E].

Nosotros nos centraremos en el problema de una sola dimensión espacial. Consideraremos por lo tanto el sistema

$$\begin{cases} u_t - u_{xx} = 0, & 0 < x < \pi, \quad t > 0 \\ u(0, t) = u(\pi, t) = 0, & t > 0 \\ u(x, 0) = \varphi(x), & 0 < x < \pi. \end{cases} \quad (2.8)$$

En este caso la solución puede obtenerse fácilmente mediante el desarrollo en series de Fourier. En efecto, las funciones trigonométricas

$$w_l(x) = \sqrt{\frac{2}{\pi}} \operatorname{sen}(jx), \quad j \geq 1 \quad (2.9)$$

constituyen una base ortonormal de $L^2(0, \pi)$ (véase Lema 2.1).

Por lo tanto, para cualquier función $\varphi \in L^2(0, \pi)$ la solución u de (2.8) se puede escribir en la forma

$$u(x, t) = \sum_{j=1}^{\infty} \hat{\varphi}_j e^{-j^2 t} w_l(x) \quad (2.10)$$

donde $\{\hat{\varphi}_j\}_{j \geq 1}$ son los coeficientes de Fourier de la función φ , i.e.

$$\hat{\varphi}_j = \int_0^\pi \varphi(x) w_l(x) dx. \quad (2.11)$$

Esta expresión de la solución en series de Fourier nos resultará de gran utilidad a la hora de abordar la aproximación numérica de la solución. En realidad, las propias sumas parciales de la serie proporcionan ya una manera sistemática de aproximar la solución. Así, para cada $M \in \mathbb{N}$ podemos introducir

$$u_M(x, t) = \sum_{j=1}^M \hat{\varphi}_j e^{-j^2 t} w_l(x), \quad (2.12)$$

y es entonces fácil comprobar que

$$\|u(t) - u_M(t)\|_{L^2(0, \pi)} \leq e^{-M^2 t/2} \|\varphi\|_{L^2(0, \pi)}, \quad \forall t \geq 0, \quad (2.13)$$

lo cual indica, efectivamente, que la aproximación de u mediante u_M mejora a medida que $M \rightarrow \infty$.

En vista de la aparente simplicidad de este método de aproximación cabe entonces preguntarse: ¿Para qué necesitamos otros métodos?

Las razones son diversas, pero hay una particularmente importante. Si bien en este caso la obtención de las funciones de base $\{w_l\}_{j \geq 1}$ (que son, en realidad, autofunciones del operador de Laplace involucrado en la ecuación del calor con condiciones de contorno de Dirichlet) es muy simple por encontrarnos en una dimensión espacial, en varias dimensiones espaciales el problema es mucho más complejo, pues pasa por calcular las autofunciones del problema:

$$\begin{cases} -\Delta w = \lambda w & \text{en } \Omega \\ w = 0 & \text{en } \partial\Omega. \end{cases} \quad (2.14)$$

Antes que nada conviene señalar que las autofunciones w_l de (2.9) se obtienen precisamente al resolver el análogo uni-dimensional de (2.14). En este caso el problema de autovalores es un sencillo problema de Sturm-Liouville que se escribe en la forma

$$\begin{cases} -w'' = \lambda w, & 0 < x < \pi \\ w(0) = w(\pi) = 0. \end{cases} \quad (2.15)$$

Los autovalores son en este caso

$$\lambda_l = l^2, \quad l \geq 1 \quad (2.16)$$

y las autofunciones correspondientes, una vez normalizadas en $L^2(0, \pi)$, las funciones trigonométricas (2.9).

Si bien la teoría espectral garantiza la existencia de una sucesión de autofunciones que constituyen una base ortogonal de $L^2(\Omega)$ ([B]), su forma depende de la geometría del dominio Ω y, por supuesto, su cálculo explícito es imposible salvo para dominios muy particulares ([J]). Por lo tanto, en varias dimensiones espaciales, la utilización de estas autofunciones exige previamente el desarrollo de métodos numéricos para su aproximación, tan elaborados (o más) como los que vamos a necesitar para aproximar la propia ecuación del calor directamente.

Este hecho, junto con otro igualmente importante como es que para muchas ecuaciones (no-lineales, coeficientes dependientes del espacio-tiempo, etc.) la resolución mediante series de

Fourier no es posible, aconsejan que desarrollemos métodos numéricos que permitan abordar sistemáticamente la ecuación del calor y sus variantes, sin pasar por la Teoría Espectral.

Sí que conviene sin embargo utilizar este formalismo de Fourier para entender las aproximaciones que los diferentes esquemas proporcionan a la ecuación del calor y el modo en que afectan a las diferentes componentes de las soluciones en función de la frecuencia de su oscilación espacial.

Volvamos entonces a la ecuación (2.8) y a su solución (2.10).

En la expresión (2.10) se observa un comportamiento de u distinto al del problema de Cauchy en \mathbb{R}^N .

En efecto, en este caso es fácil comprobar que la solución decae exponencialmente cuando $t \rightarrow \infty$:

$$\|u(t)\|_{L^2(0,\pi)}^2 = \sum_{j=1}^{\infty} |\hat{\varphi}_j|^2 e^{-2j^2 t} \leq e^{-2t} \sum_{j=1}^{\infty} |\hat{\varphi}_j|^2 = e^{-2t} \|\varphi\|_{L^2(0,\pi)}^2. \quad (2.17)$$

Esta propiedad de decaimiento puede también obtenerse directamente de la ecuación (2.8) mediante el *método de la energía*, sin hacer uso del desarrollo en serie de Fourier de la solución. En efecto, multiplicando en (2.8) por u e integrando por partes se obtiene que

$$0 = \int_0^\pi (u_t - u_{xx}) u dx = \frac{1}{2} \frac{d}{dt} \int_0^\pi u^2 dx + \int_0^\pi u_x^2 dx,$$

o, lo que es lo mismo,

$$\frac{1}{2} \frac{d}{dt} \int_0^\pi u^2 dx = - \int_0^\pi u_x^2 dx. \quad (2.18)$$

Utilizamos ahora la desigualdad de Poincaré ([B])

$$\int_0^\pi u_x^2 dx \geq \int_0^\pi u^2 dx, \quad \forall u \in H_0^1(0, \pi) \quad (2.19)$$

que, combinada con la identidad (2.18), proporciona la desigualdad

$$\frac{d}{dt} \int_0^\pi u^2 dx \leq -2 \int_0^\pi u^2 dx. \quad (2.20)$$

Integrando esta desigualdad (2.20) obtenemos exactamente la tasa exponencial de decaimiento de la solución que predijimos en (2.17).

Observación 2.1 La *desigualdad de Poincaré* (ver [B]) garantiza que

$$\int_0^\pi |a'(x)|^2 dx \geq \int_0^\pi |a(x)|^2 dx, \quad \forall a \in H_0^1(0, \pi). \quad (2.21)$$

La mejor constante de la desigualdad (2.21) viene caracterizada por el siguiente principio de minimalidad que involucra el cociente de Rayleigh:

$$\lambda_1 = \min_{a \in H_0^1(0,\pi)} \frac{\int_0^\pi |a'(x)|^2 dx}{\int_0^\pi a^2(x) dx}. \quad (2.22)$$

En este caso $\lambda_1 = 1$ puesto que se trata del primer autovalor λ_1 del operador $-d^2/dx^2$ en $H_0^1(0, \pi)$ que posee una sucesión de autovalores (2.16). ■

2.2 Semi-discretización espacial: El método de Fourier

Esta sección está dedicada a estudiar las semi-discretizaciones espaciales de la ecuación del calor $1 - d$ (unidimensional) (2.8).

Lo haremos en el caso más sencillo en el que el operador de Laplace espacial se aproxima mediante el esquema clásico y sencillo de tres puntos que ya fue objeto de estudio en capítulos anteriores. Analizaremos la convergencia del método tanto mediante series de Fourier como por estimaciones de energía.

Si bien los resultados de esta sección se refieren a un problema muy sencillo como es (2.8), en el transcurso de la misma desarrollaremos una metodología susceptible de ser adaptada a situaciones más complejas. Esto es así, muy en particular en lo referente al método de la energía, de fácil aplicación a otras condiciones de contorno, coeficientes variables, ecuaciones no-lineales, etc.

Consideramos por tanto un paso $h > 0$ del mallado espacial. Para simplificar la presentación suponemos que $h = \pi/(M + 1)$ con $M \in \mathbb{N}$, de modo que la partición que definen los nodos

$$x_j = jh, \quad j = 0, \dots, M + 1 \quad (2.23)$$

descompone el intervalo $[0, \pi]$ en $M + 1$ subintervalos de longitud $h : I_j = [x_j, x_{j+1}]$, $j = 0, \dots, M$. Obsérvese que el primer y el último nodo corresponden a los extremos del intervalo, i.e. $x_0 = 0$, $x_{M+1} = \pi$.

Utilizando la clásica aproximación de tres puntos para el operador d^2/dx^2 (que, como vimos, es de orden dos) obtenemos de manera natural la siguiente semi-discretización de (2.8):

$$\begin{cases} u'_j + \frac{2u_j - u_{j+1} - u_{j-1}}{h^2} = 0, & j = 1, \dots, M, \quad t > 0 \\ u_0 = u_{M+1} = 0, & t > 0 \\ u_j(0) = \varphi_j, & j = 1, \dots, M. \end{cases} \quad (2.24)$$

Este sistema constituye un conjunto de M ecuaciones diferenciales de orden uno, lineales, acopladas de tres en tres con M incógnitas. En vista de que, por las condiciones de contorno, $u_0 \equiv u_{M+1} \equiv 0$, las genuinas incógnitas del problema son las M funciones $u_j(t)$, $j = 1, \dots, M$.

Cada una de las funciones $u_j(t)$ proporciona una aproximación de la solución $u(\cdot, t)$ en el punto $x = x_j$. A medida que el paso h de la discretización tiende a cero tenemos más y más puntos en el mallado. Cabe por tanto esperar que obtengamos progresivamente estimaciones más finas de la solución. Sin embargo, tal y como veremos, no basta con que una aproximación parezca coherente para poder garantizar su convergencia. El objetivo principal de este capítulo es precisamente desarrollar una teoría que nos permita discernir si un método numérico es convergente o no.

Queda sin embargo por determinar una buena elección de los datos iniciales φ_j , $j = 1, \dots, M$. Las posibilidades son diversas y algunas de ellas serán analizadas a lo largo de estas notas. Cuando la función φ del dato inicial de la ecuación del calor es continua, lo más sencillo es tomar sus valores en los nodos como dato inicial del problema semi-discreto, i.e. $\varphi_j = \varphi(x_j)$. Cuando φ no es continua sino meramente integrable podemos también hacer medias del dato inicial $\varphi = \varphi(x)$ en intervalos en torno a los nodos. También es posible elegir los datos iniciales del sistema semi-discreto truncando la serie de Fourier del dato inicial de la ecuación del calor.

Las posibilidades son diversas pero, si un método es convergente, ha de serlo para cualquier elección razonable de los datos iniciales. Esto dependerá esencialmente del esquema elegido para aproximar la ecuación y las condiciones de contorno.

Frecuentemente utilizaremos una notación vectorial para simplificar las expresiones. Introducimos por tanto el vector columna $\vec{u} = \vec{u}(t)$ que representa a la incógnita del sistema (2.24):

$$\vec{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_M(t) \end{pmatrix}, \quad (2.25)$$

y la matriz tridiagonal:

$$A_h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}. \quad (2.26)$$

Así, el sistema (2.24) se escribe

$$\begin{cases} \vec{u}'(t) + A_h \vec{u}(t) = 0, & t > 0 \\ \vec{u}(0) = \vec{\varphi}. \end{cases} \quad (2.27)$$

Obviamente la solución \vec{u} de (2.27) también depende de h de modo que sería más legítimo denotarla mediante el subíndice h : \vec{u}_h . Pero, para simplificar la notación, escribiremos simplemente \vec{u} , salvo que este hecho pueda conducir a confusión.

En la sección anterior vimos que que la ecuación del calor continua verifica el *principio del máximo* que garantiza que si el dato inicial es no-negativo, la solución lo es para todo x y todo t . Es esto es cierto para el problema de Cauchy en todo el espacio pero también lo es para el problema de Dirichlet con condiciones de contorno nulas. El sistema (2.24) refleja también esta propiedad de naturaleza física. En efecto, supongamos que el dato inicial de (2.24) es positivo, i.e. $\varphi_j > 0$, $j = 1, \dots, M$. Entonces, $u_j(t) > 0$ para todo $j = 1, \dots, M$ y todo $t > 0$. Con el objeto de probar esta afirmación argumentamos del modo siguiente. En primer lugar observamos que, por continuidad, existe $\tau > 0$ tal que $u_j(t) > 0$ para todo j y todo $0 \leq t \leq \tau$. Sea τ^* el primer instante de tiempo en el que la solución se anula en alguna de sus componentes que denotaremos mediante el índice j^* . Tenemos entonces:

- $u_j(t) > 0$, $\forall j = 1, \dots, M$, $\forall 0 \leq t \leq \tau^*$.

- $u_{j^*}(\tau^*) = 0$.
- $u'_{j^*}(\tau^*) \leq 0$.
- $u_j(\tau^*) \geq 0, j = 1, \dots, M$.

Haciendose uso de estas propiedades escribimos la ecuación de (2.24) correspondiente al índice j^* en el instante $t = \tau^*$. Obtenemos $u_{j^*+1}(\tau^*) + u_{j^*-1}(\tau^*) \leq 0$, lo cual implica, en virtud de las propiedades anteriores, que $u_{j^*+1}(\tau^*) = u_{j^*-1}(\tau^*) = 0$. Iterando este argumento se obtiene que $u_j(\tau^*) = 0$ para todo $j = 1, \dots, M$. Por unicidad de las soluciones de la ecuación diferencial (2.24) esto implica que $\vec{u} \equiv 0$, lo cual está en contradicción con el hecho de que el dato inicial sea positivo. Esto prueba que el principio del máximo se verifica también para el sistema semi-discreto (2.24).

En el dato inicial de u_j hemos tomado el valor exacto del dato inicial φ de la ecuación del calor en el punto x_j . Esto exige que el dato φ sea continuo. Pero existen otras elecciones del dato inicial de la ecuación semi-discreta como decíamos anteriormente. En particular, la elección puede hacerse a través de los coeficientes de Fourier de φ .

El punto de vista de Fourier no sólo es sumamente útil a la hora de entender la teoría analítica de las EDP sino también su Análisis Numérico. En efecto, la solución de (2.27) puede escribirse en serie de Fourier en la base de autovectores de la matriz A_h . En este caso será simplemente una suma finita de M términos pues se trata efectivamente de un problema finito-dimensional de dimensión M . Para ello es preciso introducir el espectro de la matriz A_h .

Consideremos por tanto el problema de autovalores:

$$A_h \vec{W} = \lambda \vec{W}. \quad (2.28)$$

Los autovalores y autovectores solución de (2.28) pueden calcularse de forma explícita:

$$\lambda_l(h) = \frac{4}{h^2} \operatorname{sen}^2 \left(l \frac{h}{2} \right), \quad j = 1, \dots, M. \quad (2.29)$$

Los autovectores correspondientes son

$$\vec{W}_l(h) = \sqrt{\frac{2}{\pi}} \begin{pmatrix} \operatorname{sen}(lx_1) \\ \vdots \\ \operatorname{sen}(lx_M) \end{pmatrix}, \quad l = 1, \dots, M. \quad (2.30)$$

El lector interesado puede encontrar una prueba de este hecho en [I], Lema 10.5.

A partir de ahora las componentes del vector $\vec{W}_l(h)$ serán denotadas mediante $(W_{l,j})_{j=1,\dots,M}$.

En (2.29)-(2.30) se observan varias analogías con los autovalores y autofunciones del operador de Laplace expresados en (2.9) y (2.16). En efecto, para cada índice $l \geq 1$ fijo tenemos

$$\lambda_l \rightarrow l^2, \quad \text{cuando } h \rightarrow 0, \quad (2.31)$$

lo cual refleja que los autovalores del problema discreto (2.28) aproximan a los del continuo (2.15) a medida que $h \rightarrow 0$, que es a su vez consecuencia de la convergencia del esquema de tres

puntos para la aproximación del Laplaciano probada en el capítulo anterior. Por otra parte, los autovectores $\vec{W}_l(h)$ del problema discreto (2.28) no son más que una restricción a los puntos del mallado de las autofunciones (2.9) del problema continuo. Esto explica por tanto la proximidad de ambos espectros. Es oportuno indicar también que los autovectores $\vec{W}_l(h)$ dependen de h de dos maneras: Por su número de componentes $M = \pi/h - 1$ y por el valor de cada una de ellas.

Conviene sin embargo señalar que, en general, cuando se abordan problemas más sofisticados (en varias dimensiones espaciales, por ejemplo) no es frecuente que se dé la coincidencia exacta entre los autovectores del problema discreto y las autofunciones del continuo sino simplemente la convergencia a medida que $h \rightarrow 0$, si bien, para que ésto sea cierto, es indispensable que el esquema numérico elegido para la aproximación de la ecuación de Laplace sea convergente.

Por último es interesante también observar que de la expresión explícita de los autovalores $\lambda_l(h)$ del problema discreto se deduce que

$$\lambda_l(h) \leq \lambda_l = l^2. \quad (2.32)$$

Las soluciones del problema discreto, como hemos visto, son vectores columna de \mathbb{R}^M y en el caso del problema de evolución, funciones regulares del tiempo t a valores en \mathbb{R}^M . En \mathbb{R}^M consideramos la norma euclídea escalada

$$\|\vec{e}\|_h = \left[h \sum_{j=1}^M |e_j|^2 \right]^{1/2}, \quad \forall \vec{e} = (e_1, \dots, e_M)^t, \quad (2.33)$$

y su producto escalar asociado

$$\langle \vec{e}, \vec{f} \rangle_h = h \sum_{j=1}^M e_j f_j. \quad (2.34)$$

El factor de escala introducido en la norma y producto escalar (\sqrt{h} y h respectivamente) es importante pues garantiza que, cuando $h \rightarrow 0$, estas normas y producto escalar aproximan a las correspondientes de $L^2(0, \pi)$:

$$\|e\|_{L^2(0, \pi)} = \left(\int_0^\pi e^2(x) dx \right)^{1/2}, \quad (2.35)$$

$$\langle e, f \rangle_{L^2(0, \pi)} = \int_0^\pi e(x) f(x) dx. \quad (2.36)$$

En efecto, en vista de que $h = \pi/(M + 1)$, se observa inmediatamente que (2.33) y (2.34) son versiones discretas de las integrales (2.35) y (2.36), semejantes a las sumas de Riemann.

Es por tanto natural, abusando un poco del lenguaje, referirse a este producto euclídeo escalado, como el producto en L^2 .

Es fácil comprobar que los autovectores $\vec{W}_l(h)$ de (2.30) son ortonormales en el producto escalar (2.34).

En efecto, tenemos el siguiente resultado:

Lema 2.1 Para cada h de la forma $h = \pi/(M+1)$, con $M \in \mathbb{N}$ y para cada $l \in \mathbb{N}$, $1 \leq l \leq M$, se tiene

$$h \sum_{j=1}^M \text{sen}^2(ljh) = \pi/2. \quad (2.37)$$

Asimismo, si $l, l' \in \mathbb{N}$ con $1 \leq l, l' \leq M$, $l \neq l'$,

$$h \sum_{j=1}^M \text{sen}(ljh) \text{sen}(l'jh) = 0. \quad (2.38)$$

Por consiguiente,

$$\|\vec{W}_l(h)\|_h = 1, \quad l = 1, \dots, M; \quad \langle \vec{W}_l(h), \vec{W}_k(h) \rangle_h = \delta_{lk}, \quad l, k = 1, \dots, M, \quad (2.39)$$

y

$$\langle A_h \vec{W}_l(h), \vec{W}_l(h) \rangle_h = h \sum_{j=1}^M \frac{|W_{l,j+1} - W_{l,j}|^2}{h^2} = \lambda_l(h), \quad l = 1, \dots, M. \quad (2.40)$$

Demostración del Lema 2.1. Para todo par l, l' con $1 \leq l, l' \leq M$ y $l \neq l'$ se tiene:

$$\begin{aligned} h \sum_{j=1}^M \text{sen}(ljh) \text{sen}(l'jh) &= \frac{h}{2} \sum_{j=1}^M [\cos(j(l' - l)h) - \cos(j(l + l')h)] \\ &= \frac{h}{2} \text{Re} \left(\sum_{j=0}^M e^{i(l' - l)jh} - \sum_{j=0}^M e^{i(l + l')jh} \right), \end{aligned}$$

donde Re denota la parte real de un número complejo. Aplicando la fórmula de la suma para una serie geométrica tenemos

$$\sum_{j=0}^M e^{i(l' \pm l)jh} = \frac{e^{i(l' \pm l)\pi} - 1}{e^{i(l' \pm l)h} - 1} = \frac{(-1)^{(l' \pm l)} - 1}{e^{i(l' \pm l)h} - 1} = \frac{(-1)^{(l' \pm l)} - 1}{\cos(l' \pm l)h + i \sin(l' \pm l)h - 1}.$$

Aplicando las fórmulas trigonométricas

$$\cos(x) = 1 - 2 \sin^2 \frac{x}{2}, \quad \sin(x) = 2 \sin \frac{x}{2} \cos \frac{x}{2}$$

en la identidad anterior obtenemos

$$\sum_{j=0}^M e^{i(l' \pm l)jh} = \frac{(-1)^{(l' \pm l)} - 1}{-2 \sin^2 \frac{(l' \pm l)h}{2} + 2i \sin \frac{(l' \pm l)h}{2} \cos \frac{(l' \pm l)h}{2}} = \frac{(-1)^{(l' \pm l)} - 1}{2i \sin \frac{(l' \pm l)h}{2} (\cos \frac{(l' \pm l)h}{2} + i \sin \frac{(l' \pm l)h}{2})}$$

$$= \frac{\left[(-1)^{(l' \pm l)} - 1\right] e^{-i \frac{(l' \pm l)h}{2}}}{2i \sin \frac{(l' \pm l)h}{2}} = -\frac{i}{2} \left[(-1)^{(l' \pm l)} - 1\right] \cot \frac{(l' \pm l)h}{2} - \frac{1}{2} \left[(-1)^{(l' \pm l)} - 1\right].$$

Resulta que

$$\operatorname{Re} \left(\sum_{j=0}^M e^{i(l'-l)jh} \right) = -\frac{1}{2} \left[(-1)^{(l'-l)} - 1\right],$$

$$\operatorname{Re} \left(\sum_{j=0}^M e^{i(l'+l)jh} \right) = -\frac{1}{2} \left[(-1)^{(l'+l)} - 1\right]$$

y por ello

$$h \sum_{j=1}^M \sin(ljh) \sin(l'jh) = \frac{h}{2} \left[-\frac{1}{2}(-1)^{(l'-l)} + \frac{1}{2} + \frac{1}{2}(-1)^{(l'+l)} - \frac{1}{2} \right] = 0.$$

Para $1 \leq l' = l \leq M$ se tiene

$$h \sum_{j=1}^M \sin^2(ljh) = \frac{h}{2} \sum_{j=0}^M (1 - \cos(2ljh)) = \frac{h(M+1)}{2} = \frac{\pi}{2},$$

puesto que

$$\sum_{j=0}^M \cos(2ljh) = \operatorname{Re} \left(\frac{e^{i2lh(M+1)} - 1}{e^{i2lh} - 1} \right) = 0.$$

A partir de estas dos identidades se deducen automáticamente las propiedades de los autovectores de la matriz A_h . Esto concluye la prueba del Lema. ■

Este hecho permite desarrollar fácilmente las soluciones de (2.27) en series de Fourier, es decir, en la base $\{\vec{W}_l(h)\}_{l=1, \dots, M}$:

$$\vec{u}(t) = \vec{u}_h(t) = \sum_{l=1}^M \hat{\varphi}_l e^{-\lambda_l(h)t} \vec{W}_l(h), \quad (2.41)$$

donde $\hat{\varphi}_l$ son los coeficientes del vector de datos iniciales $\vec{\varphi}$ en la base de autovectores $\{\vec{W}_l\}$, i.e.

$$\hat{\varphi}_l = \langle \vec{\varphi}, \vec{W}_l(h) \rangle_h, \quad (2.42)$$

de modo que

$$\vec{\varphi} = \sum_{j=1}^M \hat{\varphi}_l \vec{W}_l(h). \quad (2.43)$$

Las analogías entre la fórmula de representación (2.41) y el desarrollo en serie de Fourier (2.10) de las soluciones del problema continuo son evidentes. En realidad sólo hay dos diferencias dignas de ser reseñadas:

- (a) En (2.41) se tiene una suma finita de M términos en lugar de la serie infinita de (2.10). Ahora bien $M \rightarrow \infty$ cuando $h \rightarrow 0$.
- (b) Las exponenciales temporales de (2.41) y (2.10) no son exactamente las mismas pues en ellas intervienen los autovalores de uno y otro problema, si bien, en virtud de (2.31), ambas son semejantes.

En vista de esta similitud existente entre las expresiones de las soluciones continuas y discretas, estas últimas presentan propiedades semejantes a las de las primeras. En particular, en lo referente a decaimiento de la solución tenemos:

$$\| \vec{u}(t) \|_h^2 = \sum_{l=1}^M | \hat{\varphi}_l |^2 e^{-2\lambda_l(h)t} \leq e^{-2\lambda_1(h)t} \| \vec{\varphi} \|_h^2. \quad (2.44)$$

Así, en el límite cuando $h \rightarrow 0$, recuperamos la tasa de decaimiento en tiempo del problema continuo (2.17) puesto que $\lambda_1(h) \rightarrow 1$ cuando $h \rightarrow 0$.

El primer resultado importante de esta sección es el siguiente y garantiza la convergencia de las soluciones del problema semi-discreto (2.15) a las soluciones del problema continuo (2.8) cuando $h \rightarrow 0$, bajo una elección adecuada de los datos iniciales del problema discreto.

Desde el punto de vista del desarrollo en serie de Fourier de las soluciones que estamos barajando, a la hora de elegir una aproximación del dato inicial φ del problema semi-discreto parece adecuado proceder del siguiente modo.

Dado $\varphi \in L^2(0, \pi)$, consideramos su desarrollo en serie de Fourier

$$\varphi(x) = \sum_{l=1}^{\infty} \hat{\varphi}_l w_l(x), \quad (2.45)$$

donde

$$\hat{\varphi}_l = \int_0^{\pi} \varphi(x) w_l(x) dx, \quad (2.46)$$

de modo que, por la identidad de Parseval, se tiene

$$\| \varphi \|_{L^2(0,\pi)} = \left[\sum_{l=1}^{\infty} (\hat{\varphi}_l)^2 \right]^{1/2}. \quad (2.47)$$

Elegimos entonces el dato inicial $\vec{\varphi}$ de la ecuación discreta de modo que tenga los mismos coeficientes de Fourier que los M primeros de $\varphi(x)$, i.e.

$$\vec{\varphi} = \vec{\varphi}(h) = \sum_{l=1}^M \hat{\varphi}_l \vec{W}_l(h). \quad (2.48)$$

En este caso los coeficientes de Fourier del desarrollo (2.41) de la solución del problema semi-discreto coinciden con los coeficientes φ_l de la función $\varphi(x)$.

