



Sección de Matemáticas
Universidad de La Laguna

Ricardo Calvo Castro

*Métodos numéricos para
la resolución de ecuaciones
diferenciales y aplicaciones*

Numerical methods for differential equations and
applications

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, Septiembre de 2019

DIRIGIDO POR

María Soledad Pérez Rodríguez

María Soledad Pérez Rodríguez
Departamento de Análisis
Matemático
Universidad de La Laguna
38200 La Laguna, Tenerife

Agradecimientos

A mi tutora María Soledad Pérez Rodríguez, a su incombustible paciencia y dedicación. A mi familia, por la perseverancia en mi educación y por lograr, gracias a su apoyo, que no me rindiera en el peor de los momentos. Dedico este trabajo a mi abuela Sabina Peña Álvarez.

Ricardo Calvo Castro
La Laguna, 12 de septiembre de 2019

Resumen · Abstract

Resumen

El objetivo de este trabajo es el estudio de algunos métodos para la resolución de problemas de valor inicial (PVI) en ecuaciones diferenciales ordinarias (EDO). En particular, nos centraremos en tres métodos básicos: el método de Euler explícito, el de Euler implícito y la regla trapezoidal. Demostraremos teoremas clásicos de convergencia de dichos métodos así como su orden y su estabilidad lineal. Posteriormente, estudiaremos la aplicación de dichos métodos a dos problemas representativos. El primero es un PVI lineal que proviene de la discretización espacial de la ecuación del calor, y el segundo es un problema no lineal que modeliza una reacción química compleja.

Palabras clave: *Métodos numéricos – Ecuaciones diferenciales ordinarias – Método de Euler – Regla Trapezoidal.*

Abstract

The study of some numerical methods for the solution of initial value problems (IVP) in ordinary differential equations (ODE) is the main goal of this work. In particular, we will focus our attention on three basic schemes: the forward Euler method, the backward Euler and the trapezoidal rule. We will show classic theorems on the convergence of these methods, as well as their properties of order and linear stability. In the last chapters we will apply these methods to two representative problems. The first one is a linear IVP that comes from the spatial discretization of the heat equation, and the second one is a non-linear problem that models a complex chemical reaction.

Keywords: *Numerical methods – Ordinary differential equations – Euler Method – Trapezoidal Rule.*

Contenido

Agradecimientos	III
Resumen/Abstract	V
Introducción	IX
1. Métodos para la resolución de problemas de valor inicial	1
1.1. Método de Euler Explícito	1
1.1.1. Consistencia y convergencia del método	3
1.1.2. Estabilidad absoluta del método	7
1.2. Problemas Stiff	10
1.3. Método de Euler Implícito	12
1.3.1. Consistencia y convergencia del método	13
1.3.2. Estabilidad absoluta del método	17
1.4. Regla Trapezoidal	18
1.4.1. Consistencia y convergencia del método	20
1.4.2. Estabilidad absoluta del método	24
1.5. Implementación de los métodos implícitos	24
1.5.1. Iteración funcional	25
1.5.2. Método de Newton	25
1.5.3. Experimentos numéricos	26
2. Problemas lineales	29
2.1. Aplicación a ecuaciones en derivadas parciales	29
2.2. Método de líneas	30
2.3. Resolución temporal	32
2.4. Experimentos numéricos	33
2.4.1. Orden de los métodos	34

3. Problemas no lineales	39
3.1. Modelo de Robertson	39
3.2. Método de Newton	40
3.2.1. Euler Implícito	41
3.2.2. Regla del Trapecio	42
3.3. Experimentos numéricos	43
3.3.1. Implementación de los métodos a paso fijo	43
3.3.2. Implementación a paso variable de Matlab	44
Bibliografía	49
Poster	51

Introducción

Los problemas de valor inicial (PVI) con sistemas de ecuaciones diferenciales ordinarias (EDOs) aparecen en muchos campos de la ciencia y la tecnología cuando se usan técnicas de modelización matemática para describir una gran variedad de fenómenos. Salvo en el caso de ciertas EDOs sencillas, escalares o de dimensión moderada, el cálculo "exacto" de la solución de estos problemas es tan complicado que los métodos numéricos resultan fundamentales para obtener aproximaciones a dicha solución de manera precisa y eficiente.

Este Trabajo Fin de Grado (TFG) es una introducción a algunos de los métodos numéricos que se suelen utilizar en la práctica para la resolución de PVI que pueden ser no lineales y de gran dimensión: el método de Euler explícito, el de Euler implícito y la regla trapezoidal.

En el Capítulo 1, introduciremos estos esquemas y veremos la demostración de algunos teoremas que aseguran su convergencia sobre PVI generales. También introduciremos una clase de PVI que hacen que los métodos explícitos no funcionen adecuadamente independientemente del orden que alcancen en teoría, los problemas stiff. Con ello se ve la necesidad práctica de utilizar métodos implícitos como el Euler implícito y la regla trapezoidal. Sin embargo, estos métodos implícitos traen consigo el inconveniente de que hay que resolver sistemas implícitos en cada paso de la integración numérica. Al final de este capítulo se proponen dos esquemas iterativos que se pueden usar para resolver estos sistemas implícitos, la iteración funcional y el método de Newton, y se testan numéricamente sobre un problema lineal stiff, con lo que se muestra claramente la mayor fiabilidad de los esquemas de tipo Newton sobre este tipo de PVI. En la elaboración de este capítulo principalmente hemos seguido los trabajos [2], [1], [3], [7].

En el Capítulo 2 presentamos una aplicación de estos métodos para la resolución de los PVI de gran dimensión que provienen de la discretización espacial de una ecuación en derivadas parciales parabólica, la ecuación del calor,

siguiendo las ideas dadas en [5]. Con los experimentos numéricos podemos ver que la stiffness del problema resultante dependerá de manera fundamental de la partición espacial que se elija, por lo que el método de Euler explícito queda superado claramente por los métodos implícitos, aunque el coste computacional de estos sea mayor.

Finalmente, en el Capítulo 3 veremos una aplicación de estos métodos a un PVI representativo de los problemas no lineales stiff, el problema de Robertson [4] que modeliza una reacción química, que es un problema complicado de resolver incluso para métodos más sofisticados que los que vemos en esta memoria.

Métodos para la resolución de problemas de valor inicial

El objetivo de esta memoria será el estudio de los métodos numéricos que aproximan la solución del Problema de Valor Inicial (PVI) en ecuaciones diferenciales ordinarias (EDOs) de primer orden

$$\begin{cases} y(a) = \alpha \\ y'(x) = f(x, y) \text{ con } x \in [a, b] \end{cases} \quad (1.1)$$

suponiendo que $f(x, y)$ es *continua* y *Lipschitz* respecto de y , lo que significa que $\exists L > 0$ tal que:

$$\|f(x, y) - f(x, \bar{y})\| \leq L\|y - \bar{y}\|, \quad \forall x \in [a, b] \quad y \quad \forall y, \bar{y} \in \mathbb{R}^m$$

El primer método que introduciremos es el método de Euler explícito. Aunque hoy día es un método poco utilizado en las aplicaciones, sigue siendo el primero que se explica en cualquier curso de este tema, ya que, por una parte, Leonhard Euler en 1768 fue el primero que propuso que con este método se podrían calcular aproximaciones a las soluciones de ecuaciones diferenciales que no se podían escribir de forma cerrada y, por otra, con este esquema se puede explicar fácilmente cómo van a aplicarse los otros métodos de un paso que veremos a continuación.

En este capítulo seguiremos principalmente la introducción de los métodos que se da en [1], complementando el texto con, [3], [4] y [7].

1.1. Método de Euler Explícito

Tomando la ecuación (1.1) dividiremos el intervalo de integración $[a, b]$ en n subintervalos de longitud $h = \frac{b-a}{n}$. Así tendremos definidas las n partes de la forma $x_i = x_0 + ih$, $i = 0, \dots, n$, con $x_0 = a$ y $x_n = b$.

Con el *Desarrollo de Taylor* se tiene, como $x_{i+1} = x_i + h$:

$$y(x_{i+1}) = y(x_i) + hy'(x_i) + \frac{1}{2}h^2y''(x_i) + \dots \quad (1.2)$$

Reescribimos la ecuación (1.2) de la siguiente forma:

$$y(x_{i+1}) = y(x_i) + hy'(x_i) + h\tau_i(h)$$

Tomando $y'(x_i) = \frac{dy}{dx}(x_i) = f(x_i, y(x_i))$ tenemos:

$$y(x_{i+1}) = y(x_i) + hf(x_i, y(x_i)) + h\tau_i(h) \quad (1.3)$$

donde $\tau_i(h) = \frac{1}{2}hy''(x_i) + \dots \rightarrow 0 \quad (h \rightarrow 0)$

Si despreciamos $\tau_i(h)$ y tomamos y_i como aproximación de $y(x_i)$, el *Método de Euler Explícito* se define como:

$$\begin{cases} y_0 = \alpha \\ y_{i+1} = y_i + hf(x_i, y_i) \text{ con } i = 0, \dots, n-1. \end{cases} \quad (1.4)$$

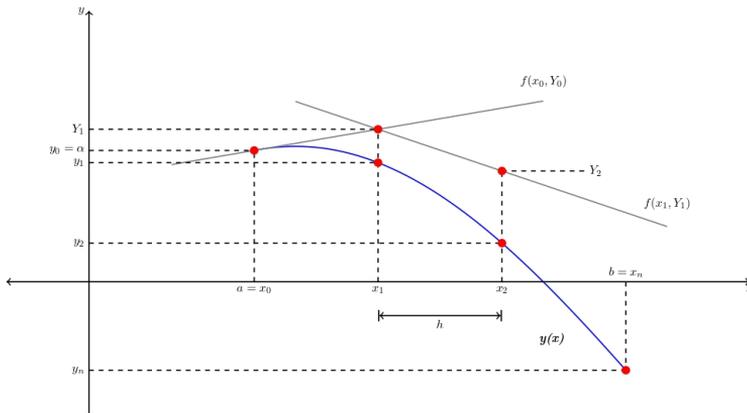


Figura 1.1: Interpretación gráfica del método de Euler

1.1.1. Consistencia y convergencia del método

Comenzamos reescribiendo el *método* (1.4) de la siguiente forma:

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i) \quad (1.5)$$

Dada una función $u(x)$, con $x \in [a, b]$ y $h > 0$, usando la *diferencia finita progresiva* se define el operador:

$$\Delta_h u(x_{i+1}) \equiv \frac{u(x_{i+1}) - u(x_i)}{h} - f(x_i, u(x_i)) \quad \text{con } i = 0, \dots, n-1$$

si tomamos $u = y_h$ como una función de malla cualquiera que toma sobre los puntos x_i el valor y_i del método (1.4):

$$\begin{aligned} \Delta_h y_h(x_{i+1}) &= \frac{y_h(x_{i+1}) - y_h(x_i)}{h} - f(x_i, y_h(x_i)) = \frac{y_{i+1} - y_i}{h} - f(x_i, y_i) = 0 \\ \text{con } i &= 0, \dots, n-1 \quad e \quad y_0 = \alpha. \quad (y_h(x_0) = \alpha) \end{aligned}$$

Para estudiar el error de este método tenemos que tener en cuenta que los errores cometidos en cada paso se van acumulando. En primer lugar, considerando $u(x) = y(x)$ la solución exacta de (1.3) tenemos que:

$$\Delta_h y(x_{i+1}) = \frac{y(x_{i+1}) - y(x_i)}{h} - f(x_i, y(x_i))$$

a estas diferencias $\Delta_h y(x_{i+1})$ se les llama *Error local de truncamiento* y por (1.3):

$$\Delta_h y(x_{i+1}) = \tau_i(h) = \mathcal{O}(h) \quad (h \rightarrow 0)$$

y por tanto el método es consistente de orden 1.

Un diseño sencillo de las aproximaciones nos llevará a aproximaciones consistentes. Sin embargo, a nosotros más que la consistencia nos interesa el estudio de la *convergencia*.

Para el estudio de la convergencia necesitamos tomar como herramienta el *error global del método*, es decir:

$$\begin{aligned} e_i(h) &= y(x_i) - y_i, \quad i = 0, \dots, n. \\ e_0(h) &= 0 \end{aligned} \quad (1.6)$$

se debe de observar que hay una relación de proporcionalidad inversa entre h y n , es decir $h \cdot n = b - a$.

En el caso del método de Euler su error global (1.6), se obtiene de (1.3) y (1.4):

$$e_{i+1}(h) = e_i(h) + h \cdot [f(x_i, y(x_i)) - f(x_i, y_i)] + h \cdot \tau_i(h)$$

Tomamos normas para la acotación del error y sabiendo que la función derivada f es *Lipschitz continua*:

$$\|e_{i+1}(h)\| \leq \|e_i(h)\| + hL\|y(x_i) - y_i\| + h\|\tau_i(h)\|$$

Por tanto:

$$\|e_{i+1}(h)\| \leq (1 + hL)\|e_i(h)\| + h\|\tau_i(h)\| \quad , \forall i = 0, \dots, n-1.$$

Utilizamos una notación más sencilla $E_i = \|e_i(h)\|$ para continuar con la búsqueda de la cota del error. Así tendríamos:

$$\begin{aligned} E_i &\leq (1 + hL)E_{i-1} + h\|\tau_{i-1}(h)\| \quad , \forall i = 1, \dots, n. \\ E_0 &= 0 \end{aligned} \quad (1.7)$$

En consecuencia,

$$\begin{aligned} E_i &\leq (1 + hL)E_{i-1} + h\|\tau_{i-1}(h)\| \\ &\leq (1 + hL)[(1 + hL)E_{i-2} + h\|\tau_{i-2}(h)\|] + h\|\tau_{i-1}(h)\| \\ &\dots \\ &\leq (1 + hL)^i E_0 + \sum_{j=0}^{i-1} (1 + hL)^j h\|\tau_{i-1-j}(h)\| \end{aligned}$$

Sabiendo que:

$$\left. \begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \dots \simeq 1 + x \quad \text{con } x \rightarrow 0 \\ x_i &= a + ih \Rightarrow ih = x_i - a \end{aligned} \right\} \Rightarrow$$

$$\Rightarrow (1 + hL)^i \leq e^{(Lh)i} \leq e^{L(x_i - a)} \quad (1.8)$$

para todo h , $0 < h \leq h^*$, siendo h^* suficientemente pequeño para que $1 + h^*L \leq e^{h^*L}$.

