

Alba del Pino Sosa Déniz

*Métodos numéricos para la resolución
de ecuaciones diferenciales y
aplicaciones. Problemas de segundo
orden.*

Numerical methods to approximate the solution
of differential equations. Second order problems.

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, Septiembre de 2021

DIRIGIDO POR

María Soledad Pérez Rodríguez

María Soledad Pérez Rodríguez

Análisis Matemático

Universidad de La Laguna

38200 La Laguna, Tenerife

Agradecimientos

En primer lugar quiero agradecer a mi tutora, M. Soledad Pérez Rodríguez, por guiarme y ayudarme en este proceso y permitirme descubrir nuevos aspectos de las matemáticas, sin ella esto no habría sido posible.

A mis padres y mi hermana, por su fundamental apoyo y cariño a lo largo de esta etapa tan importante para mí. Gracias por no permitir que me rindiera.

A mis amigas de toda la vida, que siempre están para lo que necesite, en especial para consejos y diversión.

Y a mis amigas de Tenerife, por ayudarme y acompañarme a lo largo de estos últimos años.

Alba del Pino Sosa Déniz
La Laguna, 10 de septiembre de 2021

Resumen · Abstract

Resumen

El objetivo de este trabajo es el estudio de los métodos Runge-Kutta-Nyström (RKN) para la resolución de problemas de valor inicial (PVI) en ecuaciones diferenciales ordinarias (EDO) de segundo orden. En primer lugar, profundizaremos en el estudio de los métodos Runge-Kutta (RK) para PVI de primer orden, ya que los RKN son una extensión natural de ellos. Esto es importante ya que abordar directamente la convergencia de los métodos RKN es bastante complejo sin conocer antes cómo se deducen las condiciones de orden de los métodos RK y, en particular, la teoría de las series de Butcher. A partir de esta teoría, demostraremos dos teoremas relevantes. El primero da una acotación rigurosa del error local de los RK y se usa para demostrar el segundo, que nos asegura la convergencia de dichos métodos. Establecida ya esta base sólida, pasaremos a extender estos métodos a los PVI de segundo orden, definiendo los métodos RKN y estudiando su orden y convergencia. Finalizaremos esta memoria aplicando estos métodos para la integración temporal de la ecuación de ondas y testando su comportamiento con varios experimentos numéricos.

Abstract

The aim of this work is to study Runge-Kutta-Nyström (RKN) methods for solving initial value problems (IVPs) in second-order ordinary differential equations (ODEs). First, we will delve into the study of Runge-Kutta (RK) methods for first-order IVPs, since RKNs are a natural extension of them. This is important since tackling directly the convergence of RKNs is quite difficult without first knowing how the order conditions of RK methods are deduced and, in particular, the theory of Butcher series. Then, we will prove two relevant theorems. The first one gives a rigorous bound of the local error of RK methods and is used to prove the second one, which ensures the convergence of RK methods. Having established this solid foundation, we will extend these methods to second-order IVPs, defining RKN methods and studying their order and convergence. We will conclude this report by applying these methods to the time integration of the wave equation and testing their behaviour with several numerical experiments.

Contenido

Agradecimientos	III
Resumen/Abstract	V
Introducción	IX
1. Métodos Runge-Kutta	1
1.1. Primeros métodos Runge-Kutta	1
1.2. Formulación general de los métodos Runge-Kutta	4
2. Orden de los métodos Runge-Kutta. Series de Butcher	7
2.1. Series de Butcher. Condiciones de orden de los métodos RK	12
3. Convergencia de métodos Runge-Kutta	15
3.1. Acotación del error local	16
3.2. Acotación del error global	20
4. Métodos Runge-Kutta-Nyström	23
4.1. Métodos Runge-Kutta-Nyström	24
4.2. Derivadas sucesivas y árboles ordenados	26
4.3. Derivada de la solución numérica	29
4.4. Condiciones de orden	34
5. Aplicación a la ecuación de ondas	37
Bibliografía	43
Poster	45

Introducción

Los sistemas de ecuaciones diferenciales ordinarias de segundo orden aparecen en infinidad de aplicaciones. Cuando se intenta modelizar fenómenos físicos en los que intervienen fuerzas, siempre proporcionales a la aceleración, las ecuaciones de segundo orden surgen de forma natural. Otro ámbito en el que se usan a menudo es el de la resolución numérica de ecuaciones diferenciales en derivadas parciales hiperbólicas, cuando se discretizan parcialmente usando diferencias finitas, elementos finitos, volúmenes finitos, etc. Una forma de resolver estos sistemas es transformarlo en problemas de primer orden de mayor dimensión, pero en muchas aplicaciones, sobre todo cuando las ecuaciones no dependen de las derivadas primeras, se sabe que es más ventajoso resolver directamente el sistema de segundo orden.

En este trabajo estudiamos una clase de métodos numéricos que dan aproximaciones a estos problemas con gran precisión y eficiencia, los métodos Runge-Kutta-Nystöm (RKN). Los cuales son una extensión de los métodos Runge-Kutta (RK), que se ven en la asignatura *Ecuaciones Diferenciales II* del tercer curso del Grado, a los sistemas de segundo orden. Nuestro objetivo es, en primer lugar, profundizar en algunos aspectos de estos métodos RK que no se suelen ver en dicha asignatura y que luego son necesarios para entender los RKN, introduciendo estos posteriormente.

Esta memoria se organiza de la siguiente manera. En el primer capítulo se presentan los métodos RK y la formulación general de los mismos. En el siguiente capítulo se define el orden de los métodos RK. Además, introducimos unas nuevas series denominadas series de Butcher, con las que podremos calcular las condiciones de orden de estos métodos. En el tercer capítulo se estudia la convergencia de los RK acotando el error local y el error global. En el cuarto, se introducen los métodos RKN y, al igual que se hizo en los RK, se estudia su formulación general, el orden y las condiciones necesarias y suficientes para alcanzar un orden determinado. Por último aplicaremos los métodos estudiados para aproximar numéricamente la solución de la ecuación de ondas.

Métodos Runge-Kutta

Los métodos Runge-Kutta tratan de aproximar numéricamente la solución del **problema de valores iniciales** (PVI)

$$\begin{cases} y' = f(t, y), & t \in [0, T], \quad y, f \in \mathbb{R}^m, \\ y(0) = y_0 \end{cases} \quad (1.0.1)$$

donde

$$f : [0, T] \times \mathbb{R}^m \longrightarrow \mathbb{R}^m \quad \text{es continua en } [0, T] \times \mathbb{R}^m \quad (1.0.2)$$

y Lipschitz respecto de y , i.e.,

$$\|f(t, y) - f(t, \bar{y})\| \leq L\|y - \bar{y}\| \quad \forall y, \bar{y} \in \mathbb{R}^m, \quad t \in [0, T] \quad (1.0.3)$$

Aunque en [1] y [3] se definen primero los métodos explícitos y luego se amplía el estudio a los implícitos, para simplificar la exposición hemos optado por definirlos en general, tal y como se hace en [2].

1.1. Primeros métodos Runge-Kutta

El **método de Euler**, llamado así en honor a Leonhard Euler (1707-1783), es el más simple de los métodos Runge-Kutta. Además, sirve como base para construir métodos más complejos.

Si tenemos una aproximación y_n a la solución de (1.0.1) en un punto $t_n \geq 0$, el método de Euler nos da una aproximación a la solución en $t_n + h$ mediante la fórmula:

$$y_{n+1} = y_n + hf(t_n, y_n) \quad (1.1.4)$$

Podemos reescribir el método, denotando $f_n = f(t_n, y_n)$, como

$$K_1 = f_n, \quad y_{n+1} = y_n + hK_1$$

Con esta formulación queda claro que para hallar la aproximación y_{n+1} tenemos que evaluar una vez la función derivada f , lo que supone su principal coste computacional. Por ello, se dice que este método es de una **etapa** K_1

Ejemplo 1.1.1. Queremos aproximar la solución del siguiente problema

$$\begin{cases} y_1' = y_1 + 2y_2 & y_1(0) = y_{0,1} = 1 \\ y_2' = 3y_1 + 4y_2 & y_2(0) = y_{0,2} = -3 \end{cases}$$

mediante el método de Euler. Es de la forma (1.0.1) con $f(t, y) = Ay$, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $y_0 = y(0) = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ por lo que

$$y_1 = y_0 + hAy_0 \Rightarrow \begin{pmatrix} y_{1,1} \\ y_{1,2} \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix} + h \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix} \Rightarrow \begin{cases} y_{1,1} = 1 + h(1 - 6) \\ y_{1,2} = -3 + h(3 - 12) \end{cases}$$

Si, por ejemplo, $h = 0.1$, entonces la solución aproximada será

$$y_1 = \begin{pmatrix} 1 + 0.1(-5) \\ -3 + 0.1(-9) \end{pmatrix} = \begin{pmatrix} 0.5 \\ 3.9 \end{pmatrix},$$

o si $h = 0.01$,

$$y_1 = \begin{pmatrix} 1 + 0.01(-5) \\ -3 + 0.01(-9) \end{pmatrix} = \begin{pmatrix} 0.95 \\ 3.09 \end{pmatrix}$$

Combinando pasos del método de Euler y fórmulas de cuadratura se fueron proponiendo diferentes métodos de un paso, en la literatura.

