

Adrián José Jorge Trujillo

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

Trabajo Fin de Grado  
Grado en Matemáticas  
La Laguna, Septiembre de 2018

DIRIGIDO POR  
*Soledad Pérez Rodríguez*

*Soledad Pérez Rodríguez*

*Departamento de Análisis Ma-  
temático*

*Universidad de La Laguna*

*38271 La Laguna, Tenerife*

---

## Agradecimientos

A mi familia que durante todo este periodo me ha apoyado y mantenido motivado. En especial, quiero agradecer a mi tutora Soledad Pérez Rodríguez que ha estado durante todo el año guiándome, enseñándome y por toda la ayuda que he recibido de su parte. También quiero agradecer al profesor Juan Diego Betancor Ortiz que ha posibilitado la conclusión de esta memoria con su ayuda a la hora de extraer el disco duro de mi portátil, el cual dejó de funcionar el día anterior de la entrega de esta memoria y sin él esta memoria no podría haber sido concluida.



---

## Resumen · Abstract

### *Resumen*

---

*Muchos problemas matemáticos involucran la resolución de un problema de valor inicial en ecuaciones diferenciales ordinarias o incluyen la resolución de ecuaciones en derivadas parciales. Dada la gran diversidad y complejidad de las ecuaciones diferenciales no siempre es posible determinar una solución de forma analítica. En la mayoría de los casos sólo se sabe de la existencia de una solución para dicho problema. Con objeto de determinarla se recurre al uso de los métodos numéricos.*

*En el presente trabajo se introducen, por un lado, distintos métodos numéricos como los Runge-Kutta, los cuales permiten calcular aproximaciones de la solución exacta para problemas de valor inicial.*

*Por otro lado se introduce el método de líneas, un método numérico que permite obtener soluciones de problemas de ecuaciones en derivadas parciales. En pocas palabras, este método aproxima la solución de dichos problemas mediante aproximaciones de problemas de valor inicial de ecuaciones diferenciales ordinarias.*

*Por último, como aplicación de estos métodos se estudiará un problema de polución atmosférica. Concretamente la evolución de las concentraciones de los óxidos de nitrógeno en la baja atmósfera, modelizando las reacciones químicas que tienen lugar y la difusión de las sustancias. Estos problemas de polución atmosférica son muy relevantes en el estudio del comportamiento de la atmósfera y de la evolución del cambio climático.*

**Palabras clave:** *Métodos Runge-Kutta – Stiffness – Método de líneas – Polución atmosférica.*

## **Abstract**

---

*There are a lot of mathematic problems which include initial value problems in ordinary differential equations or partial differential equations. It is not always possible to determine analytically their solutions because of the huge variety and complexity of the differential equations. In many cases, the existence of a solution is the only known information. Due to this fact, numerical methods are one way to approximate the solutions.*

*On one hand, several numerical methods are introduced in this project, like Runge-Kutta methods. These methods enable the calculation of an approximation to the exact solution of initial value problems.*

*On the other hand, the method of lines is introduced as a numerical method which allows calculate approximations for the solution of problems involving partial differential equations. Roughly speaking, this method approximates the solution of these problems by using an approximation of some related initial value problem in ordinary differential equations.*

*Finally, as an application of these methods, a pollution atmospheric problem will be studied. Specifically, the study will be focus on the evolution of the concentration of nitrogen oxides in the low atmosphere. In addition, we will model the chemical reactions which take place and the diffusion phenomena. These pollution atmospheric problems are very relevant in the atmospheric sciences, mainly in the study of climate change.*

**Keywords:** *Runge-Kutta Methods – Stiffness – Method of Lines – Atmospheric pollution.*

---

# Contenido

<b>Agradecimientos</b> .....	III
<b>Resumen/Abstract</b> .....	V
<b>Introducción</b> .....	IX
<b>1. Introducción a los métodos Runge-Kutta</b> .....	1
1.1. Introducción .....	1
1.2. Problemas de valor inicial .....	1
1.3. Método de Euler .....	2
1.4. Otros métodos de un paso .....	3
1.5. Métodos Runge-Kutta .....	6
1.6. Consistencia de los métodos Runge-Kutta .....	10
1.7. Estabilidad y convergencia de los métodos Runge-Kutta .....	16
<b>2. Problemas stiff y estabilidad lineal de los métodos</b> .....	21
2.1. El fenómeno de la stiffness .....	21
2.2. Estabilidad lineal .....	23
2.3. Un modelo de polución atmosférica .....	27
<b>3. Métodos de líneas sobre EDPs parabólicas. Problemas de polución atmosférica con difusión.</b> .....	31
3.1. Introducción .....	31
3.2. Método de líneas sobre la ecuación del calor .....	31
3.3. Condiciones de frontera de tipo Neumann .....	35
3.4. Problemas de química atmosférica con difusión .....	37
3.5. Conclusiones .....	42
<b>Bibliografía</b> .....	45

<b>Lista de Figuras</b> .....	47
<b>Poster</b> .....	49



---

## Introducción

Desde muy antiguo la predicción del tiempo ha sido uno de los grandes retos de la humanidad. Debido al ingente número de variables que influyen en los procesos meteorológicos, hasta el desarrollo de los ordenadores la predicción del tiempo de forma científica (con cierto grado de acierto) ha sido prácticamente imposible. En [7] se recoge un excelente repaso de la historia de esta ciencia.

Un área fundamental dentro del campo de la meteorología es la *química atmosférica*, esto es, la ciencia que estudia los procesos químicos que tienen lugar entre los gases que forman la atmósfera. Debido a la importancia que, desgraciadamente cada día más, tiene la emisión de gases de efecto invernadero a la atmósfera a causa de la acción humana, lo que provoca a corto plazo lluvia ácida (que afecta gravemente a la agricultura y al medio ambiente en general) y a largo plazo el calentamiento global, el desarrollo de modelos matemáticos para predecir la evolución de las concentraciones de dichos contaminantes en la atmósfera es cada vez más relevante.

En 1977 casi todos los países europeos se unieron para crear el European Monitoring and Evaluation Programme for Transboundary Long-Range Transported Air Pollutant (EMEP) [8], que sigue funcionando en la actualidad ampliado a 51 países con Estados Unidos y Canadá, con el objetivo de detectar y analizar la evolución de la lluvia ácida, recogiendo datos de cientos de químicos de toda Europa (con una resolución de  $50 \text{ km} \times 50 \text{ km}$ ). Una parte fundamental del modelo EMEP la constituyen varios modelos de advección-difusión-reacción alguno de los cuales implica hasta 1200 reacciones químicas entre 465 sustancias, lo que claramente está fuera del alcance de un trabajo de este tipo. Sin embargo, en [9, Sec.2] se recoge un problema simplificado de polución atmosférica que es el modelo de difusión-reacción cuya resolución numérica será el objetivo de esta memoria y que se abordará en el último capítulo.

Para llegar a ello, en el Capítulo 1 se presentará la clase de métodos Runge-Kutta para la resolución numérica de problemas de valor inicial en ecuaciones

diferenciales ordinarias. Además, se estudiará su consistencia, estabilidad y convergencia siguiendo [2, 3, 4, 6].

Es habitual que los modelos de química atmosférica presenten *stiffness* pues las reacciones químicas implicadas suelen tener velocidades de reacción muy diferentes entre sí, con escalas muy variables [6]. Es por eso que en el Capítulo 2 se estudian los problemas *stiff* y el comportamiento de los métodos sobre dichos problemas [5].

Para resolver numéricamente los modelos de difusión-reacción se introducirá en el Capítulo 3 el método de líneas [6] para resolver problemas de ecuaciones en derivadas parciales parabólicas, comenzando con el caso más sencillo de la ecuación del calor para, finalmente, conseguir el objetivo de implementar un método para aproximar la solución del modelo de polución atmosférica dado en [9].

# Introducción a los métodos Runge-Kutta

## 1.1. Introducción

En esta sección introduciremos una de las clases de métodos numéricos más utilizados en la actualidad para la resolución de problemas de valor inicial de ecuaciones diferenciales ordinarias, **los métodos Runge-Kutta**.

## 1.2. Problemas de valor inicial

Se considera el problema de valor inicial (PVI) de ecuaciones diferenciales ordinarias (EDO):

$$\left. \begin{array}{l} y' = f(x, y(x)) \\ y(x_0) = y_0 \end{array} \right\} \quad y, f \in \mathbb{R}^m, \quad x \in [x_0, x_N] \quad (1.1)$$

El siguiente teorema garantiza que el problema tiene solución y además esta es única.[3, p. 23]

*Teorema 1.2.1.* Si la función derivada  $f(x, y)$  del PVI (1.1) verifica las siguientes condiciones:

- (1) La función  $f \in C([x_0, x_N] \times Cl(\Omega))$  donde  $\Omega \subseteq \mathbb{R}^m$  es un abierto en el que está definida la función y  $Cl(\Omega)$  es su clausura.
- (2)  $\|f(x, y) - f(x, z)\| \leq L\|y - z\|, \forall (x, y), (x, z) \in [x_0, x_N] \times Cl(\Omega)$ , esto es, la función  $f(x, y)$  es lipschitziana respecto de  $y$ .

el problema (1.1) tiene una única solución.

Los problemas de valor inicial (1.1) no sólo incluyen problemas de ecuaciones diferenciales ordinarias de orden 1. Un PVI de orden  $n$  y dimensión  $m$ :

$$\left. \begin{aligned} y^{(n)} &= f(x, y, y', \dots, y^{(n-1)}) \\ y(x_0) &= y_0 \\ &\vdots \\ y^{(n-1)}(x_0) &= y_0^{(n-1)} \end{aligned} \right\},$$

se puede transformar en otro de la forma (1.1), considerando el cambio:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-2)} \\ y^{(n-1)} \end{pmatrix} \in \mathbb{R}^{m \cdot n} \Rightarrow v' = \begin{pmatrix} v_2 \\ v_3 \\ \vdots \\ v_n \\ f(x, v_1, \dots, v_n) \end{pmatrix} \quad v_0 = \begin{pmatrix} y_0 \\ y'_0 \\ \vdots \\ y_0^{(n-2)} \\ y_0^{(n-1)} \end{pmatrix}.$$

Asimismo es posible transformar el problema (1.1) en un problema autónomo, es decir, un PVI de la forma:

$$\left. \begin{aligned} v' &= g(v(x)) \\ v(x_0) &= v_0 \end{aligned} \right\}, \tag{1.2}$$

haciendo el cambio:

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow v' = \begin{pmatrix} 1 \\ f(v_1, v_2) \end{pmatrix} = g(v) \quad v_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Obsérvese que este PVI (1.2) tiene dimensión  $m + 1$ .

### 1.3. Método de Euler

A continuación se introducirán algunos de los métodos numéricos que se utilizan para la resolución de ecuaciones diferenciales ordinarias. Uno de los primeros métodos que se usaron para aproximar la solución del problema de valor inicial (1.1) es el **método de Euler**.

Para aplicarlo se toma una partición del intervalo  $[x_0, x_N]$  de  $n + 1$  elementos.

$$\mathcal{P} = \{x_0 < x_1 < \dots < x_n = x_N\} \tag{1.3}$$

denominando **tamaño del paso**  $h_i$  a la diferencia entre dos elementos consecutivos de la partición.

$$h_i = x_{i+1} - x_i, \quad 0 \leq i \leq n - 1.$$

Por simplicidad, en este trabajo, salvo que se especifique en algún caso concreto, vamos a considerar un tamaño de paso fijo  $h_i = h = (x_N - x_0)/n$ ,  $0 \leq i \leq n - 1$ .

El método de Euler proporciona una aproximación a los valores de la solución  $y(x)$  en los puntos de la partición  $\mathcal{P}$  (1.3) basándose en el desarrollo de Taylor de la solución exacta del problema (1.1) en torno a  $x_0$ :

$$y(x_0 + h) = y(x_0) + hy'(x_0) + \mathcal{O}(h^2) = y_0 + hf(x_0, y_0) + \mathcal{O}(h^2), \quad h \rightarrow 0$$

Así, el método de Euler toma como aproximación  $y_1$  de la solución en el punto  $x_0 + h$  al valor

$$y_1 = y_0 + hf(x_0, y_0)$$

Además, se tiene:

$$y(x_0 + h) - y_1 = \mathcal{O}(h^2), \quad h \rightarrow 0,$$

por lo que es esperable tener un error del orden de  $h^2$  en  $x_1$ . Repitiendo este proceso en cada punto  $x_i$  de la partición  $\mathcal{P}$ , el método de Euler proporciona las siguientes aproximaciones:

$$y_{i+1} = y_i + hf(x_i, y_i), \quad 0 \leq i \leq n-1 \quad (1.4)$$

Como en la integración total se dan  $n$  pasos, ( $n = (x_N - x_0)/h$ ) se espera que el error acumulado sea aproximadamente  $\mathcal{O}(h)$ , ( $h \rightarrow 0$ ).

De hecho, se demuestra por [4, p. 40]:

*Teorema 1.3.1 (Convergencia del método de Euler).* Consideremos un PVI (1.1), con  $f$  verificando las condiciones del Teorema 1.2.1 y de clase  $C^1$  en todo su dominio. Si  $\{y_i\}_{i=0}^n$  es la solución numérica dada por el método de Euler (1.4) sobre la partición  $\mathcal{P}$  para dicho PVI y se verifica

$$|f| \leq A, \quad \left| \frac{\partial f}{\partial x} \right| \leq M, \quad \left| \frac{\partial f}{\partial y} \right| \leq L,$$

tenemos que el error estimado para las aproximaciones dadas por dicho método cumple:

$$|y(x_i) - y_i| \leq \frac{M + AL}{L} \left( e^{L(x_1 - x_0)} - 1 \right) |h|$$

con  $|h|$  suficientemente pequeño.

## 1.4. Otros métodos de un paso

El método de Euler presenta un error del orden del tamaño de paso  $h$ , lo que lo hace un método muy lento. Para mejorar esto fueron surgiendo variantes de este método que presentan un error del orden de potencias más altas del tamaño de paso  $h$  como, por ejemplo:

Regla del punto medio explícita:

$$\begin{aligned} y_{i+\frac{1}{2}} &= y_i + \frac{h}{2} f(x_i, y_i) \\ y_{i+1} &= y_i + hf(x_i + \frac{h}{2}, y_{i+\frac{1}{2}}), \quad 0 \leq i \leq n-1. \end{aligned} \quad (1.5)$$

Regla trapezoidal explícita:

$$\begin{aligned} y_{i+1}^* &= y_i + hf(x_i, y_i) \\ y_{i+1} &= y_i + \frac{h}{2} (f(x_i, y_i) + f(x_i + h, y_{i+1}^*)), \quad 0 \leq i \leq n-1. \end{aligned} \quad (1.6)$$

A estos métodos se les denomina explícitos porque es posible calcular la aproximación en cada paso directamente evaluando la función  $f(x, y)$  en los pasos anteriores.

Por otro lado, también existen otros métodos denominados implícitos en los que las aproximaciones vienen definidas por un sistema de ecuaciones implícito.

Por ejemplo, existen versiones implícitas de los tres métodos citados anteriormente:

Método de Euler implícito:

$$y_{i+1} = y_i + hf(x_i + h, y_{i+1}), \quad 0 \leq i \leq n-1. \quad (1.7)$$

Regla del punto medio implícita:

$$y_{i+1} = y_i + hf\left(x_i + \frac{h}{2}, \frac{y_i + y_{i+1}}{2}\right), \quad 0 \leq i \leq n-1. \quad (1.8)$$

Regla trapezoidal implícita:

$$y_{i+1} = y_i + \frac{h}{2} (f(x_i, y_i) + f(x_i + h, y_{i+1})), \quad 0 \leq i \leq n-1. \quad (1.9)$$

Como podemos observar en estas fórmulas, para calcular cada aproximación  $y_{i+1}$  hay que resolver un sistema implícito.

Para visualizar como actúan los diferentes métodos hemos resuelto numéricamente el siguiente ejemplo de PVI.