Como $E_0 \approx 0$ y tomando (1.8):

$$E_i \leq \sum_{j=0}^{i-1} e^{L(x_i-a)} h \|\tau_{i-1-j}(h)\|$$

lo que se puede ver como una *suma inferior de Riemann* de $e^{L(x-a)}$ en $[a, x_i]$, y por tanto,

$$\sum_{j=0}^{i-1} e^{L(x_j-a)} h \leq \int_a^{x_i} e^{L(x-a)} dx = \frac{1}{L} [e^{L(x_i-a)} - 1]$$

contando con todo lo anterior, obtenemos:

$$E_i \leq \frac{1}{L} [e^{L(x_i-a)} - 1] \|\tau_{i-1-j}(h)\| \quad , i = 0, \dots, n$$

Llegados a este punto, se tiene que el *error local de truncamiento* verifica:

$$\tau_i(h) = \frac{y(x_{i+1}) - y(x_i)}{h} - f(x_i, y(x_i)) = \frac{h}{2} \cdot y''(x_i) + \frac{h^2}{3!} \cdot y'''(x_i) + \dots$$

si suponemos que la solución es suficientemente suave tal que $\|y''\|$ está acotada en $[a, b]$, entonces, por el resto de Lagrange:

$$\|\tau_i(h)\| \leq \frac{h}{2} M \quad \text{con} \quad M = \max_{x \in [a, b]} \|y''(x)\| \tag{1.9}$$

finalmente obtenemos:

$$E_i \leq \left[\frac{M}{2L} [e^{L(x_i-a)} - 1] \right] h \quad , i = 0, \dots, n$$

por lo que:

$$\begin{aligned} \max_{0 \leq i \leq n} \|e_i\| &\leq \max_{0 \leq i \leq n} \frac{M}{2L} [e^{L(x_i-a)} - 1] h \\ &\leq \frac{M}{2L} [e^{L(b-a)} - 1] h, \quad 0 < h \leq h^*. \end{aligned} \tag{1.10}$$

Por tanto, el *Método de Euler* es *convergente de orden 1*.

Ejemplo 1.1 Tomamos el problema siguiente:

$$\begin{aligned} y_1' &= -4y_1 + \frac{3}{5}y_2 + \frac{11}{5}y_3 + \frac{2}{5} & y_1(0) &= 2 \\ y_2' &= -3y_1 - \frac{14}{5}y_2 + \frac{7}{5}y_3 - \frac{1}{5} & y_2(0) &= 1 \\ y_3' &= -\frac{3}{5}y_1 + 3y_2 + \frac{2}{5}y_3 + \frac{1}{5} & y_3(0) &= 0 \end{aligned} \quad (1.11)$$

con $x \in [0, 1]$ y solución exacta:

$$\begin{aligned} y_1(x) &= 2e^{-x} + e^{-5x} - 1 \\ y_2(x) &= -e^{-x} + 2e^{-5x} \\ y_3(x) &= 3e^{-x} - e^{-5x} - 2 \end{aligned} \quad (1.12)$$

calculamos el jacobiano $f_{y_j} = \frac{\partial f}{\partial y_j}$:

$$J = \begin{pmatrix} -4 & \frac{3}{5} & \frac{11}{5} \\ -3 & -\frac{14}{5} & \frac{7}{5} \\ -\frac{3}{5} & 3 & \frac{2}{5} \end{pmatrix}$$

con $j = 1, 2, 3$.

Y por tanto la constante de Lipschitz es $L = \|J\|_\infty = \frac{34}{5}$. Hallando las derivadas segundas de las componentes de la solución exacta tenemos, por (1.9):

$$M = \max_{x \in [0,1]} \|y''(x)\|_\infty = \max_{\substack{x \in [0,1] \\ j=1,2,3}} \|y_j''(x)\|$$

$$M = \max_{j=1,2,3} |y_j''(0)| = 49$$

por lo tanto, la acotación del error es:

$$\max_{0 \leq i \leq n} \|e_i\| \leq \frac{245h}{68} [e^{34/5} - 1] \quad (1.13)$$

En la Tabla 1.1 podemos comparar el error teórico dado por (1.13) frente al error computado, en el punto final de integración $\|y_n - y(1)\|_\infty$. Está claro que esta cota teórica sobrevalora el error realmente obtenido, pero hay que tener en cuenta que (1.10) es una cota para cualquier PVI con constante de Lipschitz L y derivada segunda acotada por M . \square

n	Error teórico	Error computado
10	$3.23128804 \cdot 10^2$	$5.18416187 \cdot 10^{-2}$
20	$1.61564402 \cdot 10^2$	$2.46138212 \cdot 10^{-2}$
40	$8.07822009 \cdot 10^1$	$1.19929091 \cdot 10^{-2}$

Tabla 1.1: Comparativa de errores para (1.11).

Ejemplo 1.2 Consideramos el problema:

$$\begin{cases} y(1) = 1 \\ y'(x) = -5xy^2 + \frac{5}{x} - \frac{1}{x^2} \text{ con } x \in [1, 25] \end{cases} \quad (1.14)$$

con solución conocida $y(x) = \frac{1}{x}$.

Estimamos el jacobiano f_y :

$$f_y = \frac{\partial f}{\partial y} = -10xy$$

Por lo tanto, en un entorno cercano de la solución:

$$L = \max_{1 \leq x \leq 25} \left| \frac{\partial f}{\partial y} \right| \approx |-10| = 10$$

y como hemos denotado en (1.9), M es de la forma:

$$M \geq \frac{2}{x^3} \text{ con } x \in [1, 25] \Rightarrow M = 2$$

con estos datos acotamos el error:

$$\max_{0 \leq i \leq n} |e_i| \leq \frac{h}{10} [e^{10(x_i-1)} - 1], \quad x \in [1, 25]$$

por lo tanto $\max_{0 \leq i \leq n} |e_i| \leq \frac{h}{10} [e^{240} - 1]$ que como hemos dicho no resulta muy útil.

De hecho, cuanto mayor sea L , la cota teórica sobrevalorará más la computada. \square

1.1.2. Estabilidad absoluta del método

Contemplemos ahora la posibilidad de tomar unos tamaños de paso relativamente grandes. Así, aparece el concepto de estabilidad, que será decisivo para el buen funcionamiento de los métodos, como veremos a continuación.

Consideramos la siguiente ecuación:

$$y' = \lambda y \quad (1.15)$$

donde λ es una constante y además $\lambda \in \mathbb{C}$, dado que posteriormente corresponderá a un *autovalor* de los sistemas lineales de EDO's. Tomando el valor inicial $y(0) = c$, con $c > 0$, la solución exacta es:

$$y(x) = ce^{\lambda x}$$

Aplicamos ahora el método de Euler para la ecuación anterior:

$$y_j = y_{j-1} + h\lambda y_{j-1} = (1 + h\lambda)y_{j-1} = (1 + h\lambda)^j c \quad \text{con } j = 1, \dots, n1$$

Si representamos λ en forma binómica $\lambda = \alpha + \beta i$, con $\alpha, \beta \in \mathbb{R}$, podemos representar la solución exacta como:

$$y(x_j) = c(e^{\alpha x_j} e^{i\beta x_j}) \quad (1.16)$$

Además, tenemos que la solución numérica verifica:

$$y_j = c[1 + h(\alpha + i\beta)]^j = c[\rho e^{i\theta}]^j \quad (1.17)$$

con

$$\begin{cases} \rho = \sqrt{(1 + h\alpha)^2 + (h\beta)^2} \\ \theta = \arctan\left(\frac{h\beta}{1 + h\alpha}\right) \end{cases} \quad (1.18)$$

Dependiendo de α distinguiremos tres casos:

- $\alpha = 0$, esto implica que:

$$\begin{cases} \rho = \sqrt{1 + (h\beta)^2} \\ \theta = \arctan(h\beta) \end{cases}$$

y por lo tanto:

$$y_j = c(1 + h^2\beta^2)^{\frac{j}{2}} [\cos(j\theta) + i \sin(j\theta)]$$

la solución oscila y la distancia entre las curvas solución permanece igual.

- $\alpha > 0$, esto implica que para (1.18) tenemos que $\rho > 1$ y por lo tanto:

$$y_j = c((1 + h\alpha)^2 + (h\beta)^2)^{\frac{j}{2}} [\cos(j\theta) + i \sin(j\theta)]$$

resulta así que $|y(x)| \rightarrow \infty$ crece y la solución numérica y_j es inestable. De todos modos para algunos casos puede ser interesante emplear este método para un $e^{\alpha x_n}$ no demasiado grande, calculando así soluciones *relativas*.

- $\alpha < 0$, en este caso la solución exacta decae exponencialmente y en cuanto a la solución numérica (1.17), se tiene que:
 - $\rho > 1$ si y solo si:

$$2\alpha + h(\alpha^2 + \beta^2) > 0 \Leftrightarrow h > \frac{-2\alpha}{\alpha^2 + \beta^2} > 0$$

Por lo tanto, como aplicaremos el método para valores pequeños de h , este caso no va a considerarse en la práctica.

- $\rho \leq 1$ si y solo si:

$$h \leq \frac{-2\alpha}{\alpha^2 + \beta^2} = 2 \frac{-Re(\lambda)}{|\lambda|^2} \quad \text{con} \quad Re(\lambda) < 0 \quad (1.19)$$

es decir, aparece una restricción de paso que puede ser grande.

Este es el caso que tiene interés a la hora de trabajar con los métodos, pues aparece a menudo en la práctica. Por ello, será esencial que los métodos cumplan el siguiente requisito.

Definición 1.1 Un método es **estable** sobre el problema (1.15) si verifica:

$$|y_{i+1}| \leq |y_i|, \quad \text{con} \quad i = 0, \dots, n - 1 \quad (1.20)$$

y además, para todo método numérico se define su *región de estabilidad absoluta* como la región del z -plano complejo de manera que aplicando el método a la ecuación (1.15) con $z = \lambda h$ dentro de dicha región, produce una solución aproximada que satisface el requisito de estabilidad absoluta (1.20).

En el método de Euler este requisito se verifica si

$$|1 + h\lambda| \leq 1$$

lo que coincide con lo que vimos anteriormente en (1.19) cuando $Re\lambda \leq 0$ En

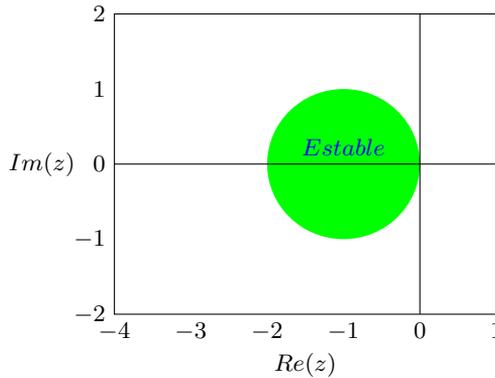


Figura 1.2: Región de estabilidad absoluta del método de Euler.

la Figura 1.2 podemos ver la región del z -plano en la que el método verifica (1.20). Si $\lambda = \alpha < 0$, es real ($\beta = 0$) tenemos una restricción de la forma:

$$h \leq -\frac{2}{\lambda}$$

En el Ejemplo 1.1 dado que el jacobiano es de la forma $f_y \approx -10 \approx \lambda$ obtenemos que esta restricción es:

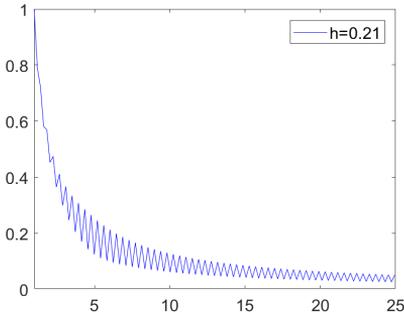
$$h \leq \frac{2}{-(-10)} = 0.2$$

En este caso la acotación no es excesivamente restrictiva, pero tenemos que tenerla en consideración. Aquí podemos observar lo que supone tener tamaños de paso muy cercanos al límite de estabilidad marcado y la diferencia tan grande que supone estar en el interior de la región de estabilidad o fuera.

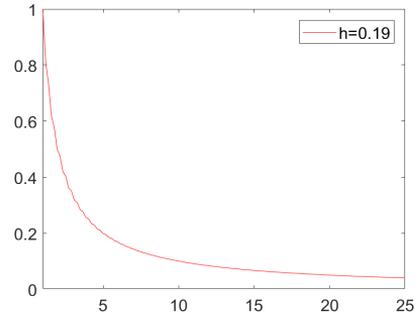
1.2. Problemas Stiff

La elección del tamaño de paso h depende de la precisión que queramos para la aproximación que obtenemos con el método, pero como vimos con el método Explícito de Euler, al entrar en juego la restricción de estabilidad absoluta debemos de tomar unos tamaños de paso suficientemente pequeños para que esto se cumpla. Un PVI es *stiff* si para ciertos métodos numéricos de resolución de ecuaciones es numéricamente inestable a menos que se considere un tamaño de paso extremadamente pequeño.

Con un ejemplo ilustraremos estos casos en el método de Euler explícito y veremos que nos llevará al planteamiento de un nuevo método para evitar el enfrentamiento con este tipo de ecuaciones.



(a) Número de pasos = 114



(b) Número de pasos = 126

Figura 1.3: Método de Euler (1.15) en el problema (1.14) con $h = 0.21$ y $h = 0.19$

Ejemplo 1.3 Consideramos el problema escalar:

$$y' = -100(y - \sin(x)), \quad x \geq 0, \quad y(0) = 1 \quad (1.21)$$

este problema tiene solución:

$$y = \left[1 + \frac{100}{100^2 + 1} \right] e^{-100x} + \frac{100}{\sqrt{100^2 + 1}} \sin(x + \alpha)$$

$$\alpha = \arctan\left(\frac{-1}{100}\right) = 9.996 \cdot 10^{-3}$$

que empieza con el valor inicial dado y después de poco tiempo, por ejemplo. $x \geq 0.03$, $y(x)$ varía de forma mucho más lenta verificando $y(x) \simeq \sin(x)$.