- *Método de Euler implícito.*

Este método de una etapa viene dado por

$$y_{n+1} = y_n + hf_{n+1} \tag{1.1.5}$$

Lo podemos reescribir como

$$K_1 = f(t_n + h, y_n + K_1), \quad y_{n+1} = y_n + hK_1$$

Si queremos aproximar la solución del PVI dado en el Ejemplo 1.1.1 con el método implícito (1.1.5) en un paso $t = h$ tenemos

$$y_1 = y_0 + hAy_1$$

Por tanto, para calcular y_1 en este caso tenemos que resolver este sistema de ecuaciones

$$y_1 - hAy_1 = y_0 \Rightarrow (I - hA)y_1 = y_0 \Rightarrow y_1 = (I - hA)^{-1}y_0$$

Si $h = 0.1$ tenemos que

$$y_1 = \begin{pmatrix} 0.9 & -0.2 \\ -0.3 & 0.6 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ -5 \end{pmatrix}$$

y si $h = 0.01$ se obtiene que

$$y_1 = \begin{pmatrix} 0.99 & -0.02 \\ -0.03 & 0.96 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0.95 \\ -3.1 \end{pmatrix}$$

Observamos así que utilizando el método (1.1.5) al igual que con el (1.1.4) para aproximar la solución necesitamos evaluar la función derivada una vez en cada paso, pero en el método (1.1.5), además, hay que resolver un sistema de ecuaciones implícitos. Por esto el método se llama **implícito**.

- *Método de Runge.*

Como bien indica su nombre este método fue obtenido por Runge en 1895 al tratar de generalizar la regla del punto medio para cuadraturas. El método resultante es

$$y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}f_n\right) \quad (1.1.6)$$

Podemos reescribir el método de la siguiente manera:

$$\begin{cases} K_1 = f(t_n, y_n) \\ K_2 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}K_1\right) \\ y_{n+1} = y_n + hK_2 \end{cases}$$

así que para aproximar la solución y_{n+1} tenemos que evaluar dos veces la función derivada f , por lo que este método se dice que es de dos etapas K_1 y K_2 .

- *Método de Heun.*

Este método al igual que el anterior tiene dos etapas y se escribe de la siguiente manera

$$\begin{cases} K_1 = f_n \\ K_2 = f\left(t_n + h, y_n + hK_1\right) \\ y_{n+1} = y_n + \frac{h}{2}(K_1 + K_2) \end{cases} \quad (1.1.7)$$

- *Regla implícita del punto medio.*

Es un método implícito de una etapa

$$\begin{cases} K_1 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}K_1\right) \\ y_{n+1} = y_n + hK_1 \end{cases} \quad (1.1.8)$$

- *Regla trapezoidal.*

Se trata de un método implícito de dos etapas que viene dado por

$$\begin{cases} K_1 = f_n \\ K_2 = f(t_n + h, y_n + \frac{h}{2}(K_1 + K_2)) \\ y_{n+1} = y_n + h(\frac{1}{2}K_1 + \frac{1}{2}K_2) \end{cases} \quad (1.1.9)$$

1.2. Formulación general de los métodos Runge-Kutta

En general, un método **Runge-Kutta** (RK) de s etapas se define como un método numérico que dada una aproximación y_n a la solución del PVI (1.0.1), con f verificando (1.0.2) y (1.0.3), en un punto $t_n \in [0, T]$, nos da una aproximación a dicha solución en el punto $t_n + h \in [0, T]$, que denotamos por y_{n+1} y obtenemos mediante la siguiente formulación

$$\begin{cases} K_1 = f(t_n + c_1 h, y_n + h \sum_{j=1}^s a_{1j} K_j) \\ K_2 = f(t_n + c_2 h, y_n + h \sum_{j=1}^s a_{2j} K_j) \\ \vdots \\ K_s = f(t_n + c_s h, y_n + h \sum_{j=1}^s a_{sj} K_j) \end{cases} \quad (1.2.10)$$

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i K_i \quad (1.2.11)$$

donde los vectores K_1, K_2, \dots, K_s se llaman **etapas** del método RK. Aunque en general no es necesario, habitualmente a estos coeficientes se les pide que cumplan

$$\sum_{i=1}^s b_i = 1 \quad y \quad \sum_{j=1}^s a_{ij} = c_i, \quad i = 1, 2, \dots, s \quad (1.2.12)$$

Además se define la tabla de Butcher asociada al RK (1.2.10)- (1.2.11) como

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

donde los coeficientes $(a_{ij})_{i,j=1}^s$ forman la matriz A que se denomina **matriz de coeficientes** del RK, el vector $c = (c_1, c_2, \dots, c_s)^T$ es el **vector de nodos o vector nodal** del RK, y el vector $b^T = (b_1, b_2, \dots, b_s)$ es el **vector de los pesos** del RK.

Utilizando esta notación las condiciones (1.2.12) se escriben como

$$b^T e = 1, \quad Ae = c \tag{1.2.13}$$

donde $e^T = (1, 1, \dots, 1) \in \mathbb{R}^s$ (vector unitario) y, como consecuencia, a un método RK definido con la matriz de coeficientes A y vectores b y c se suele denotar como $RK(A, b)$.

Por ejemplo, la tabla de Butcher de los métodos (1.1.4), (1.1.5), (1.1.6), (1.1.7), (1.1.8) y (1.1.9) son, respectivamente

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

Según la forma de la matriz A los métodos RK se suelen dividir en dos grandes grupos:

- Cuando la matriz A es triangular inferior estricta, el método RK se llama **explícito** (RKE), obteniéndose así sus etapas de manera recursiva.
- Cuando la matriz A no es inferior estricta, es decir, cuando $a_{ij} \neq 0$ para algún $j \geq i$, el método se denomina **implícito** (RKI). Como ya vimos en la sección anterior, para calcular sus etapas tendremos que resolver un sistema implícito de dimensión $s \times m$ (1.2.10).

En este último caso hay que plantear la pregunta de si el sistema implícito de las etapas (1.2.10) tiene solución y, además, su unicidad para que la fórmula de avance (1.2.11) esté bien definida.

Teorema 1.2.2 (Existencia y unicidad de los métodos RKI). [3]

Sea $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ la función derivada del PVI (1.0.1) verificando (1.0.2)-(1.0.3). Si

$$|h| < \frac{1}{L \|A\|_\infty}, \quad L = \text{constante de Lipschitz de } f,$$

el sistema implícito (1.2.10) del $RK(A, b)$ admite solución única $(K_1(h), \dots, K_s(h))$, que puede obtenerse por iteración funcional.

Aplicando el teorema de la función implícita incluso se puede garantizar que si $f \in C^p(\mathbb{R} \times \mathbb{R}^m)$,

$$K_i(h) \in C^p([0, h^*]) \text{ para cierto } h^* > 0 \text{ con } h^* < \frac{1}{L \|A\|_\infty}$$

Para acabar con este capítulo, incluimos la siguiente propiedad de los métodos RK que nos servirá en el siguiente para simplificar su estudio:

Teorema 1.2.3. [1] *Si un RK(A, b) verifica las condiciones (1.2.12), da exactamente la misma aproximación cuando se aplica al PVI no autónomo*

$$\begin{aligned} y' &= f(t, y), & y, f &\in \mathbb{R}^m \\ y(t_0) &= y_0 \end{aligned} \tag{1.2.14}$$

que cuando se aplica el problema autónomo asociado:

$$\begin{aligned} z' &= g(z), & z, g &\in \mathbb{R}^{m+1} \\ z(t_0) &= z_0 \end{aligned} \tag{1.2.15}$$

donde

$$z = \begin{pmatrix} t \\ y \end{pmatrix}, \quad g(z) = \begin{pmatrix} 1 \\ f(t, y) \end{pmatrix}, \quad z_0 = \begin{pmatrix} t_0 \\ y_0 \end{pmatrix}$$

Como consecuencia, podemos estudiar los métodos RK solamente sobre problemas autónomos.