Con esta elección de los datos iniciales es fácil probar la convergencia del esquema numérico. Tenemos el siguiente resultado:

Teorema 2.1 *Supongamos que $\varphi \in L^2(0, \pi)$ y consideremos los datos iniciales del problema semi-discreto (2.27) como en (2.48).*

Entonces, las soluciones $\vec{u}_h = \vec{u}_h(t)$ del problema semi-discreto (2.27), cuando $h \rightarrow 0$, convergen a la solución $u = u(x, t)$ del problema continuo en el sentido que

$$\| \vec{u}_h(t) - \vec{u}(t) \|_{h \rightarrow 0} \rightarrow 0, \text{ para todo } t > 0, \quad (2.49)$$

cuando $h \rightarrow 0$, donde $\vec{u}(t)$ es la restricción de la solución de la ecuación del calor a los nodos del mallado: $\underline{u}_j(t) = u(x_j, t)$.

Conviene explicar la noción de convergencia adoptada en (2.49). La cantidad que aparece en (2.49), para cada $t > 0$, representa la norma $\| \cdot \|_h$ de la diferencia entre la solución discreta $\vec{u}_h(t)$ y la continua $u(\cdot, t)$. Ahora bien, como $u(\cdot, t)$ depende continuamente de x , a la hora de compararla con la solución discreta, sólo interviene la restricción de u a los puntos x_j del mallado que denotamos mediante $\vec{u}(t)$.

Demostración del Teorema 2.1.

A la hora de estudiar la diferencia $\vec{u}_h(t) - \vec{u}(t)$ distinguimos las bajas y las altas frecuencias:

$$\vec{u}_h(t) - \vec{u}(t) = \sum_{l=1}^{M_0} \hat{\varphi}_l [e^{-\lambda_l(h)t} - e^{-\lambda_l t}] \vec{W}_l(h) + \sum_{l=M_0+1}^M \hat{\varphi}_l e^{-\lambda_l(h)t} \vec{W}_l(h) - \sum_{l=M_0+1}^{\infty} \hat{\varphi}_l e^{-\lambda_l t} \vec{w}_l. \quad (2.50)$$

El valor del parámetro de corte M_0 será fijado más adelante.

Conviene observar que en el tercer sumatorio I_3 , mediante la expresión \vec{w}_l denotamos la restricción de la autofunción continua $w_l = w_l(x)$ a los puntos del mallado, aunque para $l > M$ no corresponde a un autovector de la matriz A_h . En el caso que nos ocupa (diferencias finitas en una dimensión espacial), las expresiones son en este caso particularmente simples pues, como ya dijimos, los autovectores del problema discreto son restricciones al mallado de las autofunciones continuas.

A la hora de estimar los tres términos en los que hemos descompuesto la diferencia (I_1 para las bajas frecuencias, I_2 , I_3 para las altas) el Lema 2.1 nos será de gran utilidad.

Tomando normas $\| \cdot \|_h$ en (2.50) obtenemos

$$\| \vec{u}_h(t) - \vec{u}(t) \|_h \leq \| I_1 \|_h + \| I_2 \|_h + \| I_3 \|_h. \quad (2.51)$$

Estimamos ahora por separado las tres normas $\| I_j \|_h, j = 1, 2, 3$.

Observamos en primer lugar que

$$\| \vec{w}_l \|_h \leq \sqrt{\pi} \| \vec{w}_l \|_{\infty} = \sqrt{\pi} \max_{j=1, \dots, M} | \vec{w}_l(x_j) | = \sqrt{2}. \quad (2.52)$$

De este modo deducimos que

$$\|I_3\|_h = \sqrt{2} \sum_{j \geq M_0+1} |\hat{\varphi}_j| e^{-\lambda_j t} \leq \sqrt{2} \left[\sum_{j \geq M_0+1} |\hat{\varphi}_j|^2 \right]^{1/2} \left[\sum_{j \geq M_0+1} e^{-2\lambda_j t} \right]^{1/2}. \quad (2.53)$$

Evidentemente, este cálculo está justificado por la convergencia de la última de las series que interviene en esta desigualdad, lo cual está garantizado para todo $t > 0$.

De (2.53) tenemos

$$\|I_3\|_h^2 \leq C(t) \sum_{j \geq M_0+1} |\hat{\varphi}_j|^2, \quad (2.54)$$

con

$$C(t) = \left[\sum_{j \geq 1} e^{-2\lambda_j t} \right]^{1/2}. \quad (2.55)$$

Como el dato inicial $\varphi \in L^2(0, \pi)$, sus coeficientes de Fourier satisfacen

$$\int_0^\pi \varphi^2(x) dx = \sum_{j=1}^{\infty} |\hat{\varphi}_j|^2 \quad (2.56)$$

y, por lo tanto, dado $\varepsilon > 0$ arbitrariamente pequeño, existe $M_0 \in \mathbb{N}$ tal que

$$\sum_{j \geq M_0+1} |\hat{\varphi}_j|^2 \leq \varepsilon^2, \quad \forall M \geq M_0. \quad (2.57)$$

Dado $t > 0$, este permite por tanto fijar el valor de M_0 de modo que

$$\|I_3(t)\|_h \leq \varepsilon. \quad (2.58)$$

El término I_2 puede estimarse de manera idéntica. Pero en este caso puede incluso encontrarse una estimación uniforme para todo $t \geq 0$ gracias a las propiedades de ortonormalidad de los autovectores $\vec{w}_l(h)$. En efecto,

$$\|I_2\|_h^2 = \sum_{j=M_0+1}^M \hat{\varphi}_j^2 e^{-2\lambda_j(h)t} \leq \sum_{j=M_0+1}^M \hat{\varphi}_j^2, \quad \forall t \geq 0. \quad (2.59)$$

Dado $\varepsilon > 0$ arbitrariamente pequeño esto permite hallar M_0 tal que

$$\|I_2(t)\|_h \leq \varepsilon, \quad \forall t \geq 0. \quad (2.60)$$

Procedemos ahora a la estimación de $\|I_1\|_h$. Tenemos

$$\begin{aligned} \|I_1\|_h^2 &= \sum_{j=1}^{M_0} (\hat{\varphi}_l)^2 \left(e^{-\lambda_l(h)t} - e^{-\lambda_l t} \right)^2 \|\vec{w}_l(h)\|_h^2 \\ &= \sum_{j=1}^{M_0} (\hat{\varphi}_l)^2 \left(e^{-\lambda_l(h)t} - e^{-\lambda_l t} \right)^2. \end{aligned} \quad (2.61)$$

Nótese que en esta ocasión el valor de M_0 está ya fijado, en virtud de la elección hecha antes en función de ε . En esta ocasión es el parámetro h el que tiende a cero. Obsérvese que, para cada $l \in \{1, \dots, M_0\}$ en vista de (2.31), el sumando $(\varphi_l)^2 (e^{-\lambda_l(h)t} - e^{-\lambda_l t})$ tiende a cero cuando $h \rightarrow 0$ uniformemente en $t \in [0, T]$. Por lo tanto, en vista de que el número M_0 de sumandos está fijado, deducimos que

$$\|I_1\|_h \rightarrow 0, \quad h \rightarrow 0, \quad \text{uniformemente en } t \in [0, T].$$

En particular, eligiendo h suficientemente pequeño se puede asegurar que $\|I_1(t)\|_h \leq \varepsilon$, para todo $t \geq 0$.

Combinando estas estimaciones deducimos, que para cualquier $t > 0$ y $\varepsilon > 0$ existe h suficientemente pequeño tal que, según (2.51), $\left\| \vec{u}_h(t) - \underline{u}(t) \right\|_h \leq 3\varepsilon$.

Esto concluye la demostración del Teorema 2.1. ■

Observación 2.2 En (2.49) hemos establecido la convergencia de las soluciones del problema semi-discreto \vec{u}_h a la del problema continuo en el sentido de la norma $\| \cdot \|_h$. Pero ésto no es más que una de las posibles maneras de establecer la proximidad entre las soluciones de ambos problemas. A continuación presentamos algunas variantes. ■

Algunas variantes del Teorema de convergencia 2.1.

- *Variante 1. Datos iniciales en $H_0^1(0, \pi)$.*

Supongamos, por ejemplo, que el dato inicial φ es un poco más regular:

$$\varphi \in H_0^1(0, \pi) = \{ \varphi \in L^2(0, \pi) : \varphi' \in L^2(0, \pi), \varphi(0) = \varphi(\pi) = 0 \}.$$

En este caso, obviamente, tenemos el resultado de convergencia del Teorema 2.1. Pero, bajo esta hipótesis adicional sobre el dato inicial, se puede dar una versión más precisa y cuantitativa de este resultado.

En este caso los coeficientes de Fourier $\{\hat{\varphi}_l\}_{j \in \mathbb{N}}$ de φ satisfacen:⁵

$$\| \varphi \|_{H_0^1(0, \pi)}^2 = \sum_{l \geq 1} l^2 |\hat{\varphi}_l|^2 < \infty. \quad (2.62)$$

⁵Obsérvese que, en virtud de la desigualdad de Poincaré, la norma inducida por $H^1(0, \pi)$ sobre $H_0^1(0, \pi)$ y la norma $\| \varphi \|_{H_0^1(0, \pi)}^2 = \left[\int_0^\pi |\varphi'|^2 dx \right]^{1/2}$ son normas equivalentes. A nivel de los coeficientes de Fourier esta última se reduce a $\| \varphi \|_{H_0^1(0, \pi)} = \left[\sum_{l \geq 1} l^2 |\hat{\varphi}_l|^2 \right]^{1/2}$ que es a su vez equivalente a la inducida por la norma $\| \varphi \|_{H_0^1(0, \pi)} = \left[\sum_{l \geq 1} (1 + l^2) |\hat{\varphi}_l|^2 \right]^{1/2}$.

Gracias a esta hipótesis adicional la elección del M_0 que es preciso para que se cumpla (2.57) se puede realizar explícitamente. En efecto

$$\sum_{l \geq M_0+1} |\hat{\varphi}_l|^2 \leq \frac{1}{(M_0+1)^2} \sum_{l \geq M_0+1} l^2 |\hat{\varphi}_l| \leq \frac{\|\varphi\|_{H_0^1(0,\pi)}^2}{(M_0+1)^2} \quad (2.63)$$

y, por tanto, basta con elegir

$$M_0 = \frac{\|\varphi\|_{H_0^1(0,\pi)}}{\varepsilon} - 1 \quad (2.64)$$

para que (2.57) se cumpla.

Una vez que M_0 ha sido fijado de este modo, también h puede ser elegido de forma que $\|I_1\|_h$ sea menor que ε .

En efecto,

$$\begin{aligned} \|I_1\|_h^2 &= \sum_{j=1}^{M_0} |\hat{\varphi}_l|^2 \left(e^{-\lambda_l(h)t} - e^{-\lambda_l t} \right) = \sum_{l=1}^{M_0} |\hat{\varphi}_l|^2 t (\lambda_l - \lambda_l(h)) e^{-\mu_j(h)t} \\ &\leq \sum_{j=1}^{M_0} |\hat{\varphi}_l|^2 t (\lambda_l - \lambda_l(h)), \end{aligned} \quad (2.65)$$

donde, mediante $\mu_j(h)$, hemos denotado un número real entre $\lambda_l(h)$ y λ_l , obtenido en la aplicación del Teorema del Valor Medio.

Por otra parte,

$$\begin{aligned} \lambda_l - \lambda_l(h) &= l^2 - \frac{4}{h^2} \operatorname{sen}^2 \left(l \frac{h}{2} \right) = \frac{1}{h^2} \left[(hl)^2 - 4 \operatorname{sen}^2 \left(l \frac{h}{2} \right) \right] \\ &= \frac{1}{h^2} [hl + 2 \operatorname{sen}(lh/2)] [hl - 2 \operatorname{sen}(lh/2)] \\ &\leq \frac{2l}{h} [hl - 2 \operatorname{sen}(lh/2)] \leq \frac{l}{h} C(hl)^3 = C h^2 l^4 \leq C M_0^4 h^2 \end{aligned} \quad (2.66)$$

donde C es una constante que puede ser calculada explícitamente. El mayor valor de dicha constante viene dado por

$$C = 2 \max_{|\tau| \leq \pi} [\tau - 2 \operatorname{sen}(\tau/2)] / \tau^3 \quad (2.67)$$

siempre y cuando $h > 0$ sea lo suficientemente pequeño de modo que $lh \leq M_0 h \leq (M+1)h = \pi$, es decir que $M+1 \geq M_0$.

En virtud de (2.66) y en vista del valor explícito de M_0 dado en (2.64) el valor de h para que, según (2.65), $\|I_1\|_h \leq \varepsilon$ puede calcularse explícitamente. Esto permite cuantificar el resultado de convergencia del Teorema 2.1. Obsérvese sin embargo que este tipo de

argumentos necesita que el dato inicial este en un espacio más pequeño que el espacio $L^2(0, \pi)$ donde el resultado de convergencia ha sido probado.

Pero hay otras variantes del resultado de convergencia del Teorema 2.1 que pueden resultar incluso más elocuentes. ■

- *Variante 2. Convergencia en la norma del máximo*

Por ejemplo, puede resultar más natural estimar la distancia entre la solución \vec{u}_h del problema semi-discreto y la solución u de (2.8) en la norma del máximo:

$$\|\vec{u}_h(t) - \vec{u}(t)\|_\infty = \max_{j \in \{1, \dots, M\}} |u_j(t) - u(x_j, t)|. \quad (2.68)$$

Para la estimación de esta cantidad procedemos del modo siguiente. Descomponiendo la norma de la diferencia como en la prueba del Teorema 2.1 tenemos:

$$\begin{aligned} \|\vec{u}_h(t) - \vec{u}(t)\|_\infty &= \left\| \sum_{l=1}^M \hat{\varphi}_l e^{-\lambda_l(h)t} \vec{W}_l(h) - \sum_{l=1}^{\infty} \hat{\varphi}_l e^{-\lambda_l t} \vec{w}_l \right\|_\infty \\ &\leq \sqrt{\frac{2}{\pi}} \sum_{l=1}^{M_0} |\hat{\varphi}_l| \left| e^{-\lambda_l(h)t} - e^{-\lambda_l t} \right| + \sqrt{\frac{2}{\pi}} \sum_{l=M_0+1}^M |\hat{\varphi}_l| e^{-\lambda_l(h)t} + \sqrt{\frac{2}{\pi}} \sum_{l \geq M_0+1} |\hat{\varphi}_l| e^{-\lambda_l t} \\ &= I_1 + I_2 + I_3. \end{aligned}$$

En esta desigualdad hemos tenido en cuenta que $\|\vec{w}_l\|_\infty \leq \sqrt{2/\pi}$, para todo $l \geq 1$.

Estimamos ahora el último término:

$$I_3 = \sqrt{\frac{2}{\pi}} \sum_{l \geq M_0+1} |\hat{\varphi}_l| e^{-\lambda_l t} \leq \sqrt{\frac{2}{\pi}} \left[\sum_{l \geq M_0+1} |\hat{\varphi}_l|^2 \right]^{1/2} \left[\sum_{l \geq M_0+1} e^{-2\lambda_l t} \right]^{1/2}.$$

Como $\lambda_l = l^2$, la última serie de esta desigualdad converge para cada $t > 0$, i.e.

$$\sum_{l=1}^{\infty} e^{-2\lambda_l t} = \sum_{l=1}^{\infty} e^{-2l^2 t} < \infty$$

y, por lo tanto, con una elección adecuada de $M_0 = M_0(\varepsilon)$, puede asegurarse que

$$I_3 \leq \varepsilon \|\varphi\|_{L^2(0, \pi)}.$$

La misma acotación es válida para I_2 .

Fijado el valor de M_0 de modo que estas cotas de I_2 y I_3 sean válidas procedemos a acotar I_1 :

$$I_1 \leq \sqrt{\frac{2}{\pi}} \sum_{l=1}^{M_0} |\hat{\varphi}_l| \left| e^{-\lambda_l(h)t} - e^{-\lambda_l t} \right|.$$

Este último término tiende a cero cuando $h \rightarrow 0$ puesto que se trata de una suma finita en la que cada uno de los términos tiende a cero.

De este modo concluimos que, cuando el dato inicial $\varphi \in L^2(0, \pi)$, para todo $t > 0$ se tiene

$$\|\vec{u}_h(t) - \vec{u}(t)\|_\infty \rightarrow 0, \quad h \rightarrow 0. \quad (2.69)$$

Obsérvese que en este caso no se tiene una convergencia uniforme en tiempo $t \in [0, T]$. En efecto, la convergencia en $t = 0$ exigiría estimar los términos I_2 y I_3 de otro modo, y necesitaríamos de la hipótesis

$$\sum_{j=1}^{\infty} |\hat{\varphi}_j| < \infty,$$

cosa que no está garantizada por la condición $\varphi \in L^2(0, \pi)$. Sin embargo, sí que bastaría con suponer que $\varphi \in H_0^1(0, \pi)$, como en la variante anterior, puesto que

$$\sum_{j=1}^{\infty} |\hat{\varphi}_j| \leq \left[\sum_{j=1}^{\infty} |\hat{\varphi}_j|^2 j^2 \right]^{1/2} \left[\sum_{j=1}^{\infty} j^{-2} \right] = C \|\varphi\|_{H_0^1(0, \pi)}. \quad (2.70)$$

Volviendo al caso general $\varphi \in L^2(0, \pi)$, acabamos de comprobar que, a pesar de que el dato inicial φ se supone únicamente en $L^2(0, \pi)$, la convergencia de la solución semi-discreta a la solución continua se produce en la norma del máximo (que corresponde a la norma de $L^\infty(0, \pi)$). Esto es así gracias al efecto regularizante que, como hemos mencionado, caracteriza a la ecuación del calor y que todos los problemas semi-discretos (2.24) comparten. En efecto, la solución de la ecuación del calor es de la forma

$$u = \sum_{l=1}^{\infty} \hat{\varphi}_l e^{-l^2 t} w_l(x)$$

de modo que

$$\begin{aligned} |u(x, t)| &\leq \sum_{l=1}^{\infty} |\hat{\varphi}_l| e^{-l^2 t} |w_l(x)| \leq \sqrt{\frac{2}{\pi}} \sum_{l=1}^{\infty} |\hat{\varphi}_l| e^{-l^2 t} \\ &\leq \sqrt{\frac{2}{\pi}} \left[\sum_{l=1}^{\infty} |\hat{\varphi}_l|^2 \right]^{1/2} \left[\sum_{l=1}^{\infty} e^{-2l^2 t} \right]^{1/2} = C(t) \|\varphi\|_{L^2(0, \pi)} \end{aligned} \quad (2.71)$$

con $C(t) < \infty$ para todo $t > 0$. Esto garantiza el efecto regularizante $L^2(0, \pi) \rightarrow L^\infty(0, \pi)$ en la ecuación del calor para cualquier instante de tiempo $t > 0$. Obviamente $C(0) = \infty$, lo cual corresponde a la existencia de funciones de $L^2(0, \pi)$ que no pertenecen a $L^\infty(0, \pi)$.

Este efecto regularizante es compartido por todas las soluciones del problema semi-discreto. Para comprobarlo basta observar que existe $c > 0$ tal que

$$\lambda_l(h) \geq cl^2, \quad \forall h > 0, \quad \forall l = 1, \dots, M, \quad (2.72)$$

lo cual garantiza un control uniforme (con respecto a h) de las series que intervienen en la cuantificación de dicho efecto. ■

- *Variante 3. Datos iniciales en Fourier dependientes de h .*

En el Teorema 2.1 hemos optado por elegir en el problema semi-discreto (2.24) (o (2.27)) como dato inicial la truncamiento de la serie de Fourier del dato inicial de la ecuación del calor. Podría decirse pues que el dato inicial es independiente de h , en el sentido en que sus coeficientes de Fourier (los que se pueden imponer en el problema semi-discreto) lo son.

En realidad, el mismo método de demostración del Teorema 2.1 permite probar otro tipo de resultados en los que el dato inicial del problema semi-discreto no es necesariamente ése. En particular, permite establecer la convergencia de las soluciones a partir de informaciones mínimas sobre la convergencia de los datos iniciales.

Por ejemplo, supongamos que en el problema semi-discreto (2.24) (o (2.27)) tomamos como dato inicial

$$\vec{\varphi}(h) = \sum_{l=1}^M \hat{\varphi}_l(h) \bar{w}_l(h),$$

es decir, definimos el dato inicial mediante coeficientes de Fourier que dependen de h . Si extendemos estos coeficientes de Fourier $\hat{\varphi}_l(h)$ por cero para $l \geq M + 1$, podemos identificar, para cada $h > 0$ los coeficientes de Fourier de $\vec{\varphi}(h)$ con una sucesión de ℓ^2 (el espacio de Hilbert de las sucesiones de números reales de cuadrado sumable). El enunciado del Teorema 2.1 puede entonces generalizarse del siguiente modo:

“El resultado del Teorema 2.1 es aún cierto si, cuando $h \rightarrow 0$, $\{\hat{\varphi}_l(h)\}_{l \geq 1}$ converge en ℓ^2 a $\{\hat{\varphi}_l\}_{l \geq 1}$, donde $\{\hat{\varphi}_l(h)\}_{l \geq 1}$ (resp. $\{\hat{\varphi}_l\}_{l \geq 1}$) representa el elemento de ℓ^2 constituido por los coeficientes de Fourier del dato $\vec{\varphi}(h)$ del problema discreto (resp. del dato $\varphi \in L^2(0, \pi)$ del problema continuo)”.

En realidad, si hacemos uso del efecto regularizante, tal y como comentábamos en la variante anterior, se puede probar la convergencia del esquema bajo hipótesis más débiles sobre los datos iniciales:⁶

“Supongamos que los datos iniciales $\vec{\varphi}(h)$ del problema semi-discreto (2.24) (o (2.27)) son tales que $\{\hat{\varphi}_l(h)\}_{l \geq 1}$ converge débilmente en ℓ^2 a $\{\hat{\varphi}_l\}_{l \geq 1}$ cuando $h \rightarrow 0$. Entonces, la convergencia (2.49) es cierta uniformemente en $t \geq \delta$, para cualquier $\delta > 0$ ”.

⁶En un espacio de Hilbert H se dice que una sucesión $\{h_k\}_{k \geq 1}$ converge débilmente a un elemento $h \in H$, si $(h_k, g)_H \rightarrow (h, g)_H$ para todo $g \in H$. Obviamente, la convergencia clásica en el sentido de la norma (también denominada convergencia fuerte) implica la convergencia débil. Por otra parte, si una sucesión converge débilmente y además se tiene que $\|h_k\|_H \rightarrow \|h\|_H$, entonces se tiene también la convergencia en norma. Una de las propiedades más utilizadas de la convergencia débil es que de toda sucesión acotada en un espacio de Hilbert se puede extraer una subsucesión que converge débilmente.

■

Variante 4. Datos iniciales por restricción al mallado.

Pero en todos estos resultados la convergencia de la solución del problema discreto al continuo se establece en función del comportamiento de los coeficientes de Fourier de los datos iniciales cuando $h \rightarrow 0$. Sin embargo, desde un punto de vista estrictamente numérico, éste no es el modo más natural de proceder puesto que sería deseable disponer de un método más sencillo que no pase por el cálculo de las series de Fourier continuas y/o discretas para realizar la elección de los datos iniciales en la ecuación semi-discreta.

Supongamos por ejemplo que el dato inicial φ de la ecuación del calor es un poco más regular: $\varphi \in C([0, \pi])$. En este caso la elección más natural del dato inicial en el método numérico (2.24) (o (2.27)) es

$$u_j(0) = \varphi_j = \varphi(x_j).$$

En efecto, al elegir este dato inicial no necesitamos calcular los coeficientes de Fourier de φ (lo cual exige realizar una integral y, desde un punto de vista numérico la utilización de fórmulas de cuadratura) sino que basta evaluar el dato inicial sobre los puntos del mallado.

Si bien ésto supone una elección distinta de los datos iniciales, su valor efectivo no dista mucho del que utilizamos mediante series de Fourier. En efecto en el Teorema 2.1 hicimos la elección

$$\vec{\varphi} = \sum_{l=1}^M \hat{\varphi}_l \vec{W}_l(h)$$

que denotaremos mediante $\vec{\varphi}$ para distinguirla de la anterior. Teniendo en cuenta que la k -ésima componente del vector $\vec{w}_l(h)$ coincide con el valor de la autofunción $w_l(x)$ en el punto $x = x_k = kh$, tenemos entonces

$$\underline{\varphi}_k = \sum_{l=1}^M \hat{\varphi}_l w_l(x_k)$$

mientras que

$$\varphi_k = \varphi(x_k) = \sum_{l=1}^{\infty} \hat{\varphi}_l w_l(x_k).$$

Por tanto

$$\left| \varphi_k - \underline{\varphi}_k \right| = \left| \sum_{l=M+1}^{\infty} \hat{\varphi}_l w_l(x_k) \right| \leq \sqrt{\frac{2}{\pi}} \sum_{l \geq M+1} |\hat{\varphi}_l|.$$

Así

$$\|\vec{\varphi} - \underline{\vec{\varphi}}\|_h^2 = h \sum_{k=1}^M \left| \varphi_k - \underline{\varphi}_k \right|^2 \leq 2 \left(\sum_{l \geq M+1} |\hat{\varphi}_l| \right)^2.$$

Con el objeto de garantizar que

$$\|\vec{\varphi} - \underline{\vec{\varphi}}\|_h \rightarrow 0, \text{ cuando } h \rightarrow 0$$

basta por tanto con saber que

$$\sum_{l=1}^{\infty} |\hat{\varphi}_l| < \infty,$$

lo cual, como vimos en la variante 1, está garantizado, por ejemplo, si $\varphi \in H_0^1(0, \pi)$, lo que supone una hipótesis ligeramente más fuerte que la continuidad de φ .

Es fácil ver entonces que el resultado del Teorema 2.1 se preserva con esta elección del dato inicial del problema semi-discreto.

■

Variante 5. Datos iniciales en media.

Hay otras elecciones posibles del dato inicial en el problema semi-discreto (2.24) (o (2.27)) que no pasan por el cálculo de los coeficientes de Fourier del dato inicial. Por ejemplo, para cualquier $\varphi \in L^2(0, \pi)$, podemos elegir

$$\varphi_j = \frac{1}{h} \int_{x_j-h/2}^{x_j+h/2} \varphi(x) dx.$$

No es difícil ver que esta elección de los datos iniciales conduce a resultados semejantes de convergencia.

En realidad, y esto es comentario interesante y útil, una vez que se dispone de un resultado de convergencia para una determinada elección de los datos iniciales, no es difícil probar la convergencia para otras posibles opciones puesto que basta con estimar la diferencia de las soluciones del problema discreto o semi-discreto para las dos elecciones de los datos, sin necesidad de volver a comprobar la proximidad con el modelo continuo.