Esto se aprecia fácilmente observando las gráficas de la aproximación a la solución de la ecuación dada. Como podemos ver, si nos tomamos un $h > 0.01$ el número de pasos es más razonable, pero el método nos da una solución inestable, pero con un $h \leq 0.01$ tenemos una solución aproximada estable, pero como vemos en la gráfica (b) de la Figura 1.4 necesitamos una cantidad muy elevada de pasos, exactamente 600. \square

Así podemos observar que:

Un PVI es stiff en un intervalo $[0, b]$ si el tamaño de paso necesario para mantener la estabilidad del método de Euler explícito es mucho más pequeño que el tamaño requerido para representar la solución con precisión.

Observamos que stiffness depende, además de la ecuación diferencial en sí, de depende del criterio de precisión, de la longitud del intervalo de integración y de

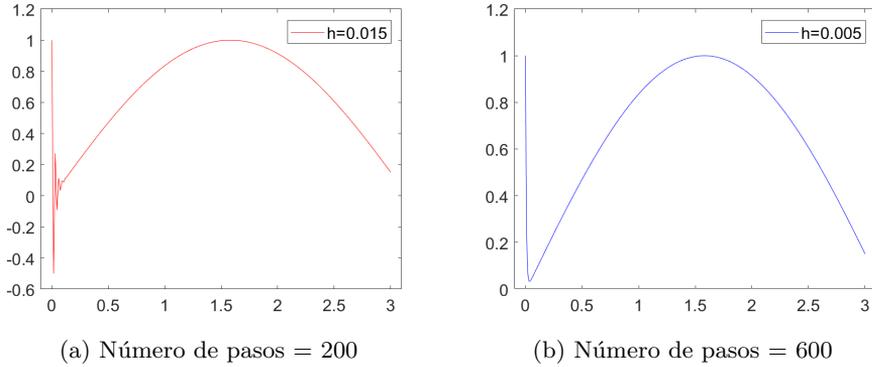


Figura 1.4: Euler (1.15) en el problema (1.21).

la región de estabilidad absoluta del método.

Para sistemas lineales, homogéneos y estables, stiffness se puede determinar por el sistema de sus autovalores.

Para la ecuación (1.15) en el intervalo $[0, b]$ decimos que el problema es stiff si:

$$bRe(\lambda) \ll -1$$

Llegados a este punto lo que nos interesa encontrar, para trabajar con él, es un método que aplicado a la ecuación (1.15) sea absolutamente estable incluso cuando $bRe(\lambda) \ll -1$.

En este caso, si suponemos por simplicidad que $\lambda \in \mathbb{R}$, $\lambda \ll -\frac{1}{b}$, como ya vimos en la sección anterior, para aplicar el método de Euler habría que tomar tamaño de paso $h \leq -\frac{2}{\lambda}$ muy pequeño.

Por tanto, si conseguimos métodos cuya región de estabilidad incuya todo el semi-plano complejo negativo $\mathbb{C}^- = \{z \in \mathbb{C} / Re(z) \leq 0\}$, no sufrirán restricción de paso cuando se aplican a problemas lineales stiff, al menos los escalares, aunque en la práctica se comprueba que también la estabilidad absoluta lineal mejora el comportamiento de los métodos sobre PVIs generales.

1.3. Método de Euler Implícito

En este caso, la diferencia que tenemos con respecto al método explícito es que desarrollamos Taylor centrado en x_{i+1} de la siguiente forma:

$$y(x_i) = y(x_{i+1}) - h \left[\frac{dy}{dx}(x_{i+1}) + \tau_i(h) \right]$$

Como $\frac{dy}{dx}(x_{i+1}) = f(x_{i+1}, y(x_{i+1}))$ por (1.1), despejamos $y(x_{i+1})$ y llegamos a la ecuación:

$$y(x_{i+1}) = y(x_i) + hf(x_{i+1}, y(x_{i+1})) + h\tau_i(h) \quad (1.22)$$

de forma análoga al caso del Euler explícito, despreciando $\tau_i(h)$ obtenemos el *Método de Euler Implícito*:

$$\begin{cases} y_0 = \alpha \\ y_{i+1} = y_i + hf(x_{i+1}, y_{i+1}) \text{ con } i = 0, \dots, n-1. \end{cases} \quad (1.23)$$

1.3.1. Consistencia y convergencia del método

Comenzamos reescribiendo el *método* (1.23) de la siguiente forma:

$$\frac{y_{i+1} - y_i}{h} = f(x_{i+1}, y_{i+1}) \quad (1.24)$$

Dada una función $u(x)$, con $x \in [a, b]$ y $h > 0$, usando la *diferencia finita regresiva* se define el operador:

$$\nabla_h u(x_{i+1}) \equiv \frac{u(x_{i+1}) - u(x_i)}{h} - f(x_{i+1}, u(x_{i+1})) \quad \text{con } i = 0, \dots, n-1$$

Si tomamos $u = y_h$ como una función de malla cualquiera que toma sobre los puntos x_i el valor y_i del método (1.23):

$$\begin{aligned} \nabla_h y_h(x_{i+1}) &= \frac{y_h(x_{i+1}) - y_h(x_i)}{h} - f(x_{i+1}, y_h(x_{i+1})) = \\ &= \frac{y_{i+1} - y_i}{h} - f(x_{i+1}, y_{i+1}) = 0 \\ &\text{con } i = 0, \dots, n-1 \quad \text{e } y_0 = \alpha. \quad (y_h(x_0) = \alpha) \end{aligned}$$

Análogamente al método de Euler Explícito para el estudio del error tenemos que tener en cuenta que los errores cometidos en cada paso se van acumulando. Considerando $u(x) = y(x)$ la solución exacta de (1.22) tenemos que:

$$\nabla_h y(x_{i+1}) = \frac{y(x_{i+1}) - y(x_i)}{h} - f(x_{i+1}, y(x_{i+1}))$$

donde $\nabla_h y(x_{i+1})$ es el *Error local de truncamiento* y por (1.22):

$$\nabla_h y(x_{i+1}) = \tau_i(h) = \mathcal{O}(h) \quad (h \rightarrow 0) \quad (1.25)$$

y por tanto el método es consistente de orden 1.

Para estudiar la convergencia partimos del *error global del método*:

$$\begin{aligned} e_i(h) &= y(x_i) - y_i, \quad i = 0, \dots, n. \\ e_0(h) &= 0 \end{aligned} \quad (1.26)$$

Para que esta fórmula tenga sentido, hay que asegurar la existencia de un único y_{i+1} que verifique (1.23) dado $\{y_i, h\}$. Si tomamos [7, Lema 4.2, p. 318-319] y lo aplicamos a este método se tiene: ($m = 1$, $a, b = 1$, $\tau_n = x_{i+1}$)

Lema 1.1 *Suponemos paso fijo h , $0 < h^* < \frac{1}{L}$, con $0 < h \leq h^*$, teniendo f continua y Lipschitz.*

a) *El sistema:*

$$Y = y + hf(x_{i+1}, Y) \quad (1.27)$$

tiene solución única $Y =: g_i(y)$ para cada $y \in \mathbb{R}^m$.

b) *Dado $Y = g_i(y)$ solución única del sistema (1.27), se define:*

$$\Phi_i(y) = f(x_{i+1}, g_i(y))$$

y verifica:

$$\|\Phi_i(y) - \Phi_i(z)\| \leq \Lambda \|y - z\|, \quad \forall y, z \in \mathbb{R}^m \text{ donde } \Lambda = \frac{L}{1 - h^*L}$$

Distinguimos ahora dos casos particulares:

- Si aplicamos este lema a $y = y_i$, tenemos que:

$$Y = y_i + hf(x_{i+1}, Y) \Rightarrow Y = y_{i+1} = g_i(y_i)$$

Definimos:

$$\Phi_i(y_i) = f(x_{i+1}, y_{i+1}) \quad (1.28)$$

y reescribimos el método:

$$y_{i+1} = y_i + h\Phi_i(y_i) \quad (1.29)$$

- Si lo aplicamos a $y = y(x_i)$ obtenemos:

$$\bar{Y} = y(x_i) + hf(x_{i+1}, \bar{Y})$$

con $\bar{Y} = g_i(y(x_i))$ única solución. Del mismo modo que antes, definimos:

$$\Phi(y(x_i)) = f(x_{i+1}, \bar{Y}) \tag{1.30}$$

y el método se reescribe de la siguiente forma:

$$\bar{Y} = y(x_i) + h\Phi(y(x_i)) \tag{1.31}$$

Teniendo en cuenta que $y(x_i)$ es la solución exacta, estudiamos ahora la relación entre $y(x_{i+1})$ e \bar{Y} ,

$$\begin{cases} y(x_{i+1}) = y(x_i) + hf(x_{i+1}, y(x_{i+1})) + h\tau_i \\ \bar{Y} = y(x_i) + hf(x_{i+1}, \bar{Y}) \end{cases}$$

$$\begin{aligned} y(x_{i+1}) - \bar{Y} &= h[f(x_{i+1}, y(x_{i+1})) - f(x_{i+1}, \bar{Y})] + h\tau_i(h) \Rightarrow \\ &\Rightarrow \|y(x_{i+1}) - \bar{Y}\| \leq Lh\|y(x_{i+1}) - \bar{Y}\| + h\|\tau_i(h)\| \underset{(0 < h < h^*)}{\leq} \end{aligned}$$

$$\leq Lh^*\|y(x_{i+1}) - \bar{Y}\| + h\|\tau_i(h)\| \Rightarrow (1 - Lh^*)\|y(x_{i+1}) - \bar{Y}\| \leq h\|\tau_i(h)\|$$

Dado que por hipótesis tenemos que $0 < h^* < \frac{1}{L}$, llegamos a:

$$\|y(x_{i+1}) - \bar{Y}\| \leq \frac{h}{1 - h^*L} \|\tau_i(h)\| \tag{1.32}$$

Por tanto, de (1.25) sabemos que:

$$y(x_{i+1}) - \bar{Y} = \mathcal{O}(h^2), \quad h \rightarrow 0 \tag{1.33}$$

Por otro lado, operamos en $y(x_{i+1})$ de la siguiente forma:

$$y(x_{i+1}) = y(x_i) + hf(x_{i+1}, \bar{Y}) + hf(x_{i+1}, y(x_{i+1})) - hf(x_{i+1}, \bar{Y}) + h\tau_i(h)$$

y así,

$$y(x_{i+1}) = y(x_i) + hf(x_{i+1}, \bar{Y}) + hS_i(h)$$

donde $S_i(h) = f(x_{i+1}, y(x_{i+1})) - f(x_{i+1}, \bar{Y}) + \tau_i(h)$

Teniendo en cuenta (1.29) y (1.31), el error global definido como $e_i = y(x_i) - y_i$ verifica:

$$e_{i+1} = e_i + h[\Phi_i(y_i) - \Phi_i(y(x_i))] + hS_i(h)$$

Tomando normas, $E_i = \|e_i\|$ y teniendo en cuenta el Lema 1.1 (b):

$$E_{i+1} \leq E_i + h\Lambda E_i + h\|S_i(h)\| = (1 + h\Lambda)E_i + h\|S_i(h)\|$$

y como:

$$\begin{aligned} \|S_i(h)\| &\leq L\|y(x_{i+1}) - \bar{Y}\| + \|\tau_i(h)\| \stackrel{(1.32)}{\leq} \frac{Lh}{1 - h^*L}\|\tau_i(h)\| + \|\tau_i(h)\| \leq \\ &\stackrel{(0 < h < h^*)}{\leq} \frac{1}{1 - h^*L}\|\tau_i(h)\| \end{aligned}$$

obtenemos:

$$E_{i+1} \leq (1 + h\Lambda)E_i + \frac{h}{1 - h^*L}\|\tau_i(h)\|$$

Observamos que es similar a lo obtenido con el Euler explícito. Continuamos:

$$1 + h\Lambda \leq e^{Ah}, \quad \forall h \quad / \quad 0 < h \leq h^* < \frac{1}{L}$$

$$E_i \leq e^{Ah}E_{i-1} + D_i \quad \text{donde} \quad D_i = \frac{h}{1 - h^*L}\|\tau_{i-1}(h)\|$$

iteramos, teniendo en cuenta $(e^{Ah})^n \leq e^{A(b-a)}$,

$$\begin{aligned} E_i &\leq e^{Ah}(e^{Ah}E_{i-2} + D_{i-1}) + D_i \leq e^{2Ah}E_{i-2} + e^{Ah}D_{i-1} + D_i \leq \\ &\leq e^{3Ah}E_{i-3} + e^{2Ah}D_{i-2} + e^{Ah}D_{i-1} + D_i \leq e^{iAh}E_0 + \sum_{k=0}^{i-1} e^{kAh}D_{i-k} \leq \\ &\leq e^{A(b-a)}E_0 + \sum_{j=1}^i e^{(i-j)Ah}D_j \end{aligned}$$

Como $E_0 \approx 0$

$$E_i \leq \sum_{j=1}^i e^{(i-j)Ah}h \left[\frac{1}{1 - h^*L}\|\tau_{j-1}(h)\| \right] \quad (1.34)$$

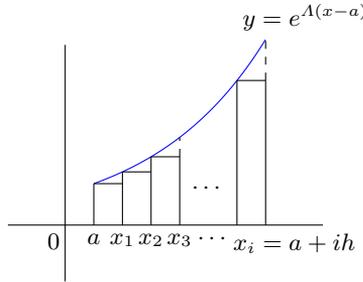
Por (1.9)

$$E_i \leq \frac{Mh}{2(1 - h^*L)} \sum_{j=1}^i e^{(i-j)Ah}h$$

tomamos

$$\sum_{j=1}^i e^{(i-j)\Lambda h} h = \sum_{k=0}^{i-1} e^{(kh)\Lambda h} h = \sum_{k=0}^{i-1} e^{(x_k - a)\Lambda} h$$

que es una suma inferior de Riemann de $y = e^{\Lambda(x-a)}$:



Entonces, tenemos que:

$$\sum_{j=1}^i e^{(i-j)\Lambda h} h \leq \int_a^{x_i} e^{\Lambda(x-a)} dx = \frac{1}{\Lambda} \left[e^{\Lambda(x_i-a)} - 1 \right] \leq \frac{1}{\Lambda} \left[e^{\Lambda(b-a)} - 1 \right]$$

Finalmente, obtenemos:

$$E_i \leq \left[\frac{M}{2(1-h^*L)} \frac{e^{\Lambda(b-a)} - 1}{\Lambda} \right] h, \quad \text{con } 0 < h \leq h^*, \quad \forall i = 0, 1, 2, \dots, n$$

Por lo tanto el método converge con orden 1.