Orden de los métodos Runge-Kutta. Series de Butcher

Para estudiar si un método $RK(A, b)$ (1.2.10)-(1.2.11) da una buena aproximación a la solución del PVI (1.0.1) o no, tenemos que estudiar su error $\|y(t_n) - y_n\|$. Este estudio se hace indirectamente usando el llamado **error local**

$$l(t_0, h) = y(t_0 + h; t_0, y_0) - y_{RK}(t_0 + h; t_0, y_0)$$

donde $y(t_0 + h; t_0, y_0)$ denota la solución exacta del PVI local $y' = f(t, y)$, $y(t_0) = y_0$, $t \in [0, T]$ e $y_{RK}(t_0 + h; t_0, y_0)$ denota la aproximación dada por el método en $t_0 + h$ a dicho PVI que parte del mismo vector inicial y_0 .

En este capítulo seguiremos el planteamiento dado en [3] y [4], hemos utilizado una notación parecida a la dada en [1] pues nos parece más simple.

Se dice que el RK tiene **orden** p (u **orden de consistencia** p) si y sólo si

$$l(t_0, h) = \mathcal{O}(h^{p+1}), \quad h \rightarrow 0$$

Por el Teorema 1.2.3, será suficiente estudiar el error local sólo sobre problemas autónomos:

$$y' = f(y), \quad y(t_0) = y_0, \quad t \in [0, T], \quad (2.0.1)$$

donde $f : U \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$, $y_0 \in U$, con f verificando (1.0.2) y (1.0.3) y al método se le exigen las condiciones (1.2.12).

Mediante el desarrollo de Taylor de la solución exacta local $y(t_0 + h; t_0, y_0)$ tenemos

$$y(t_0 + h; t_0, y_0) = \sum_{j=0}^{\infty} \frac{y^{(j)}(t_0)}{j!} h^j = \sum_{j=0}^{\infty} C_j h^j, \quad C_j := \frac{y^{(j)}(t_0)}{j!}$$

Si consiguiésemos un desarrollo igual para la aproximación y_{RK} :

$$y_{RK}(t_0 + h; t_0, y_0) = \sum_{j=0}^{\infty} C_j^{RK} h^j$$

tendríamos una condición inmediata para calcular el orden del RK:

$$l(t_0, h) = \mathcal{O}(h^{p+1}) \iff C_j = C_j^{RK}, \quad 0 \leq j \leq p$$

Esta idea fue aplicada por J.C. Butcher (Nueva Zelanda, 1933) cuando construyó un nuevo tipo de desarrollo en serie que actualmente se llama **serie de Butcher** en su honor. Para introducir este desarrollo en serie primero tendremos que introducir varias notaciones y conceptos nuevos.

Cuando tratemos con funciones vectoriales de la forma $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$, denotaremos las derivadas parciales de cada componente de f como:

$$f_i^j := \frac{\partial f_i}{\partial x_j}, \quad f_i^{jk} := \frac{\partial^2 f_i}{\partial x_j \partial x_k}, \quad f_i^{jkl} := \frac{\partial^3 f_i}{\partial x_j \partial x_k \partial x_l}, \dots, f_i^{i_1, i_2, \dots, i_n} := \frac{\partial^n f_i}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_n}}$$

Si f admite derivadas parciales k -ésimas en y_0 , se define la **derivada de Frechet k -ésima en y_0** como la aplicación

$$f_{[y_0]}^{(k)} : \mathbb{R}^m \times \mathbb{R}^m \times \dots \times \mathbb{R}^m \rightarrow \mathbb{R}^m$$

$$(u^1, u^2, \dots, u^k) \mapsto f^{(k)}(u^1, u^2, \dots, u^k) = \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix} \in \mathbb{R}^m$$

donde

$$F_i = f_i^{(k)}(u^1, u^2, \dots, u^k) = \sum_{j_1, j_2, \dots, j_k=1}^m f_i^{j_1 j_2 \dots j_k}(y_0) u_{j_1}^1 u_{j_2}^2 \dots u_{j_k}^k, \quad 1 \leq i \leq m$$

Por convenio: $f_{[y_0]}^{(0)} = f(y_0)$. Y, por simplificar la exposición, supongamos por ahora que $f \in C^\infty(U)$. Veamos cómo se usan para calcular las derivadas de funciones de varias variables.

Proposición 2.0.1. *Si $g : \mathbb{R} \rightarrow \mathbb{R}^m$, $f : \mathbb{R}^m \rightarrow \mathbb{R}$ y $z := f \circ g$, entendiéndose que las derivadas de la función g y la función z son las derivadas ordinarias y las de la función f son las derivadas de Frechet evaluadas en $g(t)$, se tiene:*

$$z(t) = f(g(t)) = f(g_1(t), \dots, g_m(t)), \quad z'(t) = \sum_{j_1=1}^m f^{j_1}(g(t)) g'_{j_1}(t) = f'(g')$$

$$z''(t) = \sum_{j_1=1}^m \left(\sum_{j_2=1}^m (f^{j_1 j_2} g'_{j_2}(t) g'_{j_1}(t)) + f^{j_1} g''_{j_1}(t) \right) = f''(g', g') + f'(g'')$$

$$z'''(t) = \sum_{j_1=1}^m \left(\sum_{j_2=1}^m \left(\sum_{j_3=1}^m (f^{j_1 j_2 j_3} g'_{j_3}(t) g'_{j_2}(t) g'_{j_1}(t)) + f^{j_1 j_2} g'_{j_2} g''_{j_1} + f^{j_1 j_2} g''_{j_2}(t) g'_{j_1}(t) \right) \right. \\ \left. + f^{j_1 j_2} g'_{j_2}(t) g''_{j_1}(t) + f^{j_1} g'''_{j_1}(t) \right) = f'''(g', g', g') + f''(g', g'') + f''(g'', g') \\ + f''(g', g'') + f'(g''')$$

En consecuencia si y es la solución exacta de (2.0.1), en función de la derivada de Frechet de f en y_0 se tiene:

$$\begin{aligned} y'(t_0) &= f \\ y''(t_0) &= f'(f) \\ y'''(t_0) &= f''(f, f) + f'(f'(f)) \\ y''''(t_0) &= f'''(f, f, f) + f''(f'(f), f) + f''(f, f'(f)) + f''(f'(f), f) \\ &\quad + f'(f''(f, f)) + f'(f'(f'(f))) \end{aligned}$$

Para poder simplificar y generalizar estos desarrollos a cualquier derivada se usan un tipo de aplicaciones, denominadas **árboles**, que se asocian a cada derivada de Frechet.

Se llama **árbol ordenado monótonamente de orden q** a cualquier aplicación

$$\tau : \{2, 3, \dots, q\} \longrightarrow \{1, 2, \dots, q - 1\} \text{ con } \tau(i) < i, \forall i = 2, 3, \dots, q. \quad (2.0.2)$$

Por definición, $\tau_0 : 1 \longrightarrow 0$ es el único árbol de orden 1 y \emptyset es el único árbol de orden 0. Se suele denotar como $\rho(\tau) = q$ el orden del árbol τ , LT_q al conjunto de árboles ordenados monótonamente de orden q y $LT = \bigcup_{q=0}^{\infty} LT_q$

Estas aplicaciones se representan gráficamente mediante los grafos de la Tabla 2.1.

Si $\tau \in LT_q$, se define la **diferencial elemental de f asociada a τ** en y_0 , y se denota por $F(\tau)(y_0)$, al vector

$$F(\tau)(y_0) = \begin{pmatrix} F_1(\tau)(y_0) \\ \vdots \\ F_m(\tau)(y_0) \end{pmatrix},$$

donde cada $F_i(\tau)(y_0)$ es la suma de productos diferenciales de f asociadas al grafo de τ . Por definición se considera

$$F_i(\emptyset)(y_0) = y_0, \quad F_i(\tau_0) = f_i, \quad 1 \leq i \leq m$$

Teorema 2.0.2. *Existe una relación biunívoca entre todos y cada uno de los términos de la derivada q -ésima de $y(t)$ con los árboles ordenados monótonamente de orden q , es decir, $y^{(q)}(t_0) = \sum_{\tau \in LT_q} F(\tau)(y_0)$.*

Corolario 2.0.3. *Si $y(t)$ es analítica en y_0 se tiene el desarrollo en potencias de h : $y(t_0 + h) = \sum_{\tau \in LT} F(\tau)(y_0) \frac{h^{\rho(\tau)}}{\rho(\tau)!}$.*