Esto es particularmente sencillo en el caso que nos ocupa puesto que si \vec{u}_h y \vec{v}_h son dos soluciones del problema semi-discreto (2.24) correspondientes a datos iniciales $\vec{\varphi}$ y $\vec{\psi}$, se tiene

$$\|\vec{u}_h(t) - \vec{v}_h(t)\|_h \leq \|\vec{\varphi} - \vec{\psi}\|_h, \quad \forall t > 0, \forall h > 0, \quad (2.73)$$

tal y como se desprende de (2.44).

Por lo tanto, el método que hemos desarrollado en esta sección, basado en series de Fourier, permite probar la convergencia del método no sólo cuando los datos iniciales han sido adaptados en función del desarrollo en serie de Fourier del dato inicial de la ecuación del calor, sino en cualquier circunstancia en la que el dato inicial de la ecuación semi-discreta haya sido elegido de manera coherente.

■

Una interpretación global del resultado de convergencia nodal.

En el Teorema 2.1 hemos probado la convergencia de la solución del problema semi-discreto a la del problema continuo en el sentido de la norma discreta $\|\cdot\|_h$, i.e. sobre los nodos del mallado. Sin embargo, en la medida en que la solución de la ecuación del calor está definida en todo el intervalo $(0, \pi)$ cabría esperar que pueda también establecerse un resultado de carácter global. Con el objeto de probar dicho resultado lo primero que tenemos que hacer es extender la función discreta \vec{u} a una función definida en todo el intervalo $(0, \pi)$. Hay diversas maneras de realizar ésto. La más sencilla es tal vez considerar la función constante a trozos:

$$[E\vec{u}_h(t)](x) = \sum_{j=1}^M u_j(t) \chi_{[x_j-h/2, x_j+h/2]}, \quad (2.74)$$

donde $\chi_{[x_j-h/2, x_j+h/2]}$ denota la función característica del intervalo $[x_j - h/2, x_j + h/2]$. Esta función extendida está definida para todo $x \in (0, \pi)$ y para todo t . Cabe por tanto plantearse su convergencia hacia la solución de la ecuación del calor cuando $h \rightarrow 0$.

Como consecuencia inmediata del Teorema 2.1 tenemos el siguiente resultado:

Corolario 2.1 *Bajo las hipótesis del Teorema 2.1 se tiene*

$$E\vec{u}_h(t) \rightarrow u(t) \quad \text{en } L^2(0, \pi), \quad \forall t > 0. \quad (2.75)$$

Demostración del Corolario 2.1.

Para probar este Corolario basta observar que

$$\|E\vec{u}_h(t) - u(t)\|_{L^2(0, \pi)}^2 = \sum_{j=1}^M \int_{x_j-h/2}^{x_j+h/2} |u_j(t) - u(x, t)|^2 dx + \int_0^{h/2} |u(x, t)|^2 dx + \int_{\pi-h/2}^{\pi} |u(x, t)|^2 dx. \quad (2.76)$$

Las dos últimas integrales, evidentemente, tienden a cero cuando $h \rightarrow 0$. Cada una de las otras integrales que interviene en el sumatorio puede ser estimada del siguiente modo:

$$\int_{x_j-h/2}^{x_j+h/2} |u_j(t) - u(x, t)|^2 dx \leq 2h |u_j(t) - u(x_j, t)|^2 + 2 \int_{x_j-h/2}^{x_j+h/2} |u(x, t) - u(x_j, t)|^2 dx. \quad (2.77)$$

Por tanto,

$$\begin{aligned} \sum_{j=1}^M \int_{x_j-h/2}^{x_j+h/2} |u_j(t) - u(x, t)|^2 dx &\leq 2h \sum_{j=1}^M |u_j(t) - u(x_j, t)|^2 + 2 \sum_{j=1}^M \int_{x_j-h/2}^{x_j+h/2} |u(x, t) - u(x_j, t)|^2 dx \\ &= 2 \|\vec{u}(t) - \underline{u}(t)\|_h^2 + 2 \sum_{j=1}^M \int_{x_j-h/2}^{x_j+h/2} |u(x, t) - u(x_j, t)|^2 dx \end{aligned} \quad (2.78)$$

El primero de estos dos términos tiende a cero en virtud del resultado de convergencia del Teorema 2.1.

Basta entonces analizar el último de ellos.

Para ello observamos que, por la desigualdad de Poincaré, cada uno de los términos que interviene en el sumatorio satisface:

$$\int_{x_j-h/2}^{x_j+h/2} |u(x, t) - u(x_j, t)|^2 dx \leq \frac{h^2}{\pi} \int_{x_j-h/2}^{x_j+h/2} |u_x(x, t)|^2 dx. \quad (2.79)$$

Por lo tanto,

$$\sum_{j=1}^M \int_{x_j-h/2}^{x_j+h/2} |u(x, t) - u(x_j, t)|^2 dx \leq \frac{h^2}{\pi^2} \int_0^\pi |u_x(x, t)|^2 dx, \quad (2.80)$$

que, evidentemente, tiende a cero cuando $h \rightarrow 0$ cuando $u(t) \in H_0^1(0, \pi)$. Pero esto es cierto para todo $t > 0$ a causa del efecto regularizante de la ecuación del calor, tal y como se puede observar en el desarrollo de Fourier de la solución u .

■

En el Corolario 2.1 hemos optado por la extensión constante a trozos de la solución numérica al intervalo $(0, \pi)$ pero el mismo resultado se cumple para cualquier otra extensión razonable, por ejemplo las funciones continuas y lineales a trozos.

2.3 Semi-discretización espacial: El método de la energía

Hemos probado, mediante métodos basados en el desarrollo en series de Fourier de las soluciones, la convergencia del problema semi-discreto (2.24) a la ecuación del calor. En lo sucesivo lo haremos mediante el *método de la energía*.

El método de la energía está basado en la identidad de energía (2.18) que las soluciones de la ecuación del calor satisfacen.

Para el problema semi-discreto se satisface una identidad de energía semejante. En efecto multiplicamos la j -ésima ecuación (2.24) por u_j y sumamos con respecto al índice $j = 1, \dots, M$. Obtenemos así

$$0 = \sum_{j=1}^M u_j u_j' - \frac{1}{h^2} \sum_{j=1}^M [u_{j+1} + u_{j-1} - 2u_j] u_j.$$

Observamos en primer lugar que

$$\sum_{j=1}^M u_j u_j' = \frac{1}{2} \frac{d}{dt} \left[\sum_{j=1}^M |u_j|^2 \right]$$

y, por otra parte,

$$\begin{aligned}
& \sum_{j=1}^M [u_{j+1} + u_{j-1} - 2u_j] u_j = \sum_{j=1}^M [(u_{j+1} - u_j) + (u_{j-1} - u_j)] u_j \\
&= \sum_{j=1}^M (u_{j+1} - u_j) u_j + \sum_{j=1}^M (u_{j-1} - u_j) u_j \\
&= \sum_{j=1}^M (u_{j+1} - u_j) u_j - \sum_{j=0}^{M-1} (u_{j+1} - u_j) u_{j+1} \\
&= - \sum_{j=0}^M (u_{j+1} - u_j)^2.
\end{aligned}$$

Obtenemos así la siguiente identidad de energía para el sistema semi-discreto (2.24):

$$\frac{d}{dt} \left[\frac{h}{2} \sum_{j=1}^M |u_j|^2 \right] = -h \sum_{j=0}^M \left| \frac{u_{j+1} - u_j}{h} \right|^2. \quad (2.81)$$

Las similitudes entre las identidades de energía (2.18) del modelo continuo y la identidad (2.81) del caso semi-discreto son obvias. En el término de la izquierda de (2.81) se observa la derivada temporal de una suma discreta que aproxima el cuadrado de la norma en $L^2(0, \pi)$ de la identidad (2.18). Por otro lado, en el miembro de la derecha de (2.81) encontramos una versión discreta de la norma de u_x en $L^2(0, \pi)$.

A partir de estas identidades podemos dar una demostración alternativa de la convergencia del problema semi-discreto (2.24) al continuo.

En primer lugar observamos que la solución del problema continuo $u = u(x, t)$ es una solución aproximada del problema discreto. En efecto, sea

$$\underline{u}_j(t) = u(x_j, t), \quad j = 1, \dots, M, t > 0, \quad (2.82)$$

siendo $u = u(x, t)$ la solución exacta de (2.8).

En efecto, $\{ \underline{u}_j(t) \}_{j=1, \dots, M}$ satisface

$$\left\{ \begin{array}{l} \underline{u}'_j + \frac{[2 \underline{u}_j - \underline{u}_{j+1} - \underline{u}_{j-1}]}{h^2} = u_{xx}(x_j, t) + \frac{[2 \underline{u}_j - \underline{u}_{j+1} - \underline{u}_{j-1}]}{h^2} = \varepsilon_j(t), \\ j = 1, \dots, M, t > 0 \\ \underline{u}_0 = \underline{u}_{M+1} = 0, \quad t > 0 \\ \underline{u}_j(0) = \varphi_j, \quad j = 1, \dots, M. \end{array} \right. \quad (2.83)$$

con $\varphi_j = \varphi(x_j)$. En el segundo miembro de (2.83) aparece el *residuo* o *error de truncación* asociado al método que, como se observa en su propia definición, se trata del error que se comete al considerar la solución del problema continuo como solución aproximada del problema

discreto. Conviene subrayar este hecho: las demostraciones de convergencia están basadas en el análisis de la solución continua como solución aproximada del problema discreto y no al revés.

Consideramos ahora la diferencia entre la solución continua \underline{u}_j sobre el mallado y la solución del problema semi-discreto:

$$\vec{v} = \{v_j\}_{j=1,\dots,M}; \quad v_j = \underline{u}_j - u_j. \quad (2.84)$$

En vista de (2.24) y (2.83), $\{v_j\}_{j=0,\dots,M}$ satisface

$$\begin{cases} v_j' + \frac{[2v_j - v_{j+1} - v_{j-1}]}{h^2} = \varepsilon_j, & j = 1, \dots, M, \quad t > 0 \\ v_0 = v_{M+1} = 0, & t > 0 \\ v_j(0) = 0, & j = 1, \dots, M. \end{cases} \quad (2.85)$$

El sistema (2.85) no es homogéneo y su dato inicial es nulo pues hemos supuesto que en el sistema semi-discreto el dato inicial del sistema continuo se toma de manera exacta sobre los puntos del mallado. Pero sería fácil adaptar las estimaciones que vamos a realizar al caso en que también hubiese un cierto error en los datos iniciales y, en particular, a las demás situaciones discutidas en el apartado anterior. Los términos ε_j del segundo miembro (el error de truncación) de (2.85) representan, tal y como se observa en su definición (2.83), la diferencia entre el laplaciano continuo y el discreto evaluado en la solución real u de (2.8). El método de la energía se adapta con facilidad a esta situación.

Retomamos la estimación de energía en el sistema (2.85). Multiplicando cada ecuación de (2.85) por v_j y sumando con respecto a $j = 1, \dots, M$, se obtiene

$$\frac{d}{dt} \left[\frac{h}{2} \sum_{j=1}^M |v_j|^2 \right] = -h \sum_{j=0}^M \left| \frac{v_{j+1} - v_j}{h} \right|^2 + h \sum_{j=1}^M \varepsilon_j v_j. \quad (2.86)$$

Necesitamos ahora la siguiente desigualdad elemental:

Lema 2.2 *Para todo $\delta > 0$ existe $h_0 > 0$ de modo que para todo $0 < h < h_0$ y para toda función discreta $\{a_0, a_1, \dots, a_M, a_{M+1}\}$ tal que $a_0 = a_{M+1} = 0$ se tiene*

$$h \sum_{j=0}^M \left| \frac{a_{j+1} - a_j}{h} \right|^2 \geq (1 - \delta) h \sum_{j=1}^M |a_j|^2. \quad (2.87)$$

Observación 2.3 Nótese que (2.87) es la versión discreta de la clásica *desigualdad de Poincaré* (2.21) que ya comentamos en la Observación 2.1. En (2.87) se establece la versión discreta de esta desigualdad con una constante $(1 - \delta)$ arbitrariamente próxima a la constante unidad de la desigualdad (2.21).

Tal y como mencionamos en aquella Observación, la mejor constante de la desigualdad de Poincaré venía dada por el principio de minimalidad que involucra en el cociente de Rayleigh.

La demostración del Lema está basada en analizar el correspondiente principio variacional en el caso discreto en función del paso del mallado h .

■

Demostración del Lema 2.2. Tal y como se indicó en (2.29) el mínimo autovalor de la matriz A_h definida en (2.26) es $\lambda_1(h) = 4 \operatorname{sen}^2(h/2)/h^2$. Como A_h es simétrica, λ_1 está caracterizado por

$$\frac{4}{h^2} \operatorname{sen}^2\left(\frac{h}{2}\right) = \lambda_1(h) = \min_{\vec{a} \in \mathbb{R}^M} \frac{\langle A_h \vec{a}, \vec{a} \rangle_h}{\|\vec{a}\|_h^2}. \quad (2.88)$$

Es fácil comprobar que

$$\langle A_h \vec{a}, \vec{a} \rangle_h / \|\vec{a}\|_h^2 = h \sum_{j=0}^M |(a_{j+1} - a_j)/h|^2 / h \sum_{j=1}^M |a_j|^2.$$

Deducimos por tanto que

$$h \sum_{j=0}^M \left| \frac{a_{j+1} - a_j}{h} \right|^2 \geq \operatorname{sen}^2\left(\frac{h}{2}\right) h \sum_{j=1}^M |a_j|^2,$$

para todo $h > 0$ y toda función discreta $\{a_0, \dots, a_{M+1}\}$, con $a_0 = a_{M+1} = 0$.

Basta por último observar que

$$\frac{4}{h^2} \operatorname{sen}^2\left(\frac{h}{2}\right) \longrightarrow 1, \quad h \rightarrow 0,$$

de modo que para todo $\delta > 0$ existe $h_0 > 0$ tal que

$$\frac{4}{h^2} \operatorname{sen}^2\left(\frac{h}{2}\right) \geq 1 - \delta, \quad \forall 0 < h < h_0.$$

■

Aplicando (2.87) con $\delta = 1/2$ en (2.86) obtenemos

$$\frac{d}{dt} \left[\frac{h}{2} \sum_{j=1}^M |v_j|^2 \right] \leq -\frac{h}{2} \sum_{j=1}^M |v_j|^2 + h \sum_{j=1}^M \varepsilon_j v_j \leq \frac{h}{2} \sum_{j=1}^M |\varepsilon_j|^2. \quad (2.89)$$

Por lo tanto

$$h \sum_{j=1}^M |v_j(t)|^2 \leq h \sum_{j=1}^M \int_0^T |\varepsilon_j|^2 dt, \quad \forall 0 < h < h_0, \quad 0 \leq t \leq T. \quad (2.90)$$

Basta por tanto con que estimemos el error producido por los términos ε_j (el error de truncación). Recordemos que

$$\varepsilon_j = u_{xx}(x_j, t) + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2}. \quad (2.91)$$

Como es bien sabido, el esquema en diferencias finitas de tres puntos proporciona una aproximación de orden dos del operador derivada segunda. Por lo tanto

$$|\varepsilon_j(t)| \leq C h^4 \|u(t)\|_{C^4([0,\pi])}^2, \forall j = 1, \dots, M, \forall 0 < h < h_0, \forall 0 \leq t \leq T. \quad (2.92)$$

Combinando (2.90) y (2.92) deducimos que

$$h \sum_{j=1}^M |v_j(t)|^2 \leq C h^2 \int_0^T \|u(t)\|_{C^4([0,\pi])}^2 dt. \quad (2.93)$$

Hemos por tanto probado el siguiente resultado.

Teorema 2.2 *Supongamos que el dato inicial $\varphi = \varphi(x)$ es tal que la solución $u = u(x, t)$ de la ecuación del calor (2.8) verifica, para todo $T < \infty$,*

$$\int_0^T \|u(t)\|_{C^4([0,\pi])}^2 dt < \infty. \quad (2.94)$$

Entonces, para todo $0 < T < \infty$ existe una constante $C_T > 0$ tal que

$$h \sum_{j=1}^M |\underline{u}_j(t) - u_j(t)|^2 = \|\underline{\tilde{u}}(t) - \vec{u}_h(t)\|_h^2 \leq C_T h^4, \quad (2.95)$$

para todo $0 \leq t \leq T$ y para todo $h > 0$, donde $\underline{\tilde{u}} = \{\underline{u}_j\}_{j=1,\dots,M}$ denota la restricción a los puntos del mallado de la solución de la ecuación del calor (2.8) y $\vec{u}_h = \{u_j\}_{j=1,\dots,M}$ representa la solución del sistema semi-discreto (2.24).

Observación 2.4 En (2.95) hemos establecido que el sistema semi-discreto (2.24) proporciona una estimación de orden dos de la ecuación del calor (2.8). Este resultado cabía ser esperado pues la única discretización que ha sido realizada es la del laplaciano espacial, al ser sustituido por el esquema de tres puntos que, como es bien sabido, es una aproximación de orden dos.

El Teorema 2.2 ha sido establecido mediante el método de energía que ha sido aplicado en una de sus versiones más simples. Muchas otras variantes son posibles. En realidad, por cada estimación de energía de la que dispongamos para la ecuación del calor (2.8) se puede establecer un resultado de convergencia distinto que, bajo hipótesis de regularidad adecuadas sobre la solución de la ecuación del calor, confirmará que el método semi-discreto es de segundo orden. Recordemos que la estimación de energía (2.18) en la que nos hemos inspirado para probar el Teorema 2.2 se obtiene multiplicando la ecuación del calor (2.8) por u . Si en lugar de multiplicar por u , multiplicamos la ecuación por $\partial^{2m}u/\partial x^{2m}$, es decir por una derivada espacial de u de orden par, se obtiene una nueva identidad de energía de la forma

$$\frac{d}{dt} \left[\frac{1}{2} \int_0^\pi \left| \frac{\partial^m u}{\partial x^m} \right|^2 dx \right] = - \int_0^\pi \left| \frac{\partial^{m+1} u}{\partial x^{m+1}} \right|^2 dx.$$

Como hemos mencionado anteriormente, esta identidad también tiene su análogo semi-discreto sobre el que podría establecerse un resultado del tipo del Teorema 2.2.

Los comentarios realizados en el Teorema 2.1 son también aplicables en el Teorema 2.2. Por ejemplo, si bien en (2.95) hemos establecido una estimación en norma L^2 , también el método de la energía nos habría permitido probar estimaciones en otras normas, por ejemplo, en la norma del máximo.

En el Teorema 2.2 hemos supuesto que la solución u de la ecuación del calor es suficientemente regular como para que (2.94) se satisfaga, i.e. que

$$u \in L^1(0, T; C^4([0, \pi])).$$

Esto, evidentemente, exige que el dato inicial sea también suficientemente regular. Bastaría por ejemplo con que el dato inicial fuese de clase C^3 , aunque esta hipótesis se podría debilitar.

Si comparamos el Teorema 2.1 y el Teorema 2.2 se observa inmediatamente que si bien en el primero, usando series de Fourier, obteníamos un resultado de convergencia bajo condiciones mínimas sobre el dato inicial ($\varphi \in L^2(0, \pi)$), en el método de la energía hemos utilizado hipótesis más fuertes sobre el dato pero, como contrapartida, hemos obtenido un resultado más fuerte puesto que hemos probado que el método es de orden dos. El método de Fourier también permite obtener ordenes de convergencia pero, tal y como se señaló en la Observación 2.2, eso exige hipótesis adicionales sobre el dato inicial, también en este caso.

Por consiguiente, el tipo de resultados que se puede obtener por ambos métodos es semejante, si bien el método de la energía es más flexible pues se puede aplicar en situaciones en las que, por la presencia de coeficientes que dependen del tiempo o de no-linealidades, las soluciones no pueden descomponerse en series de Fourier mediante el método de separación de variables.

■

2.4 Consistencia + estabilidad = Convergencia

Es habitual que en los textos dedicados al Análisis Numérico de EDP se incluya un Teorema, habitualmente atribuido a P. Lax, que garantiza que

$$\textit{Consistencia} + \textit{Estabilidad} = \textit{Convergencia}.$$

Sin embargo, tanto al interpretar el concepto de estabilidad como el de convergencia y hacer uso de este resultado, lo mismo que ocurre en el marco de las EDP, nos enfrentamos a genuinos problemas de dimensión infinita y la elección que se hace de las normas es por tanto fundamental. Así, estos tres conceptos han de ser manipulados en un mismo contexto, una vez establecidas con claridad las normas en las que trabajamos, lo cual, en realidad, consiste en determinar el criterio o distancia en la que se va a comprobar la convergencia del método.

Es por eso que, en estas notas, en lugar de incluir este enunciado como Teorema, lo presentamos simplemente como un principio general, de validez universal una vez que se han elegido con prudencia las normas, pero que conviene también utilizar con cuidado en cada caso. Es decir,

para cada EDP y cada aproximación numérica habrá que elegir de manera precisa la norma en la que se va a medir la convergencia del método.

Lo más habitual es utilizar este principio general con el objeto de probar la convergencia. De este modo el problema se reduce a verificar dos propiedades básicas del esquema: su consistencia y estabilidad.

En las secciones anteriores estos dos conceptos han surgido ya y el principio general ha sido implícitamente utilizado en la demostración de los dos Teoremas de convergencia. El objeto de esta sección es comentar y clarificar el modo en que estos conceptos han intervenido y se han combinado en las pruebas de convergencia y de este modo ilustrar este principio general fundamental del Análisis Numérico de las EDP.

Si bien estos conceptos y principios están también presentes en la prueba del Teorema 2.1 realizada mediante desarrollos en series de Fourier, son tal vez más claros en la demostración del Teorema 2.2 realizada mediante el método de la energía. Nos centraremos pues en este segundo caso, dejando para el lector la reflexión sobre la prueba del primer resultado mediante el método de Fourier que comentaremos muy brevemente al final de la sección.

- *Consistencia:*

La consistencia del método numérico hace referencia a su coherencia a la hora de aproximar la EDP. Se trata simplemente de comprobar si el esquema numérico utilizado es un esquema razonable para aproximar la EDP en cuestión o, si por el contrario, corre el riesgo de aproximar a otra EDP. Lo más habitual es comprobar la consistencia mediante el desarrollo de Taylor. El problema se reduce entonces a verificar si, cuando el paso del mallado tiende a cero, (en el caso que nos ocupa: $h \rightarrow 0$), las soluciones regulares del problema continuo son soluciones aproximadas del problema discreto en la medida en que el error de truncación tiende a cero. Cuando ésto es así con un error del orden de una potencia p del tamaño del mallado se dice que el método es de orden p .

En nuestro caso particular, como (2.24) es una aproximación semi-discreta de (2.8) en la que la variable tiempo no ha sido discretizada, la propiedad de consistencia del esquema se reduce meramente a la consistencia de la aproximación de tres puntos del operador de derivada segunda en espacio que, como sabemos y recordamos en (2.91) y (2.92), es efectivamente consistente de orden dos.

Conviene subrayar que, aunque pueda resultar paradójico, a la hora de comprobar la consistencia no verificamos hasta qué punto las soluciones del problema discreto son soluciones del problema continuo módulo un cierto error, sino que hacemos precisamente lo contrario: comprobamos si la solución del problema continuo es una solución aproximada del problema discreto.

Evidentemente ésto se hace exclusivamente a la hora de verificar la bondad del método numérico a priori puesto que, en la práctica, no disponemos de la solución del problema continuo: el objeto del cálculo numérico es precisamente aproximar la solución real de la EDP.

■

- *Estabilidad:*

Pero la consistencia por sí sola no basta para garantizar la convergencia del método. Es preciso analizar su estabilidad.

La propiedad de estabilidad consiste en asegurarse de que los esquemas discretos o semi-discretos, en su evolución temporal (discreta o continua), no amplifiquen los errores iniciales o, al menos, no lo hagan de manera creciente y descontrolada a medida que el paso del mallado tiende a cero.

En el marco del esquema (2.24) esta propiedad de estabilidad queda debidamente recogida en (2.81) donde hemos probado que, para cualquier $h > 0$, la norma L^2 de las soluciones discretas (la norma $\|\cdot\|_h$) decrece con el tiempo. Se trata pues de una situación ideal en la que los errores iniciales no sólo no se amplifican en exceso sino que decrecen en tiempo.

Si analizamos detenidamente la prueba del Teorema 2.2 mediante el método de la energía observaremos que ésta no es más que una combinación adecuada de las propiedades de consistencia y estabilidad y que por tanto obedece fielmente al principio de P. Lax antes mencionado. En efecto, con el objeto de probar la convergencia, hemos establecido en primer lugar que la solución del problema continuo es una solución aproximada del problema discreto (véase (2.83)). Es decir, hemos empezado usando la consistencia del método. Después, usando la linealidad del esquema discreto, hemos introducido la variable \vec{v} diferencia entre la solución del problema continuo y discreto, y hemos establecido que se trata de una solución aproximada del problema discreto en el que el dato inicial es nulo pero la dinámica está forzada por un segundo miembro $(\varepsilon_j, j = 1, \dots, M)$ que tiende a cero cuando $h \rightarrow 0$ (véase (2.85)). Finalmente, usando la estabilidad, hemos establecido que esta diferencia se mantiene pequeña cuando el tiempo avanza y tiende a cero cuando $h \rightarrow 0$.

Puede resultar paradójico que el concepto de estabilidad haga referencia a la propagación de errores en el dato inicial y que, sin embargo, se aplique en el sistema (2.85) donde el dato inicial es exactamente cero pero en el que está presente una fuerza externa $\vec{\varepsilon}(t)$ continua en tiempo. Pero, en vista del principio de Duhamel o de la fórmula de variación de las constantes, esto no debería de resultar extraño pues el efecto de un segundo miembro continuo en una ecuación de evolución no es otro que el de la integral temporal de los efectos que ese segundo miembro tendría, en cada instante de tiempo, como perturbación del dato inicial.

Este hecho queda claramente reflejado en la fórmula de representación de la solución de ecuaciones diferenciales no homogéneas:

$$\begin{cases} \dot{x}(t) = Ax(t) + F(t), & t > 0 \\ x(0) = x_0. \end{cases}$$

La fórmula de variación de las constantes asegura en este caso que

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}F(s)ds.$$

En esta fórmula se observa con claridad que el efecto de un segundo miembro en la ecuación es una integral temporal de los efectos que este segundo miembro tendría en el dato inicial. ■

Hemos descrito cómo los conceptos de consistencia y estabilidad han jugado un papel esencial en la prueba del resultado de convergencia del Teorema 2.2. Estos conceptos están también presentes en la prueba realizada del Teorema 2.1 mediante el método de Fourier. En efecto, la consistencia del esquema garantiza su convergencia para las soluciones que involucran un solo modo de Fourier, mientras que la estabilidad asegura que basta con probar la convergencia para datos iniciales que involucran únicamente un número finito de componentes. Nuevamente, la consistencia junto con la estabilidad proporcionan la convergencia del método.