*Ejemplo 1.4.1.*

$$y' = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} y + \begin{pmatrix} 0 \\ 10 \end{pmatrix} \quad y(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad t \in [0, 3]. \quad (1.10)$$

Es fácil comprobar que la solución exacta de este problema es:

$$y(t) = \begin{pmatrix} 10t - 10(1 - e^{-t}) \\ 10(1 - e^{-t}) \end{pmatrix}$$

Hemos aplicado los métodos anteriores en diferentes tamaños de paso fijo  $h = 3/n$  y hemos comparado la solución numérica dada por los métodos con la solución exacta. En las figuras 1.1 y 1.2 se representan los errores  $\|y_i - y(x_i)\|_\infty$ .

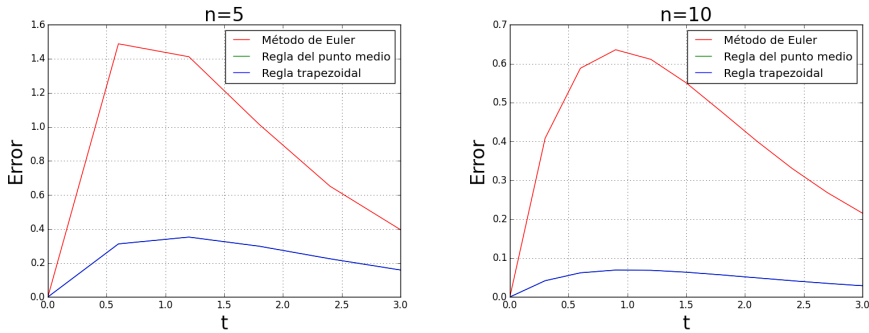


Figura 1.1. Error en los métodos explícitos del ejemplo 1.4.1.

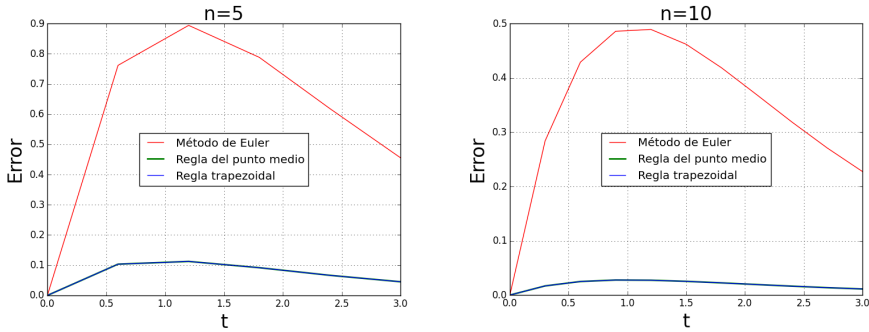


Figura 1.2. Error en los métodos implícitos del ejemplo 1.4.1.

*Nota.* El procedimiento usado para el cálculo de las aproximaciones de los métodos implícitos se explicará más adelante.

Se puede observar que la regla del punto medio y la regla trapezoidal presentan un menor error que el método de Euler, tanto en su versión explícita como en la implícita. De hecho, se obtiene que el error asociado a la regla del punto medio y la regla trapezoidal es el mismo para este ejemplo.

*Lema 1.4.1.* La regla del punto medio y la regla trapezoidal dan la misma solución numérica sobre PVI (1.1) lineales de coeficientes constantes.

*Demostración.* Sea un PVI lineal de coeficientes constantes:

$$y' = Ay + b.$$

Por inducción, si tenemos una aproximación  $y_i$  a su solución en un punto  $x_i$  veamos que las aproximaciones dadas por la regla del punto medio y la regla

trapezoidal para el valor  $x_{i+1}$  son iguales. Sean  $\tilde{y}_{i+1}$ ,  $\bar{y}_{i+1}$  las aproximaciones dadas las versiones explícitas por (1.5) y (1.6) respectivamente. Tenemos que:

$$\begin{aligned}\tilde{y}_{i+1} &= y_i + h \left( A \left( y_i + \frac{h}{2} (Ay_i + b) \right) + b \right) = \\ &= y_i + h \left( Ay_i + b + \frac{h}{2} A^2 y_i + \frac{h}{2} Ab \right) = y_i + hAy_i + hb + \frac{h^2}{2} A^2 y_i + \frac{h^2}{2} Ab = \\ &= y_i + \frac{h}{2} (Ay_i + b) + \frac{h}{2} (A(y_i + hAy_i + hb) + b) = \bar{y}_{i+1}\end{aligned}$$

La demostración para los métodos implícitos sigue el mismo razonamiento. Sean  $\tilde{y}_{i+1}$ ,  $\bar{y}_{i+1}$  las aproximaciones dadas por (1.8) y (1.9) respectivamente. Tenemos que:

$$\begin{aligned}\tilde{y}_{i+1} &= y_i + h \left( A \frac{y_i + \tilde{y}_{i+1}}{2} + b \right) \Rightarrow \left( I_m - \frac{h}{2} A \right) \tilde{y}_{i+1} = \left( I_m - \frac{h}{2} A \right) y_i + hb \\ &\Rightarrow \tilde{y}_{i+1} = y_i + h \left( I_m - \frac{h}{2} A \right)^{-1} b \\ \bar{y}_{i+1} &= y_i + \frac{h}{2} (Ay_i + b + A\bar{y}_{i+1} + b) \Rightarrow \left( I_m - \frac{h}{2} A \right) \bar{y}_{i+1} = \left( I_m - \frac{h}{2} A \right) y_i + hb \\ &\Rightarrow \bar{y}_{i+1} = y_i + h \left( I_m - \frac{h}{2} A \right)^{-1} b = \tilde{y}_{i+1}\end{aligned}$$

También se puede ver en este ejemplo que los métodos implícitos dan menor error que sus correspondientes versiones explícitas, pero hay que tener en cuenta que son más costosos computacionalmente por la resolución de los sistemas implícitos.

## 1.5. Métodos Runge-Kutta

Todos los métodos mencionados hasta ahora forman parte de una familia de métodos mucho más extensa, *los métodos Runge-Kutta*.

*Definición 1.5.1.* Dada una aproximación  $y_i$  a la solución del PVI (1.1) en  $x_i \in [x_0, x_N]$  un **Runge-Kutta de s etapas** (RK) obtiene una aproximación  $y_{i+1}$  a la solución en  $x_{i+1} = x_i + h$  por el siguiente procedimiento:

$$K_j = f \left( x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l \right), \quad 1 \leq j \leq s \quad (1.11a)$$

$$y_{i+1} = y_i + h \sum_{j=1}^s b_j K_j \quad (1.11b)$$



donde  $b = (b_1, \dots, b_s)^T$  se denomina **vector de pesos**,  $c = (c_1, \dots, c_s)^T$  es el **vector de nodos** y  $A = (a_{jl})_{j,l=1}^s$  es la **matriz de coeficientes**.

Una forma de presentar y manipular los coeficientes de los métodos Runge-Kutta es haciendo uso de las **tablas de Butcher**:

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array}$$

Escrito de manera más compacta:

$$\frac{c \mid A}{b^T} \quad (1.12)$$

y se suele denotar al método **RK(A, b, c)**.

Usualmente, los coeficientes verifican:

$$b^T e = 1, \quad Ae = c$$

donde  $e = (1, \dots, 1)^T \in \mathbb{R}^s$ , para garantizar ciertas propiedades básicas que se verán en la siguiente sección.

*Ejemplo 1.5.1.* Las tablas de Butcher de los métodos citados anteriores son:

Método de Euler explícito (1.4).

$$\frac{0 \mid 0}{\quad \mid 1}$$

Método de Euler implícito (1.7).

$$\frac{1 \mid 1}{\quad \mid 1}$$

Regla explícita del punto medio (1.5).

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}$$

Regla implícita del punto medio (1.8).

$$\frac{\frac{1}{2} \mid \frac{1}{2}}{\quad \mid 1}$$

Regla trapezoidal explícita (1.6).

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Regla trapezoidal implícita (1.9).

$$\begin{array}{c|ccc} 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

También se puede escribir el método (1.11) de forma más compacta como:

$$K = F(K) \tag{1.13a}$$

$$y_{i+1} = y_i + h(b^T \otimes I_m)K \tag{1.13b}$$

denotando

$$K = \begin{pmatrix} K_1 \\ \vdots \\ K_s \end{pmatrix}, \quad F(K) = \begin{pmatrix} f(x_i + c_1 h, y_i + h \sum_{l=1}^s a_{1l} K_l) \\ \vdots \\ f(x_i + c_s h, y_i + h \sum_{l=1}^s a_{sl} K_l) \end{pmatrix} \tag{1.14}$$

siendo  $\otimes$  el producto de Kronecker de matrices. ( $A \otimes B = (a_{jl}B)$ ) e  $I_m$  la matriz identidad de dimensión  $m$ .

Si  $a_{jl} = 0, \forall l \geq j$ , los métodos se llaman **explícitos** porque la solución de avance depende explícitamente de evaluaciones de  $f$  en puntos intermedios.

Los que no verifican esta condición se dicen **implícitos** porque hay que resolver el sistema implícito (1.13a) de las  $s$  etapas  $\{K_j\}_{j=1}^s$  en cada paso de integración. Este sistema implícito tiene solución única si se cumplen las condiciones del siguiente teorema [4, p.206]:

*Teorema 1.5.1.* Sea  $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  continua y que satisfaga la condición de Lipschitz con constante  $L$  (respecto de  $y$ ). Si

$$h < \frac{1}{L \max_j \sum_l |a_{jl}|}$$

Entonces existe una única solución de (1.13a), la cual puede ser obtenida por iteración. Si  $f$  es de clase  $C^p$ , entonces  $K_j$  también son de clase  $C^p$

*Demostración.* Probaremos la existencia mediante el Teorema del Punto Fijo, o, lo que es lo mismo, demostrando que la función  $F(K)$  dada en (1.14) es contractiva con la norma uniforme.

Sean  $K^1 = (K_1^1, \dots, K_s^1)^T, K^2 = (K_1^2, \dots, K_s^2)^T \in \mathbb{R}^{sm}$  y denotamos

$$F_j(K^r) = f \left( x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l^r \right), \quad 1 \leq j \leq s, \quad r = 1, 2.$$

En consecuencia,

$$\begin{aligned} \|F(K^1) - F(K^2)\|_\infty &= \max_{1 \leq j \leq s} \|F_j(K^1) - F_j(K^2)\| = \\ &= \max_{1 \leq j \leq s} \left\| f \left( x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l^1 \right) - f \left( x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l^2 \right) \right\| \end{aligned}$$

Aplicando la condición de Lipstchitz y la desigualdad triangular podemos ver:

$$\begin{aligned} \|F(K^1) - F(K^2)\|_\infty &\leq \max_{1 \leq j \leq s} L \left\| y_i + h \sum_{l=1}^s a_{jl} K_l^1 - \left( y_i + h \sum_{l=1}^s a_{jl} K_l^2 \right) \right\| \leq \\ &\leq hL \max_{1 \leq j \leq s} \left\| \sum_{l=1}^s a_{jl} (K_l^1 - K_l^2) \right\| \leq hL \max_{1 \leq j \leq s} \sum_{l=1}^s |a_{jl}| \|K_l^1 - K_l^2\| \leq \\ &\leq hL \max_{1 \leq j \leq s} \left[ \sum_{l=1}^s |a_{jl}| \left( \max_{1 \leq l \leq s} \|K_l^1 - K_l^2\| \right) \right] = hL \max_{1 \leq j \leq s} \sum_{l=1}^s |a_{jl}| \|K^1 - K^2\|_\infty \end{aligned}$$

En consecuencia

$$\|F(K^1) - F(K^2)\|_\infty \leq hL \max_{1 \leq j \leq s} \left[ \sum_{l=1}^s |a_{jl}| \right] \|K^1 - K^2\|_\infty$$

y por tanto, por hipótesis se tiene que  $F(K)$  es una función contractiva. Haciendo uso del Teorema del Punto Fijo, se puede afirmar que el sistema implícito (1.13a) tiene solución única.

Por otro lado, la diferenciabilidad de  $K$  es consecuencia del Teorema de la Función Implícita. La ecuación de etapas (1.13a) se puede escribir como  $G(h, K) = K - F(K) = 0$ . Cuando  $h = 0$  se tiene que  $\tilde{K} = e \otimes f(x_i, y_i)$ ,  $e = (1, \dots, 1)^T \in \mathbb{R}^s$ , verifica  $G(0, \tilde{K}) = 0$ . Además,

$$\left. \frac{\partial G}{\partial K}(h, K) \right|_{h=0} = I_{sm} - \left. \frac{\partial F}{\partial K}(h, K) \right|_{h=0} = I_{sm}$$

Esta última igualdad es debido a que para todo  $1 \leq r \leq s$ :

$$\left. \frac{\partial F_j}{\partial K_r}(h, K) \right|_{h=0} = \left[ \frac{\partial f}{\partial y} \left( x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l \right) h a_{jr} \right] \Big|_{h=0} = 0$$

por lo que  $\left. \frac{\partial G}{\partial K}(0, K) \right|_{h=0}$  es no singular para  $K = \tilde{K}$  (de hecho, para cualquier  $K \in \mathbb{R}^{sm}$ ).

Por tanto, por el Teorema de la función implícita, existe un entorno de  $h = 0$  ( $|h| < \tilde{h}$  para cierto  $\tilde{h} > 0$ ) en el que  $K = K(h)$  y es diferenciable. Además, si  $f$  es de clase  $C^p$ , todas las componentes  $K_j = K_j(h)$  son de clase  $C^p$ .  $\square$

Si la función derivada  $f$  es lineal, el sistema implícito (1.13a) se reduce a un sistema lineal. Sin embargo, cuando  $f$  no es lineal, se tiene que aplicar algún método de resolución de sistemas no lineales a (1.13a).

Uno de ellos es la **iteración funcional** que aplicada a (1.13a) resulta

$$K^{(\nu)} = F(K^{(\nu-1)}), \quad \nu = 1, 2, \dots$$

arrancando de un valor inicial  $K^{(0)}$ .

Sin embargo, en nuestro trabajo hemos preferido aplicar el **método de Newton modificado**, que es una simplificación menos costosa computacionalmente que el bien conocido método de Newton.

Para ello, se toma  $G(K) = K - F(K)$  y arrancando de un valor inicial  $K^{(0)}$  se computan las aproximaciones:

$$G'(K^{(0)})\Delta K^{(\nu)} = -G(K^{(\nu)}), \quad K^{(\nu+1)} = K^{(\nu)} + \Delta K^{(\nu)}, \quad \nu = 0, 1, 2, \dots \quad (1.15)$$

donde

$$G'(K^{(0)}) = \frac{\partial G}{\partial K}(K^{(0)}) = I_{sm} - \frac{\partial F}{\partial K}(K^{(0)})$$

Este proceso iterativo se para siguiendo alguno de los criterios usuales de parada (por ejemplo,  $\|K^{(\nu+1)} - K^{(\nu)}\| < Tol$ ), o bien, se realizan un número fijo de iteraciones, que en la práctica no suele ser muy alto, dada la buena velocidad de convergencia de este método.

## 1.6. Consistencia de los métodos Runge-Kutta

*Definición 1.6.1.* Dado un problema (1.1) donde  $f$  verifica las condiciones del Teorema 1.2.1 y un método RK  $(A, b, c)$  se denomina **error local** del método a:

$$l(x, h) = y(x+h) - \left( y(x) + h \sum_{j=1}^s b_j \tilde{K}_j \right). \quad (1.16)$$

donde  $\tilde{K}_j = f(x + c_j h, y(x) + h \sum_{l=1}^s a_{jl} \tilde{K}_l)$ ,  $1 \leq j \leq s$ .

*Ejemplo 1.6.1.* El error local del método de Euler explícito usando el desarrollo de Taylor resulta:

$$\begin{aligned} l(x, h) &= y(x+h) - (y(x) + hf(x, y(x))) = y(x+h) - y(x) - hf(x, y(x)) = \\ &= hy'(x) + \mathcal{O}(h^2) - hf(x, y(x)) = \mathcal{O}(h^2), \quad h \rightarrow 0. \end{aligned}$$

En esta sección y la siguiente se presentarán tres propiedades que son de suma importancia: la consistencia, la estabilidad y la convergencia de los métodos Runge-Kutta.