Orden	Árbol	Grafo	Derivada
1	$\sigma(1) = 0$	$\bullet 1$	$y'_{j_1} = f_{j_1}$
2	$\sigma(2) = 1$	$\begin{array}{c} \bullet 2 \\ \\ \bullet 1 \end{array}$	$y''_{j_1} = \sum_{j_2} f_{j_1}^{j_2} f_{j_2}$
3	$\sigma(2) = \sigma(3) = 1$ $\sigma(2) = 1, \sigma(3) = 2$	$\begin{array}{c} \bullet 2 \quad \bullet 3 \\ \diagdown \quad / \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 3 \\ \\ \bullet 2 \\ \\ \bullet 1 \end{array}$	$y'''_{j_1} = \sum_{j_2 j_3} f_{j_1}^{j_2 j_3} f_{j_2} f_{j_3}$ $+ \sum_{j_2 j_3} f_{j_1}^{j_2} f_{j_2}^{j_3} f_{j_3}$
4	$\sigma(2) = \sigma(3) = \sigma(4) = 1$ $\sigma(2) = \sigma(3) = 1, \sigma(4) = 2$ $\sigma(2) = \sigma(3) = 1, \sigma(4) = 3$ $\sigma(3) = 2, \sigma(2) = \sigma(4) = 1$ $\sigma(2) = 1, \sigma(3) = \sigma(4) = 2$ $\sigma(2) = 1, \sigma(3) = 2, \sigma(4) = 3$	$\begin{array}{c} \bullet 2 \quad \bullet 3 \quad \bullet 4 \\ \diagdown \quad / \quad / \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 4 \\ \\ \bullet 2 \quad \bullet 3 \\ \diagdown \quad / \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 4 \\ \\ \bullet 3 \quad \bullet 2 \\ \diagdown \quad / \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 3 \\ \\ \bullet 2 \quad \bullet 4 \\ \diagdown \quad / \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 3 \quad \bullet 4 \\ \diagdown \quad / \\ \bullet 2 \\ \\ \bullet 1 \end{array}$ $\begin{array}{c} \bullet 4 \\ \\ \bullet 3 \\ \\ \bullet 2 \\ \\ \bullet 1 \end{array}$	$y''''_{j_1} = \sum_{j_2 j_3 j_4} f_{j_1}^{j_2 j_3 j_4} f_{j_2} f_{j_3} f_{j_4}$ $+ \sum_{j_2 j_3 j_4} f_{j_1}^{j_2 j_3} f_{j_2}^{j_4} f_{j_3} f_{j_4}$ $+ \sum_{j_2 j_3 j_4} f_{j_1}^{j_3 j_2} f_{j_3}^{j_4} f_{j_2} f_{j_4}$ $+ \sum_{j_2 j_3 j_4} f_{j_1}^{j_2 j_4} f_{j_2}^{j_3} f_{j_3} f_{j_4}$ $+ \sum_{j_2 j_3 j_4} f_{j_1}^{j_2} f_{j_2}^{j_3 j_4} f_{j_3} f_{j_4}$ $+ \sum_{j_2 j_3 j_4} f_{j_1}^{j_2} f_{j_2}^{j_3} f_{j_3}^{j_4} f_{j_4}$

Tabla 2.1. Árboles ordenados asociados a las derivadas sucesivas

Se dice que dos árboles $\tau_1, \tau_2 \in LT$ son **árboles equivalentes**, y se denota por $\tau_1 \sim \tau_2$, si y sólo si son del mismo orden q y existe una permutación σ del conjunto $\{1, 2, \dots, q\}$ con $\sigma(1) = 1$ tal que

$$\sigma \circ \tau_1 = \tau_2 \circ \sigma \quad \text{sobre } \{2, 3, \dots, q\}$$

Teorema 2.0.4. *La relación " \sim " definida anteriormente es una relación de equivalencia en LT , es decir, verifica*

- (a) $\tau \sim \tau$, (b) $\tau_1 \sim \tau_2 \Rightarrow \tau_2 \sim \tau_1$, (c) $\tau_1 \sim \tau_2$ y $\tau_2 \sim \tau_3 \Rightarrow \tau_1 \sim \tau_3$.

cuyo orden es $\rho(\tau) = \rho(\tau_1) + \rho(\tau_2) + \cdots + \rho(\tau_k) + 1$. A los árboles $\tau_1, \tau_2, \dots, \tau_k$ se les llama **árboles hijos** de τ .

Teorema 2.0.6. Si $\tau = \{\tau_1, \tau_2, \dots, \tau_k\}$, entonces

$$F(\tau)(y_0) = f_0^{(k)}(F(\tau_1)(y_0), \dots, F(\tau_k)(y_0)) = \left(\sum_{j_1, \dots, j_k=1}^m f_i^{j_1 \cdots j_k} F_{j_1}(\tau_1)(y_0), \dots, F_{j_k}(\tau_k)(y_0) \right)_{i=1}^m$$

2.1. Series de Butcher. Condiciones de orden de los métodos RK

Sea a una aplicación $a : LT \rightarrow \mathbb{R}$ que verifica $a(\tau_1) = a(\tau_2)$ si $\tau_1 \sim \tau_2$. Se define la **serie de Butcher** o **B-serie** asociada a "a" en el punto y_0 como la serie formal en potencias de h :

$$B(a, y_0)(h) := \sum_{\tau \in LT} a(\tau) F(\tau)(y_0) \frac{h^{\rho(\tau)}}{\rho(\tau)!} = \sum_{\tau \in T} \alpha(\tau) a(\tau) F(\tau)(y_0) \frac{h^{\rho(\tau)}}{\rho(\tau)!}$$

Dada $g : I_0 \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ de clase $C^\infty(I_0)$ siendo I_0 un entorno del origen, se dice que g es **representable en serie de Butcher** si y sólo si existe una aplicación $a : T \rightarrow \mathbb{R}$ tal que

$$g^{(q)}(0) = \sum_{\tau \in T_q} \alpha(\tau) a(\tau) F(\tau)(y_0), \quad q = 1, 2, \dots$$

En tal caso, se tiene

$$g(h) = B(a, y_0)(h), \quad h \in I_0$$

Del Corolario 2.0.5 es evidente que la solución exacta del PVI local autónomo (2.0.1) es representable en serie de Butcher

$$y(t_0 + h) = B(a, y_0)(h), \quad \text{donde } a(\tau) = 1, \forall \tau \in T \quad (2.1.3)$$

Lo interesante es que la solución numérica de los métodos RK también se puede representar en serie de Butcher. Antes de escribir dicha serie necesitamos introducir un nuevo producto vectorial.

Dados dos vectores cualesquiera $u, v \in \mathbb{R}^s$, se define el **producto directo** de u y v como el vector $u \cdot v = (u_i v_i)_{i=1}^s$. Para distinguirlo del producto ordinario de matrices, escribiremos siempre el símbolo \cdot , mientras que para el producto ordinario no lo denotaremos.

Lema 2.1.7. Dado un RK(A, b) (1.2.10)-(1.2.11) aplicado al PVI autónomo (2.0.1), se tiene que la solución numérica en $t_0 + h$ es representable mediante la siguiente serie de Butcher:

$$y_{RK}(t_0 + h) = B(\omega, y_0)(h) \quad (2.1.4)$$

donde

$$\begin{aligned} \omega(\emptyset) &= 1, \\ \omega(\tau) &= \gamma(\tau)b^T\Phi(\tau), \quad \forall \tau \text{ con } \rho(\tau) \geq 1, \end{aligned} \quad (2.1.5)$$

siendo $\gamma(\tau)$ la aplicación $\gamma : T \rightarrow \mathbb{R}$ ("densidad de τ ") definida como

$$\begin{aligned} \gamma(\emptyset) &= \gamma(\tau_0) = 1 \\ \gamma(\tau) &= \rho(\tau)\gamma(\tau_1)\gamma(\tau_2)\cdots\gamma(\tau_k), \quad \text{si } \tau = \{\tau_1, \tau_2, \dots, \tau_k\} \end{aligned} \quad (2.1.6)$$

y $\Phi(\tau)$ se calcula recursivamente por

$$\begin{aligned} \Phi(\emptyset) &= 0, \quad \Phi(\tau_0) = e \\ \Phi(\tau) &= (A\Phi(\tau_1)) \cdot (A\Phi(\tau_2)) \cdots (A\Phi(\tau_k)), \quad \text{si } \tau = \{\tau_1, \tau_2, \dots, \tau_k\} \end{aligned} \quad (2.1.7)$$

En consecuencia, comparando la series de Butcher (2.1.3) y (2.1.4) se obtiene directamente las condiciones para que el error local de un RK sea de un orden p determinado:

Teorema 2.1.8 (Condiciones de orden de un método RK). *Un método RK(A, b) tiene orden $p \geq 1$ si y sólo si*

$$b^T\Phi(\tau) = \frac{1}{\gamma(\tau)}, \quad \forall \tau \text{ con } \rho(\tau) \leq p.$$

Además, también obtenemos una expresión del error local:

Teorema 2.1.9 (Error local de un método RK). *Un método RK(A, b) de orden p tiene un error local dado por*

$$l(t_0, h) = \sum_{\tau \in T, \rho(\tau) \geq p+1} \alpha(\tau)(1 - \gamma(\tau)b^T\Phi(\tau))F(\tau)(y_0) \frac{h^{\rho(\tau)}}{\rho(\tau)!}$$

Para simplificar más las condiciones de orden, se puede usar la aplicación densidad dada en (2.1.6) para calcular el cardinal de cada árbol.