Conviene sin embargo subrayar que los sistemas de EDO que obtenemos al realizar discretizaciones espaciales tienen un carácter *stiff*. Para ello basta observar que el ratio $\lambda_M(h)/\lambda_1(h)$ entre el máximo y mínimo autovalor de la matriz A_h es del orden de $1/h^2$. En virtud del análisis de los sistemas *stiff* realizados en capítulos previos, se trata de un hecho relevante que habremos de tener en cuenta a la hora de proceder a la discretización temporal del sistema, con el objeto de garantizar la convergencia de las discretizaciones completas espacio-tiempo.

En la sección 2.6 veremos algunos ejemplos de métodos semi-discretos y completamente discretos que, siendo consistentes, no son estables y, por tanto, no convergen. Este hecho confirma la necesidad tanto de la propiedad de consistencia como de estabilidad de un método numérico para garantizar su convergencia.

2.5 Aproximaciones completamente discretas

En las secciones 2.2 y 2.3 hemos analizado la convergencia del esquema semi-discreto (2.24) al modelo continuo (2.8). Los resultados de convergencia que hemos probado han legitimado la utilización del esquema (2.24) puesto que sus soluciones convergen a las de la ecuación del calor (2.8) cuando $h \rightarrow 0$. Desde un punto de vista práctico, estos resultados permiten concentrar nuestros esfuerzos en el cálculo de las soluciones (2.24) con h suficientemente pequeño pero fijo.

El sistema (2.24) es un sistema de M ecuaciones diferenciales acopladas con M incógnitas. Parece por tanto natural utilizar los métodos numéricos desarrollados para la resolución aproximada de ecuaciones diferenciales. Al hacerlo discretizando la variable temporal, acabamos obteniendo esquemas completamente discretos para la resolución de (2.8).

Esta sección está destinada al estudio de las dos aproximaciones completamente discretas más naturales en este contexto que son las que se obtienen al aplicar el método explícito e implícito de Euler para la resolución aproximada de EDO de primer orden. Por supuesto, la variedad de los métodos posibles es muy grande puesto que cualquier método de aproximación

del laplaciano (de la segunda derivada espacial: esquema en diferencias finitas de orden superior, método espectral o de elementos finitos, ...) junto con cualquier método de aproximación de la derivada temporal (método del trapecio, de Runge-Kutta, o multi-paso, por ejemplo) dan lugar a un método completamente discreto nuevo. Pero el análisis de estos dos métodos más básicos permite estudiar los aspectos fundamentales de la teoría que bastaría para analizar cualquier otro tipo de esquemas.

En esta ocasión el paso espacial será denotado mediante Δx (en lugar de h), mientras que el paso temporal estará denotado por Δt . Mediante u_j^k denotaremos la aproximación de la solución u de (2.8) en el punto $x = x_j = j\Delta x$ en el instante de tiempo $t = t_k = k\Delta t$.

Con esta nueva notación, si en el sistema semi-discreto (2.24) aplicamos el *esquema explícito de Euler* en la variable temporal obtenemos el sistema:

$$\begin{cases} \frac{u_j^{k+1} - u_j^k}{\Delta t} + \frac{[2u_j^k - u_{j+1}^k - u_{j-1}^k]}{(\Delta x)^2} = 0, & j = 1, \dots, M; \quad k \geq 0, \\ u_0^k = u_{M+1}^k = 0, & k \geq 0, \\ u_j^0 = \varphi_j, & j = 1, \dots, M. \end{cases} \quad (2.96)$$

Sin embargo, si aplicamos el método de Euler implícito obtenemos

$$\begin{cases} \frac{u_j^{k+1} - u_j^k}{\Delta t} + \frac{[2u_j^{k+1} - u_{j+1}^{k+1} - u_{j-1}^{k+1}]}{(\Delta x)^2} = 0, & j = 1, \dots, M; \quad k \geq 0, \\ u_0^k = u_{M+1}^k = 0, & k \geq 0, \\ u_j^0 = \varphi_j, & j = 1, \dots, M. \end{cases} \quad (2.97)$$

La diferencia principal entre (2.96) y (2.97) es la habitual entre esquemas explícitos e implícitos. Mientras que (2.97) permite “leer” explícitamente el valor de la solución discreta en el paso temporal $k + 1$ a partir de la solución en el paso temporal k , el método implícito (2.97) exige, en cada paso temporal, la resolución de un sistema tridiagonal de M ecuaciones lineales con M incógnitas.

Con el objeto de estudiar la convergencia de estos métodos conviene utilizar el *número de Courant*

$$\mu = \Delta t / (\Delta x)^2, \quad (2.98)$$

en el que queda claramente reflejada la diferencia de homogeneidad en la variable espacial y temporal del operador en derivadas parciales involucrado en la ecuación del calor, en la que una derivada temporal juega el mismo papel que dos derivadas espaciales.

Con esta nueva notación los esquemas (2.96) y (2.97) pueden reescribirse del modo siguiente:

- *Método de Euler explícito:*

$$u_j^{k+1} = u_j^k + \mu [u_{j+1}^k + u_{j-1}^k - 2u_j^k]. \quad (2.99)$$

- *Método de Euler implícito:*

$$u_j^{k+1} + \mu \left[2u_j^{k+1} - u_{j+1}^{k+1} - u_{j-1}^{k+1} \right] = u_j^k. \quad (2.100)$$

La consistencia de ambos métodos está claramente garantizada.

En efecto, hemos utilizado el esquema de tres puntos para la aproximación de la segunda derivada espacial que es consistente de orden dos, mientras que la discretización de la derivada temporal mediante la diferencia finita que involucra dos niveles temporales (k y $k+1$) proporciona una aproximación consistente de orden uno en tiempo.

El análisis de la convergencia de los métodos se hará a valor de μ fijo. En otras palabras, describiremos el rango de valores del parámetro de Courant μ para el que los métodos discretos en consideración son convergentes. Fijado este valor del parámetro de Courant tenemos $\Delta t = \mu(\Delta x)^2$. Por tanto un orden de consistencia temporal corresponde a un orden dos de consistencia espacial. Es por eso que diremos que ambos métodos (2.99) y (2.100) son consistentes de orden dos (teniendo como referencia el paso espacial). De manera más precisa, si, dada una solución u del problema continuo original (2.8), introducimos la proyección de dicha solución sobre el mallado

$$\underline{u}_j^k = u(x_j, t_k) = u(j\Delta x, k\Delta t), \quad (2.101)$$

entonces \underline{u} es solución aproximada de los esquemas discretos. En efecto se tiene

$$\underline{u}_j^{k+1} - \left[\underline{u}_j^k + \mu \left[\underline{u}_{j+1}^k + \underline{u}_{j-1}^k - 2\underline{u}_j^k \right] \right] = O \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \quad (2.102)$$

y

$$\underline{u}_j^{k+1} - \underline{u}_j^k + \mu \left[2 \underline{u}_j^{k+1} - \underline{u}_{j+1}^{k+1} - \underline{u}_{j-1}^{k+1} \right] = O \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \quad (2.103)$$

que corresponde efectivamente a la consistencia de orden dos de los métodos siempre y cuando u tenga dos derivadas temporales y cuatro derivadas espaciales acotadas.

Es natural que obtengamos el mismo tipo de hipótesis sobre dos derivadas temporales y cuatro derivadas espaciales de la solución de la ecuación del calor. En efecto, la ecuación garantiza que $u_t = u_{xx}$, lo cual implica también que $u_{tt} = u_{xxxx}$.

Conviene observar la presencia de un término adicional Δt en la estimación del error de truncación, con respecto al $(\Delta x)^2$ propio de la aproximación espacial y al Δt de la aproximación temporal puesto que (2.96) y (2.97) han sido ambas multiplicadas por Δt para obtener (2.99) y (2.100).

Pero, lo mismo que ocurría en el caso de los esquemas semi-discretos, la consistencia no basta para garantizar la convergencia del método. De hecho estos dos esquemas tienen un comportamiento bien distinto en relación con el rango de valores de μ para los que se tiene la convergencia:

Teorema 2.3 *El método explícito (2.96) es convergente si $0 < \mu \leq 1/2$ mientras que el método implícito (2.97) lo es para todo $\mu > 0$.*

En ambos casos, los métodos son convergentes de orden dos, lo que significa que, si la solución de (2.8) es suficientemente regular, se tiene, para todo $0 < T < \infty$,

$$\max_{k=0, \dots, [T/\Delta t]} \left\| \vec{u}^k - \underline{\vec{u}}^k \right\|_{\Delta x} = O\left((\Delta x)^2\right) \quad (2.104)$$

cuando $\Delta x \rightarrow 0$.

Observación 2.5 En (2.104) mediante $\| \cdot \|_{\Delta x}$ denotamos la norma L^2 en el mallado de paso Δx , i.e.

$$\| a \|_{\Delta x} = \left[\Delta x \sum_{j=1}^M |a_j|^2 \right]^{1/2},$$

de modo que

$$\left\| \vec{u}^k - \underline{\vec{u}}^k \right\|_{\Delta x} = \left[\Delta x \sum_{j=1}^M \left| u_j^k - \underline{u}_j^k \right|^2 \right]^{1/2}.$$

Conviene también observar que \vec{u}^k (resp. $\underline{\vec{u}}^k$) denota el vector de componentes $u_j^k, j = 1, \dots, M$, (resp. $\underline{u}_j^k, j = 1, \dots, M$) que, a su vez, representa la solución del problema discreto (resp. continuo) en el paso temporal k .

En (2.104) estimamos la norma de la diferencia entre la solución continua y discreta para los pasos $k = 0, 1, \dots, [T/\Delta t]$, que son los necesarios para cubrir el intervalo temporal $[0, T]$.

Por otra parte, en (2.104), se enuncia la convergencia a cero del error cuando $\Delta x \rightarrow 0$. Evidentemente, aunque no se diga explícitamente, también $\Delta t \rightarrow 0$ puesto que en el enunciado se supone que el parámetro μ de Courant está fijado dentro del rango en el que se tiene la convergencia ($\mu \in (0, 1/2)$ para el método explícito y $\mu \in (0, \infty)$ para el implícito).

En virtud de este resultado de convergencia se observa la superioridad del método implícito frente al explícito puesto que su convergencia está garantizada para un valor arbitrario del parámetro de Courant μ . Esto permite tomar pasos temporales más grandes que para el método explícito. ■

Demostración del Teorema 2.3.

En la demostración supondremos que la solución u de la ecuación continua (2.8) es de clase C^2 en tiempo y C^4 en espacio. Esto garantiza que las identidades (2.102) y (2.103) sean válidas, uniformemente en $k = 0, \dots, [T/\Delta t]$ y $j = 1, \dots, M$, a medida que $\Delta t, \Delta x \rightarrow 0$.

Introducimos ahora el error

$$e_j^k = \underline{u}_j^k - u_j^k \quad (2.105)$$

que mide la diferencia entre las soluciones del problema continuo y del problema discreto.

El error e_j^k es solución del siguiente sistema dinámico discreto:

• *Método explícito:*

$$\begin{aligned} e_j^{k+1} &= e_j^k + \mu \left[e_{j+1}^k + e_{j-1}^k - 2e_j^k \right] + O \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \\ &= (1 - 2\mu)e_j^k + \mu \left(e_{j+1}^k + e_{j-1}^k \right) + O \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right). \end{aligned} \quad (2.106)$$

• *Método implícito:*

$$\begin{aligned} e_j^{k+1} + \mu \left[2e_j^{k+1} - e_{j+1}^{k+1} - e_{j-1}^{k+1} \right] &= (1 + 2\mu)e_j^{k+1} - \mu \left(e_{j+1}^{k+1} + e_{j-1}^{k+1} \right) \\ &= e_j^k + O \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right). \end{aligned} \quad (2.107)$$

En la medida en que hemos supuesto que el dato inicial del problema discreto coincide exactamente con el del problema continuo tenemos

$$e_j^0 = 0, \quad j = 1, \dots, M. \quad (2.108)$$

Procedemos ahora a estimar el error distinguiendo los métodos explícito e implícito:

• *Método explícito*

A partir de (2.106) observamos que si

$$\varepsilon^k = \max_{j=1, \dots, M} \left| e_j^k \right|, \quad (2.109)$$

se tiene

$$\varepsilon^{k+1} \leq [|1 - 2\mu| + 2\mu] \varepsilon^k + C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right). \quad (2.110)$$

Iterando esta desigualdad y haciendo uso de (2.108) obtenemos

$$\varepsilon^k \leq C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \left[1 + \alpha_\mu + \alpha_\mu^2 + \dots + \alpha_\mu^{k-1} \right] \quad (2.111)$$

donde

$$\alpha_\mu = [|1 - 2\mu| + 2\mu]. \quad (2.112)$$

Claramente hay dos casos que distinguir: (i) $\mu \leq 1/2$; y (ii) $\mu > 1/2$. En el primero

$$\alpha_\mu = 1$$

y por tanto (2.111) se reduce a

$$\varepsilon^k \leq C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) k \leq C_T \left((\Delta x)^2 + \Delta t \right) = O \left((\Delta x)^2 \right) \quad (2.113)$$

puesto que $k\Delta t \leq T$ y que $\Delta t = \mu(\Delta x)^2$. La desigualdad (2.113) proporciona el resultado de convergencia enunciado en el Teorema para $\mu \leq 1/2$.

La situación es distinta cuando $\mu > 1/2$. En este caso

$$\alpha_\mu = 4\mu - 1 > 1$$

y por tanto el miembro de la derecha de (2.111) diverge exponencialmente. En efecto, en el último paso temporal en el que $k \sim T/\Delta t$ tenemos que

$$\begin{aligned} & C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \left(1 + \alpha_\mu + \dots + \alpha_\mu^{k-1} \right) \\ & \geq C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \alpha_\mu^{k-1} \geq C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) e^{\beta_\mu T/\Delta t} \\ & \rightarrow \infty, \text{ cuando } \Delta x \rightarrow 0 \text{ con } \Delta t = \mu(\Delta x)^2. \end{aligned} \quad (2.114)$$

Esta estimación del error no permite concluir la convergencia del método. En realidad en este caso el método es inestable y por tanto no converge. Este último aspecto será analizado con más en detalle en la próxima sección mediante el método de von Neumann.

- *Método implícito*

En este caso, de (2.107), con independencia del valor de $\mu > 0$ se deduce que

$$\varepsilon^{k+1} \leq \varepsilon^k + C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right). \quad (2.115)$$

En efecto, para comprobar este hecho basta con analizar la desigualdad (2.107) para el valor j^* de j para el que

$$\left| e_{j^*}^{k+1} \right| = \varepsilon^{k+1} = \max_{j=1, \dots, M} \left| e_j^{k+1} \right|.$$

De (2.107) con $j = j^*$ deducimos que

$$\begin{aligned} \varepsilon^{k+1} &= (1 + 2\mu)\varepsilon^{k+1} - 2\mu\varepsilon^{k+1} \\ &\leq (1 + 2\mu) \left| e_{j^*}^{k+1} \right| - \mu \left(\left| e_{j^*+1}^{k+1} \right| + \left| e_{j^*-1}^{k+1} \right| \right) \\ &\leq \left| (1 + 2\mu)e_{j^*}^{k+1} - \mu \left(e_{j^*+1}^{k+1} + e_{j^*-1}^{k+1} \right) \right| \\ &\leq \left| e_{j^*}^k \right| + C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right) \\ &\leq \varepsilon^k + C \left(\Delta t \left((\Delta x)^2 + \Delta t \right) \right). \end{aligned}$$

En este argumento el valor de j^* puede depender de k pero esto es irrelevante pues finalmente acabamos estableciendo una relación entre ε^k y ε^{k+1} que en nada depende del valor de j^* .

Por (2.115) observamos que, en el caso del método implícito, con independencia del valor de $\mu > 0$, nos encontramos exactamente en las mismas condiciones que en el método explícito cuando $\mu \in (0, 1/2)$. La convergencia del método implícito y, en particular, (2.104) quedan por tanto probados. ■

La demostración del resultado de convergencia ha sido realizada trabajando directamente y de manera elemental en las ecuaciones discretas (2.109) y (2.110) que describen cómo el error numérico se propaga. El mismo tipo de resultados podría haberse obtenido utilizando desarrollos en serie de Fourier. Este método será desarrollado con más detalle en la próxima sección donde describiremos el método de von Neumann para el estudio de la estabilidad y convergencia de métodos numéricos para problemas en toda la recta real o en un intervalo acotado con condiciones de contorno periódicas.

Pero, comentemos brevemente cómo el método de la sección 2.2 puede ser adaptado a este contexto de las aproximaciones completamente discretas, lo cual permite, en particular, comprobar el comportamiento patológico del método explícito cuando $\mu > 1/2$. En efecto, en (2.10) dimos ya el desarrollo en serie de Fourier de la solución de la ecuación del calor continua (2.8). Procedemos ahora al desarrollo de la solución del problema discreto explícito, en la base de los autovectores de la matriz $A_{\Delta x}$ ($= A_h$ definida en (2.26) con $h = \Delta x$). Tenemos que

$$\vec{u}^k = \sum_{\ell=1}^M p_{\ell}^k \vec{W}_{\ell}(\Delta x), \quad (2.116)$$

donde $\vec{W}_{\ell}(\Delta x)$ denota el ℓ -ésimo autovector de la matriz $A_{\Delta x}$ y p_{ℓ}^k el coeficiente de Fourier de este ℓ -ésimo autovector en el paso temporal k . Teniendo en cuenta que

$$A_{\Delta x} \vec{W}_{\ell}(\Delta x) = \lambda_{\ell}(\Delta x) \vec{W}_{\ell}(\Delta x), \quad (2.117)$$

la función discreta definida en (2.116) satisface la ecuación discreta (2.96) o (2.99) si y sólo si

$$p_{\ell}^{k+1} = p_{\ell}^k - \mu \lambda_{\ell}(\Delta x) (\Delta x)^2 p_{\ell}^k = [1 - \mu (\Delta x)^2 \lambda_{\ell}(\Delta x)] p_{\ell}^k. \quad (2.118)$$

Por lo tanto

$$p_{\ell}^k = [1 - \mu (\Delta x)^2 \lambda_{\ell}(\Delta x)]^k p_{\ell}^0 = (\alpha_{\ell})^k p_{\ell}^0. \quad (2.119)$$

Es decir, cada componente de Fourier de la solución del problema discreto evoluciona de manera exponencial según la ley (2.119). Evidentemente, el comportamiento de p_{ℓ}^k , i.e. su evolución a medida que k aumenta, depende de que $|\alpha_{\ell}| \leq 1$ o $|\alpha_{\ell}| > 1$. Analicemos pues el valor de $|\alpha_{\ell}|$. Tenemos

$$\alpha_{\ell} = 1 - \mu (\Delta x)^2 \lambda_{\ell}(\Delta x).$$

Por otra parte,

$$\lambda_{\ell}(\Delta x) = \frac{4}{(\Delta x)^2} \operatorname{sen}^2 \left(\frac{\ell \Delta x}{2} \right).$$

Por tanto,

$$\alpha_{\ell} = 1 - 4\mu \operatorname{sen}^2 \left(\frac{\ell \Delta x}{2} \right).$$

Cuando $0 \leq \mu \leq 1/2$, $|\alpha_{\ell}| \leq 1$ para todos los valores de ℓ . Esto asegura la estabilidad del método puesto que todas las componentes de Fourier de la solución discreta se mantienen acotadas en k , con independencia del valor Δx .

La situación es completamente distinta cuando $\mu > 1/2$. En efecto, en este caso, cuando $\ell = M = \frac{\pi}{\Delta x} - 1$, tenemos

$$\alpha_\ell = 1 - 4\mu \sin^2\left(\frac{\pi}{2} - \frac{\Delta x}{2}\right) = 1 - 4\mu \cos^2\left(\frac{\Delta x}{2}\right) \rightarrow 1 - 4\mu, \text{ cuando } \Delta x \rightarrow 0.$$

Obviamente $|1 - 4\mu| > 1$ cuando $\mu > 1/2$. Vemos por tanto que, cuando $\mu > 1/2$, para Δx suficientemente pequeño, existen índices ℓ para los que $|\alpha_\ell| > 1$, lo cual implica la inestabilidad del método.

El método explícito (2.96) (o (2.99)) con $\mu > 1/2$ es por tanto un primer y buen ejemplo de método numérico que, a pesar de ser consistente, no es convergente por no ser estable.

Esto no ocurre en el método implícito (2.97) (o (2.100)) en el que la ley de evolución de los coeficientes de Fourier es de la forma:

$$p_\ell^{k+1} + \mu(\Delta x)^2 \lambda_\ell(\Delta x) p_\ell^{k+1} = p_\ell^k,$$

es decir

$$p_\ell^k = (1 + \mu(\Delta x)^2 \lambda_\ell(\Delta x))^{-k} p_\ell^0.$$

En este caso, evidentemente, la estabilidad queda garantizada para todo valor del número de Courant $\mu > 0$ puesto que

$$\left| (1 + \mu(\Delta x)^2 \lambda_\ell(\Delta x))^{-1} \right| = \frac{1}{1 + \mu(\Delta x)^2 \lambda_\ell(\Delta x)} < 1,$$

para todo $\Delta x > 0$ y todo $\ell = 1, \dots, M$.

No sería difícil desarrollar este método de Fourier y completarlo con las técnicas desarrolladas en la sección 2.2 para dar una demostración alternativa del Teorema 2.3.

Conviene también subrayar que la inestabilidad que hemos detectado en el método explícito cuando $\mu > 1/2$ y, en particular, la interpretación que acabamos de hacer mediante el desarrollo en series de Fourier se asemeja en gran medida al tipo de análisis que hacíamos en el contexto de los problemas “stiff”. En efecto, en aquella ocasión introdujimos el concepto A -estabilidad en el que exigíamos que el esquema numérico a $h > 0$ fijo reprodujese, al avanzar el tiempo, las propiedades de estabilidad de la ecuación diferencial continua. En el caso que nos ocupa la ecuación continua en cuestión es el sistema semi-discreto (2.24) (o (2.27)) en el que sabemos que, a h fijo, todas las soluciones tienden a cero exponencialmente cuando $t \rightarrow \infty$ (para comprobarlo basta observar el desarrollo (2.41) en serie de Fourier de la solución del problema semi-discreto). Cuando $\mu > 1/2$ hemos comprobado que el esquema numérico explícito deja de reproducir este comportamiento asintótico puesto que surgen componentes de Fourier que se amplifican exponencialmente cuando $k \rightarrow \infty$.

A pesar de ello, hemos visto que el esquema explícito converge cuando $0 \leq \mu \leq 1/2$. Esto, sin embargo, no está en contradicción con el resultado que asegura la falta de A -estabilidad de los métodos de Runge-Kutta explícitos (el método de Euler explícito es efectivamente un método de Runge-Kutte explícito). En efecto, en este caso estamos analizando el sistema (2.24)

(o (2.27)) en el que la matriz involucrada, $A_{\Delta x}$, si bien es sumamente dispersa pues su autovalor mínimo es $\lambda_1(\Delta x) = 4 \operatorname{sen}^2(\Delta x/2)/(\Delta x)^2$ (muy próximo a 1 cuando $\Delta x \sim 0$) y el máximo es sin embargo $\lambda_M(\Delta x) = 4 \operatorname{cos}^2(\Delta x/2)/(\Delta x)^2$ (muy próximo a $4/(\Delta x)^2$ cuando $\Delta x \sim 0$), su grado de dispersión está limitado. Esto queda de manifiesto en el hecho que

$$(\Delta x)^2 \lambda_l(\Delta x) \leq 4, \forall \Delta x > 0, \forall j = 1, \dots, M$$

y es lo que permite garantizar la convergencia del método explícito, pero sólo en el rango $0 < \mu \leq 1/2$.

Vemos pues que el mal comportamiento del esquema explícito responde exactamente a las mismas patologías habituales de los métodos explícitos a la hora de abordar los problemas stiff, si bien, en el caso particular de la ecuación del calor que nos ocupa, ésto no es incompatible con la convergencia del método para $0 < \mu \leq 1/2$.

Sin embargo, esta limitación sobre μ obliga a tomar pasos temporales muy pequeños, lo cual hace que el coste computacional de la implementación de este método sea muy elevado.

Conviene por último observar que los métodos semi-discretos pueden verse como límite cuando $\mu \rightarrow 0$ de los métodos discretos. En efecto, si fijado Δx pasamos al límite de un esquema completamente discreto cuando $\Delta t \rightarrow 0$ recuperamos un esquema semi-discreto. Esto corresponde a considerar el número de Courant $\mu = 0$. El hecho que el método explícito sea convergente para $0 < \mu \leq 1/2$, a pesar de no serlo para $\mu > 1/2$, es pues perfectamente compatible (y lo explica) con el hecho de que el método semi-discreto (2.24) sea convergente, tal y como vimos en las secciones 2.2 y 2.3.

2.6 El análisis de von Neumann

En esta sección presentamos el análisis de von Neumann para el estudio de la estabilidad de esquemas numéricos, que es especialmente útil en problemas definidos en todo el espacio con mallados regulares. El método, basado en la utilización de la transformada discreta de Fourier, es muy semejante al desarrollado en la sección 2.2 y 2.5 mediante el desarrollo en series de Fourier.

Con el objeto de ilustrar las ideas básicas de este método consideramos el problema de Cauchy para la ecuación del calor en toda la recta real:

$$\begin{cases} u_t - u_{xx} = 0, & x \in \mathbb{R}, \quad t > 0 \\ u(x, 0) = \varphi(x), & x \in \mathbb{R}. \end{cases} \quad (2.120)$$

Suponemos que

$$\varphi \in L^2(\mathbb{R}) \quad (2.121)$$

de modo que (2.120) admite una única solución

$$u \in C([0, \infty); L^2(\mathbb{R})). \quad (2.122)$$

En este caso la identidad de energía garantiza que

$$\frac{d}{dt} \left[\frac{1}{2} \int_{\mathbb{R}} u^2(x, t) dx \right] = - \int_{\mathbb{R}} |u_x|^2 dx. \quad (2.123)$$

Como indicamos en la sección 2.1, la solución de (2.120) puede representarse por convolución con la solución fundamental

$$G(x, t) = (4\pi t)^{1/2} \exp(-|x|^2/4t), \quad (2.124)$$

de modo que

$$u(x, t) = [G(\cdot, t) * \varphi](x) = (4\pi t)^{-1/2} \int_{\mathbb{R}} \exp\left(-\frac{|x-y|^2}{4t}\right) \varphi(y) dy. \quad (2.125)$$

Dado $h > 0$ introducimos la partición

$$x_j = jh, j \in \mathbb{Z} \quad (2.126)$$

y consideramos la semi-discretización

$$\begin{cases} u'_j + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} = 0, & j \in \mathbb{Z}, \quad t > 0 \\ u_j(0) = \varphi_j, & j \in \mathbb{Z}. \end{cases} \quad (2.127)$$

En este caso (2.127) es un sistema de una infinidad de ecuaciones diferenciales de primer orden acopladas de tres en tres. Son varias las maneras en las que se puede resolver el sistema (2.127). Una de las más naturales consiste en truncar el sistema y considerarlo en un rango de índices finito $[-M, M]$ con condiciones de contorno de Dirichlet:

$$\begin{cases} u'_j + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} = 0, & -M < j < M, \quad t > 0 \\ u_{-M} = u_M = 0, & t > 0 \\ u_j(0) = \varphi_j. \end{cases} \quad (2.128)$$

Es fácil comprobar que la solución de (2.128), cuando $M \rightarrow \infty$, converge a la solución de (2.127).