*Definición 1.6.2.* Un método de un paso se dice **consistente de orden  $p \geq 1$**  para PVIIs (1.1) donde  $f$  verifica las condiciones del Teorema 1.2.1 y  $f \in C^p([x_0, x_N] \times Cl(\Omega))$ , si  $\forall y(x)$  existen constantes  $K$  y  $h^*$  tales que

$$||l(x, h)|| \leq K|h|^{p+1} \quad \forall 0 < h \leq h^* \text{ uniformemente en } x \in [x_0, x_N]$$

Esto es que  $l(x, h) = \mathcal{O}(h^{p+1})$ ,  $h \rightarrow 0$  uniformemente en  $x \in [x_0, x_N]$

Usando el desarrollo de Taylor de la función derivada, se demuestra que:

*Lema 1.6.1.* Para que un método RK( $A, b, c$ ) tenga una consistencia de orden  $p$ , con  $1 \leq p \leq 4$  se requiere que  $Ae = c$  y que cumpla las siguientes condiciones:

Orden	Condiciones
$p \geq 1$	$b^T e = 1$
$p \geq 2$	$(p \geq 1) + b^T c = \frac{1}{2}$
$p \geq 3$	$(p \geq 2) + b^T c^2 = \frac{1}{3}, b^T A c = \frac{1}{6}$
$p \geq 4$	$(p \geq 3) + b^T c^3 = \frac{1}{4}, b^T A^2 c = \frac{1}{24}, b^T A c^2 = \frac{1}{12}, b^T (A c \bullet c) = \frac{1}{8}$

donde  $\bullet$  se denota el producto componente a componente, entendiéndose que  $c^k = c \bullet c \bullet \dots \bullet c$ ,  $k \geq 2$ .

*Demostración.* Veamos la demostración para **PVIIs escalares autónomos**:

$$\left. \begin{aligned} y' &= f(y(x)) \\ y(x_0) &= y_0 \end{aligned} \right\} \quad y, f \in \mathbb{R}, \quad x \in [x_0, x_N] \quad (1.17)$$

dada en [2] para métodos Runge-Kutta ( $A, b, c$ ) explícitos de 4 etapas. Usando el desarrollo de Taylor hasta grado 4 en torno a  $y_0$ , cuando  $h \rightarrow 0$  (denotando  $f_0^{(k)} = f^{(k)}(y_0)$ ,  $k \geq 0$ )

$$K_1 = f(y_0) = f_0$$

$$K_2 = f(y_0 + ha_{21}f_0) = f_0 + f'_0 ha_{21}f_0 + \frac{f''_0}{2} h^2 a_{21}^2 f_0^2 + \frac{f'''_0}{3!} h^3 a_{21}^3 f_0^3 + \mathcal{O}(h^4)$$

$$K_3 = f(y_0 + ha_{31}f_0 + ha_{32}K_2) = f_0 + hc_3 f_0 f'_0 + h^2 R_2 + h^3 R_3 + \mathcal{O}(h^4)$$

$$K_4 = f(y_0 + ha_{41}f_0 + ha_{42}K_2 + ha_{43}K_3) = f_0 + hc_4 f_0 f'_0 + h^2 S_2 + h^3 S_3 + \mathcal{O}(h^4)$$

donde:

$$\begin{aligned}
 R_2 &= a_{32}a_{21}f_0'^2 f_0 + \frac{c_2^2}{2} f_0'' f_0^2 \\
 R_3 &= \frac{a_{32}a_{21}^2}{2} f_0' f_0'' f_0^2 + c_3 a_{32}a_{21} f_0'' f_0' f_0^2 + \frac{c_3^3}{3!} f_0''' f_0^3 \\
 S_2 &= (a_{42}a_{21} + a_{43}c_3) f_0 f_0'^2 + \frac{c_2^2}{2} f_0'' f_0^2 \\
 S_3 &= \frac{a_{42}a_{21}^2}{2} f_0' f_0^2 f_0' + a_{43}R_2 f_0' + c_4(a_{42}a_{21} + a_{43}c_3) f_0^2 f_0' f_0'' + \frac{c_3^3}{3!} f_0''' f_0^3
 \end{aligned}$$

Por tanto, la solución numérica dada por el método Runge-Kutta ( $a_{21} = c_2$ ) resulta

$$\begin{aligned}
 y_{RK}(x_0 + h) &= y_0 + hb_1K_1 + hb_2K_2 + hb_3K_3 + hb_4K_4 = \\
 &= y_0 + h[b_1 + b_2 + b_3 + b_4]f_0 + h^2[b_2c_2 + b_3c_3 + b_4c_4]f_0'f_0 + \\
 &+ h^3 \left[ b_2 \frac{c_2^2}{2} f_0'' f_0^2 + b_3 R_2 + b_4 S_2 \right] + h^4 \left[ b_2 \frac{c_2^3}{3!} f_0''' f_0^3 + b_3 R_3 + b_4 S_3 \right] + \\
 &+ \mathcal{O}(h^5)
 \end{aligned}$$

Por otro lado, haciendo uso del desarrollo de Taylor en torno a la solución exacta en el punto  $x_0$ , tenemos que dicha solución exacta verifica:

$$y(x_0 + h) = y_0 + \sum_{l=1}^4 \frac{h^l}{l!} y^{(l)}(x_0) + \mathcal{O}(h^5), \quad h \rightarrow 0.$$

Se pueden expresar las derivadas de la solución exacta evaluadas en  $x_0$  en términos de la función  $f(y)$  y sus derivadas:

$$\begin{aligned}
 y'(x_0) &= f_0 \\
 y''(x_0) &= f_0' f_0 \\
 y'''(x_0) &= f_0'' f_0^2 + (f_0')^2 f_0 \\
 y^{(4)}(x_0) &= f_0''' f_0^3 + 3f_0'' f_0 f_0' f_0 + f_0' f_0'' f_0^2 + (f_0')^3 f_0
 \end{aligned}$$

Si realizamos la diferencia:

$$\begin{aligned}
 y_{RK}(x_0 + h) - y(x_0 + h) &= h(b_1 + b_2 + b_3 + b_4 - 1)f_0 + \\
 &+ h^2 \left( b_2 c_2 + b_3 c_3 + b_4 c_4 - \frac{1}{2} \right) f_0' f_0 + \\
 &+ h^3 \left( \frac{1}{2}(b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2) - \frac{1}{6} \right) f_0'' f_0^2 + \\
 &+ h^3 \left( b_3 a_{32} c_2 + b_4 a_{42} c_2 + b_4 a_{43} c_3 - \frac{1}{6} \right) (f_0')^2 f_0 + \\
 &+ h^4 \left( \frac{1}{6}(b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3) - \frac{1}{24} \right) f_0''' f_0^3 + \\
 &+ h^4 \left( b_3 c_3 a_{32} c_2 + b_4 c_4 a_{42} c_2 + b_4 c_4 a_{43} c_3 - \frac{3}{4!} \right) f_0'' f_0 f_0' f_0 + \\
 &+ h^4 \left( \frac{1}{2}(b_3 a_{32} c_2^2 + b_4 a_{42} c_2^2 + b_4 a_{43} c_3^2) - \frac{1}{24} \right) f_0' f_0'' f_0^2 + \\
 &+ h^4 \left( b_4 a_{43} a_{32} c_2 - \frac{1}{24} \right) (f_0')^3 f_0 + \mathcal{O}(h^5)
 \end{aligned} \tag{1.18}$$

En consecuencia, las condiciones que deben cumplirse para que el método Runge-Kutta sea consistentes de orden  $1 \leq p \leq 4$ , serán:

$p \geq 1$  Para que

$$y_{RK}(x_0 + h) - y(x_0 + h) = \mathcal{O}(h^2)$$

deben de cancelarse los términos lineales de  $h$ :

$$b_1 + b_2 + b_3 + b_4 = 1 \Leftrightarrow b^T e = 1$$

$p \geq 2$  Para obtener

$$y_{RK}(x_0 + h) - y(x_0 + h) = \mathcal{O}(h^3)$$

deben de cancelarse los términos lineales y cuadráticos de  $h$ . En consecuencia, se tiene que cumplir las condiciones del caso anterior y además se debe dar:

$$b_2 c_2 + b_3 c_3 + b_4 c_4 = \frac{1}{2} \Leftrightarrow b^T c = \frac{1}{2}$$

$p \geq 3$  Se obtiene

$$y_{RK}(x_0 + h) - y(x_0 + h) = \mathcal{O}(h^4)$$

si se cancelen todos los términos de orden inferior o igual a 3 de  $h$ . Además de que deben cumplirse las condiciones de los casos anteriores, tiene que cumplirse también:

$$\frac{1}{2}(b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2) = \frac{1}{6} \Leftrightarrow b^T c^2 = \frac{1}{3}$$

$$b_3 a_{32} c_2 + b_4 a_{42} c_2 + b_4 a_{43} c_3 = \frac{1}{6} \Leftrightarrow b^T A c = \frac{1}{6}$$

$p \geq 4$  Para tener

$$y_{RK}(x_0 + h) - y(x_0 + h) = \mathcal{O}(h^5)$$

se deben de cancelar todos los términos de orden inferior a 5 de  $h$ . Además de que deben cumplirse las condiciones de los casos anteriores, tiene que cumplirse también:

$$\frac{1}{6}(b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3) = \frac{1}{24} \Leftrightarrow b^T c^3 = \frac{1}{4}$$

$$b_4 a_{43} a_{32} c_2 = \frac{1}{24} \Leftrightarrow b^T A^2 c = \frac{1}{24}$$

$$b_3 c_3 a_{32} c_2 + b_4 c_4 a_{42} c_2 + b_4 c_4 a_{43} c_3 = \frac{3}{4!} \Leftrightarrow b^T (A c \bullet c) = \frac{1}{8}$$

$$\frac{1}{2}(b_3 a_{32} c_2^2 + b_4 a_{42} c_2^2 + b_4 a_{43} c_3^2) = \frac{1}{24} \Leftrightarrow b^T A c^2 = \frac{1}{12}$$

*Nota.* Para el caso de los PVI's escalares para conseguir consistencia de orden  $p \geq 4$  se pueden unificar los sumandos sexto y séptimo de (1.18). Sin embargo, en el caso de PVI's de dimensión  $m > 1$  estos dos términos no pueden agruparse.

La demostración de este lema 1.6.1 para PVI's de dimensión  $m > 1$  requiere del desarrollo de las series de Butcher [4], lo que supondría una extensión excesiva de esta memoria. En cambio, para problemas no autónomos las condiciones de orden son las mismas debido al siguiente teorema.  $\square$

*Teorema 1.6.1.* Si un  $\text{RK}(A, b, c)$  verifica las condiciones  $Ae = c$  y  $b^T e = 1$ , da exactamente la misma aproximación cuando se aplica al PVI no autónomo (1.1) que cuando se aplica sobre el problema autónomo asociado (1.2).

*Demostración.* Consideremos las ecuaciones de etapas para cada uno de los problemas (1.1) y (1.2) respectivamente.

$$\left. \begin{aligned} K_j &= f(x_0 + c_j h, y_0 + h \sum_{l=1}^s a_{jl} K_l) \quad 1 \leq j \leq s \\ y_1 &= y_0 + h \sum_{j=1}^s b_j K_j \end{aligned} \right\}, \quad (1.19)$$

$$\left. \begin{aligned} \tilde{K}_j &= g\left(z_0 + h \sum_{l=1}^s a_{jl} \tilde{K}_l\right) \quad 1 \leq j \leq s \\ z_1 &= z_0 + h \sum_{j=1}^s b_j \tilde{K}_j \end{aligned} \right\}, \quad (1.20)$$

donde  $K_j = (K_{j1}, \dots, K_{jm}) \in \mathbb{R}^m$  y  $\tilde{K}_j = (\tilde{K}_{j1}, \dots, \tilde{K}_{j(m+1)}) \in \mathbb{R}^{m+1}$ ,  $z_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ . Veamos que las soluciones numéricas de ambos problemas son la misma, esto es, demostrar que  $z_1 = \begin{pmatrix} x_0 + h \\ y_1 \end{pmatrix}$ . Por la definición de  $g(z)$  se tiene que:

$$\tilde{K}_{j1} = g_1\left(z_0 + h \sum_{l=1}^s a_{jl} \tilde{K}_l\right) = 1, \quad 1 \leq j \leq s.$$

donde  $g_1(z)$  es la primera componente del vector  $g(z)$ . La ecuación de avance por componentes se puede escribir como:

$$z_{1k} = z_{0k} + h \sum_{j=1}^s b_j \tilde{K}_{jk}, \quad 1 \leq k \leq m + 1.$$

Para la primera componente de  $z_1$ , aplicando la condición  $b^T e = 1$  se tiene:

$$z_{11} = z_{01} + h \sum_{j=1}^s b_j \tilde{K}_{j1} = x_0 + h \sum_{j=1}^s b_j = x_0 + h.$$



Resta ver que  $z_{1,k+1} = y_{1k}$ ,  $1 \leq k \leq m$ . Para ello, analicemos por componentes el sistema de ecuaciones generado por las ecuaciones de las etapas de (1.20)(que en el peor de los supuestos es implícito).

$$\left( z_0 + h \sum_{l=1}^s a_{jl} \tilde{K}_l \right)_k = z_{0k} + h \sum_{l=1}^s a_{jl} \tilde{K}_{lk}.$$

Para  $k = 1$ , aplicando la condición  $Ae = c$  tenemos que:

$$z_{01} + h \sum_{l=1}^s a_{jl} \tilde{K}_{l1} = x_0 + h \sum_{l=1}^s a_{jl} = x_0 + c_j h.$$

Tenemos por tanto que la ecuación de etapas es:

$$\tilde{K}_{j,k+1} = f_k \left( x_0 + c_j h, y_{01} + h \sum_{l=1}^s a_{jl} \tilde{K}_{l2}, \dots, y_{0m} + h \sum_{l=1}^s a_{jl} \tilde{K}_{l,m+1} \right), \quad \begin{matrix} 1 \leq j \leq s \\ 1 \leq k \leq m \end{matrix}$$

Este sistema implícito es el mismo sistema que (1.19). Por lo tanto, como dicho sistema tiene solución única,

$$\tilde{K}_{j,k+1} = K_{jk}, \quad 1 \leq j \leq s, \quad 1 \leq k \leq m$$

En consecuencia, tenemos que:

$$z_{1,k+1} = y_{0k} + h \sum_{j=1}^s b_j K_{j,k} = y_{1k}, \quad 1 \leq k \leq m$$

que es lo que faltaba por demostrar. □

*Ejemplo 1.6.2.* A continuación se presenta una tabla en la que se recogen los órdenes de consistencia que tienen los métodos numéricos citados en las secciones anteriores aplicando el lema 1.6.1:

<i>Método numérico</i>	<i>Orden de consistencia</i>
<i>Método de Euler explícito (1.4)</i>	1
<i>Método de Euler implícito (1.7)</i>	1
<i>Regla explícita del punto medio (1.5)</i>	2
<i>Regla implícita del punto medio (1.8)</i>	2
<i>Regla trapezoidal explícita (1.6)</i>	2
<i>Regla trapezoidal implícita (1.9)</i>	2

## 1.7. Estabilidad y convergencia de los métodos Runge-Kutta

Para garantizar la convergencia de los métodos no sólo basta con la consistencia. Se necesitará también que los métodos propaguen bien los errores locales a lo largo de toda la integración. Esto lo garantiza la 0-estabilidad.

*Teorema 1.7.1.* Los métodos  $\text{RK}(A, b, c)$  (1.11) son **0-estables**, es decir, para toda partición  $\mathcal{P}$  de  $[x_0, x_N]$  con  $|\mathcal{P}| = \max_{0 \leq i \leq n-1} |h_i| \leq h \leq \bar{h}$ , las aproximaciones (1.11) del método y las secuencias definidas por:

$$\begin{cases} \bar{y}_{i+1} = \bar{y}_i + h \sum_{j=1}^s b_j \bar{K}_j + \eta_{i+1} \\ \bar{y}_0 = y_0 + \eta_0 \end{cases}, \quad 0 \leq i \leq n-1 \quad (1.21)$$

donde  $\bar{K}_j = f(x_i + c_j h, \bar{y}_i + h \sum_{l=1}^s a_{jl} \bar{K}_l)$  verifican

$$\|\bar{y}_m - y_m\| \leq C \sum_{i=0}^m \|\eta_i\|, \quad 0 \leq m \leq n$$

siendo  $C > 0$  una constante y  $\{\eta_i\}_{i=0}^n$  perturbaciones cualesquiera.