Lema 2.1.10. *Para cualquier $\tau \in T$ con $\rho(\tau) = 1$ se tiene*

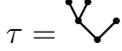
$$\alpha(\tau) = \frac{\rho(\tau)!}{\gamma(\tau)\sigma(\tau)}, \quad (2.1.8)$$

donde γ es la aplicación dada en (2.1.6) y $\sigma : T \rightarrow \mathbb{R}$ se define como

$$\begin{aligned} \sigma(\tau_0) &= 1 \\ \sigma(\tau) &= n_1!n_2!\cdots n_k!\sigma(\tau_1)^{n_1}\cdots\sigma(\tau_k)^{n_k}, \quad \text{si } \tau = \{\tau_1^{n_1}, \tau_2^{n_2}, \dots, \tau_k^{n_k}\}, \end{aligned}$$

siendo $\{\tau_1, \tau_2, \dots, \tau_k\}$ árboles distintos entre sí.

Ejemplo 2.1.11. Veamos como son las funciones γ , Φ y ω y el cardinal del árbol



$$\gamma(\tau) = \rho(\tau)\gamma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \gamma(\bullet) = 5 \cdot 3 \cdot 1 = 15$$

pues $\gamma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) = \rho\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \gamma(\bullet) \gamma(\bullet) = 3 \cdot 1 \cdot 1 = 3$.

Por otra parte $\Phi(\tau) = \left(A\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right)\right) \cdot (A\Phi(\bullet)) = (Ac^2) \cdot (Ae) = Ac^2 \cdot c$, ya que $\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) = (A\Phi(\bullet)) \cdot (A\Phi(\bullet)) = (Ae) \cdot (Ae) = c^2$.

Por lo tanto $\omega(\tau) = 15b^T(Ac^2 \cdot c)$. Además, $\sigma(\tau) = \sigma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \sigma(\bullet) = 2!\sigma(\bullet)^2 \sigma(\bullet) = 2$ por lo que el cardinal es $\alpha(\tau) = \frac{5!}{15 \cdot 2} = 4$.

En el caso del árbol $\tau =$  tenemos que:

$$\gamma(\tau) = \rho(\tau)\gamma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \gamma(\bullet) = 6 \cdot 8 \cdot 1 = 48$$

porque $\gamma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) = \rho\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \gamma(\mathbb{1}) \gamma(\bullet) = 4 \cdot 2 \cdot 1 = 8$ con $\gamma(\mathbb{1}) = \rho(\mathbb{1}) \gamma(\bullet) = 2$.

Calculamos a continuación $\Phi(\tau) = \left(A\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right)\right) \cdot (A\Phi(\bullet)) = (A(Ac \cdot c)) \cdot (Ae) = (A(Ac \cdot c)) \cdot c$, ya que $\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) = \left(A\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right)\right) \cdot (A\Phi(\bullet)) = (Ac) \cdot (Ae) = Ac \cdot c$, $\Phi\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) = A\Phi(\bullet) = Ae = c$. Por tanto $\omega(\tau) = 48b^T(A(Ac \cdot c) \cdot c)$.

Y por último calculamos el cardinal de τ $\sigma(\tau) = \sigma\left(\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}\right) \sigma(\bullet) = \sigma(\mathbb{1}) \sigma(\bullet) \sigma(\bullet) = 1 \Rightarrow \alpha(\tau) = \frac{6!}{48 \cdot 1} = 15$.

Además por el Lema 2.1.7, el Teorema 2.1.8 y el Lema 2.1.10 concluimos, por ejemplo, que para que un método RK tenga orden 3 debe cumplir las siguientes condiciones de orden:

Orden	τ	$\alpha(\tau)$	$\gamma(\tau)$	$\Phi(\tau)$	$F(\tau)(y)$	Condición de orden
1	\bullet	1	1	e	f	$b^T e = 1$
2	$\begin{array}{c} \bullet \\ \downarrow \\ \bullet \end{array}$	1	2	c	$f'(f)$	$b^T c = \frac{1}{2}$
3	$\begin{array}{c} \bullet \\ \swarrow \searrow \\ \bullet \quad \bullet \end{array}$	1	3	c^2	$f''(f, f)$	$b^T c^2 = \frac{1}{3}$
	$\begin{array}{c} \bullet \\ \downarrow \\ \bullet \\ \downarrow \\ \bullet \end{array}$	1	6	Ac	$f'(f'(f))$	$b^T Ac = \frac{1}{6}$

Tabla 2.2. Condiciones de orden

Convergencia de métodos Runge-Kutta

En el capítulo anterior hemos visto una forma sistemática de deducir el orden de consistencia de un método RK sobre PVI's autónomos. Sin embargo, el objetivo del estudio de la convergencia de los métodos es analizar su error global para PVI's en general.

Más precisamente, se considera el PVI

$$y' = f(t, y), \quad y(t_0) = y_0, \quad t \in [t_0, t_f], \quad y, f \in \mathbb{R}^m \quad (3.0.1)$$

con solución única $y(t)$ en $[t_0, t_f]$, suponiendo que f , además de ser continua y Lipschitz respecto de y , se tiene para cierto $\delta > 0$, $f \in C^p(\mathcal{T}_\delta)$, donde

$$\mathcal{T}_\delta = \{(t, y) : \|y - y(t)\| \leq \delta\}, \quad \delta > 0$$

es un tubo de amplitud δ alrededor de la solución exacta $y(t)$.

Para aproximar dicha solución tomamos una partición cualquiera de $[t_0, t_f]$:

$$\mathcal{P} = \{t_0 < t_1 < \dots < t_N = t_f\}, \quad h_j = t_{j+1} - t_j, \quad j = 0, 1, \dots, N-1$$

y aplicamos un método RK(A, b) de s etapas

$$\begin{aligned} K_i &= f(t_n + c_i h_n, y_n + h_n \sum_{j=1}^s a_{ij} K_j) \quad 1 \leq i \leq s, \\ y_{n+1} &= y_n + h_n \sum_{i=1}^s b_i K_i, \quad n = 0, 1, \dots, N-1 \end{aligned} \quad (3.0.2)$$

Se llaman **errores globales** del RK(A, b) sobre la partición \mathcal{P} a

$$e_n := \|y(t_n) - y_n\|, \quad n = 0, 1, \dots, N$$

Queremos demostrar rigurosamente que estos errores tienden a 0 para toda partición \mathcal{P} tal que $h_{\max} = \max\{h_n\}_{n=0}^{N-1} \rightarrow 0$. Para ello será necesario acotar el error local de los métodos. Hay que tener en cuenta que las condiciones de orden vistas en el capítulo anterior se basan en las series de Butcher, cuya convergencia

no se probó. Así que en este capítulo obtendremos la acotación rigurosa del error local para los métodos que cumplen las condiciones de orden dadas en el Teorema 2.1.8. En particular, demostraremos detalladamente los teoremas de acotación del error local y del error global enunciados en [3].

3.1. Acotación del error local

Teorema 3.1.1. *Si el $RK(A, b)$ (3.0.2) es de orden p y $f \in C^p(\mathcal{T}_\delta)$, existe $h^* > 0$ tal que para todo $|h| \leq h^*$ se tiene¹*

$$\|y(t+h) - y_{RK}(t+h; t, y(t))\| \leq \frac{h^{p+1}}{(p+1)!} \left[\max_{s \in [0,1]} \|y^{(p+1)}(t+sh)\| + (p+1) \sum_{i=1}^s |b_i| \max_{s \in [0,1]} \|K_i^{(p)}(sh)\| \right] \quad (3.1.3)$$

donde

$$K_i^{(p)} = \frac{\partial^p K_i}{\partial h^p}(t, h)$$

Demostración. Como $f \in C^p(\mathcal{T}_\delta)$, la solución exacta $y(t)$ de PVI (3.0.1) es derivable con continuidad hasta orden $p+1$, por lo que aplicando el desarrollo de Taylor con resto integral hasta orden p de $y(t)$

$$y(t+h) = \sum_{j=0}^p \frac{h^j}{j!} y^{(j)}(t) + \int_t^{t+h} \frac{(t+h-\xi)^p}{p!} y^{(p+1)}(\xi) d\xi$$

Haciendo el cambio de variable $\xi = t + \theta h$, $\theta \in [0, 1]$, $d\xi = h d\theta$, $t+h-\xi = (1-\theta)h$, obtenemos que

$$y(t+h) = \sum_{j=0}^p \frac{h^j}{j!} y^{(j)}(t) + \frac{h^{p+1}}{p!} \int_0^1 (1-\theta)^p y^{(p+1)}(t+\theta h) d\theta \quad (3.1.4)$$

Sabemos por el Teorema 1.2.2 que

$$y_{RK}(t+h; t, y(t)) = y(t) + h \sum_{i=1}^s b_i K_i(h) \quad \text{con } h \rightarrow 0, \quad (t \text{ fijo})$$

donde $K_i(h) = f(t+c_i h, y(t)) + h \sum_{j=1}^s a_{ij} K_j(h)$, $1 \leq i \leq s$ y $K_i(h) \in C^p([0, h^*])$.