1.3.2. Estabilidad absoluta del método

De forma análoga al método explícito, procederemos al análisis de la estabilidad absoluta partiendo del caso de PVI (1.15).

Aplicamos ahora el método de Euler implícito para dicha ecuación:

$$y_{j+1} = y_j + h\lambda y_{j+1} \Rightarrow y_{j+1} = \frac{1}{(1-h\lambda)} y_j = \dots = \left(\frac{1}{1-h\lambda} \right)^{j+1} c$$

con $j = 0, \dots, n$

Debemos de comprobar que se cumple (1.20) para que el método sea absolutamente estable.

Para que esto se cumpla debe de ocurrir:

$$|y_{i+1}| = \left| \frac{1}{1 - h\lambda} \right| |y_i| \leq |y_i|$$

si y solo si:

$$\left| \frac{1}{1 - h\lambda} \right| \leq 1 \Leftrightarrow \frac{1}{|1 - h\lambda|} \leq 1$$

y además tenemos una región de estabilidad absoluta en el z -plano complejo de manera que aplicando el método a la ecuación en dicha región z , produce una solución aproximada que satisface el requisito de estabilidad absoluta anterior. En la Figura 1.5 podemos ver la región del z -plano en la que el método verifica

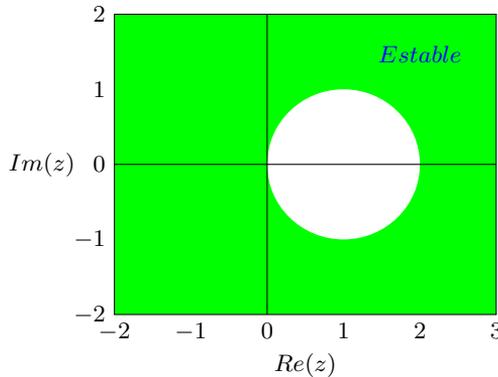


Figura 1.5: Región de estabilidad absoluta del método de Euler Implícito.

(1.20) donde $z = h\lambda$.

Por tanto, si el método cumple:

$$|1 - h\lambda| \geq 1$$

el método es absolutamente estable.

1.4. Regla Trapezoidal

Definimos el método de Euler explícito con el desarrollo de Taylor centrado en x_i y, a su vez, el método de Euler implícito con dicho desarrollo centrado en x_{i+1} . Ambos métodos son de primer orden, como comprobamos en (1.2) y (1.25).

Debido a la poca eficiencia que nos ofrecen construiremos un nuevo método basándonos en el desarrollo de Taylor centrado en $x_{i+\frac{1}{2}} = x_i + \frac{1}{2}h$.

Así,

$$y(x_{i+1}) = y(x_{i+\frac{1}{2}}) + \frac{h}{2}y'(x_{i+\frac{1}{2}}) + \frac{h^2}{8}y''(x_{i+\frac{1}{2}}) + \frac{h^3}{48}y'''(x_{i+\frac{1}{2}}) + \dots$$

$$y(x_i) = y(x_{i+\frac{1}{2}}) - \frac{h}{2}y'(x_{i+\frac{1}{2}}) + \frac{h^2}{8}y''(x_{i+\frac{1}{2}}) - \frac{h^3}{48}y'''(x_{i+\frac{1}{2}}) + \dots$$

Restando ambas expresiones se obtiene:

$$y(x_{i+1}) - y(x_i) = hy'(x_{i+\frac{1}{2}}) + \frac{h^3}{24}y'''(x_{i+\frac{1}{2}}) + \mathcal{O}(h^5) \quad (1.35)$$

Desarrollando ahora Taylor para $y'(x_{i+1})$ y $y'(x_i)$ obtenemos:

$$y'(x_{i+1}) = y'(x_{i+\frac{1}{2}}) + \frac{h}{2}y''(x_{i+\frac{1}{2}}) + \frac{h^2}{8}y'''(x_{i+\frac{1}{2}}) + \frac{h^3}{48}y''''(x_{i+\frac{1}{2}}) + \dots$$

$$y'(x_i) = y'(x_{i+\frac{1}{2}}) - \frac{h}{2}y''(x_{i+\frac{1}{2}}) + \frac{h^2}{8}y'''(x_{i+\frac{1}{2}}) - \frac{h^3}{48}y''''(x_{i+\frac{1}{2}}) + \dots$$

y si sumamos los términos:

$$y'(x_{i+1}) + y'(x_i) = 2y'(x_{i+\frac{1}{2}}) + \frac{h^2}{4}y'''(x_{i+\frac{1}{2}}) + \mathcal{O}(h^4) \Rightarrow$$

$$\Rightarrow y'(x_{i+\frac{1}{2}}) = \frac{1}{2}(y'(x_{i+1}) + y'(x_i)) - \frac{h^2}{8}y'''(x_{i+\frac{1}{2}}) + \mathcal{O}(h^4) \quad (1.36)$$

Substituimos en la ecuación (1.35) para obtener:

$$y(x_{i+1}) - y(x_i) = \frac{h}{2}(y'(x_{i+1}) + y'(x_i)) + h\tau_i(h) \quad (1.37)$$

siendo $\tau_i(h) = -\frac{h^2}{8}y'''(x_{i+\frac{1}{2}}) + \mathcal{O}(h^4) = \mathcal{O}(h^2)$, $(h \rightarrow 0)$.

Teniendo en cuenta que $y'(x) = f(x, y(x))$, y despreciando $\mathcal{O}(h^2)$, definimos el método de la *regla trapezoidal* como:

$$\begin{cases} y_0 = \alpha \\ y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})) \text{ con } i = 0, \dots, n-1. \end{cases} \quad (1.38)$$

1.4.1. Consistencia y convergencia del método

Comenzamos reescribiendo el *método* (1.38) de la siguiente forma:

$$\frac{y_{i+1} - y_i}{h} = \frac{1}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})) \quad (1.39)$$

Dada una función $u(x)$, con $x \in [a, b]$ y $h > 0$, se define el operador:

$$\delta_h u(x_{i+1}) \equiv \frac{u(x_{i+1}) - u(x_i)}{h} - \frac{1}{2}(f(x_i, u(x_i)) + f(x_{i+1}, u(x_{i+1})))$$

con $i = 0, \dots, n-1$

Si tomamos $u = y_h$ como una función de malla cualquiera que toma sobre los puntos x_i el valor y_i del método (1.38):

$$\begin{aligned} \delta_h y_h(x_{i+1}) &= \frac{y_h(x_{i+1}) - y_h(x_i)}{h} - \frac{1}{2}(f(x_i, y_h(x_i)) + f(x_{i+1}, y_h(x_{i+1}))) = \\ &= \frac{y_{i+1} - y_i}{h} - \frac{1}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})) = 0 \\ &\text{con } i = 0, \dots, n-1 \quad e \quad y_0 = \alpha. \quad (y_h(x_0) = \alpha) \end{aligned}$$

Considerando $u(x) = y(x)$ la solución exacta de (1.37) tenemos que:

$$\delta_h y(x_{i+1}) = \frac{y(x_{i+1}) - y(x_i)}{h} - \frac{1}{2}(f(x_i, y(x_i)) + f(x_{i+1}, y(x_{i+1})))$$

donde $\delta_h y(x_{i+1})$ es el *Error local de truncamiento*, denotado $\tau_i(h)$. Por (1.37):

$$\delta_h y(x_{i+1}) = \tau_i(h) = \mathcal{O}(h^2) \quad (h \rightarrow 0) \quad (1.40)$$

y por tanto, el método es consistente de orden 2.

Este método es implícito, lo que significa que el coste por cada paso de este método y el anterior es similar. Pero el método trapezoidal es más preciso, lo que nos permite usar un número menor de pasos con la seguridad de seguir satisfaciendo la cota de error.

Tomamos $m = 2$ para el [7, Lema 4.2, p. 318-319] y suponemos paso fijo h , teniendo f continua y Lipschitz.

$$\begin{cases} y_1 = y + h_n [a_{11}f(\tau_{n,1}, y_1) + a_{12}f(\tau_{n,2}, y_2)] \\ y_2 = y + h_n [a_{21}f(\tau_{n,1}, y_1) + a_{22}f(\tau_{n,2}, y_2)] \end{cases} \quad (1.41)$$

La regla trapezoidal se ajusta a esta formulación tomando $y_1 = y_i$, $y_2 = y_{i+1}$, $\tau_{n,1} = x_i$, $\tau_{n,2} = x_{i+1}$ y los valores de a_{ij} son de la forma $a_{21} =$

$$a_{22} = \frac{1}{2}, \quad a_{11} = a_{12} = 0.$$

Así, $A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 2 \end{pmatrix}$ y por tanto, como $R = (r_{ij}) = (I - Lh^*A)^{-1}$:

$$R = \begin{pmatrix} 1 & 0 \\ \frac{Lh^*}{2(2 - Lh^*)} & \frac{1}{2 - Lh^*} \end{pmatrix}$$

con $0 < h^* < \frac{2}{L}$, $0 < h \leq h^*$.

Por el apartado (a) del Lema concluimos que para cualquier y existe una única solución $\{y_1, y_2\}$ de (1.41) que denotamos:

$$\begin{cases} y_1 = y = g_{n,1}(y) \\ y_2 = g_{n,2}(y) \end{cases}$$

Para aplicar ahora el apartado (b) del mismo Lema, definimos:

$$\Phi_i(y) = \frac{1}{2}f(x_i, g_{n,1}(y)) + \frac{1}{2}f(x_{i+1}, g_{n,2}(y)) \quad (1.42)$$

que verifica:

$$\|\Phi_i(y) - \Phi_i(z)\| \leq A\|y - z\|, \quad \forall y, z \in \mathbb{R}^m \quad (1.43)$$

con A constante de Lipschitz de la forma:

$$A = L(|b_1r_{11}| + |b_1r_{12}| + |b_2r_{21}| + |b_2r_{22}|)$$

donde $b_1 = b_2 = \frac{1}{2}$ y así:

$$A = L \left(\frac{6 - Lh^*}{4(2 - Lh^*)} \right) \quad (1.44)$$

Ahora, aplicando las propiedades anteriores a $y = y_i$ obtenemos que la regla trapezoidal se puede escribir como:

$$y_{i+1} = y_i + h\Phi_i(y_i) \quad (1.45)$$

donde,

$$\Phi_i(y_i) = \frac{1}{2}f(x_i, y_i) + \frac{1}{2}f(x_{i+1}, y_{i+1})$$

Aplicando el Lema ahora a $y = y(x_i)$ obtenemos que existe un único \bar{Y} que verifica:

$$\bar{Y} = y(x_i) + h\Phi(y(x_i)) \quad (1.46)$$

donde,

$$\Phi_i(y(x_i)) = \frac{1}{2}f(x_i, y(x_i)) + \frac{1}{2}f(x_{i+1}, \bar{Y})$$

Estudiamos la relación entre $y(x_{i+1})$ e \bar{Y} , restando (1.37) – (1.46):

$$y(x_{i+1}) - \bar{Y} = \frac{h}{2}[f(x_{i+1}, y(x_{i+1})) - f(x_{i+1}, \bar{Y})] + h\tau_i(h)$$

Tomando normas:

$$\begin{aligned} \|y(x_{i+1}) - \bar{Y}\| &\leq L\frac{h}{2}\|y(x_{i+1}) - \bar{Y}\| + h\|\tau_i(h)\| \underset{(0 < h < h^*)}{\leq} \\ &\leq L\frac{h^*}{2}\|y(x_{i+1}) - \bar{Y}\| + h\|\tau_i(h)\| \\ \Rightarrow (1 - L\frac{h^*}{2})\|y(x_{i+1}) - \bar{Y}\| &\leq h\|\tau_i(h)\| \end{aligned}$$

y como $0 < h^* < \frac{2}{L}$:

$$\|y(x_{i+1}) - \bar{Y}\| \leq \frac{2h}{2 - Lh^*}\|\tau_i(h)\| \quad (1.47)$$

Si en (1.37) sumas y restas $\frac{h}{2}f(x_i, \bar{y})$, tenemos:

$$y(x_{i+1}) = y(x_i) + \frac{h}{2}f(x_i, y(x_i)) + \frac{h}{2}f(x_{i+1}, \bar{Y}) + hS_i(h) \quad (1.48)$$

donde,

$$S_i(h) = \frac{1}{2}f(x_{i+1}, y(x_{i+1})) - \frac{1}{2}f(x_{i+1}, \bar{Y}) + \tau_i(h)$$