*Demostración.* Considerando la diferencia entre (1.11b) y (1.21) aplicando la desigualdad triangular se tiene

$$\|y_{i+1} - \bar{y}_{i+1}\| \leq \|y_i - \bar{y}_i\| + h \sum_{j=1}^s |b_j| \|K_j - \bar{K}_j\| + \|\eta_{i+1}\|. \quad (1.22)$$

Desarrollando el término  $\|K_j - \bar{K}_j\|$ , aplicando la condición de Lipschitz de  $f$  y  $h < \bar{h}$ :

$$\begin{aligned} \|K_j - \bar{K}_j\| &= \left\| f\left(x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l\right) - f\left(x_i + c_j h, \bar{y}_i + h \sum_{l=1}^s a_{jl} \bar{K}_l\right) \right\| \\ &\leq L \left[ \|y_i - \bar{y}_i\| + h \sum_{l=1}^s |a_{jl}| \|K_l - \bar{K}_l\| \right] \\ &\leq L \|y_i - \bar{y}_i\| + \bar{h} L \sum_{l=1}^s |a_{jl}| \|K_l - \bar{K}_l\| \\ &\leq L \|y_i - \bar{y}_i\| + \bar{h} L \|A\|_\infty \|K - \bar{K}\|_\infty \end{aligned}$$

donde  $\|A\|_\infty = \max_{1 \leq j \leq s} \sum_{l=1}^s |a_{jl}|$  es la norma uniforme matricial y

$$\|K - \bar{K}\|_\infty = \max_{1 \leq j \leq s} \|K_j - \bar{K}_j\|.$$

Sea  $1 \leq j \leq s$  el índice tal que  $\|K_j - \bar{K}_j\| = \|K - \bar{K}\|_\infty$ . Podemos obtener una cota superior para la norma:

$$\begin{aligned} \|K - \bar{K}\|_\infty &\leq L\|y_i - \bar{y}_i\| + \bar{h}L\|A\|_\infty\|K - \bar{K}\|_\infty \\ (1 - \bar{h}L\|A\|_\infty)\|K - \bar{K}\|_\infty &\leq L\|y_i - \bar{y}_i\|. \end{aligned}$$

Por el Teorema 1.5.1, tenemos que el primer miembro es positivo estrictamente y, en consecuencia,

$$\|K - \bar{K}\|_\infty \leq \frac{L}{1 - \bar{h}L\|A\|_\infty}\|y_i - \bar{y}_i\|$$

Usando esta desigualdad en (1.22) concluimos que:

$$\begin{aligned} \|y_{i+1} - \bar{y}_{i+1}\| &\leq \|y_i - \bar{y}_i\| + \bar{h} \left( \sum_{j=1}^s |b_j| \right) \frac{L}{1 - \bar{h}L\|A\|_\infty} \|y_i - \bar{y}_i\| + \|\eta_{i+1}\| \\ \|y_{i+1} - \bar{y}_{i+1}\| &\leq \left( 1 + \bar{h} \left( \sum_{j=1}^s |b_j| \right) \frac{L}{1 - \bar{h}L\|A\|_\infty} \right) \|y_i - \bar{y}_i\| + \|\eta_{i+1}\| = \tilde{C}\|y_i - \bar{y}_i\| + \|\eta_{i+1}\| \end{aligned}$$

donde  $\tilde{C} > 0$ . Aplicando iterativamente esta última desigualdad

$$\begin{aligned} \|y_m - \bar{y}_m\| &\leq \tilde{C}\|y_{m-1} - \bar{y}_{m-1}\| + \|\eta_m\| \\ &\leq \tilde{C}^m\|y_0 - \bar{y}_0\| + \sum_{k=1}^m \tilde{C}^{m-k}\|\eta_k\| \leq \sum_{k=0}^m \tilde{C}^{m-k}\|\eta_k\| \end{aligned}$$

Si tomamos  $C = \max_{0 \leq k \leq m} \{\tilde{C}^k\}$  queda lo que se quería demostrar:

$$\|y_m - \bar{y}_m\| \leq C \sum_{k=0}^m \|\eta_k\|$$

□

En otras palabras, una ligera modificación de las condiciones iniciales del problema implica una leve variación de la solución perturbada frente a la no perturbada, así como las perturbaciones debidas a errores de cómputo.

*Definición 1.7.1.* Un método de un paso se dice **convergente de orden  $p \geq 1$**  para PVI's donde  $f$  verifica las condiciones del Teorema 1.2.1, y  $f \in C^p$ , si existen constantes  $K$  y  $\delta$  tales que para toda partición  $\mathcal{P} = \{x_i\}_{i=0}^n$  con  $|\mathcal{P}| = \max_{0 \leq i \leq n-1} |h_i| = h \leq \delta$  se verifica que:

$$\|y(x_i) - y_i\| \leq Kh^p \quad 0 \leq i \leq n$$

donde  $y(x)$  es la solución exacta. Al valor  $\max_{x_i \in \mathcal{P}} \|y(x_i) - y_i\|$  se llama **error global del método**.

*Teorema 1.7.2.* Si un RK( $A, b, c$ ) tiene orden de consistencia  $p \geq 1 \Rightarrow$  es convergente de orden  $p$

*Demostración.* Supongamos que estamos en una partición equispaciada  $\mathcal{P}$ , con  $h = (x_N - x_0)/n \leq \bar{h}$ , donde  $\bar{h}$  es la cota dada en el Teorema 1.7.1. Sea  $y(x)$  la solución exacta. De (1.16) se deduce:

$$y(x_{i+1}) = y(x_i) + h \sum_{j=1}^s b_j \tilde{K}_j + l(x_i, h)$$

Si aplicamos el teorema 1.6.1 para  $\bar{y}_i = y(x_i)$ ,  $\eta_{i+1} = l(x_i, h)$ ,  $\eta_0 = 0$  y denotamos  $e_i = y_i - y(x_i)$  se obtiene

$$\|e_m\| \leq C \sum_{i=0}^m \|\eta_i\|, \quad 0 \leq m \leq n$$

Como el método tiene consistencia de orden  $p$ , tenemos:

$$\|e_m\| \leq C \sum_{i=1}^m K h^{p+1} \leq \tilde{C} \sum_{i=1}^n h^{p+1} = \tilde{C} h^{p+1} n = \tilde{C} h^{p+1} \frac{x_N - x_0}{h} = \hat{C} h^p$$

Por tanto

$$\max_{0 \leq i \leq n} \|y_i - y(x_i)\| \leq \hat{C} h^p$$

y consecuentemente la consistencia de orden  $p \geq 1$  de un RK ( $A, b, c$ ) implica su convergencia de orden  $p$ . Si la partición no fuese equispaciada, se aplica de forma similar tomando  $|\mathcal{P}| = \max_{0 \leq i \leq n-1} |h_i| \leq h \leq \bar{h}$ .  $\square$

*Ejemplo 1.7.1.* A continuación se considera el siguiente problema de valor inicial no autónomo lineal:

$$y' = \begin{pmatrix} 0 & 4 & 0 & -1 \\ 2 & 2 & 0 & 0 \\ 2 & 4 & -1 & 8 \\ 1 & 2 & 0 & 0 \end{pmatrix} y + \begin{pmatrix} -4 + (t-1)e^{-t} \\ -2e^{-t} \\ -2 - 3t^2e^{-t} \\ -1 - te^{-t} \end{pmatrix} \quad y(0) = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad t \in [0, 1]$$

Se puede demostrar que la solución exacta es:

$$y(t) = \begin{pmatrix} -1 + e^{-t} + 2e^{-2t} \\ 1 - e^{-2t} \\ (2t + 4t^2 - t^3)e^{-t} \\ te^{-t} \end{pmatrix}$$

Como ejemplo, podemos ver en las figuras 1.3 y 1.4 el máximo error en norma uniforme obtenido para los métodos introducidos en la sección 1.4 a paso fijo

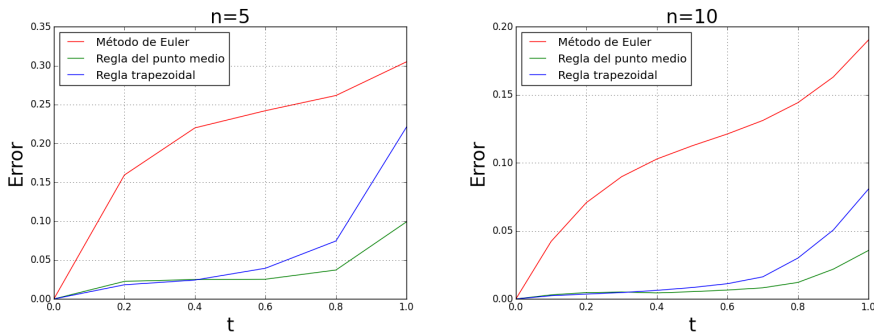


Figura 1.3. Error en los métodos explícitos del ejemplo 1.7.1.

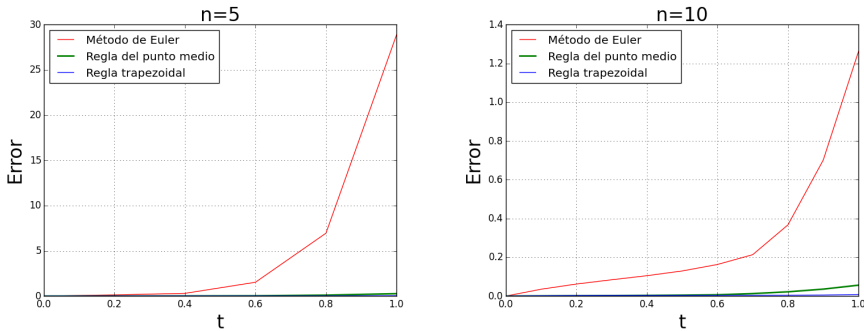


Figura 1.4. Error en los métodos implícitos del ejemplo 1.7.1.

$h = 1/n$  con  $n = 5, 10$ . En ellas podemos observar que los métodos de orden 2 (regla del punto medio y regla trapezoidal) dan menor error que los métodos de orden 1, tanto en el caso de los explícitos como en los implícitos.

En la tabla 1.1 estudiamos con más detalle los errores obtenidos cuando disminuimos el tamaño de paso fijo  $h = 1/n$ . En ella podemos observar que, tanto en los métodos explícitos como los implícitos, los errores de los métodos de orden 1 se dividen aproximadamente a la mitad cuando el paso fijo se divide a la mitad. Sin embargo, en los métodos de orden 2 los errores se dividen aproximadamente por 4. Esto refleja que los métodos alcanzan el orden predicho por la teoría cuando  $h \rightarrow 0$ . Si un método tiene orden  $p \geq 1$  y  $e(h)$  es el máximo de los errores globales obtenidos en toda la integración con paso fijo  $h$ ,  $e(h) \simeq Ch^p$ ,  $h \rightarrow 0 \Rightarrow e(h/2) \simeq Ch^p/2^p \simeq e(h)/2^p$ .

$n$	Método Exp.	Error máximo	$n$	Método Imp.	Error máximo
5	Euler	$3.049E - 1$	5	Euler	$2.886E + 1$
	Punto medio	$9.919E - 2$		Punto medio	$2.652E - 1$
	Trapezoidal	$2.2105E - 1$		Trapezoidal	$3.948E - 2$
10	Euler	$1.903E - 1$	10	Euler	$1.261E + 0$
	Punto medio	$3.565E - 2$		Punto medio	$5.620E - 2$
	Trapezoidal	$8.077E - 2$		Trapezoidal	$7.881E - 3$
20	Euler	$1.124E - 1$	20	Euler	$2.886E - 1$
	Punto medio	$1.030E - 2$		Punto medio	$1.352E - 2$
	Trapezoidal	$2.361E - 2$		Trapezoidal	$1.865E - 3$
40	Euler	$6.243E - 2$	40	Euler	$9.890E - 2$
	Punto medio	$2.737E - 3$		Punto medio	$3.348E - 3$
	Trapezoidal	$6.309E - 3$		Trapezoidal	$4.600E - 4$
80	Euler	$3.311E - 2$	80	Euler	$4.074E - 2$
	Punto medio	$7.027E - 4$		Punto medio	$8.349E - 4$
	Trapezoidal	$1.624E - 3$		Trapezoidal	$1.146E - 4$
160	Euler	$1.709E - 2$	160	Euler	$1.844E - 2$
	Punto medio	$1.779E - 4$		Punto medio	$2.086E - 4$
	Trapezoidal	$4.117E - 4$		Trapezoidal	$2.863E - 5$
320	Euler	$8.682E - 3$	320	Euler	$8.978E - 3$
	Punto medio	$4.473E - 5$		Punto medio	$5.214E - 5$
	Trapezoidal	$1.036E - 4$		Trapezoidal	$7.155E - 6$

**Tabla 1.1.** Métodos explícitos y métodos implícitos.

## Problemas stiff y estabilidad lineal de los métodos

### 2.1. El fenómeno de la stiffness

En esta sección se presentará la noción de stiffness. Este fenómeno aparece en multitud de ejemplos en las aplicaciones. En particular, suelen aparecer muchos problemas stiff cuando se modelizan reacciones químicas.

Consideremos, por ejemplo, reacciones químicas del tipo  $A + B \xrightarrow{k} C + D$ , donde  $k$  es la ratio de la reacción. La **ley de acción de masas de la química cinética** establece que la velocidad de la reacción es proporcional al producto de las concentraciones de los reactivos.

Si  $a(t)$ ,  $b(t)$ ,  $c(t)$  y  $d(t)$  son las concentraciones en un tiempo  $t$  de las sustancias  $A$ ,  $B$ ,  $C$  y  $D$  respectivamente, tendremos el siguiente sistema de EDOs que modeliza la evolución a lo largo del tiempo de dichas concentraciones:

$$\left. \begin{aligned} a'(t) &= -ka(t)b(t) \\ b'(t) &= -ka(t)b(t) \\ c'(t) &= ka(t)b(t) \\ d'(t) &= ka(t)b(t) \end{aligned} \right\}$$

Adicionalmente, para que el modelo tenga validez se debe verificar que  $a(t)$ ,  $b(t)$ ,  $c(t)$ ,  $d(t) \geq 0$ .

*Ejemplo 2.1.1.* [6, p.26] Consideramos una reacción química cíclica  $W_1 \xrightarrow{k_1} W_2 \xrightarrow{k_2} W_1$ . La variación de las concentraciones  $w_1(t)$  y  $w_2(t)$  de las sustancias  $W_1$  y  $W_2$  respectivamente pueden ser modelizadas por el siguiente sistema:

$$\left. \begin{aligned} w_1'(t) &= -k_1w_1(t) + k_2w_2(t) \\ w_2'(t) &= k_1w_1(t) - k_2w_2(t) \end{aligned} \right\}. \quad (2.1)$$

Como es un sistema de EDOs lineales de coeficientes constantes se puede calcular la solución exacta. Si  $w_1(0)$ ,  $w_2(0)$  son las concentraciones iniciales de las sustancias  $W_1$  y  $W_2$  se tiene que la solución exacta de (2.1) es:

$$w_1(t) = \frac{k_2}{k_1 + k_2} (w_1(0) + w_2(0)) + \frac{e^{-(k_1+k_2)t}}{k_1 + k_2} (k_1 w_1(0) - k_2 w_2(0))$$

$$w_2(t) = \frac{k_1}{k_1 + k_2} (w_1(0) + w_2(0)) - \frac{e^{-(k_1+k_2)t}}{k_1 + k_2} (k_1 w_1(0) - k_2 w_2(0))$$

Tomamos los valores:  $w_1(0) = 0.1, w_2(0) = 0.9, T = 1, h = 1/50, k_1 = 1$  y hacemos experimentos para valores diferentes de  $k_2 = 10, 100, 1000$ .

Como es conocida la solución exacta del problema es posible graficar los errores asociados a los métodos explícitos e implícitos (1.4), (1.6), (1.7), (1.9) en las figuras 2.1, 2.2 y 2.3.