Por lo tanto aplicando Taylor con resto integral a $K_i(h)$ hasta orden $p-1$ obtenemos:

¹ Podemos considerar cualquier norma $\|\cdot\|$ en \mathbb{R}^m , aunque habitualmente se usa la norma euclídea o la uniforme

$$K_i(h) = \sum_{j=0}^{p-1} \frac{h^j}{j!} K_i^{(j)}(h) + \int_0^h K_i^{(p)}(\xi) \frac{(h-\xi)^{p-1}}{(p-1)!} d\xi.$$

Realizando el cambio de variable $\xi = \theta h$, $\theta \in [0, 1]$ obtenemos que

$$K_i(h) = \sum_{j=0}^{p-1} \frac{h^j}{j!} K_i^{(j)}(0) + \frac{h^{p+1}}{(p-1)!} \int_0^1 K_i^{(p)}(\theta h) (1-\theta)^{p-1} d\theta,$$

por lo que

$$\begin{aligned} y_{RK}(t+h; t, y(t)) &= y(t) + h \sum_{i=1}^s b_i \sum_{j=0}^{p-1} \frac{h^j}{j!} K_i^{(j)}(0) \\ &\quad + \frac{h^{p+1}}{(p-1)!} \sum_{i=1}^s b_i \int_0^1 (1-\theta)^{p-1} K_i^{(p)}(\theta h) d\theta \end{aligned} \quad (3.1.5)$$

Por ser un método de orden p , $y(t+h) - y_{RK}(t+h; t, y(t)) = \mathcal{O}(h^{p+1})$, por lo que al restar (3.1.4)-(3.1.5) se eliminan los términos hasta orden h^p obteniéndose así

$$\begin{aligned} y(t+h) - y_{RK}(t+h; t, y(t)) &= \frac{h^{p+1}}{p!} \int_0^1 (1-\theta)^p y^{(p+1)}(t+\theta h) d\theta \\ &\quad - \frac{h^{p+1}}{(p-1)!} \sum_{i=1}^s b_i \int_0^1 (1-\theta)^{p-1} K_i^{(p)}(\theta h) d\theta \end{aligned} \quad (3.1.6)$$

Tomando normas

$$\begin{aligned} \| y(t+h) - y_{RK}(t+h; t, y(t)) \| &\leq \frac{h^{p+1}}{(p-1)!} \left\{ \frac{1}{p} \int_0^1 (1-\theta)^p \| y^{(p+1)}(t+\theta h) \| d\theta \right. \\ &\quad \left. + \sum_{i=1}^s |b_i| \int_0^1 (1-\theta)^{p-1} \| K_i^{(p)}(\theta h) \| d\theta \right\} \\ &\leq \frac{h^{p+1}}{(p-1)!} \left\{ \frac{1}{p} \int_0^1 (1-\theta)^p d\theta \max_{\theta \in [0,1]} \| y^{(p+1)}(t+\theta h) \| \right. \\ &\quad \left. + \sum_{i=1}^s |b_i| \int_0^1 (1-\theta)^{p-1} d\theta \max_{\theta \in [0,1]} \| K_i^{(p)}(\theta h) \| \right\} \end{aligned}$$

Como $\int_0^1 (1-\theta)^p d\theta = \frac{1}{p+1}$ y $\int_0^1 (1-\theta)^{p-1} d\theta = \frac{1}{p}$ se tiene

$$\begin{aligned} \| y(t+h) - y_{RK}(t+h; t, y(t)) \| &\leq \frac{h^{p+1}}{(p+1)!} \left[\max_{\theta \in [0,1]} \| y^{(p+1)}(t+\theta h) \| \right. \\ &\quad \left. + (p+1) \sum_{i=1}^s |b_i| \max_{\theta \in [0,1]} \| K_i^{(p)}(\theta h) \| \right] \end{aligned}$$

□

Para acotar el error global no será suficiente con acotar el error local. Necesitaremos, además, demostrar la siguiente propiedad de los métodos RK dada en el Lema 3.1.2, que demostraremos seguidamente. Para ello, expresamos los métodos RK en la **notación de Henrici** de la siguiente manera:

$$y_{n+1} = y_n + h\Phi(t_n, y_n, h) \quad (3.1.7)$$

donde

$$\Phi(t_n, y_n, h) = \sum_{i=1}^s b_i K_i \quad (3.1.8)$$

se llama **función incremento**, entendiéndose que $K_i = K_i(t_n, y_n, h)$ son las etapas del método dadas en (1.2.10).

Lema 3.1.2. *La función incremento $\Phi(t, y, h)$ tiene una constante de Lipschitz Λ respecto de y en $[0, T] \times \mathbb{R}^m \times [0, h_0]$ dada por*

$$\Lambda = L \|b\|^T (I - h_0 L \|A\|)^{-1} e, \quad (3.1.9)$$

donde L es la constante de Lipschitz de f ,

$$\|b\|^T = (\|b_1\|, \dots, \|b_s\|), \quad \|A\| = (\|a_{ij}\|)_{i,j=1}^s, \quad h_0 < \frac{1}{L\rho(\|A\|)},$$

y $\rho(\|A\|)$ denota el radio espectral de la matriz $\|A\|$.

Para demostrar este Lema, previamente necesitamos el siguiente resultado:

Lema 3.1.3. *Dados dos vectores $u, v \in \mathbb{R}^n$, se define la desigualdad ' \leq ' **componente a componente** como*

$$u \leq v \text{ componente a componente} \iff u_i \leq v_i, \forall i = 1, \dots, n.$$

Esta desigualdad verifica las siguientes propiedades:

- Sean $u, v, w \in \mathbb{R}^n$ cualesquiera. Si $u \leq v + w$ componente a componente, entonces $u - v \leq w$ componente a componente.
- Sean $u, v \in \mathbb{R}^n$, $A \in \mathcal{M}_{n \times n}$ inversible. Si $Au \leq v$ componente a componente y A^{-1} tiene todos sus elementos ≥ 0 , entonces $u \leq A^{-1}v$ componente a componente

Demostración del Lema 3.1.3. La demostración del primer apartado de este lema es inmediata. Para el segundo, basta observar que todos los elementos de $v - Au$ son no negativos, por lo que $A^{-1}(v - Au) \geq 0$. □

Demostración del Lema 3.1.2. Consideramos $t \in [0, T]$, $y, \bar{y} \in \mathbb{R}^m$ y $h \in [0, h_0]$, y denotamos para simplificar

$$\begin{aligned}\Phi &= \sum_{i=1}^s b_i K_i, & K_i &= f \left(t + c_i h, y + h \sum_{i=1}^s a_{ij} K_j \right) \\ \bar{\Phi} &= \sum_{i=1}^s b_i \bar{K}_i, & \bar{K}_i &= f \left(t + c_i h, \bar{y} + h \sum_{i=1}^s a_{ij} \bar{K}_j \right)\end{aligned}$$

Aplicando normas

$$\| \Phi - \bar{\Phi} \| \leq \sum_{i=1}^s |b_i| \| \Delta K_i \|, \quad \Delta K_i := K_i - \bar{K}_i$$

Por la condición de Lipschitz de f ,

$$\begin{aligned}\| \Delta K_i \| &\leq \left\| f \left(t + c_i h, y + h \sum_{i=1}^s a_{ij} K_j \right) - f \left(t + c_i h, \bar{y} + h \sum_{i=1}^s a_{ij} \bar{K}_j \right) \right\| \leq \\ &\leq L \| \Delta y \| + L h_0 \sum_{j=1}^s |a_{ij}| \| \Delta K_j \|, \quad \forall i = 1, \dots, s, \quad \Delta y = y - \bar{y}\end{aligned}$$