Teniendo en cuenta el *error global del método* se calcula como $e_i = y(x_i) - y_i$, y así, restando (1.48)-(1.45):

$$e_{i+1} = e_i + h[\Phi_i(y(x_i)) - \Phi_i(y_i)] + hS_i(h)$$

y usando normas, $E_i = \|e_i\|$, por (1.43), tenemos:

$$E_{i+1} \leq E_i + h\Lambda E_i + h\|S_i(h)\| = (1 + h\Lambda)E_i + h\|S_i(h)\|$$

donde:

$$\begin{aligned} \|S_i(h)\| &\leq \frac{L}{2} \|y(x_{i+1}) - \bar{Y}\| + \|\tau_i(h)\| \stackrel{(1.47)}{\leq} \frac{2Lh}{4-2h^*L} \|\tau_i(h)\| + \|\tau_i(h)\| \leq \\ &\stackrel{(0 < h < h^*)}{\leq} \frac{2}{2-h^*L} \|\tau_i(h)\| \end{aligned}$$

y así, obtenemos:

$$E_{i+1} \leq (1+h\Lambda)E_i + \frac{2h}{2-h^*L} \|\tau_i(h)\|$$

Análogamente al método de Euler implícito tenemos:

$$1+h\Lambda \leq e^{\Lambda h}, \quad \forall h \quad / \quad 0 < h \leq h^* < \frac{2}{L}$$

$$E_i \leq e^{\Lambda h} E_{i-1} + D_i \quad \text{donde} \quad D_i = \frac{2h}{2-h^*L} \|\tau_{i-1}(h)\|$$

iteramos de la misma forma que en el Euler implícito, teniendo en cuenta $(e^{\Lambda h})^n \leq e^{\Lambda(b-a)}$, para obtener una cota similar a (1.34):

$$E_i \leq \sum_{j=1}^i e^{(i-j)\Lambda h} h \left[\frac{2}{2-h^*L} \|\tau_{j-1}(h)\| \right]$$

Por el resto de Lagrange del polinomio de Taylor y como $\|y'''\|$ está acotada en $[a, b]$ tenemos que:

$$\|\tau_{j-1}\| \leq \frac{h^2}{8} M \quad \text{donde} \quad M = \max_{x \in [a, b]} \|y'''(x)\|$$

y por tanto:

$$E_i \leq \frac{Mh^2}{4(2-h^*L)} \sum_{j=1}^i e^{(i-j)\Lambda h} h$$

y además:

$$\sum_{j=1}^i e^{(i-j)\Lambda h} h \leq \int_a^{x_i} e^{\Lambda(x-a)} dx = \frac{1}{\Lambda} \left[e^{\Lambda(x_i-a)} - 1 \right] \leq \frac{1}{\Lambda} \left[e^{\Lambda(b-a)} - 1 \right]$$

Finalmente, obtenemos:

$$E_i \leq \left[\frac{M}{4(2-h^*L)} \frac{e^{\Lambda(b-a)} - 1}{\Lambda} \right] h^2, \quad \text{con} \quad 0 < h \leq h^*, \quad \forall i = 0, 1, 2, \dots, n$$

Por lo tanto el método converge con orden 2.

1.4.2. Estabilidad absoluta del método

Para el estudio de la estabilidad absoluta del método debemos de comprobar que se cumple (1.8), de la misma forma que en los anteriores métodos.

Partimos del caso de PVI (1.15) y aplicamos el método:

$$y_{i+1} = y_i + \frac{h\lambda}{2} y_{i+1} + \frac{h\lambda}{2} y_i \Rightarrow y_{i+1} = \frac{2 + h\lambda}{2 - h\lambda} y_i$$

y por tanto,

$$|y_{i+1}| = \left| \frac{2 + h\lambda}{2 - h\lambda} \right| |y_i| \leq |y_i| \Leftrightarrow \left| \frac{2 + h\lambda}{2 - h\lambda} \right| \leq 1 \tag{1.49}$$

A continuación ilustramos la región del z -plano que delimita la estabilidad del método.

Dado $z = h\lambda$ con $z \in \mathbb{C}$ y por tanto de la forma $z = a + bi$, teniendo en cuenta (1.49):

$$\begin{aligned} |2 + z|^2 \leq |2 - z|^2 &\Rightarrow |(2 + a) + bi|^2 \leq |(2 - a) + bi|^2 \Rightarrow \\ &\Rightarrow (2 + a)^2 + b^2 \leq (2 - a)^2 + b^2 \Rightarrow a \leq 0 \end{aligned}$$

y dado que $a = Re(h\lambda)$, la región de estabilidad absoluta de la regla trapezoidal es todo \mathbb{C}^- como podemos ver en la Figura 1.6.

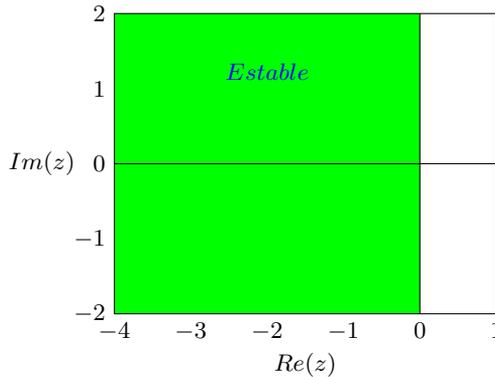


Figura 1.6: Región de estabilidad absoluta de la regla trapezoidal.

1.5. Implementación de los métodos implícitos

Tanto el método de Euler implícito como la regla trapezoidal son métodos implícitos, es decir, en cada paso de integración hay que resolver un sistema no

lineal para calcular el valor de avance $\omega = y_{i+1}$.
 En el caso de Euler implícito el sistema es:

$$F(\omega) = \omega - y_i - hf(x_{i+1}, \omega) = 0 \tag{1.50}$$

Para la resolución de la ecuación anterior podemos emplear dos métodos.

1.5.1. Iteración funcional

Para aplicar la iteración funcional hay que expresar (1.50) en la forma $\omega = g(\omega)$, donde:

$$g(\omega) = y_i + hf(x_{i+1}, \omega)$$

Por tanto, se calculan las iteraciones sucesivas $\omega^{(\nu)}$, $j = 0, 1, 2, \dots$ mediante:

$$\omega^{(\nu+1)} = g(\omega^{(\nu)}), \quad \nu = 0, 1, 2, \dots$$

donde ν es el contador de iteraciones y con $\omega^{(0)}$ valor inicial dado. En este caso, se suele tomar $\omega^{(0)} = y_i$.

Para garantizar la convergencia de esta iteración viendo el [2, Teorema 2.7] se tiene que satisfacer que $\|g'(\omega)\| < 1$, es decir:

$$h \left\| \frac{\partial f}{\partial y}(x_{i+1}, \omega) \right\| < 1$$

lo que puede suponer una restricción importante del tamaño de paso cuando integramos problemas stiff.

1.5.2. Método de Newton

Otro de los métodos con el que se puede resolver el sistema implícito (1.50) es el método de Newton.

Este método, para aproximar la solución de una ecuación no lineal $g(x) = 0$, construye una sucesión de aproximantes $x^{(\nu)}$, $\nu = 1, 2, \dots$ a partir de un valor inicial dado $x^{(0)}$.

Para cada iteración $x^{(\nu)}$ se tiene por Taylor:

$$0 = g(x) = g(x^{(\nu)}) + g'(x^{(\nu)})(x - x^{(\nu)}) + \dots$$

Se desprecian todos los términos restantes de este desarrollo y se toma como siguiente iteración $x^{(\nu+1)}$ el valor que cumple la ecuación lineal:

$$0 = g(x^{(\nu)}) + g'(x^{(\nu)})(x^{(\nu+1)} - x^{(\nu)}), \quad \nu = 0, 1, \dots \tag{1.51}$$

Esto se generaliza para un sistema de ecuaciones no lineales $F(\omega) = 0$, teniendo en cuenta que la derivada que aparece en (1.51) es ahora el jacobiano $F'(\omega^{(\nu)}) = \frac{\partial F}{\partial \omega}(\omega^{(\nu)})$:

$$F'(\omega^{(\nu)})(\omega^{(\nu+1)} - \omega^{(\nu)}) = -F(\omega^{(\nu)}), \quad \nu = 0, 1, \dots \quad (1.52)$$

En el caso de Euler implícito, con $F(\omega)$ dada en (1.50), resulta:

$$\left[I - h \frac{\partial f}{\partial y}(x_{i+1}, \omega^{(\nu)}) \right] (\omega^{(\nu+1)} - \omega^{(\nu)}) = -\omega^{(\nu)} + y_i + hf(x_{i+1}, \omega^{(\nu)}), \quad \nu = 0, 1, \dots$$

con $\omega^{(\nu)}$ dado. Normalmente, se suele tomar $\omega^{(\nu)} = y_i$, pero a veces se usan, por ejemplo, valores interpolados de pasos anteriores.

1.5.3. Experimentos numéricos

Consideramos de nuevo el Ejemplo (1.14) pero en forma general:

$$\begin{cases} y(0) = 1 \\ y'(x) = -5 \left(xy^2 - \frac{1}{x} \right) - \frac{1}{x^2} \quad \text{con } x > 1 \end{cases} \quad (1.53)$$

con solución exacta $y = \frac{1}{x}$.

Implementamos con Matlab la integración de este problema con el método de Euler implícito, con iteración funcional y con el método de Newton para comparar.

En ambos esquemas iterativos dan $nmax = 25$ iteraciones. Como conocemos la solución exacta del problema computamos en cada caso:

$$error = \max_{1 \leq i \leq n} |y_i - y(x_i)|$$

En las Figura 1.7, se puede ver el comportamiento de la integración con iteración funcional para dos valores distintos de h .

Claramente, en estas gráficas se ve que con $h = 0.11$ el método presenta unas oscilaciones considerables que no tiene la solución, mientras que con $h = 0.10$, la integra correctamente. Se puede apreciar que los errores aumentan significativamente variando poco el h .

Este es un comportamiento similar al que presenta el método de Euler explícito sobre este problema, por lo que la iteración funcional no parece recomendable cuando se integran problemas stiff.

En cambio, en las Figuras 1.8 se ve que, con el método de Newton, el método de Euler implícito integra este PVI correctamente para valores moderados de h con errores más que aceptables, como se muestra en la Tabla 1.2.

Por tanto, para problemas stiff se preferirá aplicar métodos de Newton (o variantes) con los métodos implícitos.

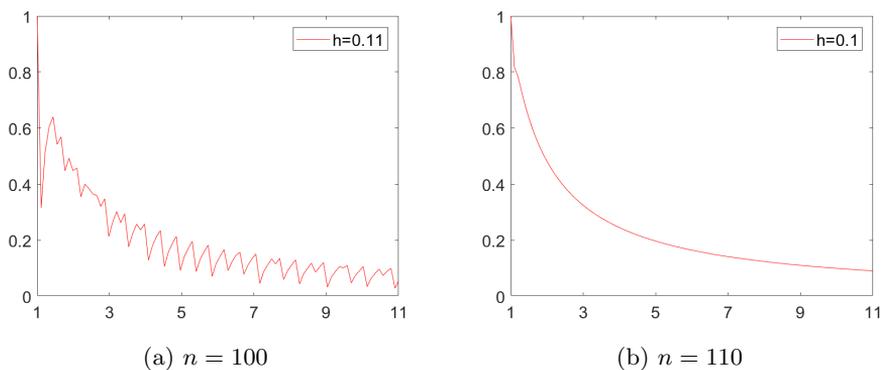


Figura 1.7: Euler implícito con iteración funcional a paso fijo en (1.53).

Iteración funcional		Newton	
h	error	h	error
0.10	9.18539e-02	0.10	5.21219e-03
0.11	5.85474e-01	0.50	1.83090e-02

Tabla 1.2: Error máximo Euler implícito.

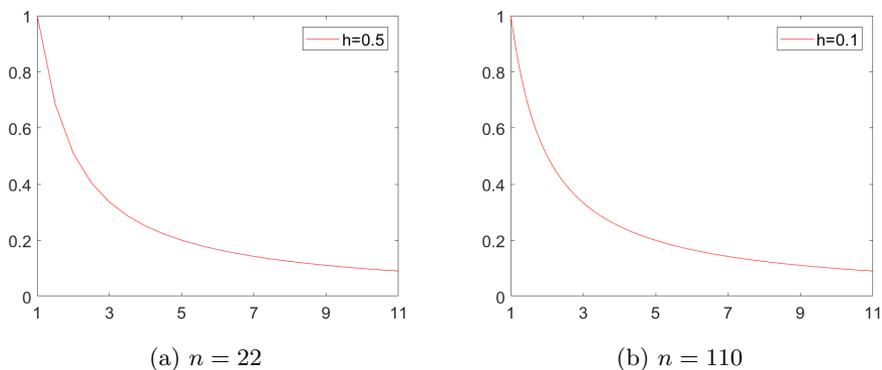


Figura 1.8: Euler implícito con Newton a paso fijo en (1.53).

Problemas lineales

Una de las principales fuentes de problemas de valores iniciales lineales interesantes es la resolución de ecuaciones en derivadas parciales mediante el método de líneas, en particular, las parabólicas. La mejor forma de explicar este método de líneas es ver un ejemplo.

2.1. Aplicación a ecuaciones en derivadas parciales

Consideramos la siguiente ecuación parabólica en derivadas parciales:

$u(x, t)$, con $x \in [0, 1]$, $t \in [0, 1]$.

$$\begin{cases} u_t(x, t) = du_{xx}(x, t) + g(x, t), & d > 0 \\ u(0, t) = u(1, t) = 0 \\ u(x, 0) = f_0(x) \end{cases} \quad (2.1)$$

donde

$$\begin{cases} f_0 = x(1 - x) \\ g(x, t) = 2d \cos(t) - x(1 - x) \sin(t) \end{cases} \quad (2.2)$$

Una de las interpretaciones más habituales de esta ecuación es la de la distribución del calor en una barra de longitud 1, de material y sección uniformes, en la que en sus extremos hay un refrigerador que mantiene su temperatura a cero grados, a la que se aplica una fuente de calor externa que viene representada por la función $g(x, t)$ y que tiene una distribución inicial de temperaturas dada por f_0 . Sin embargo, las ecuaciones de este tipo aparecen también en la modelización

de muchos fenómenos de difusión.

Contamos, además, con la expresión de la solución exacta dada por:

$$u(x, t) = x(1 - x) \cos(t) \quad (2.3)$$

lo que nos servirá para comprobar el comportamiento de los métodos sobre este problema.