Obsérvese que los métodos (1.5) y (1.8) dan los mismos resultados que los (1.6)

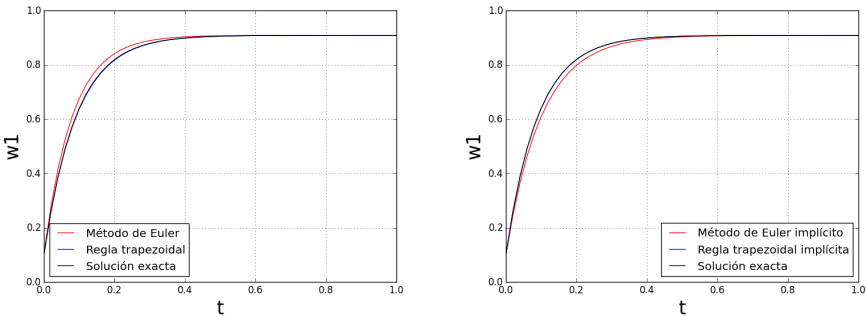


Figura 2.1.  $w_1(t)$  del ejemplo 2.1.1 para  $k_2 = 10$

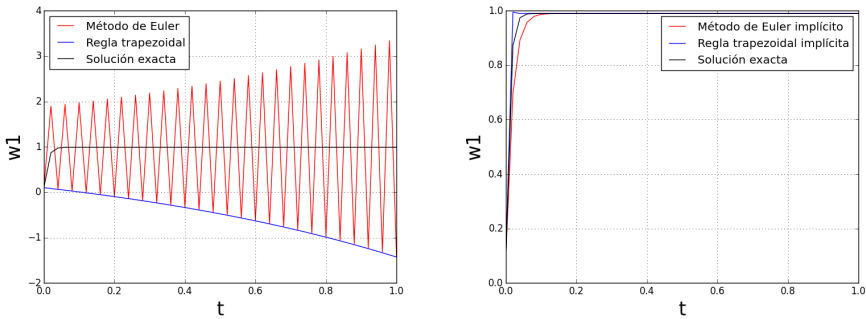


Figura 2.2.  $w_1(t)$  del ejemplo 2.1.1 para  $k_2 = 100$

y (1.9) respectivamente por el lema 1.4.1.



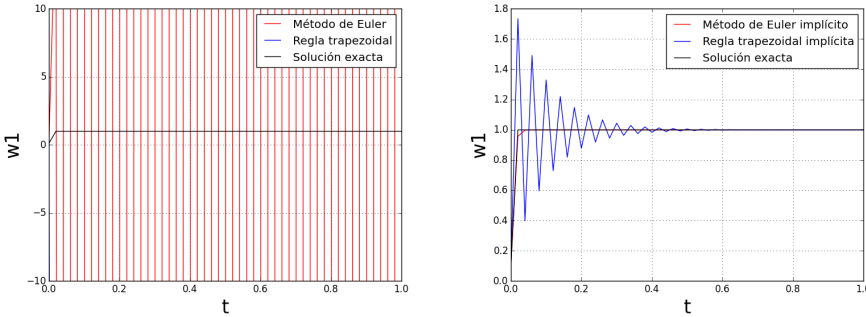


Figura 2.3.  $w_1(t)$  del ejemplo 2.1.1 para  $k_2 = 1000$

Podemos concluir que en este problema sencillo los métodos explícitos no presentan una convergencia práctica. Dicho de otra forma, para que los métodos explícitos den una buena aproximación a la solución exacta se tendrá que reducir drásticamente el tamaño de paso. Sin embargo, los métodos implícitos sí proporcionan una aproximación buena con un tamaño de paso aceptable.

Y esto ocurre incluso con métodos explícitos con un mayor orden de convergencia en teoría. Esta diferencia de comportamiento entre los métodos explícitos y los implícitos sobre este problema es lo que hace que se incluya dentro de la clase de **problemas stiff**.

“Las ecuaciones stiff son problemas sobre los que los métodos explícitos no funcionan” [5]. Esta es la “definición” más aproximada de lo que son los problemas de esta clase. No existe una definición rigurosa de este fenómeno de la stiffness. El estudio de los métodos sobre problemas lineales de coeficientes constantes nos da una visión más clara de lo que hace stiff a un problema de este tipo, siendo a la vez una ayuda fundamental para diseñar métodos implícitos que integran eficientemente estos problemas.

## 2.2. Estabilidad lineal

Sea el PVI lineal escalar:

$$y' = \lambda y, \quad y(0) = y_0 \quad \lambda, y \in \mathbb{C} \tag{2.2}$$

con solución exacta  $y(x) = y_0 e^{\lambda x}$ ,  $\forall x$ . Tenemos que la solución de este PVI es estable si y sólo si  $Re\lambda \leq 0$ .

Si aplicamos un RK( $A, b, c$ ) de  $s$  etapas (1.11) sobre este PVI (2.2) obtenemos de (1.11a) que

$$K_j = \lambda \left[ y_i + h \sum_{l=1}^s a_{jl} K_l \right], \quad 1 \leq j \leq s, \quad K_j \in \mathbb{R}$$

lo que escrito matricialmente resulta

$$K = \lambda e y_i + h \lambda A K \Rightarrow (I_s - h \lambda A) K = \lambda e y_i \Rightarrow K = \lambda (I_s - h \lambda A)^{-1} e y_i$$

Llevándolo a (1.11b) se obtiene:

$$y_{i+1} = y_i + h b^T K = y_i + h \lambda b^T (I_s - h \lambda A)^{-1} e y_i = [1 + h \lambda b^T (I_s - h \lambda A)^{-1} e] y_i$$

Así, de forma natural se define:

**Definición 2.2.1.** Sea un método Runge-Kutta  $(A, b, c)$  (1.11) de  $s$  etapas, aplicado a (2.2) con paso  $h > 0$ . Entonces, la solución de avance del método es  $y_i = R(h\lambda)^i y_0$ , donde  $R(z)$  es la función:

$$R(z) = 1 + z b^T (I_s - z A)^{-1} e,$$

y se la denomina **función de estabilidad lineal**.

Sería deseable que cuando el PVI (2.2) sea estable, o sea, cuando  $Re \lambda \leq 0$ , la solución numérica verifique una propiedad similar a la de la solución exacta, por lo que será necesario que  $|R(z)| \leq 1$  (en caso contrario,  $y_i \rightarrow \infty$  si  $i$  aumenta). Por ello, se puede definir a partir de la función  $R(z)$  una región  $\mathcal{S}$  del plano complejo en el que el método es estable.

$$\mathcal{S} = \{z \in \mathbb{C} / |R(z)| \leq 1\}.$$

A este conjunto se le denomina **dominio de estabilidad absoluta lineal**.

Un método RK  $(A, b, c)$  es **A-estable** si  $\mathbb{C}^- = \{z \in \mathbb{C} / Re z \leq 0\} \subseteq \mathcal{S}$ . Es decir, un método es A-estable si lo es sobre problemas lineales de coeficientes constantes escalares.

**Ejemplo 2.2.1.** A continuación, presentaremos las funciones de estabilidad lineal de los métodos citados en las anteriores secciones:

$r$	Método numérico	Función de estabilidad lineal
1	Método de Euler explícito (1.4)	$R(z) = 1 + z$
2	Método de Euler implícito (1.7)	$R(z) = 1 + \frac{z}{1-z} = \frac{1}{1-z}$
3	Regla explícita del punto medio (1.5)	$R(z) = 1 + z + \frac{z^2}{2}$
4	Regla implícita del punto medio (1.8)	$R(z) = 1 + \frac{z}{1-\frac{z}{2}} = \frac{1+\frac{z}{2}}{1-\frac{z}{2}}$
5	Regla trapezoidal explícita (1.6)	$R(z) = 1 + z + \frac{z^2}{2}$
6	Regla trapezoidal implícita (1.9)	$R(z) = 1 + \frac{z}{1-\frac{z}{2}} = \frac{1+\frac{z}{2}}{1-\frac{z}{2}}$

(2.3)

*Nota.* La regla del punto medio y la regla trapezoidal tienen la misma función de estabilidad lineal porque presentan las mismas aproximaciones para problemas lineales de coeficientes constantes como vimos en el lema 1.4.1 y, por lo tanto, también tendrán el mismo dominio de estabilidad lineal.

Tenemos por tanto, que los dominios de estabilidad lineal de los métodos son:

$$\begin{aligned} \mathcal{S}_1 &= \{z \in \mathbb{C} / |1 + z| \leq 1\} \\ \mathcal{S}_2 &= \{z \in \mathbb{C} / |1 - z| \geq 1\} \supseteq \mathbb{C}^- \\ \mathcal{S}_3 = \mathcal{S}_5 &= \{z \in \mathbb{C} / \left|1 + z + \frac{z^2}{2}\right| \leq 1\} \\ \mathcal{S}_4 = \mathcal{S}_6 &= \{z \in \mathbb{C} / \left|\frac{1 + \frac{z}{2}}{1 - \frac{z}{2}}\right| \leq 1\} = \{z \in \mathbb{C} / \operatorname{Re} z \leq 0\} = \mathbb{C}^- \end{aligned}$$

donde  $\mathcal{S}_r$  es el dominio de estabilidad lineal del método  $r$  (Ver (2.3)).

En consecuencia, los tres métodos implícitos son A-estables. En cambio, ningún explícito lo es pues  $\mathcal{S}$  es una región acotada en  $\mathbb{C}^-$ .

Seguidamente, se reconsidera los resultados del ejemplo 2.1.1 desde el punto de vista de la estabilidad lineal. No es un PVI escalar, por lo que tendremos que transformarlo para aplicar lo anterior.

Como el problema es lineal de coeficientes constantes  $w' = \mathcal{A}w$  podemos estudiar los autovalores de la matriz

$$\mathcal{A} = \begin{pmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{pmatrix}$$

El espectro de esta matriz es:  $\sigma(\mathcal{A}) = \{0, -(k_1 + k_2)\}$ , por lo que es posible diagonalizar esta matriz y transformar este problema en otro lineal cuyas variables estén desacopladas.

Es decir, existe una matriz no singular  $P$  tal que

$$A = PDP^{-1}, \quad D = \begin{pmatrix} 0 & 0 \\ 0 & -(k_1 + k_2) \end{pmatrix}$$

Por tanto,  $w' = PDP^{-1}w \Rightarrow$  si denotamos  $u = P^{-1}w$  se tiene el sistema equivalente  $u' = \mathcal{D}u$  o, lo que es lo mismo,

$$\left. \begin{aligned} u'_1(t) &= 0 \\ u'_2(t) &= -(k_1 + k_2)u_2(t) \end{aligned} \right\}$$

que son dos PVI de tipo (2.2). En consecuencia, un método será estable sobre este problema si tanto  $z = 0$  como  $z = -h(k_1 + k_2)$  pertenecen a su dominio de

estabilidad.

Veamos qué ocurre con los métodos testados sobre este ejemplo.

Método de Euler explícito (1.4).

$$|1+z| \leq 1 \Rightarrow -1 \leq 1-h(k_1+k_2) \leq 1 \Rightarrow -2 \leq -h(k_1+k_2) \leq 0 \Rightarrow 0 \leq h \leq \frac{2}{k_1+k_2}$$

Regla trapezoidal explícita (1.6).

$$\left| 1 + z + \frac{z^2}{2} \right| \leq 1 \Rightarrow 0 \leq h - \frac{(k_1+k_2)h^2}{2} \leq \frac{2}{k_1+k_2}$$

O lo que es lo mismo, se debe verificar simultáneamente:

$$0 \leq h - \frac{(k_1+k_2)h^2}{2} \Rightarrow h \in \left[ 0, \frac{2}{k_1+k_2} \right]$$

y

$$h - \frac{(k_1+k_2)h^2}{2} \leq \frac{2}{k_1+k_2} \Rightarrow h \in (-\infty, \infty)$$

Por lo tanto, el método es estable si:

$$h \in \left[ 0, \frac{2}{k_1+k_2} \right]$$

Método de Euler implícito (1.7).

$$|1-z| \geq 1 \Rightarrow 1+h(k_1+k_2) \geq 1 \Rightarrow h \geq 0$$

Regla trapezoidal implícita (1.9).

$$\operatorname{Re} z \leq 0 \Rightarrow -h(k_1+k_2) \leq 0 \Rightarrow h \geq 0$$

Por lo tanto, tenemos para los distintos valores de  $k_2$ , los implícitos no tienen problemas de estabilidad. En cambio para que los explícitos vayan bien será necesario que

$$\begin{cases} k_2 = 10 & 0 \leq h \leq \frac{2}{11} \\ k_2 = 100 & 0 \leq h \leq \frac{2}{101} \\ k_2 = 1000 & 0 \leq h \leq \frac{2}{1001} \end{cases}$$

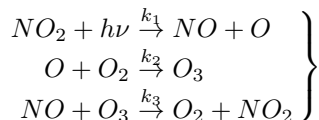
lo que coincide con lo observado en las figuras 2.1, 2.2 y 2.3.

Por otra parte, en este ejemplo vemos también que la presencia del autovalor negativo  $-(k_1+k_2)$  y su módulo determina la stiffness de dicho problema. Esta es una de las características más habituales que provocan este fenómeno: la presencia de autovalores de la matriz jacobiana del PVI en el semiplano complejo negativo con módulos muy grandes.

### 2.3. Un modelo de polución atmosférica

En el estudio de la atmósfera es muy importante conocer la evolución de las reacciones químicas que ocurren entre los distintos gases que la componen. [6, p. 7]

*Ejemplo 2.3.1.* Vamos a reproducir el comportamiento de algunas sustancias de la atmósfera ( $O$ ,  $NO$ ,  $NO_2$  y  $O_3$ ) cuyas concentraciones vienen determinadas por las siguientes reacciones químicas:



Cabe destacar que la primera de las reacciones es una fotodisociación, es decir, para que esta reacción tenga lugar debe haber presencia de radiación electromagnética  $h\nu$ . Si queremos predecir la concentración de estas sustancias en un tiempo concreto es preciso establecer de una manera aproximada el ratio  $k_1$  para un tiempo  $t$  (en segundos) que será diferente si es de día a si es de noche. Una forma de simular esto es tomar como ratio  $k_1(t)$  a la función definida a trozos:

$$k_1(t) = \begin{cases} 10^{-5} e^{7c(t)}, & \hat{t}_h \in [4, 20) \\ 10^{-40}, & \hat{t}_h \in [0, 4) \cup [20, 24) \end{cases}$$

donde

$$c(t) = \left( \sin \left( \frac{\pi}{16} (\hat{t}_h - 4) \right) \right)^{0.2}, \quad \hat{t}_h = \frac{t}{3600} - 24 \left\lfloor \frac{t}{3600 \cdot 24} \right\rfloor$$

y  $[x]$  denota a la parte entera de  $x$ . Además, tomamos  $k_2 = 10^5$  y  $k_3 = 10^{-16}$ . Llamaremos  $u_1(t)$ ,  $u_2(t)$ ,  $u_3(t)$  y  $u_4(t)$  a las concentraciones de  $O$ ,  $NO$ ,  $NO_2$  y  $O_3$  para un cierto tiempo  $t$ , respectivamente.

Se considerarán como supuestos adicionales que la concentración de  $O_2$  es constante y que hay una fuente constante que emite  $NO$  a la atmósfera a un ritmo de  $\sigma_2 = 10^6$ .

En definitiva, tenemos el sistema de EDOs:

$$\left. \begin{array}{l} u_1'(t) = k_1(t)u_3 - k_2u_1 \\ u_2'(t) = k_1(t)u_3 - k_3u_2u_4 + \sigma_2 \\ u_3'(t) = k_3u_2u_4 - k_1(t)u_3 \\ u_4'(t) = k_2u_1 - k_3u_2u_4 \end{array} \right\}. \quad (2.4)$$

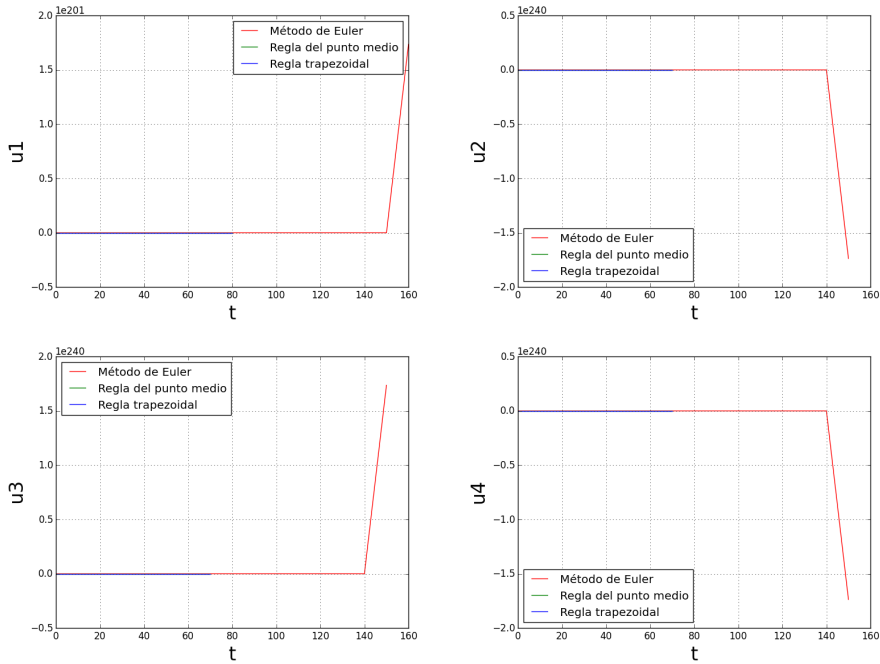
Los valores iniciales de las concentraciones serán:

$$u(0) = (u_1(0), u_2(0), u_3(0), u_4(0))^T = (0, 1.3 \cdot 10^8, 5.0 \cdot 10^{11}, 8.0 \cdot 10^{11})^T.$$

teniendo en cuenta que dichas concentraciones se miden en moléculas por  $cm^3$ .