Pero, admitiendo que el sistema (2.127) ha sido ya resuelto (problema sobre el que volveremos más adelante) analicemos la convergencia de las soluciones de (2.127) a las de (2.120) cuando $h \rightarrow 0$.

Para el sistema (2.127) se verifica la siguiente identidad de energía

$$\frac{d}{dt} \left[\frac{h}{2} \sum_{j \in \mathbb{Z}} |u_j|^2 \right] = -h \sum_{j \in \mathbb{Z}} \left| \frac{u_{j+1} - u_j}{h} \right|^2, \quad (2.129)$$

que es claramente una versión discreta de (2.123). En (2.129) estamos implícitamente asumiendo que $\vec{u}(t) = \{u_j(t)\}_{j \in \mathbb{Z}} \in \ell^2 = \ell^2(\mathbb{Z})$ lo cual equivale a suponer que $\vec{\varphi} = \{\varphi_j\}_{j \in \mathbb{Z}} \in \ell^2$. Esta última condición se verifica puesto que el dato inicial φ del problema continuo (2.120) pertenece a $L^2(\mathbb{R})$, como indicamos en (2.121).

El esquema (2.127) es consistente de orden dos. La prueba de este hecho, basada en el desarrollo de Taylor, es independiente de que el intervalo en el que se verifique la ecuación sea acotado o no. El esquema será por tanto convergente si comprobamos que es estable. Es

en la verificación de la propiedad de estabilidad donde el método de von Neumann basado en la utilización de la transformada discreta de Fourier resulta sumamente útil. En este caso sin embargo la estabilidad es también inmediata a partir de la identidad de energía (2.129).

Pero existen otros esquemas semi-discretos sumamente naturales que, siendo consistentes, carecen de la propiedad de estabilidad y, por tanto, no son convergentes. Para observar este hecho basta considerar la siguiente familia de métodos:

$$\alpha u'_{j+1} + (1 - 2\alpha)u'_j + \alpha u'_{j-1} + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} = 0, \quad (2.130)$$

donde $0 \leq \alpha \leq 1/2$. Cuando $\alpha = 0$ recuperamos el sistema anterior que, como vimos, es convergente.

Para esta familia de métodos la consistencia es fácil de verificar. Sin embargo la estabilidad no está garantizada y, de hecho, sólo se verifica cuando $0 \leq \alpha \leq 1/4$. Como consecuencia de este hecho, el método es divergente cuando $\alpha > 1/4$. Para comprobar este hecho lo más sencillo es utilizar el método de von Neumann que introducimos en el caso completamente discreto. Volveremos sobre el ejemplo (2.130) más adelante.

Introducimos ahora los dos esquemas completamente discretos más naturales en la aproximación de la ecuación del calor, correspondientes a utilizar el método de Euler explícito e implícito para la discretización de la derivada temporal.

El *método explícito* se escribe en este caso

$$\begin{cases} \frac{u_j^{k+1} - u_j^k}{\Delta t} + \frac{[2u_j^k - u_{j+1}^k - u_{j-1}^k]}{(\Delta x)^2} = 0, & j \in \mathbb{Z}, \quad t > 0, \\ u_j^0 = \varphi_j, & j \in \mathbb{Z}, \end{cases} \quad (2.131)$$

mientras que el *método implícito* es

$$\begin{cases} \frac{u_j^{k+1} - u_j^k}{\Delta t} + \frac{[2u_j^{k+1} - u_{j+1}^{k+1} - u_{j-1}^{k+1}]}{(\Delta x)^2} = 0, & j \in \mathbb{Z}, \quad t > 0, \\ u_j^0 = \varphi_j, & j \in \mathbb{Z}. \end{cases} \quad (2.132)$$

Nuevamente u_j^k representa una aproximación de $u(j\Delta x, k\Delta t)$.

Los esquemas (2.131) y (2.132) son ambos consistentes de orden dos tal y como vimos en la sección 2.5. La solución en el paso temporal k es una sucesión $\{u_j^k\}_{j \in \mathbb{Z}}$ que denotamos de manera simplificada mediante la notación vectorial \vec{u}^k .

En el estudio de la estabilidad utilizamos la *transformada discreta de Fourier*. Así a cada sucesión

$$\vec{w} = \{w_l\}_{l \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \quad (2.133)$$

le asignamos la función continua

$$\check{w}(\theta) = \sum_{j=-\infty}^{\infty} w_j e^{-ij\theta}. \quad (2.134)$$

Recíprocamente, tenemos la fórmula de inversión

$$w_l = \frac{1}{2\pi} \int_0^{2\pi} \check{w}(\theta) e^{ij\theta} d\theta. \quad (2.135)$$

La aplicación que a \vec{w} le asocia \check{w} es una isometría de $\ell^2(\mathbb{Z})$ en $L^2(0, 2\pi)$ si dotamos a estos espacios con las normas

$$\|\vec{w}\|_{\ell^2(\mathbb{Z})} = \sum_{j=-\infty}^{\infty} |w_j|^2 \quad (2.136)$$

y

$$\|\check{w}\|_{L^2(0, \pi)} = \left[\frac{1}{2\pi} \int_0^{2\pi} |\check{w}(\theta)|^2 d\theta \right]^{1/2}. \quad (2.137)$$

Para comprobar estas propiedades basta con tener en cuenta las identidades:

$$\int_0^{2\pi} e^{ij\theta} e^{-ik\theta} d\theta = \begin{cases} 0 & \text{si } j \neq k \\ 2\pi & \text{si } j = k. \end{cases} \quad (2.138)$$

Mediante esta transformación de Fourier, a cada solución $\{\vec{u}^k\}_{k \geq 0}$ de (2.131) o (2.132) le podemos asignar funciones continuas $\{\check{u}^k(\theta)\}_{k \geq 0}$ donde

$$\check{u}^k(\theta) = \sum_{j \in \mathbb{Z}} u_j^k e^{-ij\theta}. \quad (2.139)$$

La clave en el estudio de la estabilidad de los esquemas numéricos mediante esta transformada de Fourier reside en la siguiente identidad:

$$\begin{aligned} \sum_{j \in \mathbb{Z}} a_{j+1} e^{-ij\theta} &= \sum_{j \in \mathbb{Z}} a_{j+1} e^{-i(j+1)\theta} e^{i\theta} \\ &= e^{i\theta} \sum_{j \in \mathbb{Z}} a_{j+1} e^{-i(j+1)\theta} = e^{i\theta} \sum_{j \in \mathbb{Z}} a_j e^{-ij\theta} = e^{i\theta} \check{a}(\theta). \end{aligned} \quad (2.140)$$

De modo análogo tenemos

$$\sum_{j \in \mathbb{Z}} a_{j-1} e^{-ij\theta} = e^{-i\theta} \check{a}(\theta). \quad (2.141)$$

Así, tomando la transformada de Fourier en las ecuaciones (2.131) y (2.132) obtenemos las siguientes ecuaciones para la transformada de Fourier $\check{u}^k(\theta)$.

- *Método explícito:*

$$\frac{\check{u}^{k+1}(\theta) - \check{u}^k(\theta)}{\Delta t} + \frac{a(e^{i\theta})}{(\Delta x)^2} \check{u}^k(\theta) = 0, \quad k \geq 0, \theta \in [0, 2\pi); \quad (2.142)$$

• *Método implícito:*

$$\frac{\tilde{u}^{k+1}(\theta) - \tilde{u}^k(\theta)}{\Delta t} + \frac{a(e^{i\theta})}{(\Delta x)^2} \tilde{u}^{k+1}(\theta) = 0, \quad k \geq 0, \theta \in [0, 2\pi), \quad (2.143)$$

donde

$$a(e^{i\theta}) = 2 - e^{i\theta} - e^{-i\theta}. \quad (2.144)$$

En vista de la equivalencia de las normas de $\ell^2(\mathbb{Z})$ y de $L^2(0, 2\pi)$ para las sucesiones de $\ell^2(\mathbb{Z})$ y sus transformadas de Fourier, para analizar la evolución de la norma de \tilde{u}^k en $\ell^2(\mathbb{Z})$ basta con analizar la norma $L^2(0, 2\pi)$ de $\tilde{u}^k(\theta)$. Ahora bien, en vista de (2.142) y (2.143) tenemos que

$$\tilde{u}^{k+1}(\theta) = \left[1 - \mu a(e^{i\theta})\right] \tilde{u}^k(\theta) \quad (2.145)$$

y

$$\tilde{u}^{k+1}(\theta) = \frac{1}{[1 + \mu a(e^{i\theta})]} \tilde{u}^k(\theta) \quad (2.146)$$

en el método explícito e implícito respectivamente. Aquí y en lo que sigue μ denota el número de Courant $\mu = \Delta t / (\Delta x)^2$.

Iterando esta regla de recurrencia obtenemos para cada uno de estos esquemas

$$\tilde{u}^k(\theta) = \left[1 - \mu a(e^{i\theta})\right]^k \hat{\varphi}(\theta) \quad (2.147)$$

y

$$\tilde{u}^k(\theta) = \left[1 + \mu a(e^{i\theta})\right]^{-k} \hat{\varphi}(\theta). \quad (2.148)$$

Tenemos por tanto

$$\sum_{j \in \mathbb{Z}} |u_j^k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |\tilde{u}^k(\theta)|^2 d\theta = \frac{1}{2\pi} \int_0^{2\pi} |1 - \mu a(e^{i\theta})|^{2k} |\hat{\varphi}(\theta)|^2 d\theta \quad (2.149)$$

y

$$\sum_{j \in \mathbb{Z}} |u_j^k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |\tilde{u}^k(\theta)|^2 d\theta = \frac{1}{2\pi} \int_0^{2\pi} |1 + \mu a(e^{i\theta})|^{-2k} |\hat{\varphi}(\theta)|^2 d\theta. \quad (2.150)$$

La comprobación de la estabilidad del esquema numérico consiste en probar cotas sobre el crecimiento de la norma $L^2(0, 2\pi)$ de $\tilde{u}^k(\theta)$ de manera controlada, independientemente del paso del mallado. En la medida que para recorrer un intervalo temporal $[0, T]$ el número de pasos temporales que hemos de dar es del orden de $k \sim T/\Delta t$, es fácil comprobar que la condición necesaria y suficiente para la estabilidad de cada uno de los métodos es que

$$\left|1 - \mu a(e^{i\theta})\right| \leq 1, \quad \forall \theta \in [0, 2\pi) \quad (2.151)$$

y

$$\left|1 + \mu a\left(e^{i\theta}\right)\right|^{-1} \leq 1, \forall \theta \in [0, 2\pi) \quad (2.152)$$

respectivamente.

La suficiencia de las condiciones (2.151) y (2.152) para la estabilidad del método explícito e implícito respectivamente es evidente pues, bajo estas condiciones, en virtud de (2.149) y (2.150), se tiene

$$\left\|u^{\vec{k}}\right\|_{\ell^2(\mathbb{Z})} \leq \|\vec{\varphi}\|_{\ell^2(\mathbb{Z})}, \forall k \geq 0. \quad (2.153)$$

El hecho de que las condiciones (2.151) o (2.152) sean también necesarias exige un poco más de reflexión. Pero para entenderlo basta observar que si para algún $\theta_0 \in [0, 2\pi)$ el módulo de alguna de estas cantidades es estrictamente mayor que uno, por continuidad lo es en un intervalo de la forma $[\theta_0 - \delta, \theta_0 + \delta]$ para algún $\delta > 0$. Basta entonces considerar funciones φ tales que el soporte de $\widehat{\varphi}$ esté contenido en $[\theta_0 - \delta, \theta_0 + \delta]$ para ver que la norma de la solución $u^{\vec{k}}$ correspondiente crece exponencialmente en k . En este punto conviene observar que, en vista de la fórmula de inversión (2.135), a cada elección $\widehat{\varphi}$ le corresponde una del dato inicial discreto $\vec{\varphi} \in \ell^2(\mathbb{Z})$.

Por lo tanto para analizar la estabilidad de los métodos discretos en consideración basta estudiar el rango de números de Courant μ para el que se satisface (2.151) y (2.152), respectivamente.

- *Método explícito*

Conviene observar que

$$a\left(e^{i\theta}\right) = 2 - e^{i\theta} - e^{-i\theta} = 2(1 + \cos \theta). \quad (2.154)$$

Por tanto,

$$\left|1 - \mu a\left(e^{i\theta}\right)\right| = |1 - 2\mu(1 + \cos \theta)|.$$

Es fácil comprobar que

$$|1 - 2\mu(1 + \cos \theta)| \leq 1, \forall \theta \in [0, 2\pi) \Leftrightarrow 0 \leq \mu \leq 1/2.$$

Deducimos por tanto que el método explícito (2.131) es estable y por tanto convergente (puesto que, como dijimos, el método es consistente de orden dos) si y sólo si $0 < \mu \leq 1/2$.

- *Método implícito*

En vista de (2.154) es evidente que

$$1 + \mu a\left(e^{i\theta}\right) \geq 1, \forall \theta \in [0, 2\pi), \forall \mu > 0.$$

Por lo tanto (2.152) se satisface para todo $\mu > 0$ y por tanto el método implícito (2.132) es estable (y, por consiguiente, convergente) para todo $\mu > 0$.

De este modo vemos que en el caso del problema de Cauchy en toda la recta real se tienen los mismos resultados que en el caso de un intervalo acotado con condiciones de contorno de

Dirichlet. Volvamos ahora sobre el sistema semi-discreto (2.130). Recordemos que, tal y como se comprueba con facilidad mediante el desarrollo de Taylor, el método es consistente si $0 \leq \alpha \leq 1/2$.

Con el objeto de estudiar la estabilidad observamos que si \vec{u} satisface (2.130), entonces, \check{u} verifica:

$$(1 - 2(\alpha - \cos \theta))\check{u}'(\theta, t) + \frac{a(e^{i\theta})}{h^2}\check{u}(\theta, t) = 0, \quad (2.155)$$

o, de otro modo,

$$\check{u}'(\theta, t) + \frac{b(\theta)}{h^2}\check{u}(\theta, t) = 0, \quad (2.156)$$

donde

$$b(\theta) = \frac{a(e^{i\theta})}{1 - 2(\alpha - \cos \theta)}. \quad (2.157)$$

En este caso, por tanto, el esquema es estable, sí y sólo sí, $b(\theta) \geq 0$ para todo $\theta \in [0, 2\pi)$ y esto ocurre sólo cuando $0 \leq \theta \leq 1/4$.

Vemos por tanto que, gracias al análisis de von Neumann, es fácil estudiar también la estabilidad de los métodos semi-discretos y que, tal y como indicábamos al inicio de esta sección, permite comprobar la existencia métodos consistentes y no estables y, por tanto, no convergentes. A la hora de elegir los datos iniciales en el sistema semi-discreto (2.127) o en los sistemas completamente discretos (2.131) y (2.132) tenemos varias posibilidades:

- (a) Cuando $\varphi \in L^2(\mathbb{R}) \cap C(\mathbb{R})$ podemos elegir simplemente

$$\varphi_j = \varphi(jh);$$

- (b) Cuando $\varphi \in L^2(\mathbb{R})$ podemos sustituir la evaluación puntual de φ por el conjunto de sus medias

$$\varphi_j = \frac{1}{h} \int_{(j-1/2)h}^{(j+1/2)h} \varphi(x) dx;$$

- (c) Cuando $\varphi \in L^2(\mathbb{R})$ podemos también definir los datos iniciales utilizando la transformada de Fourier. Sea $\mathcal{F}(\varphi) \in L^2(\mathbb{R})$ la transformada de Fourier continua de φ , i.e.

$$\mathcal{F}(\varphi)(\xi) = \int_{\mathbb{R}} \varphi(x) e^{-ix\xi} dx.$$

Es natural aproximar esta transformada de Fourier por una función de banda limitada

$$\psi_h = \mathcal{F}(\varphi) 1_{[-\pi/h, \pi/h]},$$

donde $1_{[-\pi/h, \pi/h]}(\xi)$ denota la función característica del intervalo $\xi \in [-\pi/h, \pi/h]$. Basta entonces elegir el dato inicial φ_j del problema semi-discreto o completamente discreto como el

valor de la antitransformada de Fourier de ψ_h en el punto jh (o $j\Delta_x$ en el caso completamente discreto).

Conviene por último señalar que la transformada discreta de Fourier puede ser utilizada para probar la existencia y unicidad de las soluciones del problema semi-discreto y discreto.

Consideremos por ejemplo el problema semi-discreto (2.127). Se trata de un sistema de una infinidad de ecuaciones diferenciales de orden uno acopladas de tres en tres. Por lo tanto, los resultados clásicos de la teoría de EDO no pueden ser aplicados por tratarse de un problema en dimensión infinita. La transformada discreta de Fourier puede, efectivamente, utilizarse para resolver (2.127). En efecto, $\{\vec{u}(t)\} = \{u_j(t)\}_{j \in \mathbb{Z}}$ es la solución de (2.127) si y sólo si la transformada de Fourier correspondiente $\check{u}(\theta, t)$ verifica

$$\begin{cases} \check{u}_t(\theta, t) + \frac{1}{h^2} a(e^{i\theta}) \check{u}(\theta, t) = 0, & \theta \in [0, 2\pi), \quad t > 0, \\ \check{u}(\theta, 0) = \widehat{\varphi}(\theta), & \theta \in [0, 2\pi). \end{cases} \quad (2.158)$$

De (2.158) deducimos que

$$\check{u}(\theta, t) = \exp \left[-\frac{1}{h^2} a(e^{i\theta}) t \right] \widehat{\varphi}(\theta). \quad (2.159)$$

Como la transformada de Fourier define una isometría de $\ell^2(\mathbb{Z})$ en $L^2(0, \pi)$ sabemos que, como $\vec{\varphi} \in \ell^2(\mathbb{Z})$ entonces $\widehat{\varphi}(\theta) \in L^2(0, 2\pi)$. Como $a(e^{i\theta}) \geq 0$ para todo $\theta \in [0, 2\pi)$ tenemos que

$$|\check{u}(\theta, t)| \leq |\widehat{\varphi}(\theta)|, \quad \forall \theta \in [0, 2\pi), \quad \forall t > 0.$$

Por tanto, $\check{u}(\theta, t) \in L^2(0, 2\pi)$, para todo $t > 0$. Por consiguiente, la transformada inversa de Fourier $\vec{u}(t) \in \ell^2(\mathbb{Z})$. De este modo deducimos que para todo $\vec{\varphi} \in \ell^2(\mathbb{Z})$ el sistema (2.127) admite una única solución $\vec{u}(t) \in C([0, \infty); \ell^2(\mathbb{Z}))$ que es la transformada de Fourier inversa de la solución de (2.158).

Vemos por tanto que la transformada discreta de Fourier no sólo es un método para analizar la estabilidad y convergencia de los métodos numéricos sino incluso para resolver las ecuaciones semi-discreta y discreta que en ellos intervienen.

2.7 El método de elementos finitos

Tal y como vimos en el capítulo anterior en el estudio de la ecuación de Laplace, uno de los métodos más frecuentemente utilizados en la actualidad es el *método de elementos finitos* (MEF).

El MEF, si bien en una dimensión espacial y en problemas con coeficientes constantes como los que estamos estudiando proporciona esquemas numéricos de aproximación muy semejantes a los que hemos barajado mediante diferencias finitas, conceptualmente es bien distinto. Como vimos en el contexto de la ecuación de Laplace, el MEF consiste en buscar “soluciones” en un espacio de dimensión finita. No adoptamos pues el punto “de vista nodal” sino el del *método de Galerkin*: habiendo elegido una base de funciones y definido subespacios de dimensión finita, elegimos en él la función que es solución de la EDP en el sentido más preciso posible. Para ello comprobamos si el producto escalar del operador diferencial con todas las funciones de dicho espacio es nulo (para lo cual es preciso adoptar una formulación variacional del problema que

se obtiene integrando por partes). De este modo obtenemos una proyección de la solución real de la EDP sobre el espacio finito-dimensional considerado.

Esta metodología puede también ser aplicada en el marco de las ecuaciones de evolución y, en particular, en el de la ecuación del calor analizada en esta sección.

Presentemos pues brevemente la adaptación del MEF al problema que nos ocupa. Lo haremos en el caso de las aproximaciones semi-discretas si bien la adaptación al caso de aproximaciones completamente discretas (explícitas o implícitas en tiempo) es automática.

Consideramos la misma partición $\{x_j\}_{j=1,\dots,M}$, $x_j = jh$ del intervalo $(0, \pi)$. A cada nodo interno x_j , $j = 1, \dots, M$ le asociamos una función de base $\phi_j(x)$, continua y lineal a trozos tal que $\phi_j(x_l) = \delta_{jl}$, $j = 1, \dots, M; l = 0, \dots, M + 1$, siendo δ la delta de Kronecker.

Introducimos ahora el subespacio vectorial de dimensión M generado por las funciones $\{\phi_j\}_{j=1,\dots,M}$:

$$V_h = \text{span} \{\phi_j : j = 1, \dots, M\}. \quad (2.160)$$

Buscamos ahora soluciones aproximadas de la ecuación del calor en el subespacio $C([0, T]; V_h)$ de modo que

$$u_h(x, t) = \sum_{j=1}^M u_j(t) \phi_j(x). \quad (2.161)$$

Observese que se trata de una función que depende tanto de la variable espacial x como de la temporal. Por tanto, en este caso, el método numérico proporciona automáticamente una función continua. Es sin embargo evidente que la función u_h así obtenida, para cada instante de tiempo t , varía con respecto a x como una función continua y lineal a trozos. Es pues imposible que pueda reproducir exactamente la dinámica de cualquier solución $u = u(x, t)$ de la ecuación del calor. Sin embargo, como veremos, se puede demostrar con bastante facilidad que esta función puede ser elegida de modo que converja a la solución del calor cuando $h \rightarrow 0$.

Pero para que la función u_h esté unívocamente determinada es imprescindible definir con precisión sus coeficientes $u_j(t)$. Conviene señalar en este punto que, debido a la elección de las funciones de base $\{\phi_j\}$ realizada, $u_j(t)$ es también el valor en el punto $x = x_j$ de la función u_h . Sin embargo, en el marco de los elementos finitos la interpretación más natural es que $u_j(t)$ son los coeficientes de u_h como función que, en cada instante t , pertenece al subespacio V_h .

Conviene también señalar que el método de elementos finitos no es sólo una herramienta para la aproximación numérica de EDP sino que en realidad permite modelizar fenómenos de la Mecánica de Medios continuos a través de modelos discretos, que pueden ser manipulados mucho más fácilmente tanto de un punto de vista conceptual como computacional. Es por eso que el MEF está tan extendido en la literatura de las diversas ramas de la Ingeniería y que fue introducido paralelamente tanto en el ámbito de la modelización como de las aproximaciones numéricas hasta convertirse en una disciplina unificada y bien establecida.⁷

Debemos por tanto obtener las ecuaciones más naturales que gobiernan la dinámica de los coeficientes $u_j(t)$ de la solución aproximada. Como decíamos lo hacemos a través del método de Galerkin. Para ello es necesario, en primer lugar, introducir la formulación variacional de

⁷El lector interesado en una descripción de los orígenes del método de elementos finitos podrá consultar ([Zi]).

la ecuación del calor (2.8). Recordemos que la solución de esta ecuación pertenece al espacio $u \in C([0, T]; L^2(0, \pi)) \cap L^2(0, T; H_0^1(0, \pi))$. La formulación variacional de (2.8) es por tanto la siguiente: *Hallar $u \in C([0, T]; L^2(0, \pi)) \cap L^2(0, T; H_0^1(0, \pi))$ que satisfaga*

$$\frac{d}{dt} \int_0^\pi u(x, t) \Phi(x) dx = \int_0^\pi u_x(x, t) \Phi_x(x) dx, \text{ p. c. t. } t > 0, \forall \Phi \in H_0^1(0, \pi), \quad (2.162)$$

junto con la condición inicial

$$u(x, 0) = \varphi(x), \quad \text{en } (0, \pi). \quad (2.163)$$

Conviene comentar brevemente el sentido de esta formulación variacional. La condición inicial de (2.8) ha sido, evidentemente, tenida en cuenta en (2.163). Esta última tiene sentido puesto que u depende continuamente en tiempo a valores en el espacio $L^2(0, \pi)$. Se evaluación en un instante t dado y, en particular, en $t = 0$ está pues plenamente justificada. Por otra parte, en (2.162) se da una versión débil o distribucional de la ecuación del calor (2.8). Los dos términos de (2.162) tienen sentido. El segundo miembro es una función de $L^2(0, T)$ puesto que la solución u pertenece al espacio $L^2(0, T; H_0^1(0, \pi))$. Sin embargo, a causa del efecto regularizante de la ecuación del calor, se trata también de una función de $C([0, T]; H_0^1(0, \pi))$ por lo que podría exigirse que (2.163) se verifique no sólo para casi todo t en el intervalo $[0, T]$ sino para todo $t > 0$. Por otra parte, como $u \in C([0, T]; L^2(0, \pi))$ tenemos que la función $\int_0^\pi u(x, t) \Phi(x) dx$ es continua en tiempo. Su derivada temporal ha de ser por tanto entendida en el sentido de las distribuciones. Sin embargo, la solución de la ecuación calor también pertenece al espacio $H^1(0, T; H^{-1}(0, \pi))$, siendo $H^{-1}(0, \pi)$ el dual de $H_0^1(0, \pi)$, por lo que esta derivada temporal tiene también sentido en $L^2(0, T)$ o, por el efecto regularizante, un sentido puntual para cada $t > 0$. Por último señalar que la condición de contorno que se impone en (2.8) está incorporada en la formulación variacional al haber exigido a la función u que pertenezca al espacio $L^2(0, T; H_0^1(0, \pi))$. Pero no vamos a profundizar más en este terreno. El lector interesado en un estudio más detallado del punto de vista variacional para la resolución de EDP de evolución podrá consultar los textos [E] y [L].