2.2. Método de líneas

Con este método es posible resolver la ecuación (2.1) como un sistema lineal de ecuaciones diferenciales ordinarias de coeficientes constantes.

Primero se elige un número de puntos de la barra sobre los que queremos aproximar la solución de la ecuación, que se suelen llamar nodos. En este caso tomaremos $N \geq 1$ nodos igualmente espaciados $x_i = i\Delta x$, $0 \leq i \leq N + 1$, siendo $\Delta x = \frac{1}{N + 1}$.

Denotamos $u_i(t) = u(x_i, t)$, $t \in [0, 1]$. Por (2.1) los extremos de la barra tendrán el valor $u_0 = u_{N+1} = 0$.

Sobre cada punto x_i la derivada parcial respecto a t es de la siguiente forma:

$$\frac{\partial u}{\partial t}(x_i, t) = \frac{\partial u}{\partial t}(t) = u'_i(t)$$

usando el desarrollo de Taylor obtenemos las siguientes expresiones:

$$\begin{aligned} u(x_{i+1}, t) = u(x_i + \Delta x, t) &= u(x_i, t) + \frac{\partial u}{\partial x}(x_i, t)\Delta x + \frac{\partial^2 u}{\partial x^2}(x_i, t)\frac{\Delta x^2}{2!} + \\ &+ \frac{\partial^3 u}{\partial x^3}(x_i, t)\frac{\Delta x^3}{3!} + \dots \end{aligned}$$

$$\begin{aligned} u(x_{i-1}, t) = u(x_i - \Delta x, t) &= u(x_i, t) - \frac{\partial u}{\partial x}(x_i, t)\Delta x + \frac{\partial^2 u}{\partial x^2}(x_i, t)\frac{\Delta x^2}{2!} - \\ &- \frac{\partial^3 u}{\partial x^3}(x_i, t)\frac{\Delta x^3}{3!} + \dots \end{aligned}$$

y, dado que (2.3) es un polinomio de grado 2 respecto a la variable x , $\frac{\partial^k u}{\partial x^k}(x, t) = 0$, $\forall k \geq 3$, se tiene:

$$u(x_{i+1}, t) = u(x_i, t) + \frac{\partial u}{\partial x}(x_i, t)\Delta x + \frac{\partial^2 u}{\partial x^2}(x_i, t)\frac{\Delta x^2}{2!} \quad (2.4)$$

$$u(x_{i-1}, t) = u(x_i, t) - \frac{\partial u}{\partial x}(x_i, t)\Delta x + \frac{\partial^2 u}{\partial x^2}(x_i, t) \frac{\Delta x^2}{2!} \quad (2.5)$$

por último, restando (2.4) – (2.5) obtenemos:

$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x_{i+1}, t) - 2u(x_i, t) + u(x_{i-1}, t)}{\Delta x^2} \quad (2.6)$$

Reescribimos el problema (2.1) con los resultados obtenidos:

$$\begin{cases} u'_i(t) = \frac{d}{\Delta x^2} [u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)] + g(x_i, t), & i = 1, 2, \dots, N \\ u_0(t) = u_{N+1}(t) = 0 \end{cases} \quad (2.7)$$

Para manejar vectorialmente todos los valores $u_i(t)$, $1 \leq i \leq N$, definimos

$$U = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_N(t) \end{pmatrix} \in \mathbb{R}^N \text{ y escribimos matricialmente este sistema como:}$$

$$\begin{cases} U' = AU + b(x), & x \in [0, 1] \\ U(0) = U_0 \end{cases} \quad (2.8)$$

donde:

$$A = \frac{d}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ & & \ddots & & \\ 0 & \cdots & 1 & -2 & 1 \\ 0 & \cdots & 0 & 1 & -2 \end{bmatrix} \in \mathbb{M}_{N \times N}, \quad b(x) = \begin{pmatrix} g(x_1, t) \\ g(x_2, t) \\ \vdots \\ g(x_N, t) \end{pmatrix} \in \mathbb{R}^N$$

$$U(0) = \begin{pmatrix} u_1(0) \\ u_2(0) \\ \vdots \\ u_N(0) \end{pmatrix} = \begin{pmatrix} f_0(x_1) \\ f_0(x_2) \\ \vdots \\ f_0(x_N) \end{pmatrix} = \begin{pmatrix} x_1(1 - x_1) \\ x_2(1 - x_2) \\ \vdots \\ x_N(1 - x_N) \end{pmatrix} = U_0$$

Evidentemente, si usáramos otra fórmula en diferencias finitas para aproximar a las derivadas parciales segundas obtendríamos otros PVI. Por eso, el método de líneas no se debe entender como un método único sino como una forma de abordar la discretización de las EDPs. En [5] se recogen otros tipos de discretizaciones espaciales, y se hace un estudio de sus características dependiendo del tipo de problema que se esté tratando.

2.3. Resolución temporal

La fase final del método de líneas consiste en integrar temporalmente el PVI (2.8) resultante de la discretización de la EDP (2.1).

Trabajaremos ahora con el método de Euler explícito e implícito y la regla del trapecio en el sistema (2.8). Tendremos que tener en cuenta el carácter más o menos *stiff* del problema, que depende de los autovalores de la matriz A que analizaremos a continuación.

La matriz A , por su estructura, es una matriz Toeplitz tridiagonal de la siguiente forma:

$$A = \begin{bmatrix} b & a & 0 & \cdots & 0 \\ c & b & a & \cdots & 0 \\ & & \ddots & & \\ 0 & \cdots & c & b & a \\ 0 & \cdots & 0 & c & b \end{bmatrix} \in \mathbb{M}_{N \times N}, \quad a \neq 0 \neq c$$

Podemos ver en [6] que los autovalores de las matrices de este tipo vienen dados por la siguiente expresión:

$$\lambda_j = b + 2a\sqrt{\frac{c}{a}} \cos\left(\frac{j\pi}{N+1}\right), \quad j = 1, \dots, N \quad (2.9)$$

En la matriz de nuestro PVI (2.8) las componentes de la matriz A son $b = -\frac{2d}{\Delta x^2}$

y $a = c = \frac{d}{\Delta x^2}$, por lo que sus autovalores son:

$$\lambda_j = -\frac{2d}{\Delta x^2} + 2\frac{d}{\Delta x^2} \cos\left(\frac{j\pi}{N+1}\right) = -\frac{2d}{\Delta x^2} \left[1 - \cos\left(\frac{j\pi}{N+1}\right)\right]$$

y así:

$$\lambda_j = -\frac{4d}{\Delta x^2} \sin^2\left(\frac{j\pi}{2(N+1)}\right), \quad j = 1, \dots, N \quad (2.10)$$

Como $\sin^2\left(\frac{j\pi}{2(N+1)}\right) \in (0, 1)$ podemos acotar los autovalores de la forma:

$$-\frac{4d}{\Delta x^2} < \lambda_j < 0$$

Sabemos así que para que un método sea estable sobre este problema, $h\lambda$ para todos los autovalores λ de la matriz jacobiana A , tiene que caer dentro de la región de estabilidad del método correspondiente. Así que para los tres métodos que hemos considerado se tiene:

- Para el método de Euler Explícito la restricción de estabilidad viene dada por:

$$|1 + h\lambda| \leq 1$$

y por tanto:

$$-2 \leq h \left(-\frac{4d}{\Delta x^2} \right) \leq 0 \Rightarrow 0 < h \leq \frac{\Delta x^2}{2d} \quad (2.11)$$

lo que es una restricción importante del tamaño de paso temporal.

- Para el método de Euler Implícito la región de estabilidad se encuentra en:

$$|1 - h\lambda| \geq 1$$

y así tenemos que:

$$h \left(-\frac{4d}{\Delta x^2} \right) \leq 0 \quad \text{ó} \quad h \left(-\frac{4d}{\Delta x^2} \right) \geq 2 \Rightarrow h \geq 0$$

por lo que el método no sufre ninguna restricción de paso por estabilidad.

- Por último, para la regla del Trapecio tomábamos $h\lambda = z = a + bi$ y así obteníamos estabilidad en el método si nos encontramos en la región en la que $a \leq 0$, y por tanto:

$$h \left(-\frac{4d}{\Delta x^2} \right) \leq 0 \Rightarrow h \geq 0$$

por lo que este método también es estable sobre este problema.

2.4. Experimentos numéricos

A continuación mostraremos y compararemos los resultados obtenidos con el método de líneas usando para la integración temporal los tres métodos dados anteriormente. Para ello hemos elaborado una rutina de Matlab en la que el usuario puede elegir el coeficiente de difusión d así como el número de nodos espaciales N y el número de pasos temporales n que da cada uno de los tres métodos considerados.

En primer lugar, presentamos en la Tabla 2.1 los errores obtenidos para cada uno de los métodos temporales sobre el PVI (2.8), esto es, la norma infinito del vector $u(1) - u_{metodo}$ evaluado en el punto final de integración temporal $t = 1$, para diferentes valores de N y de n y con $d = 1$. En ella observamos cómo Euler implícito y la regla trapezoidal integran bien el problema considerado, siendo este último el que es más preciso de los dos, debido a su orden más alto. En cambio, Euler explícito es incapaz de dar una solución válida con estos valores,

N	n	Euler Explícito	Euler Implícito	Regla Trapecio
10	10	7.2410114e+08	8.3125276e-04	1.6843710e-05
	20	4.8011982e+18	4.0918963e-04	4.2080219e-06
	40	1.1584267e+32	2.0288351e-04	1.0518242e-06
20	10	2.6817207e+13	8.3290528e-04	1.6892343e-05
	20	1.3830647e+29	4.0997986e-04	4.2201700e-06
	40	1.1816941e+55	2.0327019e-04	1.0548606e-06
40	10	1.2538993e+18	8.3335950e-04	1.6905778e-05
	20	4.7713879e+39	4.1019698e-04	4.2235259e-06
	40	2.8858400e+77	2.0337641e-04	1.0556994e-06
80	10	6.8386845e+22	8.3347916e-04	1.6909322e-05
	20	2.2093778e+50	4.1025418e-04	4.2244113e-06
	40	1.0303067e+100	2.0340440e-04	1.0559207e-06

Tabla 2.1: Errores en el punto final del PVI (2.8) obtenidos con $d = 1$.

ya que sufre una severa restricción de paso por su falta de estabilidad absoluta. Para reflejar mejor esta restricción por estabilidad, en la Tabla 2.2, para $d = 1$ y distintos valores de N , se recoge el mínimo número de pasos que debe dar para conseguir una aproximación razonable.

Por ejemplo, cuando $N = 10$ necesita dar al menos 216 pasos, pues si toma 215, la aproximación obtenida no es buena.

En cambio, los métodos implícitos dan soluciones con errores muy bajos para tamaños de paso relativamente grandes.

N	n	Euler Explícito	Euler Implícito	Regla Trapezoidal
10	215	2.4185535e+00	3.7478607e-05	3.6405094e-08
	216	1.4269305e-01	3.7304811e-05	3.6068789e-08
20	857	1.3597075e+00	9.4085038e-06	2.2978849e-09
	858	1.1755486e-01	9.3975337e-06	2.2925315e-09
40	3337	4.2173969e+00	2.4167666e-06	1.5166593e-10
	3338	5.7381944e-01	2.4160425e-06	1.5160492e-10
80	13097	1.5801337e+00	6.1580525e-07	9.8882846e-12
	13098	6.8131956e-01	6.1575827e-07	9.8252240e-12

Tabla 2.2: Errores en el punto final del PVI (2.8) obtenidos para $d=1$.

2.4.1. Orden de los métodos

Como en el problema anterior los errores solo provienen de la integración temporal y hemos computado los tres métodos a paso fijo, podemos comprobar en la práctica si los métodos realmente alcanzan el orden que predice la teoría.

Si denotamos $E(n)$ al módulo del error de un método en el punto final de integración, cuando se ha integrado con n pasos temporales de tamaño fijo $h = \frac{(b-a)}{n}$, sabemos que si el método es de orden p , para cualquier h suficientemente pequeño:

$$E(n) \approx ch^p \quad \text{para cierta constante } c > 0.$$

y por tanto, si realizamos otra integración con el doble número de pasos,:

$$E(2n) \approx c \left(\frac{h}{2}\right)^p = \frac{ch^p}{2^p}$$

así:

$$E(2n) \approx \frac{E(n)}{2^p}$$

$$2^p \approx \frac{E(n)}{E(2n)}$$

y usando la función logaritmo, llegamos a la expresión:

$$p \approx \frac{\log(E(n)) - \log(E(2n))}{\log 2} \quad (2.12)$$

n	Euler Implícito	Regla Trapecio
8	1.0270151	2.0015550
16	1.0148887	2.0003886
32	1.0077602	2.0000971
64	1.0039539	2.0000243
128	1.0019947	2.0000061
256	1.0010017	2.0000016
512	1.0005019	2.0000015

Tabla 2.3: Aproximación del orden de los métodos con n pasos.

En la Tabla 2.3 vemos estos valores para distintas integraciones de los dos métodos implícitos para $d=1$ y $N=10$, Ahí claramente vemos que ambos métodos reflejan en la práctica el orden predecido por la teoría. En el caso del método de Euler explícito esto ocurre de forma similar pero siempre que se reduzca el paso por estabilidad lo suficiente para que el método integre el problema apropiadamente.

Con la integración a paso fijo y las gráficas logarítmicas podemos construir gráficas de eficiencia de cada método colocando $\log(E(2^k n))$ en el eje de ordenadas y $\log(2^k n)$ en el de abscisas, con $k = 3, \dots, 10$, de la forma:

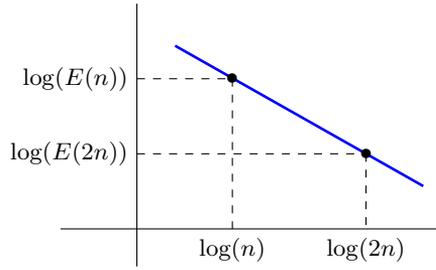


Figura 2.1: Gráfica logarítmica.