Debido a que el problema (2.4) es no lineal resulta complicado (o no es posible) hallar la solución de forma analítica. En consecuencia, no se puede realizar una gráfica con gran exactitud del error de los métodos. Sin embargo, podemos verificar si la aproximación dada por los métodos es buena si comparamos con otros métodos que posean un mayor control del error.

En la figura 2.4 vemos las soluciones numéricas obtenidas por los tres métodos explícitos (1.4), (1.5) y (1.6) con paso fijo  $h = 10$ .

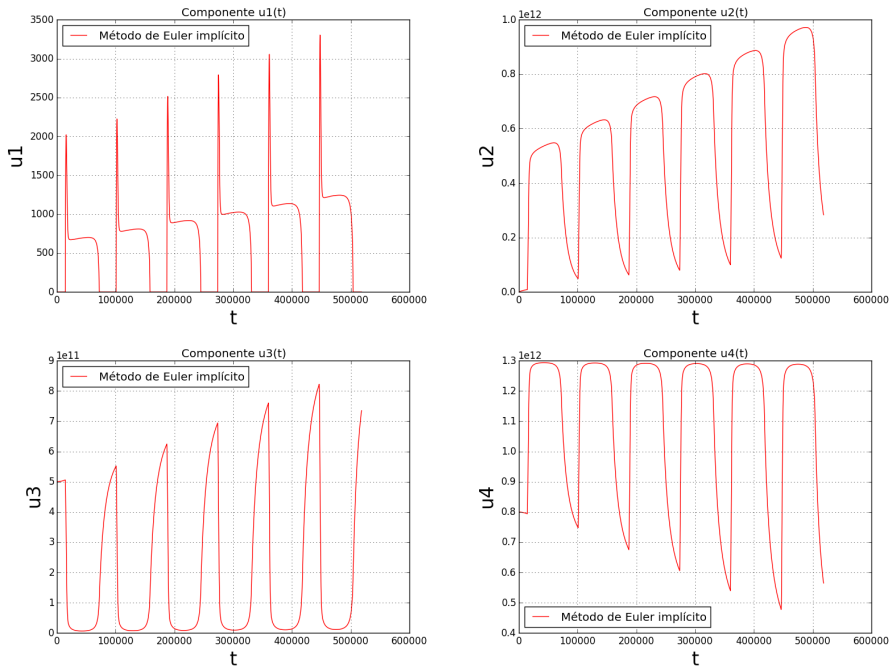


**Figura 2.4.** Soluciones numéricas con métodos explícitos para (2.4) con  $h = 10$

Tenemos que los métodos explícitos no presentan una buena aproximación de la solución real. Se puede observar que las componentes toman valores del orden de  $10^{200}$  en un tiempo de 160 segundos. Estas aproximaciones están bastante alejadas del comportamiento real de estos gases en la atmósfera.

En cambio, si se aplica el método de Euler implícito (1.7) se consiguen las siguientes gráficas dadas en la figura 2.5, de las concentraciones de las cuatro sustancias consideradas.

En estas gráficas podemos observar una evolución más realista de las concentra-



**Figura 2.5.** Soluciones numéricas con el método de Euler implícito para (2.4) con  $h = 10$ .

ciones, viéndose claramente las variaciones debidas al ciclo diario y observando el aumento gradual de las concentraciones de los óxidos de nitrógeno altamente contaminantes.

Hay que resaltar el hecho de que al ser este problema (2.4) no lineal, en este caso hay que resolver un sistema implícito por paso de integración. Para ello se ha usado el método de Newton modificado visto en (1.15), lo que incrementa su coste computacional.

Por otro lado, desde del punto de vista numérico este problema (2.4) es claramente stiff, lo que suele ser habitual en muchos problemas de polución atmosférica, por lo que el uso de métodos implícitos será fundamental en este campo.





## Métodos de líneas sobre EDPs parabólicas. Problemas de polución atmosférica con difusión.

### 3.1. Introducción

Otro fenómeno que aparece en el estudio de la evolución de la polución atmosférica es la difusión de las concentraciones a lo largo del tiempo. Para modelizarlo tendremos que utilizar ecuaciones en derivadas parciales (EDP).

El campo de los métodos numéricos para resolver ecuaciones en derivadas parciales es demasiado amplio para verlo en un solo TFG. Por ello, en este trabajo solo veremos el método de líneas sobre problemas de ecuaciones en derivadas parciales parabólicas sencillas.

En la siguiente sección introduciremos dicho método sobre el caso de la ecuación del calor unidimensional y posteriormente lo aplicaremos a un problema de polución atmosférica con difusión.

### 3.2. Método de líneas sobre la ecuación del calor

*Ejemplo 3.2.1.* Sea una barra de longitud  $L = 4$  con una constante de difusión del calor  $\alpha^2$  y sea  $u = u(t, x)$  la temperatura en cada punto  $x$  de la barra en un tiempo  $t$ . La ecuación que gobierna la evolución de  $u(t, x)$  es:

$$u_t = \alpha^2 u_{xx}, \quad 0 \leq x \leq L, \quad t \geq 0 \quad (3.1)$$

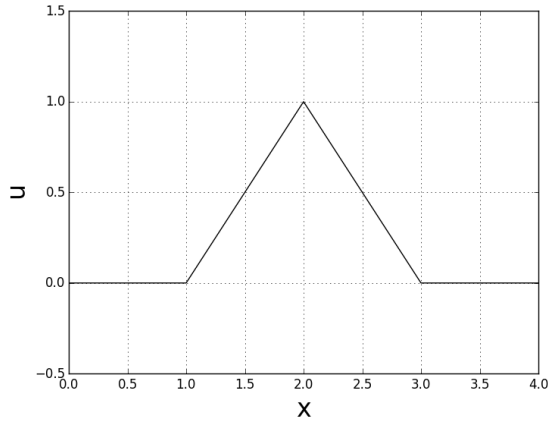
Esta función  $u(t, x)$  tendrá que verificar una condición inicial, que tomaremos

$$u(0, x) = u_0(x) = \max\{0, 1 - |x - 2|\}.$$

*Nota.* El perfil de temperaturas inicial en la barra será como la figura 3.1:

Además, se considera que verifica condiciones de frontera de tipo Dirichlet homogéneas, esto es,

$$u(t, 0) = u(t, L) = 0$$



**Figura 3.1.** Distribución de temperaturas inicial en  $\mathcal{P}$

Más que un método propiamente dicho, el método de líneas es un procedimiento que sigue varios pasos. El primer paso del procedimiento es discretizar la componente espacial. En este ejemplo, consideremos la partición equiespaciada:

$$\mathcal{P} = \{x_j = jh, \quad j = 1, \dots, m\} \quad \text{donde } h = \frac{L}{m+1}$$

El objetivo del siguiente paso es encontrar unas funciones  $v_j(t)$  que aproximen a  $u(t, x_j)$  en cada punto de la partición.

Del desarrollo de Taylor de la solución sobre cada  $x_j$  se tiene

$$u(t, x_j \pm h) = u(t, x_j) \pm hu_x(t, x_j) + \frac{h^2}{2}u_{xx}(t, x_j) \pm \frac{h^3}{3!}u_{xxx}(t, x_j) + \mathcal{O}(h^4) \Rightarrow$$

$$u(t, x_j + h) + u(t, x_j - h) = 2u(t, x_j) + h^2u_{xx}(t, x_j) + \mathcal{O}(h^4) \Rightarrow$$

$$u_{xx}(t, x_j) = \frac{u(t, x_j + h) + u(t, x_j - h) - 2u(t, x_j)}{h^2} + \mathcal{O}(h^2), \quad h \rightarrow 0 \quad (3.2)$$

Para cada  $x = x_j$ , por (3.2):

$$u_{tt}(t, x_j) = \alpha^2 u_{xx}(t, x_j) = \frac{\alpha^2}{h^2} (u(t, x_j - h) - 2u(t, x_j) + u(t, x_j + h)) + \mathcal{O}(h^2) \quad (3.3)$$

Si llamamos  $v_j(t)$  a la aproximación buscada de  $u(t, x_j)$  se tiene que  $v'_j(t) \simeq u_t(t, x_j)$ ,  $v_{j-1}(t) \simeq u(t, x_j - h)$  y  $v_{j+1}(t) \simeq u(t, x_j + h)$ .

Por tanto, si tomamos dicha aproximación  $v_j$  verificando

$$v'_j(t) = \frac{\alpha^2}{h^2}(v_{j-1} - 2v_j + v_{j+1}) \tag{3.4}$$

es esperable de (3.3) que  $v_j(t) - u(t, x_j) = \mathcal{O}(h^2)$ ,  $h \rightarrow 0$ .

Hay que tener cuidado en esta discretización con las condiciones de frontera, que en este caso son Dirichlet homogéneas y que se incorporan al sistema de EDOs de la siguiente forma:

Para  $j = 1$ :

$$v'_1(t) = \frac{\alpha^2}{h^2}(v_0 - 2v_1 + v_2), \quad v_0 = u(t, x_0) = u(t, 0) = 0$$

Para  $j = m$ :

$$v'_m(t) = \frac{\alpha^2}{h^2}(v_{m-1} - 2v_m + v_{m+1}), \quad v_{m+1} = u(t, x_{m+1}) = u(t, L) = 0$$

Por tanto, obtenemos el sistema de EDOs siguiente:

$$v' = \frac{\alpha^2}{h^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 \\ 0 & 1 & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 \end{pmatrix} v = \mathcal{A}v \quad v \in \mathbb{R}^m \tag{3.5}$$

Además, la condición inicial de la EDP se incorpora como la condición inicial del sistema de EDOs:

$$v_0 = \begin{pmatrix} u(0, x_1) \\ \vdots \\ u(0, x_m) \end{pmatrix} = \begin{pmatrix} u_0(x_1) \\ \vdots \\ u_0(x_m) \end{pmatrix}$$

Este PVI es un problema lineal y de coeficientes constantes que puede ser más o menos stiff en función de los autovalores de la matriz  $\mathcal{A}$ .

$\mathcal{A}$  es una matriz tridiagonal y de coeficientes constantes de tipo Toeplitz. Existe una expresión explícita de los autovalores y autovectores de matrices de este tipo [1, p.237].

Como resultado, tenemos que los autovalores de la matriz  $\mathcal{A}$  son:

$$\lambda_j = -2\frac{\alpha^2}{h^2} \left( 1 - \cos \left( \frac{j\pi}{m+1} \right) \right), \quad 1 \leq j \leq m$$

Aplicando la identidad trigonométrica  $1 - \cos \alpha = \frac{1}{2} \sin^2 \left( \frac{\alpha}{2} \right)$  se obtiene

$$\lambda_j = -\frac{\alpha^2}{h^2} \sin^2 \left( \frac{j\pi}{2(m+1)} \right) \leq 0, \quad 1 \leq j \leq m$$

y como  $\sin^2(\alpha) \leq 1$  se tiene que todos los autovalores de  $\mathcal{A}$  verifican

$$-\frac{\alpha^2}{h^2} \leq \lambda_j \leq 0, \quad 1 \leq j \leq m \quad (3.6)$$

y hay valores muy cercanos a  $-\frac{\alpha^2}{h^2}$ , lo que puede ser muy pequeño si el número de elementos de la partición espacial es grande.

Como último paso del procedimiento del método de líneas, se aplica un método numérico para la integración temporal del PVI (3.5) con  $m = 40$  usando los métodos vistos en el capítulo 1 a paso fijo con un tamaño de paso temporal  $k = 1/40$ . En la figura 3.2 se recoge el perfil de temperaturas de la barra transcurrido 1 segundo para los cuatro métodos considerados.

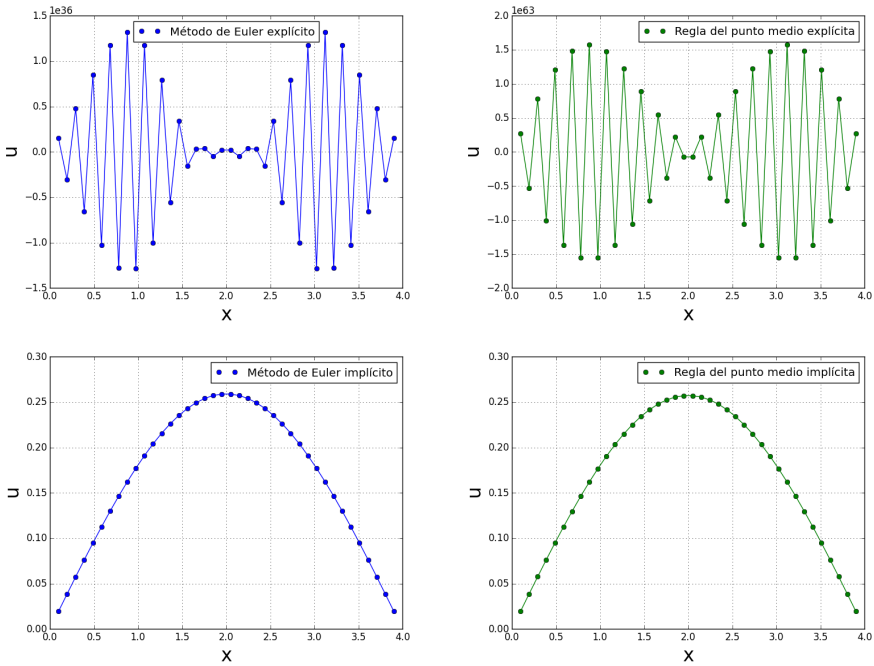


Figura 3.2. Distribución de temperaturas en  $\mathcal{P}$

En estas gráficas podemos observar que, de nuevo, los métodos explícitos no funcionan bien pues, como vimos en el capítulo anterior, para que sean estables  $k\lambda_j \in \mathcal{S}$ ,  $1 \leq j \leq m$ , donde  $k$  es el tamaño de paso temporal y  $\mathcal{S}$  el dominio de estabilidad lineal. Como vimos que los autovalores son reales negativos, en los

tres métodos considerados esto quiere decir que es necesario que  $k\lambda_j \in [-2, 0]$ . De (3.6) es fácil ver que esto solo se garantiza si

$$k \leq \frac{2h^2}{\alpha^2}$$

Cuando  $m = 40$ ,  $h = 10^{-1}$ ,  $\alpha = 1$ , se tiene que para que funcionen bien los explícitos se necesitará que  $k \leq 0.02$ . Obsérvese que esto justifica que en el ejemplo dado en la figura 3.2 los métodos explícitos fallan ya que se ha tomado  $k = 0.025$ , un valor superior aunque muy cercano a la cota teórica, por lo que los resultados numéricos avalan claramente la teoría vista.