Inspirándonos en la formulación variacional de la ecuación del calor es fácil introducir el análogo discreto que caracteriza a la solución que el método de elementos finitos proporciona. La formulación variacional para el cálculo de la solución aproximada u_h es pues la siguiente: *Hallar $u_h \in C^1([0, T]; V_h)$ tal que*

$$\frac{d}{dt} \int_0^\pi u_h(x, t) \Phi_j(x) dx = \int_0^\pi u_{h,x}(x, t) \Phi_{j,x}(x) dx, \forall t > 0, \forall j = 1, \dots, M, \quad (2.164)$$

⁸El espacio $C([0, T]; L^2(0, \pi)) \cap L^2(0, T; H_0^1(0, \pi))$ es un espacio de Banach, intersección de $C([0, T]; L^2(0, \pi))$ y de $L^2(0, T; H_0^1(0, \pi))$, en el que la norma viene dada por la suma de las normas en cada uno de los espacios. Las funciones de $C([0, T]; L^2(0, \pi))$, dependen continuamente de $t \in [0, T]$ y toman valores en $L^2(0, \pi)$, y su norma viene dada por $\|f\|_{C([0, T]; L^2(0, \pi))} = \max_{t \in [0, T]} \|f(t)\|_{L^2(0, \pi)}$ mientras que las de $L^2(0, T; H_0^1(0, \pi))$ son funciones medibles y de cuadrado integrable de $t \in [0, T]$ a valores en el espacio de Sobolev $H_0^1(0, \pi)$. La norma correspondiente es entonces $\|f\|_{L^2(0, T; H_0^1(0, \pi))} = \left[\int_0^T \int_0^\pi f_x^2(x, t) dx dt \right]^{1/2}$.

junto con la condición inicial

$$u_h(x, 0) = \varphi_h(x), \quad \text{en } (0, \pi). \quad (2.165)$$

Las analogías entre la formulación variacional del problema continuo (2.162)-(2.163) y la del problema discreto (2.164)-(2.165) son evidentes. En esta ocasión, como $u_h(t)$ “vive” en cada instante de t en el espacio de dimensión finita V_h , suponemos que u_h es de clase C^1 en tiempo y eso permite escribir (2.165) para cada instante t . Por otra parte, en (2.164) exigimos que la versión débil de la ecuación del calor discretizada se verifique para todas las funciones de base Φ_j , $j = 1, \dots, M$. Esto es totalmente equivalente a suponer que se verifica para cualquier función test Φ del espacio V_h . En (2.165) hemos tomado un dato inicial $u_{h,0}$. Tal y como comentamos en el marco de las diferencias finitas, son varias las maneras en las que este dato inicial se puede tomar de manera que aproxime el dato inicial u_0 de la ecuación del calor. La manera más sencilla en este contexto es tal vez elegir φ_h como la proyección ortogonal de u_0 sobre V_h .

El sistema (2.164)-(2.165) es un sistema de M ecuaciones diferenciales lineales acopladas en las incógnitas $u_j(t)$, $j = 1, \dots, M$.

Este sistema se puede escribir de manera mucho más explícita usando las matrices de masa \mathcal{M}_h y de rigidez \mathcal{R}_h del método de elementos finitos.

Recordemos que los elementos m_{jk} de la matriz de masa \mathcal{M}_h son precisamente

$$m_{jk} = \int_0^\pi \Phi_j(x) \Phi_k(x) dx, \quad (2.166)$$

y lo elementos (r_{jk}) de la matriz de rigidez \mathcal{R}_h :

$$r_{jk} = \int_0^\pi \Phi_{j,x}(x) \Phi_{k,x}(x) dx. \quad (2.167)$$

Todos estos coeficientes pueden calcularse explícitamente con facilidad y constituyen en ambos casos matrices simétricas tridiagonales, con coeficientes constantes a lo largo de las diagonales y de las sub y super-diagonales. De manera más precisa se tiene:

$$\mathcal{M}_h = h \begin{pmatrix} 2/3 & 1/6 & 0 & 0 \\ 1/6 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 1/6 \\ 0 & 0 & 1/6 & 2/3 \end{pmatrix}, \quad (2.168)$$

y

$$\mathcal{R}_h = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}. \quad (2.169)$$

Observese que, en particular,

$$\mathcal{R}_h = hA_h, \quad (2.170)$$

donde A_h es la matriz (2.26) que interviene en la discretización de tres puntos del Laplaciano mediante elementos finitos.

El sistema (2.164)-(2.165), con la notación vectorial empleada en el método de diferencias finitas puede entonces escribirse en la forma:

$$\begin{cases} \mathcal{M}_h \vec{u}'(t) + \mathcal{R}_h \vec{u}(t) = 0, & t > 0 \\ \vec{u}(0) = \vec{\varphi}. \end{cases} \quad (2.171)$$

Ambas matrices \mathcal{M}_h y \mathcal{R}_h comparten los autovectores de la matriz A_h . Así los autovectores de estas matrices son $\{\vec{w}_l(h)\}_{j=1,\dots,M}$ definidos (2.30). Esto permite desarrollar con facilidad las soluciones de (2.171) en la base de estos autovectores. En efecto, teniendo en cuenta que la ecuación en (2.171) puede también escribirse en la forma

$$\vec{u}'(t) + \mathcal{M}_h^{-1} \mathcal{R}_h \vec{u}(t) = 0, \quad (2.172)$$

calculando los autovalores de $\mathcal{M}_h^{-1} \mathcal{R}_h$ que vienen dados por

$$\mu_j(h) = \frac{6}{h^2} \frac{1 - \cos(jh)}{2 + \cos(jh)}, \quad j = 1, \dots, M \quad (2.173)$$

Obtenemos así la siguiente expresión en series de Fourier de las soluciones del problema semi-discreto:

$$\vec{u}(t) = \vec{u}_h(t) = \sum_{j=1}^M \hat{\varphi}_j e^{-\mu_j(h)t} \vec{w}_l(h). \quad (2.174)$$

A partir de esta expresión no es difícil adaptar los resultados de convergencia que hemos probado para la aproximación mediante diferencias finitas al caso presente del MEF. Nuevamente, en este caso simple unidimensional nos encontramos con que los autovectores del esquema discreto coinciden en los nodos del mallado con los del problema continuo. Y nuevamente también los autovalores $\mu_j(h)$ del problema discreto convergen a los del continuo cuando $h \rightarrow 0$. En efecto, de (2.173) se deduce inmediatamente que

$$\mu_j(h) \rightarrow j^2, \quad \text{cuando } h \rightarrow 0, \quad (2.175)$$

para cada j fijo.

También el método de energía puede ser desarrollado en este caso sin dificultad. En efecto, multiplicando la ecuación (2.171) componente a componente por $u_j(t)$, obtenemos la identidad:

$$\frac{d}{dt} \mathcal{E}_h(t) = -h \sum_{j=0}^M \left| \frac{u_{j+1} - u_j}{h} \right|^2 \quad (2.176)$$

donde en esta ocasión la energía viene dada por

$$\mathcal{E}_h(t) = \frac{h}{6} \sum_{j=1}^M |u_j|^2 + \frac{h}{12} \sum_{j=0}^M |u_j + u_{j+1}|^2. \quad (2.177)$$

A partir de esta ley de disipación de la energía discreta (que es un análogo de la energía L^2 de la ecuación del calor continua), se establece con facilidad resultados de convergencia similares a los probados en el caso del método de diferencias finitas.

3 La ecuación de ondas

Tal y como hemos mencionado en la introducción, la ecuación de ondas es otro de los modelos más relevantes que se escribe en términos de EDP puesto que interviene, de uno u otro modo, en infinidad de problemas de la Mecánica, de la Física y de la Ingeniería. Así, se trata de un modelo ubicuo en elasticidad y vibraciones de estructuras pero también en el ámbito de la propagación de ondas acústicas o electromagnéticas.

Desde un punto de vista matemático la ecuación de ondas es el opuesto exacto de la del calor pues se trata de un sistema reversible en tiempo, conservativo, carente de efectos regularizantes y en el que la velocidad de propagación es finita.

En esta sección seguiremos el guión de la anterior. En primer lugar recordaremos algunas propiedades básicas de la ecuación de ondas que nos servirán de guía a la hora de analizar los modelos discretizados correspondientes. Después estudiaremos la convergencia de las aproximaciones semi-discretas y por último las completamente discretas.

Por una cuestión de espacio nos ceñiremos a la ecuación de ondas en una sola dimensión espacial. Sin embargo, muchos de los conceptos y resultados que veremos y desarrollaremos se adaptan con relativa facilidad a la ecuación de ondas en varias dimensiones espaciales, al sistema de Lamé en elasticidad, a las ecuaciones de placas que involucran habitualmente el operador biarmónico, las ecuaciones de Schrödinger de la Mecánica Cuántica o incluso a otros modelos como las ecuaciones de Korteweg-de-Vries para las olas en canales poco profundos. Todos estos temas quedan para desarrollos posteriores.

3.1 Propiedades básicas de la ecuación de ondas 1 - d

En primer lugar consideramos el problema de Cauchy en toda la recta:

$$\begin{cases} u_{tt} - u_{xx} = 0, & x \in \mathbb{R}, \quad t > 0 \\ u(x, 0) = \varphi(x), u_t(x, 0) = \psi(x), & x \in \mathbb{R}. \end{cases} \quad (3.1)$$

El operador en derivadas parciales involucrado $\partial_t^2 - \partial_x^2$ se denota frecuentemente mediante el símbolo \square y se denomina d'Alembertiano. Su versión multi-dimensional es

$$\square = \partial_t^2 - \Delta_x = \partial_t^2 - \sum_{i=1}^N \partial_{x_i}^2.$$

Pero en esta sección nos limitaremos al caso unidimensional.

En una dimensión espacial la ecuación de ondas es un modelo simplificado para las vibraciones de pequeña amplitud de una cuerda y mediante $u = u(x, t)$ se describen las deformaciones verticales de la misma.

La solución de (3.1) puede calcularse de forma explícita. En efecto, es fácil comprobar que la solución de (3.1) es

$$u(x, t) = \frac{1}{2} [\varphi(x+t) + \varphi(x-t)] + \frac{1}{2} \int_{x-t}^{x+t} \psi(s) ds. \quad (3.2)$$

Conviene también observar que u es de la forma

$$u(x, t) = \mathcal{F}(x+t) + \mathcal{G}(x-t) \quad (3.3)$$

donde

$$\mathcal{F}(s) = \frac{1}{2} \varphi(s) + \frac{1}{2} \int_0^s \psi(\sigma) d\sigma, \quad (3.4)$$

$$\mathcal{G}(s) = \frac{1}{2} \varphi(s) + \frac{1}{2} \int_s^0 \psi(\sigma) d\sigma. \quad (3.5)$$

Es también digno de mención que cualquier función de la forma (3.3) es solución de (3.1). Este hecho corresponde a la siguiente factorización del *operador de d'Alembert*

$$\square = \partial_t^2 - \partial_x^2 = (\partial_t - \partial_x)(\partial_t + \partial_x) = (\partial_t + \partial_x)(\partial_t - \partial_x), \quad (3.6)$$

según la cual el *operador de ondas* es la composición de los dos *operadores de transporte* de orden uno $\partial_t \pm \partial_x$, cuyas soluciones son efectivamente *ondas viajeras* de la forma $\mathcal{F}(x+t)$ o $\mathcal{G}(x-t)$.

En la fórmula (3.2) se observa también otra de las propiedades fundamentales de la ecuación de ondas: *la velocidad finita de propagación*. Así el valor de la solución u en el punto (x, t) depende exclusivamente del valor de los datos iniciales en el *intervalo de dependencia* $[x-t, x+t]$. Por otra parte, una perturbación de los datos iniciales en el instante $t = 0$ en el punto x_0 sólo afecta al valor de la solución en el cono de influencia $|x - x_0| < |t|$.

Otra de las propiedades que se deduce de la fórmula de representación (3.2) es la *ausencia de efecto regularizante*. En efecto, de (3.2) se deduce que la solución u , en cualquier instante $t > 0$, es tan regular como el dato inicial φ para la posición y gana una derivada con respecto a la velocidad inicial ψ . Del mismo modo, la velocidad u_t tiene la misma regularidad que ψ y pierde una derivada con respecto a φ .

Al tratarse de una ecuación de orden dos en tiempo, las genuinas incógnitas del problema son tanto u como u_t . Es por eso que en (3.1) hemos de proporcionar los datos iniciales de ambas incógnitas para garantizar la existencia y unicidad de soluciones. Desde este punto de vista es habitual escribir la ecuación (3.1) como un sistema de la forma

$$\begin{cases} u_t = v \\ v_t = u_{xx} \end{cases}$$

o bien como un sistema de leyes de conservación hiperbólica

$$\begin{cases} u_t = w_x \\ w_t = u_x \end{cases}$$

Como hemos dicho anteriormente algunas de estas propiedades de la ecuación se preservan en más de una dimensión espacial. En particular, se tienen fórmulas de representación explícitas semejantes a (3.2) aunque algo más complejas ([E], [J]).

Otra de las propiedades importantes de la ecuación de ondas es la *ley de conservación de la energía*. En este caso la energía correspondiente es

$$E(t) = \frac{1}{2} \int_{\mathbb{R}} \left[|u_x(x, t)|^2 + |u_t(x, t)|^2 \right] dx \quad (3.7)$$

y se tiene

$$\frac{dE}{dt}(t) = 0, \quad \forall t \geq 0, \quad (3.8)$$

lo cual es fácil de comprobar multiplicando la ecuación de (3.1) por u_t e integrando en \mathbb{R} .

Esta ley de conservación de la energía sugiere que el espacio $H^1(\mathbb{R}) \times L^2(\mathbb{R})$ es un marco funcional adecuado para la resolución de la ecuación de ondas. Y, efectivamente, es así. Por tanto, para todo dato inicial $(\varphi, \psi) \in H^1(\mathbb{R}) \times L^2(\mathbb{R})$ existe una única solución de (3.1) en la clase $u \in C([0, \infty); H^1(\mathbb{R})) \cap C^1([0, \infty); L^2(\mathbb{R}))$. Consiguientemente vemos que el vector incógnita preserva, con continuidad en tiempo, la regularidad de los datos iniciales.

Son diversos los métodos que permiten probar este tipo de resultados de existencia y unicidad: método de Galerkin, teoría de semigrupos, transformada de Fourier, etc. Pero en el caso que nos ocupa puede deducirse inmediatamente de la fórmula de d'Alembert (3.2).

De la fórmula de representación (3.2) se deduce que la ecuación (3.1) está bien puesta en infinitud de otros espacios. Pero el más natural para resolverlo y el que se extiende de manera natural a otras situaciones como problemas de frontera o situaciones multidimensionales es precisamente el marco hilbertiano $H^1(\mathbb{R}) \times L^2(\mathbb{R})$.

Consideramos ahora la ecuación de ondas en un intervalo acotado:

$$\begin{cases} u_{tt} - u_{xx} = 0, & 0 < x < \pi, \quad t > 0 \\ u(0, t) = u(\pi, t) = 0, & t > 0 \\ u(x, 0) = \varphi(x), \quad u_t(x, 0) = \psi(x), & 0 < x < \pi. \end{cases} \quad (3.9)$$

Se trata de un modelo simplificado para las vibraciones de una cuerda de longitud π . En este caso hemos impuesto condiciones de contorno de Dirichlet que indican que la cuerda está fija en sus extremos, si bien los resultados que aquí describimos se adaptan con facilidad a otras.

Las soluciones de (3.9) se pueden representar fácilmente en series de Fourier. En efecto, si los datos iniciales admiten el desarrollo en serie de Fourier

$$\varphi(x) = \sum_{l \geq 1} \hat{\varphi}_l w_l(x); \quad \psi(x) = \sum_{l \geq 1} \hat{\psi}_l w_l(x), \quad 0 < x < \pi, \quad (3.10)$$

donde $w_l(x)$ viene dado por (2.9), la solución de (3.9) se escribe del siguiente modo

$$u(x, t) = \sum_{l=1}^{\infty} \left(\hat{\varphi}_l \cos(lt) + \frac{\hat{\psi}_l}{l} \operatorname{sen}(lt) \right) w_l(x). \quad (3.11)$$

Esta expresión puede ser simplificada utilizando exponenciales complejas:

$$u(x, t) = \sum_{l=-\infty}^{+\infty} \hat{\theta}_l e^{ilt} w_l(x), \quad (3.12)$$

donde

$$w_{-l}(x) = w_l(x), \quad l \geq 1 \text{ y } \hat{\theta}_l = \frac{l\hat{\varphi}_l - i\hat{\psi}_l}{2l}, \quad \hat{\theta}_{-l} = \hat{\theta}_l + \frac{i\hat{\psi}_l}{l}, \quad \forall l \geq 1. \quad (3.13)$$

Como veremos más adelante, las soluciones de los problemas semi-discretos y completamente discretos que consideramos admiten desarrollos en serie de Fourier semejantes.

Nuevamente, la energía de las soluciones de (3.9) se conserva en tiempo.

En efecto

$$E(t) = \frac{1}{2} \int_0^{\pi} \left[|u_t(x, t)|^2 + |u_x(x, t)|^2 \right] dx, \quad (3.14)$$

satisface

$$\frac{dE}{dt}(t) = 0, \quad \forall t \geq 0, \quad (3.15)$$

lo cual puede comprobarse fácilmente de dos maneras. La primera, por el método de la energía, multiplicando la ecuación (3.9) por u_t e integrando con respecto a $x \in (0, \pi)$. La segunda mediante simple inspección del desarrollo en serie de Fourier (3.11). El espacio natural para resolver (3.9) es $H_0^1(0, \pi) \times L^2(0, \pi)$. De ese modo, cuando $(\varphi, \psi) \in H_0^1(0, \pi) \times L^2(0, \pi)$, (3.9) admite una única solución $u \in C([0, \infty); H_0^1(0, \pi)) \cap C^1([0, \infty); L^2(0, \pi))$.

La energía equivale al cuadrado de la norma canónica del espacio $H_0^1(0, \pi) \times L^2(0, \pi)$. El hecho de que la energía permanezca constante en tiempo equivale a que la trayectoria de la solución permanece indefinidamente en una esfera del espacio $H_0^1(0, \pi) \times L^2(0, \pi)$, la que corresponde a los datos iniciales del sistema.

El hecho de que los datos iniciales del problema pertenezcan a $H_0^1(0, \pi) \times L^2(0, \pi)$ equivale a que los coeficientes $\{\vec{\varphi}_l\}_{k \geq 1}$ y $\{\vec{\psi}_l\}_{k \geq 1}$ del desarrollo en serie de Fourier (3.11) de u satisfagan

$$\sum_{k \geq 1} \left[l^2 |\vec{\varphi}_l|^2 + |\vec{\psi}_l|^2 \right] < \infty, \quad (3.16)$$

o, equivalentemente,

$$\sum_{k=-\infty}^{+\infty} |\vec{\theta}_l|^2 l^2 < \infty. \quad (3.17)$$

3.2 Semi-discretización espacial: El método de Fourier

Con las notaciones de la sección 2 dedicada a la ecuación del calor, la semi-discretización espacial más natural de la ecuación de ondas (3.9) es la siguiente:

$$\begin{cases} u_j'' + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} = 0, & j = 1, \dots, M, \quad t > 0 \\ u_0 = u_{M+1} = 0, & t > 0 \\ u_j(0) = \varphi_j, u_j'(0) = \psi_j, & j = 1, \dots, M. \end{cases} \quad (3.18)$$

Se trata de un sistema acoplado de M ecuaciones diferenciales de orden dos con M incógnitas. Las M incógnitas $u_1(t), \dots, u_M(t)$ proporcionan aproximaciones de la solución continua $u = u(x, t)$ de la ecuación de ondas (3.9) en los puntos x_1, \dots, x_M del mallado. En vista de las condiciones de contorno de Dirichlet es natural imponer que los desplazamientos u_0 y u_{M+1} que proporcionan aproximaciones de u en los extremos $x_0 = 0$ y $x_{M+1} = \pi$ sean nulos.

Los datos iniciales $(\varphi_j, \psi_j)_{j=1}^M$ son típicamente una aproximación de los datos iniciales $(\varphi(x), \psi(x))$ de la ecuación de ondas (3.9) sobre el mallado. Tal y como vimos en la sección anterior, son diversas las elecciones posibles. En cada uno de los resultados de convergencia que probaremos indicaremos explícitamente la elección de los datos iniciales realizada.

Con la notación vectorial de la sección 2.2 el sistema (3.18) puede escribirse en la forma

$$\begin{cases} \vec{u}''(t) + A_h \vec{u}(t) = 0, & t > 0, \\ \vec{u}(0) = \vec{\varphi}, \quad \vec{u}'(0) = \vec{\psi}. \end{cases} \quad (3.19)$$

Las soluciones de (3.19) también pueden desarrollarse en serie de Fourier. Así, los datos iniciales admiten el desarrollo

$$\vec{\varphi} = \sum_{l=1}^M \hat{\varphi}_l(h) \vec{W}_l(h); \quad \vec{\psi} = \sum_{l=1}^M \hat{\psi}_l(h) \vec{W}_l(h), \quad (3.20)$$

donde

$$\hat{\varphi}_l(h) = \langle \vec{\varphi}, \vec{W}_l(h) \rangle_h; \quad \hat{\psi}_l(h) = \langle \vec{\psi}, \vec{W}_l(h) \rangle_h. \quad (3.21)$$

Entonces, la solución de (3.19) puede escribirse del siguiente modo

$$\vec{u}_h(t) = \sum_{l=1}^M \left(\hat{\varphi}_l(h) \cos(\mu_l(h)t) + \frac{\hat{\psi}_l(h)}{\mu_l(h)} \operatorname{sen}(\mu_l(h)t) \right) \vec{W}_l(h), \quad (3.22)$$

donde

$$\mu_l(h) = \sqrt{\lambda_l(h)}. \quad (3.23)$$

Aquí y en lo sucesivo $\vec{W}_l(h)$ y $\lambda_l(h)$ denotan los autovectores y autovalores de la matriz A_h introducidos en (2.29)-(2.30).

Nuevamente la expresión puede simplificarse:

$$\vec{u}_h(t) = \sum_{l=-M}^M \hat{\theta}_l(h) e^{i\mu_l(h)t} \vec{W}_l(h), \quad (3.24)$$

con la convención

$$\vec{W}_{-l}(h) = \vec{W}_l(h), \quad l = 1, \dots, M; \quad \mu_{-l}(h) = -\mu_l(h), \quad l = 1, \dots, M, \quad (3.25)$$

donde

$$\hat{\theta}_l(h) = \frac{\mu_l(h)\hat{\varphi}_l(h) - i\hat{\psi}_l(h)}{2\mu_l(h)}, \quad \hat{\theta}_{-l}(h) = \hat{\theta}_l(h) + \frac{i\hat{\psi}_l(h)}{2\mu_l(h)}, \quad \forall l \geq 1. \quad (3.26)$$

Si tenemos en cuenta que, para l fijo, $\mu_l(h) \rightarrow l$ cuando $h \rightarrow 0$, a la vez la analogía entre las expresiones (3.13) y (3.24) son evidentes.

El sistema (3.18) y (3.19) tiene también la propiedad de conservación de la energía. En este caso la energía conservada admite la expresión

$$E_h(t) = \frac{h}{2} \sum_{k=0}^M \left[\left| \frac{u_{j+1} - u_j}{h} \right|^2 + |u'_j|^2 \right]. \quad (3.27)$$

Esta ley de conservación de la energía puede obtenerse al menos de dos maneras: a) A partir del desarrollo en serie de Fourier de las soluciones (3.22); b) Multiplicando cada una de las ecuaciones de (3.18) por u'_j y sumando con respecto al índice j .

La energía discreta E_h es evidentemente una aproximación de la energía continua E de (3.14) y las dos formas que hemos indicado de probar su carácter conservativo son también versiones discretas de las pruebas habituales en la ecuación de ondas continua.

De todas estas observaciones se deduce que el sistema semi-discreto (3.18) es una aproximación natural del sistema continuo (3.9).

Con el objeto de establecer la convergencia cuando $h \rightarrow 0$ de las soluciones del sistema semi-discreto a las de la ecuación de ondas hemos de introducir normas que midan la distancia entre ambas. Tenemos por una parte la norma L^2 discreta introducida en la sección 2:

$$\|\vec{a}\|_h = \left[h \sum_{j=1}^M |a_j|^2 \right]^{1/2}. \quad (3.28)$$

Introducimos también la norma H^1 -discreta correspondiente

$$\|\vec{a}\|_{1,h} = \left[h \sum_{j=0}^M \left| \frac{a_{j+1} - a_j}{h} \right|^2 \right]^{1/2}, \quad (3.29)$$

siempre con la convención $a_0 = a_{M+1} = 0$.

Pero es también natural medir la proximidad de las soluciones en función de sus coeficientes de Fourier. Tal y como vimos en la sección 3.1, las soluciones de energía finita de la ecuación de ondas pueden ser identificadas, en virtud de (3.17), con sus coeficientes de Fourier pertenecientes al espacio \mathcal{H}^1 :

$$\mathcal{H}^1 = \left\{ \{\hat{a}_l\}_{j \in \mathbb{Z}} : \sum_{l \in \mathbb{Z}} l^2 |\hat{a}_l|^2 < \infty \right\} \quad (3.30)$$

que es un espacio de Hilbert con la norma

$$\|\{\hat{a}_l\}_{l \in \mathbb{Z}}\|_{\mathcal{H}^1} = \left[\sum_{l \in \mathbb{Z}} l^2 |\hat{a}_l|^2 \right]^{1/2}. \quad (3.31)$$

En efecto, tal como observamos en (3.17), las soluciones de energía finita de la ecuación de ondas corresponden a datos iniciales que se representan a través de los coeficientes $\{\hat{\theta}_l\}_{l \in \mathbb{Z}} \in \mathcal{H}^1$.

En vista de la representación de Fourier (3.12) de las soluciones de la ecuación de ondas, éstas pueden escribirse del siguiente modo

$$u(x, t) = \sum_{l=-\infty}^{+\infty} \hat{u}_l(t) w_l(x), \quad (3.32)$$

donde

$$\hat{u}_l(t) = \hat{\theta}_l e^{ilt}. \quad (3.33)$$

De este modo se observa que la regularidad propia de las soluciones de energía finita, i.e. el hecho de que

$$u \in C([0, \infty); H_0^1(0, \pi)) \cap C^1([0, \infty); L^2(0, \pi)), \quad (3.34)$$

equivale a que sus coeficientes de Fourier $\{\hat{u}_l(t)\}_{l \in \mathbb{Z}}$ pertenezcan al espacio

$$\{\hat{u}_l(t)\}_{l \in \mathbb{Z}} \in C([0, \infty); \mathcal{H}^1) \cap C^1([0, \infty); \ell^2). \quad (3.35)$$

Asimismo, las soluciones del problema semi-discreto, en vista de (3.24), pueden escribirse como

$$\vec{u}(t) = \sum_{l=-M}^M \hat{u}_{l,h}(t) \vec{W}_l(h) \quad (3.36)$$

donde

$$\hat{u}_{l,h}(t) = \hat{\theta}^l(h) e^{i\mu_l(h)t}, \quad l = -M, \dots, M. \quad (3.37)$$

Si extendemos por cero estos coeficientes de Fourier para los índices $|l| > M$, nos encontramos nuevamente con que, para las soluciones del problema semi-discreto, se tiene

$$\{\hat{u}_{l,h}(t)\}_{l=-M}^M \in C([0, \infty); \mathcal{H}^1) \cap C^1([0, \infty); \ell^2). \quad (3.38)$$

Es por tanto natural medir la convergencia de las soluciones del problema semi-discreto al continuo a través de la convergencia de sus coeficientes de Fourier en el espacio $C([0, \infty); \mathcal{H}^1) \cap C^1([0, \infty); \ell^2)$.