Esto se debe a que si el método va con el orden p predicho por la teoría, por (2.12) se tiene que el cociente que aparece a la derecha debe ser el mismo para todos los n . Así si dibujamos la recta que une los puntos $(\log(n), \log(E(n)))$ y $(\log(2n), \log(E(2n)))$, como vemos en la Fig. 2.1, su pendiente es :

$$m = \frac{\log(E(2n)) - \log(E(n))}{\log(2n) - \log(n)}$$

como $\log(2n) - \log(n) = \log\left(\frac{2n}{n}\right) = \log 2$, la pendiente será:

$$m = \frac{\log(E(2n)) - \log(E(n))}{\log 2}$$

Por tanto, comparando con (2.12) vemos que si los métodos van con orden p , la gráfica logarítmica en la que se pone en el eje de abscisas los valores $n = 2^k$, para $k = 3, \dots, 10$, frente a los errores $E(2^k)$ en el eje vertical, debe salir una línea recta de pendiente $-p$. Esto lo podemos comprobar con el Euler implícito y la regla trapezoidal en la Fig. 2.2.

Como podemos observar, la pendiente de la recta obtenida con el método de Euler implícito tiende a $m \approx -1$ y en la regla del trapecio tenemos una pendiente aproximada de $m \approx -2$.

En esta gráfica también se puede ver que la regla trapezoidal es más eficiente que el Euler implícito. Si pensamos que cada paso de la regla trapezoidal tiene un coste computacional doble que cada paso de Euler implícito, el coste de dar n pasos de la regla trapezoidal es equivalente al de dar $2n$ pasos con el EI. En la Fig. 2.2 anterior se puede ver que si comparamos el error obtenido con ambos métodos en n y en $2n$ pasos, la regla trapezoidal le saca al menos una cifra significativa al método de Euler implícito.

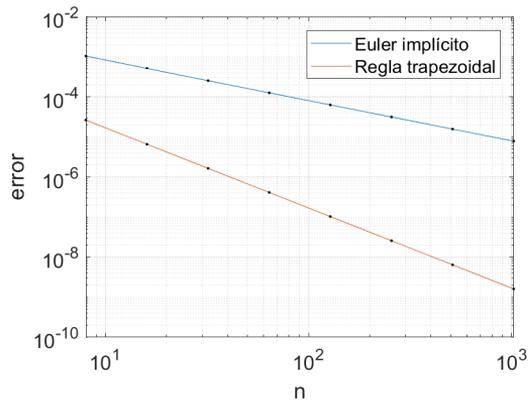


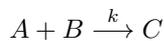
Figura 2.2: Comparativa de orden de los métodos implícitos.

Problemas no lineales

Como hemos visto en el capítulo anterior, los problemas de valor inicial lineales aparecen en muchas aplicaciones. Sin embargo, en el momento en que tratamos de modelizar problemas más realistas, la no-linealidad surge de manera natural. Además, existen muchos ámbitos de la ciencia y la tecnología en los que los modelos que se utilizan para simular un amplio y variado grupo de fenómenos resultan ser problemas no lineales stiff, como, por ejemplo, el siguiente modelo de química cinética.

3.1. Modelo de Robertson

Antes de presentar el modelo con el que vamos a trabajar, veamos cómo se modeliza una reacción química sencilla, recogida en [5], de la forma:



Este esquema describe una reacción con ratio constante k entre dos compuestos, A y B , de la que resulta el compuesto producto C . A continuación denotamos a , b y c como las concentraciones de los tres compuestos, y por la ley de acción de masas se tiene que la velocidad de la reacción es proporcional a ab , lo que se expresa matemáticamente con el siguiente sistema de ecuaciones diferenciales ordinarias:

$$a'(t) = -ka(t)b(t),$$

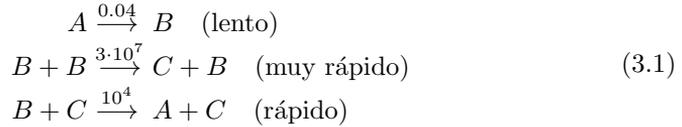
$$b'(t) = -ka(t)b(t),$$

$$c'(t) = ka(t)b(t).$$

Como podemos observar $a'(t) + c'(t) = 0$ y $b'(t) + c'(t) = 0$. Esto es debido a que nos encontramos ante un sistema de reacciones químicas cerrado y por lo tanto

el número total de átomos debe de ser constante.

Sabiendo lo anterior, podemos considerar el ejemplo de H.H. Robertson (1966) de [4], el cual se hizo muy popular en la década de los 70 en el análisis numérico:



lo que nos lleva a la ecuación:

$$\begin{aligned} A : \quad y_1' &= -0.04y_1 + 10^4 y_2 y_3 & y_1(0) &= 1 \\ B : \quad y_2' &= 0.04y_1 - 10^4 y_2 y_3 - 3 \cdot 10^7 y_2^2 & y_2(0) &= 0 \\ C : \quad y_3' &= & 3 \cdot 10^7 y_2^2 & y_3(0) = 0 \end{aligned} \quad (3.2)$$

con $t \in [0, T]$.

Así obtenemos un problema de valor inicial no lineal de dimensión 3, que conjuga tres reacciones químicas cuyas velocidades son muy diferentes en escala. Este acoplamiento de distintas ecuaciones con escalas de valores muy diferentes entre sí es un indicativo práctico de que es bastante probable que el problema sea stiff. Para la resolución de este problema con los métodos estudiados en los capítulos anteriores tenemos que tener en cuenta que dos de esos métodos, el Euler implícito y la regla trapezoidal, tienen carácter implícito, y por tanto, debemos de trabajar con el método de Newton para poder implementarlos como vimos en la Sección 1.5.

Como no es posible obtener una expresión aplicable de la solución de este problema en los experimentos numéricos que presentamos, usaremos como solución de referencia para el cálculo de los errores la obtenida con la rutina ode23s [9] de Matlab, con tolerancia absoluta de 10^{-8} y relativa de 10^{-6} .

3.2. Método de Newton

Para aplicar el método de Newton sobre los métodos implícitos aplicados a este PVI, comenzamos reescribiendo el problema en forma vectorial:

$$\begin{cases} y' = f(y) \\ y(0) = y_0 \end{cases} \quad (3.3)$$

con $y, f \in \mathbb{R}^3$, $t \in [0, T]$ y tal que:

$$f(y) = \begin{pmatrix} -0.04y_1 + 10^4y_2y_3 \\ 0.04y_1 - 10^4y_2y_3 - 3 \cdot 10^7y_2^2 \\ 3 \cdot 10^7y_2^2 \end{pmatrix}, \quad y_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

3.2.1. Euler Implícito

Recordamos la fórmula del método de Euler implícito, que para el problema autónomo (3.3) resulta:

$$y_{i+1} = y_i + hf(y_{i+1}) \quad \text{con } i = 0, \dots, n-1$$

siendo el tamaño de paso $h = \frac{T}{n}$.

En cada paso de integración, por tanto, tenemos que resolver un sistema implícito de ecuaciones. Denotando $y_{i+1} = \omega$:

$$\omega = y_i + hf(\omega)$$

el sistema implícito es:

$$F(\omega) = \omega - y_i - hf(\omega) = 0$$

Para aplicar el método de Newton necesitamos computar la matriz jacobiana:

$$F'(\omega) = I - h \frac{\partial f}{\partial y}(\omega)$$

donde

$$\frac{\partial f}{\partial y}(y) = \begin{pmatrix} -0.04 & 10^4y_3 & 10^4y_2 \\ 0.04 & -10^4y_3 - 6 \cdot 10^7y_2 & -10^4y_2 \\ 0 & 6 \cdot 10^7y_2 & 0 \end{pmatrix}$$

es la matriz jacobiana de $f(y)$.

En consecuencia, el método de Newton queda de la forma:

$$\omega^{(0)} = y_i$$

$$F'(\omega^{(\nu)})(\omega^{(\nu+1)} - \omega^{(\nu)}) = -F(\omega^{(\nu)})$$

$$\text{Si } \|\omega^{(\nu+1)} - \omega^{(\nu)}\|_\infty \leq Tol \Rightarrow y_{n+1} = \omega^{(\nu)}$$

o, lo que es lo mismo, arrancando del valor inicial $\omega^{(0)} = y_i$, por cada iteración tenemos que resolver el sistema lineal:

$$\left(I - h \frac{\partial f}{\partial y}(\omega^{(\nu)}) \right) \Delta^{(\nu)} = -\omega^{(\nu)} + y_i + hf(\omega^{(\nu)}) \quad (3.4)$$

$$\omega^{(\nu+1)} = \omega^{(\nu)} + \Delta^{(\nu)}$$

con $\nu = 0, \dots, nmax$, siendo $nmax$ el número máximo de iteraciones para Newton que le marquemos, que en el caso óptimo no llegarían a ser alcanzadas aplicando el criterio de parada:

$$\text{Si } \|\Delta^{(\nu)}\| \leq Tol \Rightarrow y_{i+1} = \omega^{(\nu)}$$

siendo Tol una tolerancia de error prefijada por el usuario.

3.2.2. Regla del Trapecio

De forma análoga al método de Euler implícito tenemos que aplicar también el método de Newton para resolver el sistema implícito que aparece en cada paso de la regla del trapecio cuya fórmula sobre el problema (3.2) es:

$$y_{i+1} = y_i + \frac{h}{2}[f(y_i) + f(y_{i+1})] \quad \text{con } i = 0, \dots, n-1$$

Si denotamos $y_{i+1} = \omega$ a la solución del sistema, tenemos $y_{i+1} = \omega$ y tenemos:

$$\omega = y_i + \frac{h}{2}[f(y_i) + f(\omega)]$$

y así, tenemos que resolver el sistema implícito:

$$F(\omega) = \omega - y_i - \frac{h}{2}f(y_i) - \frac{h}{2}f(\omega) = 0$$

Ahora, en este caso, Newton resulta:

$$\omega^{(0)} = y_i$$

$$\left(I - \frac{h}{2} \frac{\partial f}{\partial y}(\omega^{(\nu)}) \right) \Delta^{(\nu)} = -\omega^{(\nu)} + y_i + \frac{h}{2}f(y_i) + \frac{h}{2}f(\omega^{(\nu)}) \quad (3.5)$$

$$\omega^{(\nu+1)} = \omega^{(\nu)} + \Delta^{(\nu)}$$

donde ν recorre desde 0 hasta el número máximo de iteraciones que, como en el caso anterior, no llegarían a ser alcanzadas si se cumple la tolerancia marcada.

3.3. Experimentos numéricos

3.3.1. Implementación de los métodos a paso fijo

Hemos construido un código Matlab que integra el problema de Robertson (3.2) a paso fijo con los tres métodos estudiados: Euler explícito, Euler implícito y regla trapezoidal.

En primer lugar, si aplicamos Euler explícito para $T = 1$ obtenemos que para un número de pasos fijos $n = 729$ el código no consigue una solución válida. En cambio, con $n = 730$, se obtiene un error de $3.49433 \cdot 10^{-3}$. Esto claramente refleja una severa restricción de paso por estabilidad. Como, además, este problema tiene interés para valores de T grandes, es evidente que el método de Euler explícito no da buenos resultados.

Para integrar el problema con los métodos implícitos hemos implementado el método de Newton para resolver los sistemas implícitos de ambos, mediante (3.4) y (3.5) respectivamente, teniendo como número máximo de iteraciones $nmax = 20$ y una tolerancia de error $Tol = 10^{-4}$.

Recogemos en las Tablas 3.1-3.4 los errores absolutos y relativos en el último punto de integración obtenidos con los métodos implícitos, de la forma

$$\left\| \left(\frac{y_k(T) - y_k}{y_k(T)} \right)_{k=1}^3 \right\|_{\infty}$$

n	Euler Implícito	Regla Trapezoidal
20	2.46940e-04	4.61096e-03
40	1.27652e-04	8.62978e-05
80	6.24863e-05	2.97118e-04
160	3.09340e-05	4.52546e-06
320	1.54122e-05	2.70220e-08

n	Euler Implícito	Regla Trapezoidal
20	7.36925e-03	3.58563e-01
40	3.80941e-03	1.89461e-02
80	1.86473e-03	8.86666e-03
160	9.23140e-04	1.35050e-04
320	4.59935e-04	1.18715e-06

Tabla 3.1: Errores absolutos ($T = 1$). Tabla 3.2: Errores relativos ($T = 1$).

n	Euler Implícito	Regla Trapezoidal
800	1.74119e-04	4.75177e-04
1600	8.71086e-05	8.46637e-06
3200	4.30912e-05	2.72813e-05
6400	2.11661e-05	1.11635e-06
12800	1.02239e-05	7.12530e-07

n	Euler Implícito	Regla Trapezoidal
800	7.62744e-04	1.98232e-03
1600	3.94301e-04	2.97941e-05
3200	2.08031e-04	9.60059e-05
6400	1.15257e-04	2.09753e-05
12800	6.89577e-05	2.26850e-05

Tabla 3.3: Errores absolutos ($T = 40$). Tabla 3.4: Errores relativos ($T = 40$).

En primer lugar, de estos experimentos podemos concluir de forma clara que, a pesar de ser más cara computacionalmente, la regla trapezoidal es más eficiente sobre este problema que el método de Euler implícito.

En segundo lugar, comparando los errores absolutos y relativos vemos la importante diferencia de escala entre las tres componentes (la primera y la tercera toman valores sobre 10^{-4} mientras que la cuarta toma valores del orden de 10^{-8}), lo que es una de las características habituales que presentan los problemas stiff.