### 3.3. Condiciones de frontera de tipo Neumann

*Ejemplo 3.3.1.* Reconsidérese el ejemplo 3.2.1 pero en esta ocasión consideremos que las condiciones de contorno son de tipo Neumann homogéneas, esto es:

$$u_x(t, 0) = u_x(t, L) = 0$$

En este caso, se adapta la partición espacial para incorporar estas condiciones. Para ello se considera la partición  $\mathcal{P}$ :

$$\mathcal{P} = \left\{ x_j = \left( j - \frac{1}{2} \right) h, \quad j = 1, \dots, m \right\} \quad \text{donde } h = \frac{L}{m}$$

Igual que en la sección anterior nuestro objetivo es encontrar unas funciones  $v_j(t)$  que aproximen a  $u(t, x_j)$  en cada punto de la partición, y se obtiene (3.4) de forma similar.

Sin embargo, las condiciones de frontera fuerzan a que se evalúe el caso  $j = 1$  y el  $j = m$  de forma diferente el caso de la sección anterior.

Para  $j = 1$ :

$$v_1'(t) = \frac{\alpha^2}{h^2}(v_0 - 2v_1 + v_2), \quad v_0 = u(t, x_0) = u\left(t, -\frac{h}{2}\right)$$

Como no conocemos este valor  $v_0$  como pasaba con las condiciones de Dirichlet, se usa otra discretización en los bordes. Por ejemplo, podemos tomar diferencias centrales de orden dos, esto es:

$$u_x(t, x) = \frac{u(t, x + h) - u(t, x - h)}{2h} + \mathcal{O}(h^2)$$

Por lo tanto, se concluye:

$$u_x(t, 0) \simeq \frac{u\left(t, \frac{h}{2}\right) - u\left(t, -\frac{h}{2}\right)}{2\frac{h}{2}} = 0 \Rightarrow u\left(t, \frac{h}{2}\right) \simeq u\left(t, -\frac{h}{2}\right) \Rightarrow v_0 \simeq v_1$$

$$v_1'(t) = \frac{\alpha^2}{h^2}(-v_1 + v_2)$$

Para  $j = m$ :

$$v_m'(t) = \frac{\alpha^2}{h^2}(v_{m-1} - 2v_m + v_{m+1}), \quad v_{m+1} = u(t, x_{m+1}) = u\left(t, L + \frac{h}{2}\right)$$

Análogamente, se llega a que  $v_{m+1} \simeq v_m$ . En consecuencia se tiene:

$$v_m'(t) = \frac{\alpha^2}{h^2}(v_{m-1} - v_m)$$

Tenemos el sistema de EDO siguiente:

$$v' = \frac{\alpha^2}{h^2} \begin{pmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 \\ 0 & 1 & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix} v, \quad v \in \mathbb{R}^m.$$

La condición inicial de la EDP se incorpora como condición inicial de la EDO al igual que en la sección anterior

$$v_0 = \begin{pmatrix} u(0, x_1) \\ \vdots \\ u(0, x_m) \end{pmatrix} = \begin{pmatrix} u_0(x_1) \\ \vdots \\ u_0(x_m) \end{pmatrix}.$$

El PVI resultante se parece en forma al ejemplo anterior, aunque sus autovalores son algo diferentes.

Si aplicamos los distintos métodos Runge-Kutta citados previamente a este PVI obtenemos los perfiles de temperatura de la barra tras 1s dados en la figura 3.3.

Podemos observar nuevamente que los métodos implícitos presentan aproximaciones mucho más aceptables que las soluciones dadas por los métodos explícitos. Así que las condiciones de frontera van a determinar el tipo de discretización espacial y el tipo de PVI resultante.

*Nota.* La elección de la partición no es única. Se pueden considerar también como incógnitas los valores de la temperatura en los extremos, a coste de obtener un sistema de EDO de mayor dimensión.

El método de líneas es, por tanto, un procedimiento por el que se discretiza una EDP de forma que se obtiene una aproximación numérica a la solución de dicha EDP mediante la resolución de grandes sistemas de EDOs.

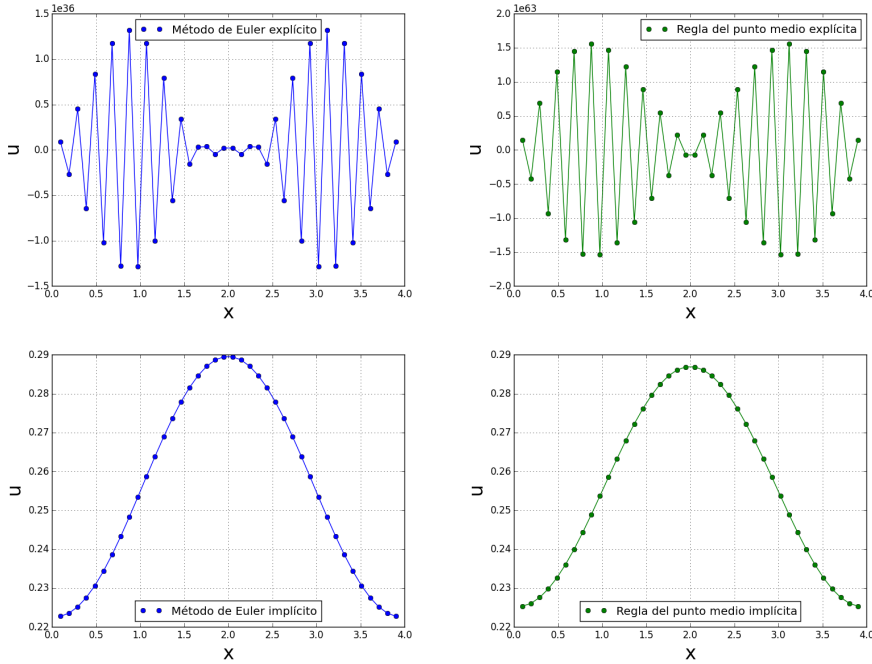


Figura 3.3. Distribución de temperaturas en  $\mathcal{P}$

### 3.4. Problemas de química atmosférica con difusión

En la sección 2.3 vimos un modelo de polución atmosférica que solo consideraba las reacciones químicas de los óxidos de nitrógeno. Un modelo más realista para predecir la evolución de la polución atmosférica es un modelo en el cual se considere el fenómeno de la difusión de las sustancias. Esta difusión normalmente viene en función de la altura de la atmósfera en la que se encuentren. Consideremos de nuevo las concentraciones  $u = (u_1, u_2, u_3, u_4)^T$  de  $O$ ,  $NO$ ,  $NO_2$  y  $O_3$  respectivamente, pero considerando ahora que cada componente depende también de la altura  $\sigma$  en la atmósfera a la que se localizan, esto es,  $u_i = u_i(t, \sigma)$ , para  $0 \leq t \leq T$  (en segundos) y  $0 \leq \sigma \leq \sigma_H = 200$  (medido en metros). Como vimos en la sección 2.3, las reacciones químicas entre estas sustancias evolucionan en el tiempo siguiendo las ecuaciones (2.4), en cada posición espacial fija en que se encuentran. Pero ahora consideremos que las sustancias se difunden en la atmósfera dependiendo de la altura a la que estén, con una velocidad de difusión dada por cierta función  $K(\sigma)$ . En consecuencia, la variación de las concentraciones de estas sustancias va a modelizarse por el siguiente sistema de EDPs [9]:

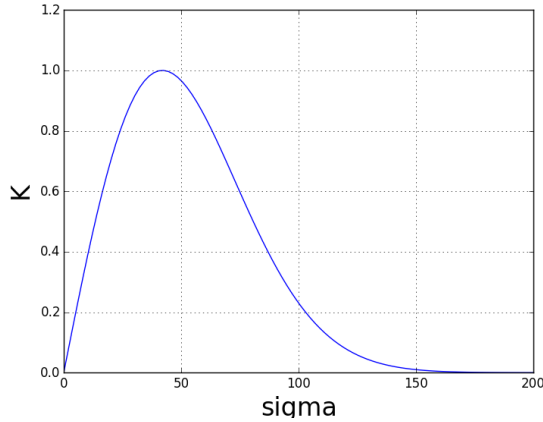
$$u_t = (K(\sigma)u_\sigma)_\sigma + r(t, u), \quad 0 \leq \sigma \leq \sigma_H = 200, \quad 0 \leq t \leq T. \quad (3.7)$$

$$r(t, u) = \begin{pmatrix} k_1(t)u_3 - k_2u_1 \\ k_1(t)u_3 - k_3u_2u_4 + \sigma_2 \\ k_3u_2u_4 - k_1(t)u_3 \\ k_2u_1 - k_3u_2u_4 \end{pmatrix}$$

Habitualmente, la difusión suele ser máxima a una cierta altura, disminuyendo su velocidad en las capas superiores de la atmósfera. Este fenómeno pueda simularse con una función de difusión  $K$  dada por:

$$K(\sigma) = \frac{1}{42}\sigma e^{0.5\left(1 - \frac{\sigma^2}{0.21^2\sigma_H^2}\right)}$$

En la figura 3.4 podemos ver el perfil de  $K(\sigma)$ ,  $0 \leq \sigma \leq 200$



**Figura 3.4.** Función de difusión  $K(\sigma)$

En [9] se estudian este tipo de problemas de polución con difusión, aunque en dicho trabajo se aplica a un sistema de más de 140 sustancias diferentes, lo que está fuera de nuestro alcance. El método que proponen en dicho trabajo con algunas mejoras posteriores y un largo proceso para incluir los datos de cientos de sustancias químicas en la atmósfera recogidas en Europa cada año, es el que se usa actualmente en el centro de investigación de Oslo, el Meteorological Synthetizing Centre-West para monitorizar la evolución de contaminantes a nivel global. [8]

Siguiendo las ideas dadas en [9], las condiciones de frontera que se considerarán serán de Neumann homogéneas:

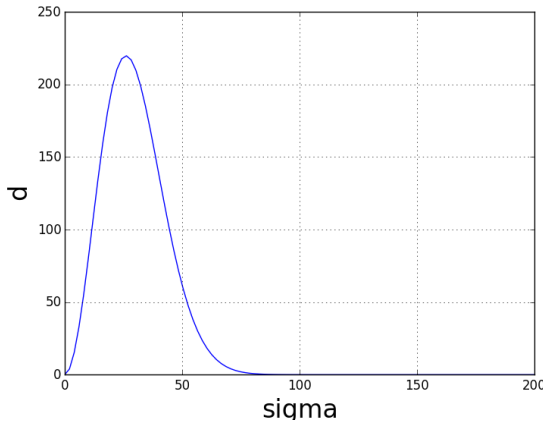


$$K(0) \frac{\partial u}{\partial \sigma}(t, 0) = 0, \quad K(\sigma_H) \frac{\partial u}{\partial \sigma}(t, \sigma_H) = 0$$

Por otro lado, las condiciones iniciales que se consideran son:

$$\begin{aligned} u(0, \sigma) &= \frac{\sigma^2(\sigma_H - \sigma)^2}{\sigma_H^2} e^{-\frac{\sigma^2}{4\sigma_H}} (0, 1.3 \cdot 10^8, 5.0 \cdot 10^{11}, 8.0 \cdot 10^{11})^T \\ &= d(\sigma)(0, 1.3 \cdot 10^8, 5.0 \cdot 10^{11}, 8.0 \cdot 10^{11})^T \end{aligned}$$

En la figura 3.5 podemos ver la función  $d(\sigma)$ ,  $0 \leq \sigma \leq 200$ . Dicha función inicial refleja el hecho de que las concentraciones de las sustancias en las partes más altas son bastante menores que en las capas bajas de la atmósfera.



**Figura 3.5.** Perfil de  $d(\sigma)$

Para aplicar el método de líneas como se hace en [9], primero discretizamos la componente espacial. Debido a las condiciones de Neumann, como vimos en la sección 3.3, usaremos la partición  $\Omega$ :

$$\Omega = \left\{ \sigma_j = \left( j - \frac{1}{2} \right) h, \quad 1 \leq j \leq N \right\} \quad \text{donde} \quad h = \frac{\sigma_H}{N} \quad (3.8)$$

Como la función de difusión depende de  $\sigma$ , no se puede aplicar la discretización de segundo orden (3.2). En este caso se aplicará de forma sucesiva las diferencias finitas centrales de orden 2.

$$g_x(t, x) = \frac{g(t, x + h) + g(t, x - h)}{2h} + \mathcal{O}(h^2) \quad (3.9)$$

Para cada  $t$ , definimos la función:

$$f(\sigma) = K(\sigma) \frac{\partial u}{\partial \sigma}(t, \sigma)$$

Realizamos la primera discretización, que consistirá en la aproximación de la derivada espacial de la función  $f(\sigma)$ :

$$\begin{aligned} (f(\sigma))_{\sigma}(\sigma_j, t) &\simeq \frac{1}{h} (f(\sigma_{j+1/2}, t) - f(\sigma_{j-1/2}, t)) \\ &\simeq \frac{1}{h} \left( K(\sigma_{j+1/2}) \frac{\partial u}{\partial \sigma}(t, \sigma_{j+1/2}) - K(\sigma_{j-1/2}) \frac{\partial u}{\partial \sigma}(t, \sigma_{j-1/2}) \right) \end{aligned}$$

A continuación, realizamos la segunda discretización usando (3.9). Esta consistirá en aproximar la derivada espacial de  $u$  mediante:

$$(f(\sigma))_{\sigma}(\sigma_j, t) \simeq \frac{1}{h} \left( K(\sigma_{j+1/2}) \frac{u(t, \sigma_{j+1}) - u(t, \sigma_j)}{h} - K(\sigma_{j-1/2}) \frac{u(t, \sigma_j) - u(t, \sigma_{j-1})}{h} \right)$$

Si definimos  $c_j(t)$  como la función que aproxima a  $u(t, \sigma_j)$  tenemos que  $c_{j+1}(t) = u(t, \sigma_{j+1})$ ,  $c_{j-1}(t) = u(t, \sigma_{j-1})$  y  $c'_j(t) = u_t(t, \sigma_j)$ . También para simplificar la escritura, adoptaremos algunos aspectos de la notación de [9, p.193]

$$K_j^{\pm} = K(\sigma_{j\pm 1/2}).$$

Tenemos, por lo tanto:

$$c'_j(t) = \frac{1}{h} \left( K_j^+ \frac{c_{j+1} - c_j}{h} - K_j^- \frac{c_j - c_{j-1}}{h} + r(t, c_j) \right) \quad 1 \leq j \leq N$$

A continuación introducimos las condiciones de frontera, del mismo modo que en el ejemplo 3.3.1. Para  $j = 1$ :

$$c'_1(t) = \frac{1}{h} \left( K_1^+ \frac{c_2 - c_1}{h} - K_1^- \frac{c_1 - c_0}{h} \right)$$

Si discretizamos la condición de contorno haciendo uso de (3.9).

$$K(0) \frac{\partial \rho}{\partial \sigma}(t, 0) = 0 \Rightarrow K_1^- \frac{c_1 - c_0}{h} = 0$$

Por lo tanto,

$$c'_1(t) = \frac{1}{h} \left( K_1^+ \frac{c_2 - c_1}{h} \right)$$

Para  $j = N$ :

$$c'_N(t) = \frac{1}{h} \left( K_N^+ \frac{c_{N+1} - c_N}{h} - K_N^- \frac{c_N - c_{N-1}}{h} \right)$$

El razonamiento es análogo al realizado anteriormente y se llega a que:

$$c'_N(t) = \frac{1}{h} \left( -K_N^- \frac{c_N - c_{N-1}}{h} \right)$$

Seguidamente, para compactar la notación definimos para cada  $1 \leq j \leq N$  (fijando  $\alpha_N = 0$  y  $\beta_1 = 0$ ):

$$\alpha_j = \frac{K_j^+}{h^2}, \quad \beta_j = \frac{K_j^-}{h^2}, \quad \gamma_j = \alpha_j + \beta_j \geq 0$$

Con esto el sistema de EDOs queda de la siguiente forma:

$$c'_j(t) = \alpha_j c_{j+1} - \gamma_j c_j + \beta_j c_{j-1} + r(t, c_j), \quad 1 \leq j \leq N$$

Si escribimos componente a componente se tiene:

$$c'_{jk}(t) = \alpha_j c_{j+1k} - \gamma_j c_{jk} + \beta_j c_{j-1k} + r_k(t, c_j), \quad 1 \leq j \leq N, \quad 1 \leq k \leq 4$$

Cabe destacar que este sistema está fuertemente acoplado debido a  $r(t, c_j)$ . En lo que sigue denotamos  $(c_{j1}, c_{j2}, c_{j3}, c_{j4})^T$ ,  $1 \leq j \leq N$ , mientras que llamamos  $\tilde{c}_k = (c_{1k}, c_{2k}, \dots, c_{Nk})$ ,  $1 \leq k \leq 4$ . Por lo tanto, tendremos que para cada  $\tilde{c}_k$  se cumple el siguiente sistema de EDOs:

$$\tilde{c}'_k(t) = \mathcal{M}\tilde{c}_k + \mathcal{R}_k(t, \tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4), \quad 1 \leq k \leq 4$$

siendo

$$\mathcal{M} = \begin{pmatrix} -\gamma_1 & \alpha_1 & 0 & \dots & 0 & 0 & 0 \\ \beta_2 & -\gamma_2 & \alpha_2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \beta_{N-1} & -\gamma_{N-1} & \alpha_{N-1} \\ 0 & 0 & 0 & \dots & 0 & \beta_N & -\gamma_N \end{pmatrix}, \quad \mathcal{R}_k(\cdot) = \begin{pmatrix} r_k(t, c_{11}, c_{12}, c_{13}, c_{14}) \\ r_k(t, c_{21}, c_{22}, c_{23}, c_{24}) \\ \vdots \\ r_k(t, c_{N1}, c_{N2}, c_{N3}, c_{N4}) \end{pmatrix}$$

donde  $r_k$  es la componente  $k$ -ésima de  $r$ . Si definimos  $C = (\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4)^T$  podemos expresar el sistema de EDOs de manera más compacta:

$$C'(t) = (I_4 \otimes \mathcal{M})C + \mathcal{R}(t, C) = F(t, C), \quad C(0) = C^0 \quad (3.10)$$

donde  $\mathcal{R}(t, C) = (\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \mathcal{R}_4)^T$ .