Con estas notaciones y con el objeto de establecer la convergencia de las soluciones del problema semi-discreto al continuo realizamos la siguiente elección de los datos iniciales de (3.19):

$$\vec{\varphi} = \sum_{l=-M}^M \hat{\varphi}_l \vec{W}_l(h); \quad \vec{\psi} = \sum_{l=-M}^M \hat{\psi}_l \vec{W}_l(h), \quad (3.39)$$

donde $\{\hat{\varphi}_l\}_{l \in \mathbb{Z}}$ y $\{\hat{\psi}_l\}_{l \in \mathbb{Z}}$ son los coeficientes de Fourier del dato inicial continuo $(\varphi(x), \psi(x)) \in H_0^1(0, \pi) \times L^2(0, \pi)$. Es decir, elegimos los datos iniciales del problema semi-discreto de modo que coincidan con los del problema continuo para los índices $-M \leq k \leq M$. De este modo, la solución de (3.19) admite el desarrollo (3.24) (o (3.36)-(3.37)) con los coeficientes de Fourier

$$\hat{\theta}_l(h) = \frac{\mu_l(h)\hat{\varphi}_l - i\hat{\psi}_l}{2\mu_l(h)}, \quad l = -M, \dots, M. \quad (3.40)$$

Con esta elección de los datos iniciales en el sistema semi-discreto tenemos el siguiente resultado de convergencia.

Teorema 3.1 *Supongamos que $(\varphi, \psi) \in H_0^1(0, \pi) \times L^2(0, \pi)$ y que elegimos los datos iniciales del problema semi-discreto como en (3.39).*

Entonces, las soluciones del problema semi-discreto convergen a las del continuo en el sentido que

$$\begin{cases} \{\hat{u}_{\ell, h}(t)\}_{\ell=-M}^M \rightarrow \{\hat{u}_{\ell}(t)\}_{\ell=-\infty}^{\infty} & \text{en } C([0, T]; \mathcal{H}^1) \cap C^1([0, T]; \ell^2) \\ \text{cuando } h \rightarrow 0. \end{cases} \quad (3.41)$$

para todo $0 < T < \infty$.

Observación 3.1

- Hemos enunciado la convergencia de las soluciones en términos de sus coeficientes de Fourier. En efecto, en (3.41) se establece que los coeficientes de Fourier de las soluciones semi-discretas, junto con sus derivadas temporales, convergen, cuando $h \rightarrow 0$, uniformemente en tiempo, a los de la solución continua en el espacio $\mathcal{H}^1 \times \ell^2$. Como hemos dicho antes, este es el espacio natural al que pertenecen los coeficientes de Fourier de las soluciones de energía finita de la ecuación de ondas, por lo que el resultado es óptimo.
- Este resultado puede ser también interpretado a través de la convergencia de las soluciones en el espacio de la energía. Volveremos sobre este punto más adelante.
- Tal y como habíamos anunciado en (3.41) hemos abusado de la notación puesto que hemos interpretado que $\{\hat{u}_{\ell, h}\}_{\ell=-M}^M$, que es un vector finito, es una sucesión. Como mencionábamos, ha de sobreentenderse que extendemos el valor de $\hat{u}_{\ell, h}$ por cero para todos los índices $|\ell| > M$.

■

Demostración del Teorema 3.1.

Probamos exclusivamente la convergencia en $C([0, T]; \mathcal{H}^1)$. La convergencia en $C^1([0, T]; \ell^2)$ puede ser probada de manera análoga.

Sea

$$\hat{v}_{\ell, k}(t) = \hat{u}_{\ell}(t) - \hat{u}_{\ell, h}(t) = \hat{\theta}_{\ell} e^{i\ell t} - \hat{\theta}_{\ell}(h) e^{i\mu_{\ell}(h)t}, \quad (3.42)$$

donde los coeficientes $\hat{\theta}_\ell$ y $\hat{\theta}_\ell(h)$ vienen dados por (3.13) y (3.40) respectivamente.

Tenemos

$$\begin{aligned} \|\{\hat{v}_{\ell,h}(t)\}\|_{\mathcal{H}^1}^2 &= \sum_{\ell=-\infty}^{\infty} |\hat{v}_{\ell,h}(t)|^2 \ell^2 = \sum_{\ell=-\infty}^{\infty} |\hat{u}_\ell(t) - \hat{u}_{\ell,h}(t)|^2 \ell^2 \\ &= \sum_{\ell=-m}^M \left| \hat{\theta}_\ell e^{i\ell t} - \hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} \right|^2 \ell^2 + \sum_{|\ell|>M} \left| \hat{\theta}_\ell \right|^2 \ell^2 = I_1(h) + I_2(h) \end{aligned} \quad (3.43)$$

El término $I_2(h)$ es fácil de estimar. En efecto, en virtud de (3.17), y en la medida en que los coeficientes que intervienen en la serie $I_2(h)$ no dependen de h es fácil ver que, dado $\varepsilon > 0$ arbitrario, existe $M > 0$ de modo que

$$|I_2(h)| \leq \varepsilon, \quad \forall h > 0.$$

Una vez fijado el valor de M es fácil ver que $I_1(h) \rightarrow 0$. De hecho, esta convergencia es uniforme en $t \in [0, T]$. En efecto, como, una vez fijado el valor de M , $I_1(h)$ es una suma finita, basta comprobar que, para cada ℓ fijo,

$$\hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} \rightarrow \hat{\theta}_\ell e^{i\ell t}, \quad h \rightarrow 0, \quad \text{uniformemente en } t \in [0, T],$$

lo cual es evidentemente cierto puesto que

$$\mu_\ell(h) \rightarrow \ell, \quad h \rightarrow 0$$

y de que, en vista de la elección de los catos iniciales para el sistema discreto realizado en el enunciado del Teorema 3.1,

$$\hat{\theta}_\ell(h) \rightarrow \hat{\theta}_\ell, \quad h \rightarrow 0.$$

■

Tal y como hemos mencionado anteriormente, este Teorema no proporciona un resultado muy intuitivo puesto que garantiza la convergencia de los coeficientes de Fourier pero no de las soluciones en el espacio físico. En este sentido, sería natural analizar la convergencia de la solución discreta \vec{u}_h hacia los valores de la solución continua $u = u(x, t)$ sobre los puntos del mallado (que denotamos mediante \vec{u}) tal como hicimos en el Teorema 2.1 en el caso de la ecuación del calor.

En este caso, en vista de la ausencia de efectos disipativos necesitamos hipótesis adicionales de regularidad sobre los datos iniciales de la ecuación de ondas:

Teorema 3.2 *Supongamos que $(\varphi, \psi) \in [H^2 \cap H_0^1(0, \pi)] \times H_0^1(0, \pi)$ y que elegimos los datos iniciales del problema semi-discreto como en (3.39).*

Entonces

$$\|\vec{u}_h(t) - \vec{u}(t)\|_{1,h} + \|\vec{u}'_h(t) - \vec{u}'(t)\|_h \rightarrow 0 \quad (3.44)$$

cuando $h \rightarrow 0$ uniformemente en $0 \leq t \leq T$ para todo $0 < T < \infty$.

Observación 3.2

- El espacio $H^2 \cap H_0^1(0, \pi)$ es el constituido por las funciones $\varphi \in H_0^1(0, \pi)$ tales que $\varphi_{xx} \in L^2(0, \pi)$. Dotado de la norma

$$\|\varphi\|_{H^2 \cap H_0^1(0, \pi)} = \left[\int_0^\pi |\varphi_{xx}|^2 dx \right]^{1/2}$$

constituye un espacio de Hilbert.

- El espacio $[H^2 \cap H_0^1(0, \pi)] \times H_0^1(0, \pi)$ es un marco funcional natural para resolver la ecuación de ondas. En efecto, si $(\varphi, \psi) \in [H^2 \cap H_0^1(0, \pi)] \times H_0^1(0, \pi)$, la ecuación de ondas (3.1) admite una única solución

$$u \in C([0, \infty); H^2 \cap H_0^1(0, \pi)) \cap C^1([0, \infty); H_0^1(0, \pi)). \quad (3.45)$$

Además la energía

$$E_+(t) = \frac{1}{2} \int_0^\pi [|u_{xx}(x, t)|^2 + |u_{tx}(x, t)|^2] dx \quad (3.46)$$

que es equivalente al cuadrado de la norma en este espacio se conserva en el tiempo (para comprobarlo basta con multiplicar la ecuación de ondas por $-u_{xxt}$ e integrar por partes en x).

- Cuando $\varphi \in H^2 \cap H_0^1(0, \pi)$ sus coeficientes de Fourier satisfacen

$$\sum_{\ell=1}^{\infty} |\hat{\varphi}_\ell|^2 \ell^4 < \infty. \quad (3.47)$$

Este hecho jugará un papel decisivo en la prueba del Teorema. ■

Demostración del Teorema 3.2.

Analizamos el primer término de (3.44) puesto que el segundo puede ser tratado del mismo modo.

Procediendo como en la demostración del Teorema 2.1 y distinguiendo bajas y altas frecuencias descomponemos $\vec{u}_h - \vec{u}$ del siguiente modo:

$$\begin{aligned} \vec{u}_h(t) - \vec{u}(t) &= \sum_{\ell=-M_0}^{M_0} \left[\hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} - \hat{\theta}_\ell e^{i\ell t} \right] \vec{W}_\ell(h) \\ &+ \sum_{M_0+1 \leq |\ell| \leq M} \hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} \vec{W}_\ell(h) - \sum_{|\ell| > M_0} \hat{\theta}_\ell e^{i\ell t} \vec{W}_\ell = I_1 + I_2 + I_3. \end{aligned} \quad (3.48)$$

Analizamos ahora la norma $\|I_k\|_{1,h}^2$, para cada $k = 1, 2, 3$.

- *Término I_1 .*

Tomando normas $\|\cdot\|_{1,h}$ en I_1 y utilizando las propiedades de los autovectores $\vec{W}_\ell(h)$ enunciados en el Lema 2.1 y la desigualdad (2.32) para los autovalores discretos se tiene:

$$\|I_1\|_{1,h}^2 = \sum_{\ell=-M_0}^{M_0} \lambda_\ell(h) \left| \hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} - \hat{\theta}_\ell e^{i\ell t} \right|^2 \leq \sum_{\ell=-M_0}^{M_0} \ell^2 \left| \hat{\theta}_\ell(h) e^{i\mu_\ell(h)t} - \hat{\theta}_\ell e^{i\ell t} \right|^2. \quad (3.49)$$

- *Término I_2 .*

Tenemos, por el mismo argumento,

$$\|I_2\|_{1,h}^2 \leq \sum_{M_0+1 < |\ell| \leq M} \lambda_\ell(h) |\hat{\theta}_\ell(h)|^2, \quad (3.50)$$

en virtud de la elección (3.40) de los coeficientes de Fourier de los datos iniciales del problema semi-discreto deducimos por tanto que

$$\|I_2\|_{1,k}^2 \leq C \sum_{M_0+1 < |\ell| \leq M} \left[\ell^2 |\hat{\varphi}_\ell|^2 + |\hat{\psi}_\ell|^2 \right]. \quad (3.51)$$

La convergencia de esta serie está garantizada por la hipótesis de que los datos iniciales considerados son de energía finita, i.e. $(\varphi, \psi) \in H_0^1(0, \pi) \times L^2(0, \pi)$.

- *Término I_3 .*

Este término ha de ser analizado con algo más de cuidado. En efecto, \vec{W}_ℓ representa en este caso la restricción a los puntos del mallado de las autofunciones continuas que no se obtienen en el espectro de la matriz. Por lo tanto se pierden las propiedades de ortogonalidad enunciadas en el Lema 2.1. Por tanto, en este caso aplicamos nuevamente la desigualdad triangular y obtenemos

$$\|I_3\|_{1,h} \leq \sum_{|\ell| > M_0} |\hat{\theta}_\ell| \| \vec{W}_\ell \|_{1,h} \leq \sum_{|\ell| > M_0} |\hat{\theta}_\ell| \| W_{\ell,x} \|_{L^2(0,\pi)} = \sum_{|\ell| > M_0} |\ell| |\hat{\theta}_\ell|. \quad (3.52)$$

En esta última desigualdad hemos utilizado la desigualdad elemental

$$h \sum_{j=0}^M \frac{|f(x_{j+1}) - f(x_j)|^2}{h^2} \leq \int_0^\pi f_x^2(x) dx \quad (3.53)$$

que es válida para cualquier función de $H_0^1(0, \pi)$ y cualquier $h > 0$, y por otra parte que

$$\|W_{\ell,x}\|_{L^2(0,\pi)}^2 = \int_0^\pi |W_{\ell,x}|^2 dx = \ell^2. \quad (3.54)$$

Con el objeto de poder estimar la serie que se obtiene en esta estimación es indispensable suponer que

$$\sum_{\ell} |\ell| |\hat{\theta}_{\ell}| < \infty, \quad (3.55)$$

lo cual queda plenamente garantizado si los datos iniciales (φ, ψ) pertenecen a $[H^2 \cap H_0^1(0, \pi)] \times H_0^1(0, \pi)$ puesto que, en virtud de (3.16), (3.17) y (3.47), se tiene

$$\sum_{\ell} |\ell| |\hat{\theta}_{\ell}| \leq \left(\sum_{\ell} \ell^2 |\hat{\theta}_{\ell}|^2 \right)^{1/2} \left(\sum_{\ell} \ell^{-2} \right)^{1/2} < \infty.$$

De estas desigualdades es fácil concluir la demostración del Teorema 3.2. En efecto, de (3.51), (3.52) y (3.55) vemos que dado $\varepsilon > 0$ arbitrario existe $M_0 > 0$ tal que

$$\| I_2 \|_{1,h} + \| I_3 \|_{1,h} \leq \varepsilon, \quad \forall h > 0.$$

Una vez fijado el valor de M_0 el hecho que $\| I_1 \|_{1,h} \rightarrow 0$ cuando $h \rightarrow 0$ resulta evidente de (3.49), puesto que se trata de una suma de un número finito de términos en los que cada uno tiende a cero puesto que $\hat{\theta}_{\ell}(h) \rightarrow \hat{\theta}_{\ell}$ y $\mu_{\ell}(h) \rightarrow \ell$ respectivamente.

Se observa asimismo que la convergencia es uniforme en intervalos compactos de tiempo $[0, T]$.

Esto concluye la demostración de Teorema 3.2. ■

Observación 3.3

- Un análisis cuidadoso de la demostración del Teorema 3.2 permite establecer una estimación sobre el orden de convergencia.
- El Teorema 3.2 sólo demuestra un posible resultado de convergencia entre los varios que podrían probarse mediante argumentos semejantes. Muchas otras variantes, similares a las descritas en la sección 2.2 en el contexto de la ecuación del calor, son posibles. ■

Los resultados de convergencia que hemos presentado en los dos Teoremas anteriores hacen referencia a los coeficientes de Fourier o a la restricción de las soluciones a los puntos del mallado. Sin embargo, tal y como ocurría en el caso de la ecuación del calor, estos resultados admiten también una interpretación global. Para ello es necesario introducir un operador de extensión que a una función discreta definida sobre el mallado asocie una función continua de la variable x . Para ello, tal y como se hizo en la sección 2.7 dedicada al estudio de la ecuación del calor

mediante el método de elementos finitos, dada una solución discreta \vec{u}_h de (3.18), podemos definir su extensión

$$[E\vec{u}_h](x, t) = \sum_{j=1}^N u_j(t) \phi_j(x), \quad (3.56)$$

donde $\{\phi_j\}_{j=1, \dots, N}$ son precisamente las funciones de base del método de elementos finitos que, tal y como indicamos en la sección 2.7, son continuos, lineales a trozos y satisfacen $\phi_j(x_\ell) = \delta_{j\ell}$.

Es fácil comprobar que, dada una solución discreta \vec{u}_h de (3.18), $E\vec{u}_h$ es una función perteneciente al espacio $C([0, \infty); H_0^1(0, \pi)) \cap C^1([0, \infty); L^2(0, \pi))$. Cabe por tanto plantearse la convergencia de $E\vec{u}_h$ hacia la solución $u = u(x, t)$ del problema continuo (3.9) en el espacio de la energía.

El siguiente resultado responde a esta cuestión:

Teorema 3.3 *Bajo las hipótesis el Teorema 3.2 se tiene*

$$E\vec{u}_h \rightarrow u \text{ en } C([0, T]; H_0^1(0, \pi)) \cap C^1([0, T]; L^2(0, \pi)) \quad (3.57)$$

para todo intervalo compacto $[0, T]$

Demostración

Procedemos en dos etapas.

Etapa 1. Convergencia en $C([0, T]; H_0^1(0, \pi))$.

Observamos que

$$\begin{aligned} \|E\vec{u}_h(t) - u(t)\|_{H_0^1(0, \pi)}^2 &= \int_0^\pi |\partial_x (E\vec{u}_h)(x, t) - u_x(x, t)|^2 dx & (3.58) \\ &= \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \frac{u_{j+1}(t) - u_j(t)}{h} - u_x(x, t) \right|^2 dx \\ &\leq 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \frac{u_{j+1}(t) - u_j(t)}{h} - \frac{u(x_{j+1}, t) - u(x_j, t)}{h} \right|^2 dx \\ &\quad + 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| u_x(x, t) - \frac{u(x_{j+1}, t) - u(x_j, t)}{h} \right|^2 dx \\ &= 2h \sum_{j=0}^N \left| \frac{[u_{j+1}(t) - u(x_{j+1}, t)] - [u_j(t) - u(x_j, t)]}{h} \right|^2 \\ &\quad + 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| u_x(x, t) - \frac{1}{h} \int_{x_j}^{x_{j+1}} u_x(s, t) ds \right|^2 dx = I_1 + I_2. \end{aligned}$$

Por otra parte, $I_1 = 2 \|\vec{u}_h(t) - \vec{u}(t)\|_{1, h}^2$ que, tal y como vimos en el Teorema 3.2, converge a cero uniformemente en $0 \leq t \leq T$ cuando $h \rightarrow 0$.

Además,

$$\begin{aligned}
I_2 &= 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \frac{1}{h} \int_{x_j}^{x_{j+1}} (u_x(x, t) - u_x(s, t)) ds \right|^2 dx \\
&\leq \frac{2}{h^2} \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^{x_{j+1}} (u_x(x, t) - u_x(x, t)) ds \right|^2 dx \\
&= \frac{2}{h^2} \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^{x_{j+1}} \int_s^x u_{xx}(\sigma, t) d\sigma dx \right|^2 dx \\
&\leq \frac{2}{h^2} \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} |u_{xx}(\sigma, t)| d\sigma ds \right|^2 dx \\
&= 2h \sum_{j=0}^N \left| \int_{x_j}^{x_{j+1}} |u_{xx}(\sigma, t)| d\sigma \right|^2 \\
&\leq 2h^2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} |u_{xx}(\sigma, t)|^2 d\sigma = 2h^2 \int_0^\pi |u_{xx}(x, t)|^2 dx \\
&= 2h^2 \|u(t)\|_{H^2 \cap H_0^1(0, \pi)}^2.
\end{aligned}$$

Este último término tiende a cero con un orden $O(h^2)$ cuando $h \rightarrow 0$ puesto que, tal y como se indicó en la Observación 3.2, bajo la hipótesis de regularidad de los datos iniciales del Teorema 3.2, la solución u del problema continuo (3.9) pertenece a la clase $C([0, \infty); H^2 \cap H_0^1(0, \pi))$. En realidad la convergencia es uniforme para todo $t \geq 0$.

Etapa 2. Convergencia en $C^1([0, T]; L^2(0, \pi))$.

En este caso

$$\begin{aligned}
&\|E\vec{u}'_h(t) - u_t(t)\|_{L^2(0, \pi)}^2 = \int_0^\pi \left| \sum_{k=1}^N u'_k(t) \phi_k(x) - u_t(x, t) \right|^2 dx \\
&\leq 2 \int_0^\pi \left| \sum_{k=1}^N (u'_k(t) - u_t(x_k, t)) \phi_k(x) \right|^2 dx \\
&\quad + 2 \int_0^\pi \left| \sum_{k=1}^N u_t(x_k, t) \phi_k(x) - u_t(x, t) \right|^2 dx = I_1 + I_2.
\end{aligned}$$

El primer término I_1 , en virtud de la expresión de la matriz de masa que aparece en el método de elementos finitos, admite la expresión

$$I_1 = 2 \sum_{j, k=1}^N (u'_k(t) - u_t(x_k, t)) (u'_j(t) - u_t(x_j, t)) \int_0^\pi \phi_k(x) \phi_j(x) dx$$

$$\begin{aligned}
&= 2h \sum_{k=1}^N \left[\frac{2}{3} |u'_k(t) - u_t(x_k, t)|^2 + \frac{1}{3} (u'_k(t) - u_t(x_k, t)) (u'_{k+1}(t) - u_t(x_{k+1}, t)) \right] \\
&\leq C \|\bar{u}'_h(t) - \bar{u}'(t)\|_h^2
\end{aligned}$$

que tiende a cero uniformemente en $[0, T]$ cuando $h \rightarrow 0$, por el Teorema 3.2.

Por otra parte,

$$\begin{aligned}
I_2 &= 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| u_t(x, t) - \frac{(u_t(x_{j+1}, t) - u_t(x_j, t))}{h} (x - x_j) - u_t(x_j, t) \right|^2 dx \\
&= 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^x \left[u_{tx}(s, t) - \frac{(u_t(x_{j+1}, t) - u_t(x_j, t))}{h} \right] ds \right|^2 dx \\
&= 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^x \left[u_{tx}(s, t) - \frac{1}{h} \int_{x_j}^{x_{j+1}} u_{tx}(\sigma, t) d\sigma \right] ds \right|^2 dx \\
&= 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| \int_{x_j}^x \left[\frac{1}{h} \int_{x_j}^{x_{j+1}} [u_{tx}(s, t) - u_{tx}(\sigma, t)] d\sigma ds \right] \right|^2 dx \\
&\leq 2 \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \left| 2 \int_{x_j}^{x_{j+1}} |u_{tx}(s, t)| ds \right|^2 dx \leq 8h \sum_{j=0}^N \int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} |u_{tx}(s, t)|^2 ds dx \\
&= 8h^2 \int_0^\pi |u_{tx}(s, t)|^2 ds.
\end{aligned}$$

Este término tiende a cero con un orden $O(h^2)$ puesto que $u_t \in C([0, T]; H_0^1(0, \pi))$, tal y como señalamos en la Observación 3.2. ■

3.3 Semi-discretización espacial: El método de la energía

La prueba de la convergencia de las soluciones del problema discreto a las del continuo está basada en el hecho que la energía (3.12) (resp. (3.27)) se conserva para las soluciones del problema continuo (3.9) (rep. para las del problema discreto (3.18)).

Procedemos como en la sección 2.3 en el caso de la ecuación del calor y por tanto consideramos la solución $u = u(x, t)$ de la ecuación de ondas continua (3.9) como una solución aproximada de la ecuación discreta (3.18). Tenemos entonces

$$\left\{ \begin{array}{l} \underline{u}_j''(t) + \frac{[2 \underline{u}_j(t) - \underline{u}_{j+1}(t) - \underline{u}_{j-1}(t)]}{h^2} = u_{xx}(x_j, t) + \frac{[2 \underline{u}_j(t) - \underline{u}_{j+1}(t) - \underline{u}_{j-1}(t)]}{h^2} \\ = \varepsilon_j(t), \quad j = 1, \dots, M, \quad t > 0 \\ \underline{u}_0 = \underline{u}_{M+1} = 0, \quad t > 0 \\ \underline{u}_j(0) = \varphi_j, \quad \underline{u}'_j(0) = \psi_j, \quad j = 1, \dots, M \end{array} \right. \quad (3.59)$$

con $\varphi_j = \varphi(x_j)$ y $\psi_j = \psi(x_j)$ si, por ejemplo, los datos (φ, ψ) son continuos, lo cual está garantizado en las hipótesis de los Teoremas 3.2 y 3.3 cuando $(\varphi, \psi) \in [H^2 \cap H_0^1(0, \pi)] \times H_0^1(0, \pi)$.

En el segundo miembro de (3.59) aparece un *residuo* o *error de truncación* $\vec{\varepsilon}(t) = (\varepsilon_j(t))_{j=1, \dots, M}$.

Consideramos ahora la diferencia de las soluciones continua y discreta

$$\vec{v}_h(t) = \vec{u}(t) - \vec{u}_h(t), \quad (3.60)$$

que satisface

$$\begin{cases} v_j''(t) + \frac{[2v_j(t) - v_{j+1}(t) - v_{j-1}(t)]}{h^2} = \varepsilon_j(t), & j = 1, \dots, M, t \geq 0 \\ v_0 = v_{M+1} = 0, & t \geq 0 \\ v_j(0) = v_j'(0) = 0, & j = 1, \dots, M. \end{cases} \quad (3.61)$$

Multiplicando en (3.61) por $v_j'(t)$ y sumando en j , como es propio para la obtención de la identidad de energía para las soluciones del sistema semi-discreto, obtenemos que

$$\frac{d}{dt} \left[\frac{h}{2} \sum_{j=1}^N |v_j'(t)|^2 + \frac{h}{2} \sum_{j=0}^N \left| \frac{v_{j+1}(t) - v_j(t)}{h} \right|^2 \right] = h \sum_{j=1}^N \varepsilon_j(t) v_j'(t).$$

y por tanto

$$\left| \frac{d}{dt} \left[\frac{h}{2} \sum_{j=1}^N \left[|v_j'(t)|^2 + \left| \frac{v_{j+1}(t) - v_j(t)}{h} \right|^2 \right] \right] \right| \leq \frac{h}{2} \sum_{j=1}^N [|\varepsilon_j(t)|^2 + |v_j'(t)|^2].$$

Aplicando el Lema de Gronwall deducimos que

$$\frac{h}{2} \sum_{j=1}^N \left[|v_j'(t)|^2 + \left| \frac{v_{j+1}(t) - v_j(t)}{h} \right|^2 \right] \leq \frac{T e^T}{2} \max_{0 \leq t \leq T} \|\vec{\varepsilon}(t)\|_h^2, \quad \forall t \in [0, T].$$

Basta por tanto que estimemos el error de truncación $\vec{\varepsilon}(t)$. Tenemos

$$|\varepsilon_j(t)| \leq Ch^2 \|u(t)\|_{C^4([0, \pi])}, \quad \forall j = 1, \dots, M, \forall 0 < h < h_0, \forall t \geq 0.$$

De estas dos últimas desigualdades deducimos que

$$\frac{1}{2} [\|\vec{v}(t)\|_{1,h}^2 + \|\vec{v}'(t)\|_h^2] \leq Ch^4 \|u(t)\|_{L^\infty(0,T; C^4(0,\pi))}, \quad \forall t \in [0, T].$$

Hemos probado el siguiente resultado:

Teorema 3.4 *Supongamos que los datos iniciales (φ, ψ) de la ecuación de ondas (3.9) son tales que la solución $u = u(x, t)$ satisface*

$$u \in C([0, T]; C^4([0, \pi])). \quad (3.62)$$

Entonces, para todo $0 < T < \infty$ existe una constante $C - T > 0$ tal que

$$\|\vec{u}_h(t) - \vec{u}(t)\|_{1,h}^2 + \|\vec{u}'_h(t) - \vec{u}'(t)\|_h^2 \leq C_T h^4 \quad (3.63)$$

para todo $0 \leq t \leq T$ y todo $h > 0$, donde \vec{u} denota la restricción a los puntos del mallado de la solución de la ecuación de ondas (3.9) y \vec{u}_h representa la solución del sistema semi-discreto (3.18).