Esta desigualdad se puede reescribir definiendo

$$V = \begin{pmatrix} \| \Delta K_1 \| \\ \vdots \\ \| \Delta K_s \| \end{pmatrix},$$

como

$$\begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_s \end{pmatrix} \leq \begin{pmatrix} L \| \Delta y \| + L h_0 (|a_{11}| V_1 + |a_{12}| V_2 + \dots + |a_{1s}| V_s) \\ L \| \Delta y \| + L h_0 (|a_{21}| V_1 + |a_{22}| V_2 + \dots + |a_{2s}| V_s) \\ \vdots \\ L \| \Delta y \| + L h_0 (|a_{s1}| V_1 + |a_{s2}| V_2 + \dots + |a_{ss}| V_s) \end{pmatrix},$$

es decir, la desigualdad componente a componente,

$$V \leq L \| \Delta y \| e + L h_0 |A| V \quad \text{donde} \quad e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Aplicando el primer ítem del Lema 3.1.3,

$$(I - Lh_0 | A |)V \leq L \| \Delta y \| e \quad \text{componente a componente}$$

Por las propiedades elementales de las matrices convergentes, existe $(I - Lh_0 | A |)^{-1}$ si y sólo si $\rho(Lh_0 | A |) < 1$, lo que tenemos por la hipótesis $h_0 < \frac{1}{L\rho(| A |)}$. Además, se tiene

$$(I - Lh_0 | A |)^{-1} = \sum_{r=0}^{\infty} L^r h_0^r | A |^r$$

por lo que tiene todos sus elementos no negativos. Por tanto, aplicando el segundo ítem del Lema 3.1.3

$$V \leq L \| \Delta y \| (I - Lh_0 | A |)^{-1} e$$

Finalmente, como estamos tratando con vectores con elementos no negativos, llegamos a

$$\| \Phi - \bar{\Phi} \| \leq \| b |^T V \leq (L | b |^T (I - Lh_0 | A |)^{-1} e) \| \Delta y \| \leq \Lambda \| \Delta \|,$$

donde $\Lambda = L | b |^T (I - Lh_0 | A |)^{-1} e$, obteniendo así la tesis del Lema. \square

3.2. Acotación del error global

Teorema 3.2.4. *Si un RK(A, b) (3.0.2) es de orden p y $f \in C^p(\mathcal{T}_\delta)$, y consideramos una partición cualquiera*

$$\mathcal{P} = \{t_0 < t_1 < t_2 < \dots < t_N = t_f\}$$

de $[t_0, t_f]$

$$h_{\max} := \max_j h_j \leq \bar{h} = \min\{h^*, h'\}$$

donde h^* es el valor dado por el Teorema 3.1.1 y

$$h' < \frac{1}{L\rho(| A |)}, \quad \text{siendo } L \text{ la constante de Lipschitz de } f,$$

entonces se tiene que

$$\| y(t_n) - y_n \| \leq h_{\max}^p \frac{C}{\Lambda} [e^{A(t_n - t_0)} - 1], \quad n = 0, 1, \dots, N,$$

donde

$$C = \frac{1}{(p+1)!} \left\{ \begin{array}{l} \text{máx}_{t \in [t_0, t_f]} \| y^{(p+1)}(t) \| + (p+1) \sum_{i=1}^s |b_i| \text{máx}_{\substack{t \in [t_0, t_f] \\ h \in [0, h^*]}} \| K_i^{(p)}(t, h) \| \end{array} \right\},$$

$$K_i(t, h) = f(t + c_i h, y(t) + h \sum_{j=1}^s a_{ij} K_j(t, h)), \quad 1 \leq i \leq s,$$

siendo Λ la constante de Lipschitz (3.1.9) de la función incremento Φ del RK (3.0.2).

Demostración. Denotamos para $n = 0, 1, \dots, N$,

$$e_n := \| y(t_n) - y_n \|, \quad \bar{y}_n := y_{RK}(t_n; t_{n-1}, y(t_{n-1}))$$

Sumando y restando la solución del RK en e_n obtenemos para $n \geq 1$ que

$$e_n = \| y(t_n) - \bar{y}_n + \bar{y}_n - y_n \| \leq l_n + \| \bar{y}_n - y_n \|^2$$

ya que el error local en t_n es $l_n = \| y(t_n) - \bar{y}_n \|$. Veamos que ocurre con el segundo sumando. Sabemos que

$$\begin{aligned} \bar{y} &= y_{RK}(t_n; t_{n-1}, y(t_{n-1})) = y(t_{n-1}) + h_{n-1} \Phi(t_{n-1}, y(t_{n-1}), h_{n-1}) \\ y &= y_{RK}(t_n; t_{n-1}, y_{n-1}) = y_{n-1} + h_{n-1} \Phi(t_{n-1}, y_{n-1}, h_{n-1}) \end{aligned}$$

luego, por el lema 3.1.2

$$\begin{aligned} \| \bar{y}_n - y_n \| &= \| y(t_{n-1}) - y_{n-1} + h_{n-1} (\Phi(t_{n-1}, y(t_{n-1}), h_{n-1}) - \Phi(t_{n-1}, y_{n-1}, h_{n-1})) \| \\ &\leq \| y(t_{n-1}) - y_{n-1} \| + h_{n-1} \Lambda \| y(t_{n-1}) - y_{n-1} \| \leq (1 + h_{n-1} \Lambda) e_{n-1} \end{aligned}$$

Por lo tanto

$$\begin{cases} e_n \leq (1 + h_{n-1} \Lambda) e_{n-1} + l_n, & n \geq 1 \\ e_0 = 0 \end{cases}$$

Si $x \simeq 0$ entonces aplicando Taylor tenemos que $e^x = 1 + x + \frac{x^2}{2!} + \dots$ por lo que $1 + x \leq e^x$. En nuestro caso resulta que $1 + h_{n-1} \Lambda \leq e^{h_{n-1} \Lambda}$, luego

$$\begin{cases} e_n \leq e^{h_{n-1} \Lambda} e_{n-1} + l_n, & n \geq 1 \\ e_0 = 0 \end{cases}$$

Realizando iteraciones en el proceso

$$\begin{aligned} e_n &\leq e^{h_{n-1} \Lambda} e_{n-1} + l_n \leq e^{h_{n-1} \Lambda} (e^{h_{n-2} \Lambda} e_{n-2} + l_{n-1}) + l_n \leq e^{A(h_{n-1} + h_{n-2})} e_{n-2} \\ &\quad + e^{A h_{n-1}} l_{n-1} + l_n \leq \left[\sum_{j=1}^{n-1} e^{A(h_{n-1} + h_{n-2} + \dots + h_j)} l_j \right] + l_n \Rightarrow \\ e_n &\leq \left[\sum_{j=1}^{n-1} e^{A(t_n - t_j)} l_j \right] + l_n \leq e^{A t_n} \left[\sum_{j=1}^{n-1} e^{-A t_j} l_j + e^{-A t_n} l_n \right] = e^{A t_n} \sum_{j=1}^n e^{-A t_j} l_j \\ &= e^{A t_n} \sum_{j=0}^{n-1} e^{-A t_{j+1}} l_{j+1} \end{aligned}$$

Por el teorema 3.1.1 $l_{j+1} \leq Ch_j^{p+1}$. Además, por hipótesis $h_j^{p+1} = h_j^p h_j \leq h_{\max}^p h_j$ de modo que

$$e_n \leq Ce^{At_n} \sum_{j=0}^{n-1} e^{-At_{j+1}} h_j^{p+1} \leq Ce^{At_n} h_{\max}^p \left[\sum_{j=0}^{n-1} e^{-At_{j+1}} h_j \right]$$

La suma $\sum_{j=0}^{n-1} e^{-At_{j+1}} h_j$ es una suma inferior de Riemann de e^{-At} en $[t_0, t_n]$, por lo que la acotamos por su integral, completando la demostración:

$$e_n \leq Ce^{At_n} h_{\max}^p \int_{t_0}^{t_n} e^{-At} dt = Ce^{At_n} h_{\max}^p \left[-\frac{e^{-At_n} - e^{-At_0}}{\Lambda} \right] = h_{\max}^p \frac{C}{\Lambda} [e^{\Lambda(t_n - t_0)} - 1]$$

□

Métodos Runge-Kutta-Nyström

En muchas aplicaciones hay que resolver sistemas de segundo orden

$$y'' = f(t, y, y') \quad y(t_0) = y_0, \quad y'(t_0) = y'_0 \quad y, f \in \mathbb{R}^m, \quad t \in [t_0, T] \quad (4.0.1)$$