Por lo tanto, para obtener una aproximación a la solución del problema de polución atmosférica (3.7) sobre los puntos de la malla espacial (3.8), tenemos que aproximar la solución de este PVI (3.10) de dimensión  $4N$ . Además, debido tanto a la discretización de la difusión como a la parte de reacción  $\mathcal{R}(t, C)$  (que ya vimos en el capítulo anterior que es stiff), presenta un alto grado de stiffness, por lo que será necesario aplicar métodos implícitos.

En particular, se ha usado el método de Euler implícito (1.7), resolviendo en cada paso un sistema implícito de dimensión  $4N$  con el esquema de Newton modificado (1.15). Para ilustrar su comportamiento, hemos representado en la Figura 3.6 los resultados obtenidos con una partición equiespaciada de  $N = 40$  puntos y con paso temporal fijo  $k = 10$ , para cada una de las cuatro sustancias ( $O$ ,  $NO$ ,  $NO_2$ ,  $O_3$ ) a ciertas alturas fijas  $\sigma = 2.5, 12.5, 32.5, 72.5, 152.5, 197.5$ . En [9] este problema se integra durante un intervalo de 112 horas pero separándolo en 56 intervalos de 2 horas, recomenzando en cada uno de ellos la integración con el Euler implícito con un tamaño de paso diez veces más pequeño que en el intervalo anterior. Como nuestro objetivo es solo ilustrativo, hemos tomado solo un primer intervalo de dos horas, esto es, con tiempo final  $T = 2 \times 3600$ s.

En las gráficas de cada sustancia podemos ver que el método de líneas implementado produce una solución aceptable, reflejándose claramente el fenómeno de difusión de las sustancias a lo largo del tiempo en cada nivel de altura. Obsérvese, sin embargo, que la componente  $u_2$  (correspondiente al  $NO$ ) aumenta a lo largo del tiempo debido a la presencia de la fuente constante representada por  $\sigma_2$ .

### 3.5. Conclusiones

En esta memoria nos hemos introducido en el campo de los métodos numéricos para resolver modelos de química atmosférica. En particular, hemos estudiado las principales propiedades teóricas y prácticas del método de líneas. Debido a la presencia habitual de stiffness en los modelos semi-discretos obtenidos en este ámbito, hemos profundizado en el uso de los métodos implícitos que son indispensables para la resolución de dichos problemas stiff.

Además, en todos los capítulos se ilustran los resultados teóricos con diversos experimentos numéricos llevados a cabo usando Python, obteniendo finalmente un código que implementa el método de líneas sobre el problema de difusión-reacción (3.7) propuesto en [9], como ejemplo de los modelos similares que se utilizan en la actualidad en las predicciones de la evolución de la contaminación atmosférica en Europa [8].

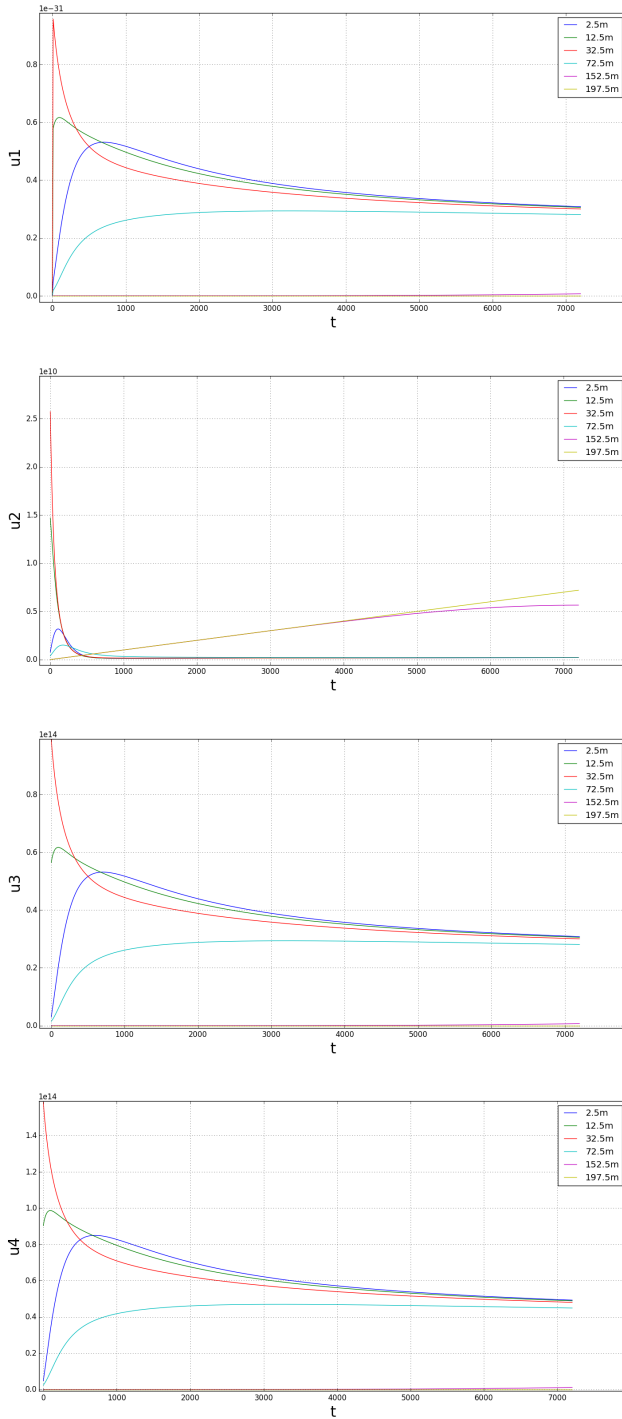


Figura 3.6. Componentes obtenidas a partir del método de Euler implícito



---

## Bibliografía

- [1] BALDÍA CONTELLES J. M. ,VIDAL MACIÁ A. M. *Cálculo de los valores propios de matrices tridiagonales simétricas mediante la iteración de Laguerre. Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería. Volumen 12,6, pág 227-249 (2000)*
- [2] BUTCHER J. C. *Numerical Methods for Differential Equations and Applications*, Arab. J. Sci. Eng. 22(2C) (1997).
- [3] BUTCHER J. C. *Numerical Methods for Ordinary Differential Equations. Second Edition. John Wiley & Sons, 2008*
- [4] HAIRER E., NORSETT S. P., WANNER G. *Solving Ordinary Differential Equations I. Second Edition. Springer. 1993*
- [5] HAIRER E., WANNER G. *Solving Ordinary Differential Equations II. Stiff and differential-algebraic Problems. Second Revised Edition. Springer 1996*
- [6] HUNDSORFER W., VERWER J. *Numerical Solution of Time-Dependent Advection-Diffusion-Reactions Equations*
- [7] LEZAUN M., *Predicciones del tiempo y matemáticas. Bol. Soc. Esp. Mat. Apl. 22 (2002) 61-100.*
- [8] SIMPSON D. ET AL. *The EMEP MSC-W chemical transport model-technical description. Atmospheric Chemistry and Physics,12 (2012) 7825-7865*
- [9] VERWER J.G., BLOM J.G., HUNDSORFER W. *An implicit-explicit approach for atmospheric transport-chemistry problems. Applied Numerical Mathematics 20 (1996) 191-209*





---

## Lista de Figuras

1.1. Error en los métodos explícitos del ejemplo 1.4.1. ....	5
1.2. Error en los métodos implícitos del ejemplo 1.4.1. ....	5
1.3. Error en los métodos explícitos del ejemplo 1.7.1. ....	19
1.4. Error en los métodos implícitos del ejemplo 1.7.1. ....	19
2.1. $w_1(t)$ del ejemplo 2.1.1 para $k_2 = 10$ .....	22
2.2. $w_1(t)$ del ejemplo 2.1.1 para $k_2 = 100$ .....	22
2.3. $w_1(t)$ del ejemplo 2.1.1 para $k_2 = 1000$ .....	23
2.4. Soluciones numéricas con métodos explícitos para (2.4) con $h = 10$	28
2.5. Soluciones numéricas con el método de Euler implícito para (2.4) con $h = 10$ . ....	29
3.1. Distribución de temperaturas inicial en $\mathcal{P}$ .....	32
3.2. Distribución de temperaturas en $\mathcal{P}$ .....	34
3.3. Distribución de temperaturas en $\mathcal{P}$ .....	37
3.4. Función de difusión $K(\sigma)$ .....	38
3.5. Perfil de $d(\sigma)$ .....	39
3.6. Componentes obtenidas a partir del método de Euler implícito ...	43



# Numerical methods for the solution of differential equations and applications in

### Abstract

In this work the numerical solution by the method of lines of atmospheric pollution problems with diffusion is considered.

### 1. Introduction

The atmospheric chemistry is an important area in the subject of the weather forecast and the study of the evolution of air pollution. A diffusion-reaction problem is used in [5] that is a good example of the models that nowadays are being used for monitoring the acid deposition in Europe [4]. In order to solve numerically this model, in the first section of this work Runge-Kutta methods (RK) are introduced. In the second section the phenomenon of stiffness is introduced by an example and in the last section the method of lines is used to describe the time evolution of nitrogen oxides.

### 2. Introduction to Runge-Kutta methods

Let us consider an initial value problem (IVP)  $y' = f(x, y(x))$ ,  $y(x_0) = y_0$ ,  $y, f \in \mathbb{R}^m$ . If  $y_i$  is an approximation for the solution of the IVP at  $x_i \in [x_0, x_N]$ , a **s-stage RK method** gives an approximation  $y_{i+1}$  at  $x_{i+1} = x_i + h$  by:

$$K_j = f(x_i + c_j h, y_i + h \sum_{l=1}^s a_{jl} K_l), \quad 1 \leq j \leq s$$

$$y_{i+1} = y_i + h \sum_{j=1}^s b_j K_j$$

If  $a_{ji} = 0$ ,  $\forall l \geq j$ , the method is **explicit** whereas, in other cases, the method is called **implicit**.

In this section the convergence of these methods is studied, obtaining that

**Lemma:** A RK(A, b, c) is convergent of order  $p \geq 1$ , con  $1 \leq p \leq 4$ , that is,  $\|y(x_i) - y_i\| \leq Kh^p$ ,  $0 \leq i \leq n$ ,  $h \rightarrow 0$ , if

$p \geq 1$	$b^T e = 1$
$p \geq 2$	$(p \geq 1) + b^T c = \frac{1}{2}$
$p \geq 3$	$(p \geq 2) + b^T c^2 = \frac{1}{3}, b^T A c = \frac{1}{6}$
$p \geq 4$	$(p \geq 3) + b^T c^3 = \frac{1}{4}, b^T A^2 c = \frac{1}{24}, b^T A c^2 = \frac{1}{12}, b^T (A c \cdot c) = \frac{1}{24}$

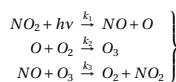
where  $\cdot$  is the component-by-component product.

### 3. Stiff problems and linear stability

"Stiff equations are problems for which explicit methods don't work" [2].

In real-life pollution models the large differences between the rate coefficients of

the chemical reactions involved usually cause stiffness, so the use of implicit schemes is compulsory in this framework. For example, let us consider this chain of chemical reactions which involves the nitrogen oxides [3]:



which is modelled by

$$\left. \begin{aligned} u_1'(t) &= k_1(t)u_3 - k_2u_1 \\ u_2'(t) &= k_1(t)u_3 - k_3u_2u_4 + \sigma_2 \\ u_3'(t) &= k_2u_1u_4 - k_1(t)u_3 \\ u_4'(t) &= k_2u_1 - k_3u_2u_4 \end{aligned} \right\} \quad (1)$$

where  $k_2 = 10^5$ ,  $k_3 = 10^{-16}$ ,  $\sigma_2 = 10^6$  and the ratio  $k_1(t)$  depends on the diurnal cycle, and the initial values of the concentrations are  $u(0) = (0, 1.3e8, 5.0e11, 8.0e11)^T$ .

The explicit methods do not provide an appropriate approximation for the exact solution of this problem. However, using an implicit method such as the *backward Euler* [1] we obtain a good approximation given in Figure 1

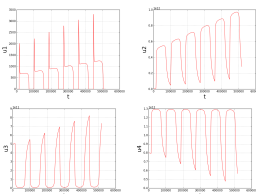


Figure 1: Backward Euler with  $h = 10$

### 4. Method of lines for parabolic PDEs

In the previous pollution model, the concentrations of the chemicals only depend on  $t$ , but in atmospheric models they also depend on the vertical position  $\sigma$  in the atmosphere where the chemicals are. i.e.,  $u = u(t, \sigma)$ . In [5] they consider problems of type

$$u_t = (K(\sigma)u_\sigma)_\sigma + r(t, u)$$

where homogeneous Neumann boundary conditions are imposed. For our tests, the reaction part  $r(t, u)$  is the same as in the previous problem (1).

The method of lines (MoL) is a procedure which firstly discretize this PDE on the spatial mesh

$$\Omega = \left\{ \sigma_j = \left(j - \frac{1}{2}\right)h, 1 \leq j \leq N \right\}, \quad h = \frac{\sigma_H}{N},$$

and, including the boundary conditions, obtain an IVP of this type

$$C'(t) = (I_4 \otimes \mathcal{M})C + \mathcal{R}(t, C) = F(t, C)$$

where  $I_4$  is the identity matrix of dimension 4,  $\mathcal{M}$  is a tridiagonal matrix of dimension  $N$ , and  $\mathcal{R}(t, C)$  is a non-linear function of  $4N$  variables, since  $C(t) = (c_j(t))_{j=1}^N$ , with  $c_j(t) \approx u(t, \sigma_j) \in \mathbb{R}^4$ .

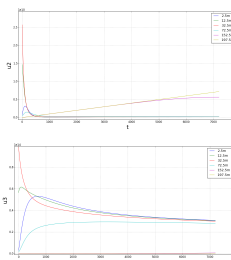


Figure 2: MoL with Backward Euler

### References

- [1] HAIRER E., NORSETT S. P., WANNER G. *Solving Ordinary Differential Equations I. Second Edition.* Springer, 1993
- [2] HAIRER E., WANNER G. *Solving Ordinary Differential Equations II. Stiff and differential-algebraic Problems. Second Revised Edition.* Springer 1996
- [3] HUNSDORFER W., VERWER J. *Numerical Solution of Time-Dependent Advection-Diffusion-Reactions Equations*
- [4] SIMPSON D. ET AL. *The EMEP MSC-W chemical transport model-technical description.* Atmospheric Chemistry and Physics, 12 (2012) 7825-7865
- [5] VERWER J.G., BLOM J.G., HUNSDORFER W. *An implicit-explicit approach for atmospheric transport-chemistry problems.* Applied Numerical Mathematics 20 (1996) 191-209