Observación 3.4

- La estimación de error (3.63) garantiza que la ecuación semi-discreta (3.18) proporciona una aproximación de orden 2 de la ecuación de ondas (3.9). Se trata de un resultado natural puesto que la única discretización realizada en el esquema (3.18) es la del laplaciano a través del esquema de tres puntos que, como hemos visto en capítulos anteriores, es un esquema de aproximación de orden 2.
- El resultado del Teorema 3.4 es sólo uno de los posibles que pueden ser obtenidos mediante el método de la energía. En particular, el método de la energía puede ser utilizado para probar la convergencia del método bajo hipótesis de regularidad de la solución mucho más débiles que (3.62).
- Tal y como se observó en la sección 2.3 la ventaja del método de la energía frente al de descomposición en series de Fourier es que puede ser aplicado en un contexto mucho más amplio: coeficientes variables dependientes del espacio y del tiempo, problemas no-lineales,...

■

Observación 3.5 “Consistencia+Estabilidad=Convergencia”.

Aunque no se haya mencionado explícitamente, las demostraciones de convergencia de los resultados anteriores están inspiradas en el principio de Lax discutido en la sección 2.4 según el cual la propiedad de convergencia equivale a la de consistencia más la de estabilidad. Comentemos este aspecto brevemente.

Las demostraciones de los Teoremas 3.1 y 3.2 mediante series de Fourier reflejan perfectamente este principio. La consistencia garantiza que los autovectores y autovalores del esquema discreto se aproximan a los del continuo. La estabilidad permite truncar las series de Fourier y reducir el problema a la convergencia de una suma finita (garantizada a su vez por la consistencia) manteniendo un control uniforme en tiempo de la cola de la serie.

La demostración del Teorema 3.3, basada en el método de la energía, también refleja esta filosofía. La consistencia del esquema garantiza que una solución regular de la ecuación continua es una solución aproximada del sistema semi-discreto con una cota (del orden de $O(h^2)$) sobre el error de truncamiento. La estabilidad del esquema se ve reflejada en la propiedad de conservación de la energía para el sistema semi-discreto.

3.4 Aproximaciones completamente discretas

Utilizamos las mismas notaciones que en la sección 2.5 dedicada al estudio de las aproximaciones completamente discretas de la ecuación del calor.

El esquema completamente discreto más habitual para la aproximación numérica de la ecuación de ondas es el conocido como “leap-frog”:

$$\begin{cases} \frac{u_j^{k+1} - 2u_j^k + u_j^{k-1}}{(\Delta t)^2} = \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{(\Delta x)^2}, & j = 1, \dots, M; k \geq 0 \\ u_0^k = u_{M+1}^k = 0, & k \geq 0 \\ u_j^0 = \varphi_j, u_j^1 = \psi_j, & j = 1, \dots, M. \end{cases} \quad (3.64)$$

Obviamente, el esquema (3.64) se obtiene aplicando el clásico esquema de tres puntos centrado para aproximar tanto la segunda derivada temporal como espacial.

Se trata de un esquema de dos pasos por lo que es indispensable para su inicialización dar el valor de la solución discreta $\{u_j^k\}$ en los dos primeros niveles temporales $k = 0, 1$. Como u_j^0 es una aproximación de la solución de la ecuación de ondas $u = u(x, t)$ en el punto $(x, t) = (x_j, 0)$ es natural tener como dato inicial en el nivel $k = 0$ para el problema discreto

$$u_j^0 = \varphi_j = \varphi(x_j). \quad (3.65)$$

Por otra parte u_j^1 es una aproximación de $u = u(x, t)$ en el punto $(x, t) = (x_j, \Delta t)$. Por lo tanto, por el desarrollo de Taylor, es natural tomar como dato inicial en el esquema discreto para $h = 1$

$$\xi_j = \varphi_j + \Delta t \psi_j = \varphi(x_j) + \Delta t \psi(x_j). \quad (3.66)$$

En efecto, si u es suficientemente regular se tiene

$$u(x_j, \Delta t) = u(x_j, 0) + \Delta t u_t(x_j, 0) + O((\Delta t)^2),$$

lo cual justifica la elección (3.66).

Obviamente, para que (3.66) sea posible es indispensable que tanto φ como ψ sean funciones continuas. Cuando se eligen datos iniciales (φ, ψ) de energía finita, i.e. $(\varphi, \psi) \in H_0^1(0, \pi) \times L^2(0, \pi)$, la continuidad de φ está garantizada pero no la de ψ . En ese caso, el valor de $\psi_j = \psi(x_j)$ puede ser sustituido, por ejemplo, por una media de ψ entorno al punto $x = x_j$.

El esquema (3.64) es claramente explícito y permite calcular fácilmente el valor de la solución discreta en el paso $k + 1$ a partir de los valores en los dos pasos anteriores $k - 1$ y k .

Por otra parte, el esquema es consistente de orden dos puesto que, como indicábamos anteriormente, nos hemos limitado a aplicar el esquema centrado de tres puntos a la hora de discretizar la derivada segunda en tiempo y en espacio.

El número de Courant en este caso viene dado por

$$\mu = \Delta t / \Delta x. \quad (3.67)$$

Con esta notación el esquema (3.64) puede ser escrito de la siguiente manera:

$$u_j^{k+1} = 2u_j^k - u_j^{k-1} + \mu^2 [u_{j+1}^k - 2u_j^k + u_{j-1}^k]. \quad (3.68)$$

Conviene señalar que el modo en que los pasos de espacio y tiempo intervienen en la definición del número de Courant μ en (3.67) es muy distinto a cómo lo hacen en el caso de la ecuación del calor ((2.95)). En este caso en el número de Courant se refleja el hecho de que en la ecuación de ondas las derivadas temporales y espaciales juegan un papel completamente simétrico, lo cual queda claramente de manifiesto en la propia expresión de la ecuación de ondas (dos derivadas temporales coinciden con dos derivadas espaciales) o de la energía de la misma.

Como decíamos anteriormente, el método es consistente de orden y , por consiguiente, si $u = u(x, t)$ es una solución suficientemente regular (de clase C^4) de la ecuación de ondas y denotamos mediante \underline{u}_j^k sus restricciones a los puntos del mallado, al insertar esos valores en el esquema discreto obtenemos un error de truncatura del orden de $O((\Delta t)^2((\Delta x)^2 + (\Delta t)^2))$. Es decir \underline{u}_j^k es una solución aproximada de (3.68) que verifica

$$\underline{u}_j^{k+1} = 2 \underline{u}_j^k - \underline{u}_j^{k-1} + \mu^2 \left[\underline{u}_{j+1}^k - 2 \underline{u}_j^k + \underline{u}_{j-1}^k \right] + O((\Delta t)^2((\Delta x)^2 + (\Delta t)^2)). \quad (3.69)$$

Sin embargo, la consistencia del esquema no garantiza su convergencia, sino que es necesaria también una propiedad de estabilidad.

La estabilidad de los métodos multi-paso ha sido ya estudiada en el marco de los sistemas de ecuaciones diferenciales ordinarias. En aquél contexto veíamos que la estabilidad necesaria podía garantizarse a través de la *condición de la raíz*: El polinomio característico del método lineal multi-paso ha de tener todas las raíces de módulo menor o igual que uno y, en caso de tener alguna raíz de módulo unidad, ésta ha de ser simple.

Con el objeto de adaptar este concepto al marco del sistema discreto (3.69) conviene utilizar el desarrollo en serie de Fourier. Así, la solución discreta puede descomponerse como

$$\vec{u}^k = \sum_{\ell=1}^M \rho_\ell^k \vec{W}_\ell(\Delta x), \quad (3.70)$$

donde, como es habitual, $\vec{W}_\ell(\Delta x)$ denota el ℓ -ésimo autovector de la matriz $A_{\Delta x}$ de la discretización mediante el esquema de tres puntos del laplaciano ((2.26) con $h = \Delta x$) y ρ_ℓ^k el coeficiente de Fourier correspondiente en el paso temporal k .

Aplicando el esquema discreto en la expresión (3.70) obtenemos

$$\rho_\ell^{k+1} = 2\rho_\ell^k - \rho_\ell^{k-1} - \mu^2(\Delta x)^2 \lambda_\ell(\Delta x) \rho_\ell^k. \quad (3.71)$$

Cada una de las expresiones (3.71) es un esquema de evolución discreto en tiempo en dos pasos para cada valor de $\ell = 1, \dots, M$. Su polinomio característico es en este caso

$$P_\ell(\lambda) = \lambda^2 - [2 - \mu^2(\Delta x)^2 \lambda_\ell(\Delta x)] \lambda + 1. \quad (3.72)$$

Sus raíces son

$$\lambda_\ell^\pm = \frac{2 - \mu^2(\Delta x)^2 \lambda_\ell(\Delta x) \pm \sqrt{(2 - \mu^2(\Delta x)^2 \lambda_\ell(\Delta x))^2 - 4}}{2}. \quad (3.73)$$

En vista de esta expresión es fácil comprobar que la condición de estabilidad se satisface si y sólo si

$$\mu \leq 1. \quad (3.74)$$

Es lo que se denomina la condición de Courant-Friedrichs-Lewy.

En efecto, para comprobarlo conviene distinguir dos casos:

Caso 1: Autovalores complejos.

Esto ocurre cuando

$$(2 - \mu^2(\Delta x)\lambda_\ell(\Delta x))^2 - 4 \leq 0$$

i.e.

$$|2 - \mu^2(\Delta x)^2\lambda_\ell(\Delta x)| \leq 2. \quad (3.75)$$

Teniendo en cuenta que

$$c(\Delta x)^2 \leq (\Delta x)^2\lambda_\ell(\Delta x) \leq 4 \quad (3.76)$$

para $c > 0$ adecuado, es obvio que la condición (3.75) se verifica siempre y cuando $\mu \leq 1$. Cuando $\mu > 1$, como la cota superior en (3.76) es óptima, i.e. a medida que $h \rightarrow 0$, $\lambda_M(\Delta x) \rightarrow 4$, se deduce la existencia de coeficientes ℓ para los cuales la condición (3.75) es violada. Estos serán analizados en el segundo caso.

Volviendo al caso $\mu \leq 1$, en el que (3.75) se cumple, observamos que

$$(\lambda_\ell^\pm)^2 = \frac{1}{4} \left[|2 - \mu^2(\Delta x)^2\lambda_\ell(\Delta x)|^2 + 4 - |2 - \mu^2(\Delta x)^2\lambda_\ell(\Delta x)|^2 \right] = 1.$$

Además, las raíces son simples siempre y cuando $\mu > 0$, cosa que, evidentemente, siempre se cumple.

Caso 2: Autovalores reales:

Consideremos ahora el caso en que alguno de los autovalores es real. Esto ocurre cuando

$$(2 - \mu^2(\Delta x)^2\lambda_\ell(\Delta x))^2 \geq 4$$

lo cual exige que

$$\mu^2(\Delta x)^2\lambda_\ell(\Delta x) \geq 4. \quad (3.77)$$

Tal y como veíamos en el caso 1, esto ocurre, por ejemplo, en cuanto $\mu > 1$, para el último autovalor $\lambda_M(\Delta x)$, a condición que $h > 0$ sea suficientemente pequeño.

En este caso los dos autovalores λ_ℓ^\pm son reales y el de mayor módulo es el que corresponde al signo negativo para el cual se tiene

$$|\lambda_\ell^-| = \frac{\mu^2(\Delta x)^2\lambda_\ell(\Delta x) - 2 + \sqrt{(2 - \mu^2(\Delta x)^2\lambda_\ell(\Delta x))^2 - 4}}{2}$$

que es, evidentemente, estrictamente mayor que 1 en virtud de (3.77).

De este análisis deducimos que el método lineal multipaso que el método de leap-frog genera en cada uno de los componentes de Fourier del problema discreto es estable si y sólo se verifica la condición $\mu \leq 1$.

Como consecuencia de este análisis deducimos que

Teorema 3.5 *El esquema “leap-frog” (3.64) es un esquema convergente de orden 2 para la ecuación de ondas si y sólo si $\mu \leq 1$.*

Volveremos más adelante sobre la demostración de este resultado.

Un aspecto muy importante de la condición de estabilidad (3.74) es el relacionado con los dominios de dependencia.

En la sección 3.1 vimos, por medio de la fórmula de d’Alembert, que la solución de la ecuación de ondas continua en toda la recta real depende en el punto (x, t) de los valores de los datos iniciales en el intervalo de dependencia $[x - t, x + t]$. Para el esquema discreto (3.68) es fácil comprobar que la solución discreta en el punto $(x, t) = (j\Delta x, k\Delta t)$ depende del valor de los datos iniciales en intervalo $[(j - k)\Delta x, (j + k)\Delta x] = [x - k\Delta x, x + k\Delta x] = \left[x - \frac{k\Delta t}{\mu}, x + \frac{k\Delta t}{\mu} \right] = \left[x - \frac{t}{\mu}, x + \frac{t}{\mu} \right]$. Vemos por tanto que la condición de estabilidad $\mu \leq 1$ del esquema numérico garantiza simplemente que “el dominio de dependencia en el esquema discreto contenga al dominio de dependencia de la ecuación de ondas continua.” Esto es una condición perfectamente natural e indispensable para que el esquema numérico sea convergente. En efecto, en caso de que la condición no se cumpla, el esquema numérico ignora información esencial de los datos iniciales para la determinación de la solución del problema continuo y esto hace que la convergencia sea imposible.

Volvamos ahora sobre la demostración de Teorema 3.5. En primer lugar es fácil comprobar que cuando $\mu > 1$ el esquema diverge. En efecto, para verlo basta con utilizar que cuando $\mu > 1$, existen componentes de Fourier ℓ para los que el esquema multipaso correspondiente viola la condición de estabilidad. Esto permite construir fácilmente soluciones en variables separadas para el sistema discreto que, a medida que $h \rightarrow 0$, divergen.

La prueba de la convergencia del esquema cuando $\mu \leq 1$ puede realizarse de diversas maneras. La primera prueba es la basada en el desarrollo en serie de Fourier. No es difícil adaptar la prueba del Teorema 3.1. Basta con proceder del mismo modo y utilizar la propiedad de estabilidad del esquema para mantener controladas uniformemente en tiempo las colas de las series truncadas.

La segunda prueba de la convergencia está basada en el método de la energía. En este caso la energía de las soluciones en el nivel temporal k viene dada por

$$E^k = \frac{\Delta x}{2} \sum_{j=0}^M \left[\left[\frac{u_j^{k+1} - u_j^k}{\Delta t} \right]^2 + \left[\frac{u_{j+1}^{k+1} - u_j^{k+1}}{\Delta x} \right] \left[\frac{u_{j+1}^k - u_j^k}{\Delta x} \right] \right]. \quad (3.78)$$

No es difícil comprobar que la energía E^k se conserva en tiempo para las soluciones del problema discreto (3.64), i.e.

$$E^{k+1} = E^k, \quad \forall k \geq 0. \quad (3.79)$$

Para ello basta con multiplicar en la ecuación discreta por $\frac{1}{2\Delta t} (u_j^{k+1} - u_j^{k-1})$ y sumar con respecto al índice espacial j . Esta prueba es análoga a la de la conservación de la energía en la ecuación de ondas continua. En efecto, en la ecuación de ondas multiplicamos la ecuación por u_t e integramos con respecto a $x \in (0, \pi)$. El procedimiento presentado por el sistema discreto es la versión discreta del método continuo. Conviene sin embargo subrayar que, en vista de la

estructura simétrica del esquema de “leap-frog” la discretización introducida de u_t es también centrada y simétrica.

Pero la conservación de la energía por si sola no garantiza la estabilidad del método. Para poder garantizar que se cumple esta propiedad es preciso comprobar que la energía definida una cantidad definida positiva, cosa que no es obvio en su definición. No se difícil comprobar que esto último es cierto si y sólo si $\mu \leq 1$.

3.5 El análisis de von Neumann

En la sección 2.6 veíamos como el análisis de von Neumann permitía analizar con facilidad la estabilidad de los esquemas numéricos de aproximación, sobre todo en ausencia de condiciones de contorno.

Consideremos pues el problema de Cauchy en toda la recta real para la ecuación de ondas (3.1) y sus aproximaciones semi-discreta y completamente discretas siguientes:

• *Aproximación semi-discreta:*

$$\begin{cases} u_j'' + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} = 0, & j \in \mathbb{Z}, t \geq 0 \\ u_j(0) = \varphi_j, u_j'(0) = \psi_j, & j \in \mathbb{Z} \end{cases} \quad (3.80)$$

• *Aproximación completamente discreta:*

$$\begin{cases} \frac{u_j^{k+1} - 2u_j^k + u_j^{k-1}}{(\Delta t)^2} = \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{(\Delta x)^2}, & j \in \mathbb{Z}, k \geq 0 \\ u_j^0 = \varphi_j, u_j^1 = \xi_j, & j \in \mathbb{Z}. \end{cases} \quad (3.81)$$

No es difícil adaptar los desarrollos de la sección anterior para probar que:

- El método semi-discreto (3.80) converge y es de orden 2, cuando $\Delta x \rightarrow 0$.
- El método completamente discreto (3.81) es convergente si y sólo si $\mu = \Delta t / \Delta x \leq 1$

Comprobemos como, efectivamente, el análisis de von Neumann permite detectar las propiedades de estabilidad que estos resultados exigen.

Con las notaciones de la sección 2.6, dada la solución de (3.80) o (3.81) definimos sus transformada de Fourier $\tilde{w}(\theta, t)$ o $\tilde{w}^k(\theta)$ respectivamente. Las ecuaciones que gobiernan su evolución son entonces:

• *Método semi-discreto:*

$$\tilde{w}''(\theta, t) + \frac{a(e^{i\theta})}{(\Delta x)^2} \tilde{w}(\theta, t) = 0, \quad t \geq 0, \theta \in [0, 2\pi) \quad (3.82)$$

donde, $a(\cdot)$ está definida como en (2.140), i.e.

$$a(e^{i\theta}) = 2 - e^{i\theta} - e^{-i\theta}. \quad (3.83)$$

En (3.82) y en lo que sigue ' denota la derivada con respecto al tiempo.

• *Método completamente discreto:*

$$\frac{\check{u}^{k+1}(\theta) - 2\check{u}^k(\theta) + \check{u}^{k-1}(\theta)}{(\Delta t)^2} + \frac{a(e^{i\theta})}{(\Delta x)^2}\check{u}^k(\theta) = 0, \quad k \geq 0, \theta \in [0, 2\pi). \quad (3.84)$$

En el caso de la ecuación semi-discreta, multiplicando por $\check{u}'(\theta, t)$ se obtiene que

$$\frac{d}{dt} \left[\frac{1}{2} |\check{u}'(\theta, t)|^2 + \frac{a(e^{i\theta})}{2(\Delta x)^2} |\check{u}(\theta, t)|^2 \right] = 0.$$

Esto garantiza la estabilidad del método.

En el caso completamente discreto, para cada valor de θ obtenemos un esquema discreto de dos pasos semejante al que obteníamos en (3.71) al utilizar la separación de variables para estudiar la estabilidad del método en el intervalo finito con condiciones de contorno de Dirichlet. En este caso la propiedad de estabilidad se reduce a comprobar que las raíces del polinomio característico de (3.84), para cada valor de θ , son de módulo menor o igual a uno y, en caso de ser de módulo unidad, se trata de una raíz simple. No es difícil comprobar que esto ocurre si y sólo si $\mu \leq 1$.

3.6 El método de elementos finitos

No es difícil de adaptar el método de elementos finitos, tal y como lo desarrollamos en la sección 2.7 para la ecuación del calor, al caso de la ecuación de ondas.

Para hacerlo, conviene en primer lugar dar una formulación variacional de la ecuación de ondas (3.9).

En este caso, tal y como indicamos en la sección 3.1, el espacio de energía natural es

$$u \in C([0, \infty); H_0^1(0, \pi)) \cap C^1([0, \infty); L^2(0, \pi)) \quad (3.85)$$

y la formulación variacional correspondiente es

$$\frac{d^2}{dt^2} \int_0^\pi u(x, t) \phi(x) dx + \int_0^\pi u_x(x, t) \phi_x(x) dx = 0, \quad \forall \phi \in H_0^1(0, \pi). \quad (3.86)$$

Conviene interpretar (3.86) con cautela puesto que la regularidad indicada en (3.85) no garantiza que el primer término de (3.86) tenga un sentido clásico puesto que de (3.85) no se deduce que la función $\int_0^\pi u(x, t) \phi(x) dx$ sea de clase C^2 con respecto al tiempo. Sin embargo lo es puesto que las soluciones de la ecuación de ondas con la propiedad de regularidad (3.85) verifican también que

$$u \in C^2([0, \infty); H^{-1}(0, \pi)), \quad (3.87)$$

donde $H^{-1}(0, \pi)$ es el espacio de Sobolev dual de $H_0^1(0, \pi)$. Esto permite interpretar (3.86) como una ecuación diferencial clásica que se verifica para cada tiempo $t \geq 0$ y cada función test $\phi \in H_0^1(0, \pi)$.

Obviamente (3.85) y (3.86) han de ser completadas con las condiciones iniciales

$$u(x, 0) = \varphi(x), u_t(x, 0) = \psi(x), 0 < x < \pi. \quad (3.88)$$

De esta formulación débil de la ecuación de ondas es fácil obtener las ecuaciones de la aproximación por elementos finitos.

Para ello introducimos el espacio V_h de funciones lineales a trozos y continuas de $H_0^1(0, \pi)$, de dimensión M , asociado al mallado de paso h .

La formulación variacional más natural es entonces: *Hallar $u_h \in C^2([0, \infty); V_h)$ tal que*

$$\frac{d^2}{dt^2} \int_0^\pi u_h(x, t) \phi_j(x) dx + \int_0^\pi u_{h,x}(x, t) \phi_{j,x}(x) dx = 0, \forall t > 0, \forall j = 1, \dots, M \quad (3.89)$$

junto con las condiciones iniciales.

$$u_h(x, 0) = \varphi_h(x), u_h'(x, 0) = \psi_h(x), 0 < x < \pi. \quad (3.90)$$

Aquí y en lo sucesivo, como en la sección 2.7, $\phi_j = \phi_j(x)$ denota la j -ésima función de base del espacio V_h , con centro en el punto $x_j = jh$ del mallado. Las funciones φ_h y ψ_h son por otra parte aproximaciones del dato inicial continuo (φ, ψ) en V_h .

Utilizando la notación vectorial habitual y las matrices de masa y rigidez introducidas en el marco de la ecuación del calor, es fácil reescribir el problema como un sistema de M ecuaciones diferenciales lineales de orden 2, acopladas, con M incógnitas:

$$\begin{cases} \mathcal{M}_h \vec{u}''(t) + \mathcal{R}_h \vec{u}(t) = 0, t > 0 \\ \vec{u}(0) = \vec{\varphi}_h, \vec{u}'(0) = \vec{\psi}_h. \end{cases} \quad (3.91)$$

Los métodos desarrollados en las secciones anteriores, tanto los basados en las descomposiciones espectrales como en el método de la energía, permiten comprobar que se trata de un método convergente de orden 2.

Como es habitual en el método de elementos finitos, en una dimensión espacial, (3.91) proporciona un esquema discreto que también puede interpretarse como una variación del método semi-discreto en diferencias finitas (3.18). En efecto, (3.91) puede reescribirse como

$$h \left[\frac{2}{3} u_j''(t) + \frac{1}{6} u_{j+1}''(t) + \frac{1}{6} u_{j-1}''(t) \right] = \frac{1}{h} [2u_j(t) - u_{j+1}(t) - u_{j-1}(t)]. \quad (3.92)$$

Pero, la ventaja del método de elementos finitos reside precisamente no en este hecho sino en su versatilidad para adaptarse a situaciones más complejas donde el método de diferencias finitas es de difícil utilización (problemas no-lineales, geometrías complejas en varias variables espaciales, coeficientes variables, etc.).

La identidad de energía es también fácil de obtener. De (3.91) y teniendo en cuenta que tanto la matriz de masa como de rigidez son simétricas y definidas positivas la energía correspondiente viene dada por

$$E_h(t) = \frac{1}{2} \langle \mathcal{M}_h \vec{u}'(t), \vec{u}'(t) \rangle + \frac{1}{2} \langle \mathcal{R}_h \vec{u}_h(t), \vec{u}_h(t) \rangle, \quad (3.93)$$

donde $\langle \cdot, \cdot \rangle$ denota le producto escalar euclideo.

La energía puede también escribirse componente a componente. Obtendríamos entonces:

$$E_h(t) = \frac{h^2}{2} \sum_{j=0}^M \left[\frac{2}{3} |u'_j(t)|^2 + \frac{1}{3} u'_{j+1}(t) u'_j(t) + \left| \frac{u_{j+1}(t) - u_j(t)}{h} \right|^2 \right], \quad (3.94)$$

cuya conservación es también fácil deducir de la escritura del sistema componente a componente como en (3.92). En efecto, multiplicando en (3.92) por $u'_j(t)$ y sumando en j no es difícil verificar que la energía (3.94) se conserva.

En la expresión (3.94) de la energía y con el objeto de comprobar su carácter definido-positivo es conveniente observar que los dos primeros términos del sumatorio pueden agruparse dando lugar al cuadrado perfecto

$$\frac{1}{3} |u'_{j+1}(t) + u'_j(t)|^2.$$

Una vez que se ha obtenido el esquema de aproximación semi-discreto (3.91) y probado su convergencia, no es difícil introducir esquemas completamente discretos basados en el método de elementos finitos. Bastaría por ejemplo utilizar diferencias finitas centradas en la discretización de la segunda derivada temporal de (3.91). Pero el análisis de la convergencia de este método y, en particular, de la condición de estabilidad en función del número de Courant queda fuera del contenido de este texto y constituye un excelente ejercicio para que el lector evalúe su dominio de las materias abordadas en estas notas.

Bibliografía

- [B] H. Brezis, *Analyse Fonctionnelle, Théorie et Applications*, Masson, Paris, 1983.
- [E] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Vol.19, AMS, 1998
- [I] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 1996.
- [IK] E. Isaacson and H.B. Keller, *Analysis of Numerical Methods*, John Wiley & Sons, 1966.
- [J] F. John, *Partial differential Equations*, (4. ed), Springer, 1982.
- [L] J.L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris, 1969.
- [Q] F. Quirós, *Notas de Cálculo Numérico*. 2002.
- [Zi] O. C. Zienkiewicz, Achievements and some unsolved problems of the finite element method, *Int. J. Numer. Meth. Engng.* **47** (2000), 9–28.