3.3.2. Implementación a paso variable de Matlab

En la sección anterior hemos visto que el método de Euler explícito es incapaz de dar una solución razonable para este problema. Cabría la duda de si esto se debe a que es explícito o si es debido a que se tiene una implementación pobre de este método. Para comprobar que este mismo comportamiento se va a dar también para métodos explícitos mejores y más sofisticados, en esta sección vamos a probar con dos rutinas de Matlab mucho mejores que nuestros esquemas, que además usan una estrategia de paso variable que es más eficiente en general. El programa Matlab tiene varias funciones internas para aproximar soluciones de ecuaciones diferenciales. Para trabajar con este problema no lineal y poder analizarlo con mayor precisión, usaremos dos de ellas:

- *ode45* [8], basado en el par explícito Runge-Kutta DOPRI(4,5) de orden 4, muy versátil, es el *solver* de Matlab recomendado como primera prueba para la resolución de problemas. Sin embargo, si nos encontramos ante un problema stiff o que requiera una alta precisión existen otras rutinas mejores.
- *ode23s*, está basado en un RK semi-implícito de tipo Rosenbrock de orden 2 con un nivel de eficiencia superior a la rutina anterior y la mejor elección si nos encontramos ante un problema stiff.

En las Tablas 3.5 - 3.9 vemos una comparativa de los resultados obtenidos por las rutinas de Matlab analizando su error con respecto a *ode23s* con $AbsTol = 10^{-8}$ y $RelTol = 10^{-6}$. En ellas NPA significa el número de pasos aceptados del método, es decir, el número de pasos en total dados para llevar a cabo la integración. NPR significa el número de pasos rechazados, es decir, el número de pasos que rechaza por no conseguir la tolerancia de error pedida. Error denota la norma infinita del error obtenido en el punto final de integración.

AbsTol	RelTol	NPA ODE45	NPR ODE45	Error ODE45	NPA ODE23S	NPR ODE23S	Error ODE23S
10^{-4}	10^{-2}	–	–	–	14	2	5.7500e-08
10^{-5}	10^{-3}	–	–	–	15	2	6.6918e-08
10^{-6}	10^{-4}	678	148	4.5244e-07	16	2	6.4828e-08

Tabla 3.5: $T = 1$.

AbsTol	RelTol	NPA ODE45	NPR ODE45	Error ODE45	NPA ODE23S	NPR ODE23S	Error ODE23S
10^{-4}	10^{-2}	–	–	–	16	3	3.7725e-05
10^{-5}	10^{-3}	–	–	–	17	3	2.2171e-05
10^{-6}	10^{-4}	7276	1392	2.2423e-05	21	3	1.0320e-05

Tabla 3.6: $T = 10$.

AbsTol	RelTol	NPA ODE45	NPR ODE45	Error ODE45	NPA ODE23S	NPR ODE23S	Error ODE23S
10^{-4}	10^{-2}	–	–	–	19	3	4.3561e-04
10^{-5}	10^{-3}	–	–	–	22	3	1.6186e-04
10^{-6}	10^{-4}	106175	18115	6.4702e-04	28	3	6.0021e-05

Tabla 3.7: $T = 100$.

AbsTol	RelTol	NPA ODE45	NPR ODE45	Error ODE45	NPA ODE23S	NPR ODE23S	Error ODE23S
10^{-4}	10^{-2}	–	–	–	23	3	8.8866e-04
10^{-5}	10^{-3}	–	–	–	28	3	2.8144e-04
10^{-6}	10^{-4}	1.71976e+06	367552	0.0049	37	3	9.6662e-05

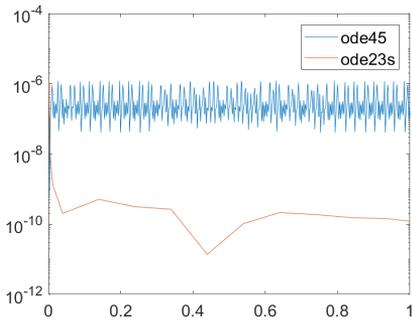
Tabla 3.8: $T = 1000$.

AbsTol	RelTol	NPA ODE45	NPR ODE45	Error ODE45	NPA ODE23S	NPR ODE23S	Error ODE23S
10^{-4}	10^{-2}	–	–	–	27	3	6.0328e-04
10^{-5}	10^{-3}	–	–	–	35	3	1.9376e-04
10^{-6}	10^{-4}	2.39595e+07	5.34415e+06	0.0467	50	3	5.9183e-05

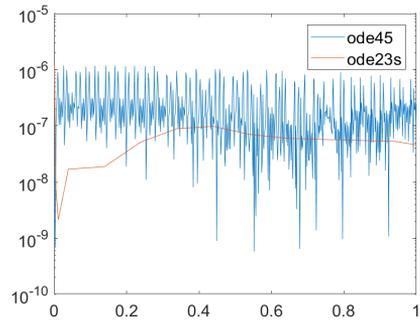
Tabla 3.9: $T = 10000$.

En estas Tablas podemos observar que la rutina basada en el método explícito es incapaz de acabar la integración de este problema o, cuando lo hace, la aproximación que da no sirve.

En las Figuras 3.1 y 3.2 se puede ver las aproximaciones obtenidas por dichos métodos a las componentes 2 y 3 del problema.

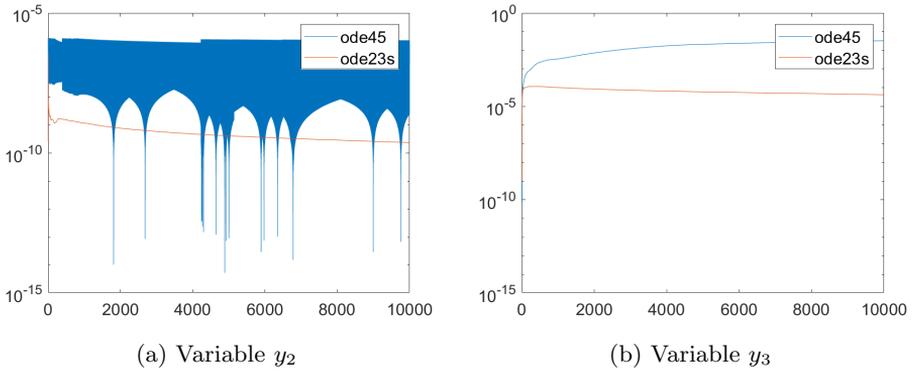


(a) Variable y_2



(b) Variable y_3

Figura 3.1: $T = 1$.

Figura 3.2: $T = 10000$.

En estas gráficas se ve que es principalmente la segunda componente la que hace que el método explícito vaya mal, y la gran diferencia de escala entre la segunda y la tercera. Así vemos que la stiffness hace imprescindible el uso de métodos implícitos.

Bibliografía

- [1] ASCHER, U.M., PETZOLD L.R. (1998) *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia, Pennsylvania, United States: SIAM.
- [2] ATKINSON, K.E. (1988). *An introduction to numerical analysis*. (2nd ed.) New York City, United States: Wiley.
- [3] HAIRER E., NORSETT S.P., WANNER G. (1993) *Solving Ordinary Differential Equations I: Nonstiff Problems*. (2nd ed.). New York, United States: Sprintger.
- [4] HAIRER E., WANNER G. (1996) *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems* (2nd ed.). New York, United States: Sprintger.
- [5] HUNSDORFER W., VERWER J.G. (1996) *Numerical Solution of Advection-Diffusion-Reaction Equations*.
- [6] MEYER C.D. (2000) *Matrix analysis and applied linear algebra*. Philadelphia, Pennsylvania, United States: SIAM.
- [7] SPIJKER, M.N. (1996) *Error propagation in Runge-Kutta methods*. Amsterdam, The Netherlands: Elsevier.
- [8] *ode45* Recuperado de <https://es.mathworks.com/help/matlab/ref/ode45.html>
- [9] *ode23s* Recuperado de <https://es.mathworks.com/help/matlab/ref/ode23s.html>

Numerical methods for differential equations and applications

Sección de Matemáticas
Universidad de La Laguna

Ricardo Calvo Castro
Facultad de Ciencias · Sección de Matemáticas
Universidad de La Laguna
alu0101072159@ull.edu.es

Abstract

THE study of some numerical methods for the solution of initial value problems (IVP) in ordinary differential equations (ODE) is the main goal of this work. In particular, we will focus our attention on three basic schemes: the forward Euler method, the backward Euler and the trapezoidal rule. We will show classic theorems on the convergence of these methods, as well as their properties of order and linear stability. In the last chapters we will apply these methods to two representative problems.

1. Introduction

INITIAL value problems (IVPs) with systems of ordinary differential equations (ODEs) appear in many fields of science and technology when mathematical modeling techniques are used to describe a wide variety of phenomena. Except in the case of certain simple, scalar or moderate-sized ODEs, the "exact" calculation of the solution to these problems is so complicated that numerical methods are essential to obtain approximations to that solution accurately and efficiently. This work is an introduction to some of the numerical methods that are usually applied for the resolution of IVPs that can be non-linear and large-scale: forward Euler, backward Euler and trapezoidal rule.

2. Outline

CHAPTER 1. For the numerical solution of the IVP $y' = f(x, y)$, $y(0) = y_0$, $x \in [a, b]$, $y, f \in \mathbb{R}^m$

we consider the following methods:

- Forward Euler (FE): $y_{i+1} = y_i + hf(x_i, y_i)$
 - Backward Euler (BE): $y_{i+1} = y_i + hf(x_i, y_{i+1})$
 - Trapezoidal Rule (TR): $y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1}))$
- with $x_i = a + ih$, $i = 0, \dots, n-1$.

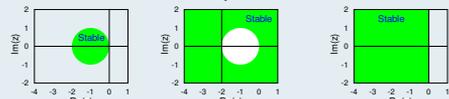
Theorems of Convergence:

$$\text{FE: } \|e_i\| \leq \frac{Mh}{2L} [e^{L(b-a)} - 1], \quad 0 < h \leq h^*, \quad \forall i = 0, 1, 2, \dots, n.$$

$$\text{BE: } \|e_i\| \leq \left[\frac{M}{2(1-h^*L)} \frac{e^{L(b-a)} - 1}{L} \right] h, \quad 0 < h \leq h^*, \quad \forall i = 0, 1, 2, \dots, n.$$

$$\text{TR: } \|e_i\| \leq \left[\frac{M}{4(2-h^*L)} \frac{e^{L(b-a)} - 1}{L} \right] h^2, \quad 0 < h \leq h^*, \quad \forall i = 0, 1, 2, \dots, n.$$

Linear Stability: $|y_{i+1}| \leq |y_i|, \forall i$, on the test equation $y' = \lambda y$, $\lambda \in \mathbb{C}$, if $z = h\lambda$ is on the stability domains



$$\text{FE: } |1 + h\lambda| \leq 1 \quad \text{BE: } |1/(1 - h\lambda)| \leq 1 \quad \text{TR: } |2 + h\lambda|/|2 - h\lambda| \leq 1$$

Stiff Problems: IVPs for which an extremely small step size must be taken to achieve stability.

Explicit methods do not work on stiff problems \Rightarrow implicit methods

Implicit methods implementation:

- Functional iteration
- Newton method

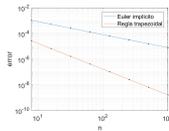
CHAPTER 2. **Linear Problems:** partial differential equation: $u_t = d u_{xx} + g(x, t)$, $u(0, t) = u(1, t) = 0$, $u(x, 0) = f_0(x)$, $x, t \in [0, 1]$

Method of lines: spatial discretization $x_i = i\Delta x$, $1 \leq i \leq N$, $\Delta x = 1/(N+1)$, $U(t) = (u(x_i, t))_{i=1}^N$:

$$U'(t) = AU + b(x), \quad x \in [0, 1]$$

$$A = \frac{d}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & \dots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -2 \end{bmatrix} \in \mathbb{M}_{N \times N}, \quad b(x) = \begin{pmatrix} g(x_1, t) \\ g(x_2, t) \\ \vdots \\ g(x_N, t) \end{pmatrix} \in \mathbb{R}^N$$

Order of the Methods: tests of the performance of the methods are given and the computed order:



CHAPTER 3. **Non-linear problems: Robertson's Problem,** which models a chemical reaction, $t \in [0, T]$:

$$\begin{aligned} y_1' &= -0.04y_1 + 10^4 y_2 y_3 & y_1(0) &= 1 \\ y_2' &= 0.04y_1 - 10^4 y_2 y_3 - 3 \cdot 10^7 y_2^2 & y_2(0) &= 0 \\ y_3' &= & 3 \cdot 10^7 y_2^2 & y_3(0) &= 0 \end{aligned}$$

Newton Method for implicit schemes, $\omega^{(0)} = y_i$,

$$\text{BE: } \left(I - h \frac{\partial f}{\partial y}(\omega^{(v)}) \right) \Delta \omega^{(v)} = -\omega^{(v)} + y_i + hf(\omega^{(v)})$$

$$\text{TR: } \left(I - \frac{h \partial f}{2 \partial y}(\omega^{(v)}) \right) \Delta \omega^{(v)} = -\omega^{(v)} + y_i + \frac{h}{2} f(y_i) + \frac{h}{2} f(\omega^{(v)}), \quad v = 0, 1, \dots$$

Numerical experiments:

n	BE	TR	n	BE	TR
800	1.74119e-04	4.75177e-04	800	7.62744e-04	1.98232e-03
1600	8.71086e-05	8.46637e-06	1600	3.94301e-04	2.97941e-05
3200	4.30912e-05	2.72813e-05	3200	2.08031e-04	9.60059e-05
6400	2.11661e-05	1.11635e-06	6400	1.15257e-04	2.09753e-05
12800	1.02239e-05	7.12530e-07	12800	6.89577e-05	2.26850e-05

Absolute errors

Relative errors

References

- [1] ASCHER U.M., PETZOLD L.R. (1998) *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia, Pennsylvania, United States: SIAM.
- [2] HAIRER E., NORSETT S.P., WANNER G. (1993) *Solving Ordinary Differential Equations I: Nonstiff Problems*. (2nd ed.). New York, United States: Springer.
- [3] HAIRER E., WANNER G. (1996) *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems* (2nd ed.). New York, United States: Springer.
- [4] HUNSDORFER W., VERWER J.G. (1996) *Numerical Solution of Advection-Diffusion-Reaction Equations*.
- [5] SPIJKER M.N. (1996) *Error propagation in Runge-Kutta methods*. Amsterdam, The Netherlands: Elsevier.