Para aproximar numéricamente su solución se puede transformar en una ecuación diferencial de primer orden de dimensión $2m$, considerando el vector $z = (y, y')^T$

$$z' = \begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} y' \\ f(t, y, y') \end{pmatrix} = g(t, z) \quad z(t_0) = z_0 = \begin{pmatrix} y_0 \\ y'_0 \end{pmatrix} \quad (4.0.2)$$

Como hemos visto en el capítulo anterior, si resolvemos el problema (4.0.2) mediante un método RK(A, b), denotando $K_i = \begin{pmatrix} k_i \\ k'_i \end{pmatrix}$ y $z_1 = \begin{pmatrix} y_1 \\ y'_1 \end{pmatrix}$,

$$K_i = g\left(t_0 + c_i h, z_0 + h \sum_{j=1}^s a_{ij} K_j\right) = \begin{pmatrix} y'_0 + h \sum_{j=1}^s a_{ij} k'_j \\ f\left(t_0 + c_i h, y_0 + h \sum_{j=1}^s a_{ij} k_j, y'_0 + h \sum_{j=1}^s a_{ij} k'_j\right) \end{pmatrix}$$

$$z_1 = z_0 + h \sum_{i=1}^s b_i K_i = \begin{pmatrix} y_0 + h \sum_{i=1}^s b_i k_i \\ y'_0 + h \sum_{i=1}^s b_i k'_i \end{pmatrix} \quad (4.0.3)$$

pues $z_0 + h \sum_{j=1}^s a_{ij} K_j = \begin{pmatrix} y_0 \\ y'_0 \end{pmatrix} + h \sum_{j=1}^s a_{ij} \begin{pmatrix} k_j \\ k'_j \end{pmatrix}$.

Podemos simplificar estas fórmulas ya que tenemos que $k_j = y'_0 + h \sum_{l=1}^s a_{jl} k'_l$, por tanto,

Definición 4.1.2. Un método RKN (4.0.5) tiene orden p si para el problema (4.0.1) suficientemente suave se cumple que

$$y(t_0 + h) - y_1 = \mathcal{O}(h^{p+1}), \quad y'(t_0 + h) - y'_1 = \mathcal{O}(h^{p+1}) \quad (4.1.7)$$

Si estos métodos se aplican a PVIIs de la forma

$$y'' = f(t, y) \quad (4.1.8)$$

se simplifica obviado la matriz A por lo que la tabla de Butcher quedará de la forma

$$\begin{array}{c|c} c & \bar{A} \\ \hline & \bar{b} \\ \hline & b \end{array}$$

Ejemplo 4.1.3. El método RKN del Ejemplo 4.1.1 sobre el problema (4.1.8) se reduce a

$$\begin{array}{c|c} 0 & \\ \frac{1}{2} & \frac{1}{8} \\ 1 & 0 \quad \frac{1}{2} \\ \hline & \frac{1}{6} \quad \frac{1}{3} \quad 0 \\ \hline & \frac{1}{6} \quad \frac{4}{6} \quad \frac{1}{6} \end{array}$$

Por otra parte un método RKN, usando la notación vectorial (4.0.2)-(4.0.3) se puede escribir en notación de Henrici (3.1.7), como los RK, de la siguiente forma

$$z_1 = z_0 + h\Phi(t_0, z_0, h) \quad (4.1.9)$$

donde

$$\Phi(t_0, z_0, h) = \begin{pmatrix} y'_0 + h \sum_i^s \bar{b}_i k'_i \\ \sum_i^s b_i k'_i \end{pmatrix}$$

Por lo tanto, los RKN son un caso particular de un método de un paso.

Si tenemos que el método es de orden p , entonces

$$z(t_0 + h) - z_1 = \begin{pmatrix} y(t_0 + h) - y_1 \\ y'(t_0 + h) - y'_1 \end{pmatrix} = \begin{pmatrix} \mathcal{O}(h^{p+1}) \\ \mathcal{O}(h^{p+1}) \end{pmatrix}$$

por lo tanto $z(t_0 + h) - z_1 = \mathcal{O}(h^{p+1})$. En consecuencia, podríamos aplicar el Teorema 3.2.4 para probar su convergencia. Para ello es necesario pedir que $g(t, z)$ en (4.0.2) sea continua y Lipschitz respecto de z , es decir

$$\|g(t, z) - g(t, \bar{z})\| \leq L\|z - \bar{z}\|, \quad \forall z, \bar{z} \in \mathbb{R}^{2m}, \quad \forall t \in [0, T] \quad (4.1.10)$$

Consideramos la norma $\|z\|_\infty = \max\{\|z_1\|, \|z_2\|\}$ en \mathbb{R}^{2m} , siendo $\|\cdot\|$ la norma elegida en \mathbb{R}^m . Entonces, como

$$\|g(t, z)\| = \left\| \begin{pmatrix} z_2 \\ f(t, z_1, z_2) \end{pmatrix} \right\| = \max\{\|z_2\|, \|f(t, z_1, z_2)\|\},$$

(4.1.10) es equivalente a que

$$\|f(t, y, y') - f(t, \bar{y}, \bar{y}')\| \leq L \max\{\|y - \bar{y}\|, \|y' - \bar{y}'\|\}, \quad \forall y, y', \bar{y}, \bar{y}' \in \mathbb{R}^m, \forall t \in [0, T] \quad (4.1.11)$$

Por tanto, si f es continua en $[0, T] \times \mathbb{R}^m \times \mathbb{R}^m$ y cumple la condición (4.1.11), la función incremento Φ del RKN verifica la condición de Lipschitz por el Lema 3.1.2 y, en consecuencia, aplicando el Teorema 3.2.4 el método RKN es convergente de orden p si es consistente de orden p .

Nuestro objetivo ahora es deducir las condiciones de orden de estos métodos. Para ello se realiza una extensión de la teoría de las series de Butcher.

4.2. Derivadas sucesivas y árboles ordenados

Al igual que con los RK, la teoría de las series de Butcher se desarrolla sobre problemas autónomos:

$$y'' = f(y, y') \quad (4.2.12)$$

Siguiendo una notación análoga a la usada para deducir las condiciones de los RK, denotaremos

$$f_i^j := \frac{\partial f_i}{\partial y_j}, \quad f_i^{j'} := \frac{\partial f_i}{\partial y_j'}, \quad f_i^{jk} := \frac{\partial^2 f_i}{\partial y_j \partial y_k}, \quad f_i^{j'k'} := \frac{\partial^2 f_i}{\partial y_j' \partial y_k'}, \quad f_i^{jk'} := \frac{\partial^2 f_i}{\partial y_j \partial y_k'}, \quad \dots$$

teniendo en cuenta que

$$y_i'' = f_i(y_1, \dots, y_m, y_1', \dots, y_m') \quad i = 1, \dots, m, \quad (4.2.13)$$

donde el subíndice i denota la posición i -ésima del vector y . De forma análoga a los RK, se intenta asociar a una aplicación (árbol), a cada uno de los términos de las derivadas sucesivas de la solución exacta $y(t)$.

Derivando $y(t)$ se tiene

$$y_i^{(3)} = \sum_{j=1}^m \frac{\partial f_i}{\partial y_j}(y, y') y'_j + \sum_j \frac{\partial f_i}{\partial y'_j}(y, y') y''_j = \sum_j f_i^j y'_j + \sum_j f_i^{j'} f_j \quad (4.2.14)$$

$$\begin{aligned} y_i^{(4)} = & \sum_{j,k} f_i^{jk} y'_j y'_k + \sum_{j,k} f_i^{jk'} y'_j f_k + \sum_j f_i^j f_j + \sum_{j,k} f_i^{j'k} f_j y'_k + \sum_{jk} f_i^{j'k'} f_j f_k \\ & + \sum_{jk} f_i^{j'} f_j^k y'_k + \sum_{jk} f_i^{j'} f_j^{k'} f_k \end{aligned} \quad (4.2.15)$$

Observamos que hay que diferenciar las derivadas respecto de y_j e y'_j , por lo que a cada sumatorio se le asocia un árbol que tendrá dos tipos de vértices: “vértices flacos”¹ (si la derivada es respecto de y) y “vértices gordos”² (si la derivada es respecto de y').

Definición 4.2.4. *Un N -árbol ordenado monótonamente de orden $q \geq 1$ (labeled N -tree) es un árbol ordenado monótonamente que esta definido mediante la aplicación*

$$\tau : \{2, 3, \dots, q\} \longrightarrow \{1, 2, 3, \dots, q\}$$

junto con la correspondencia

$$\tau' : \{1, 2, 3, \dots, q\} \longrightarrow \{\text{“flaco”}, \text{“gordo”}\}$$

que satisface:

1. la raíz de τ es siempre gordo, es decir, $\tau'(1) = \text{“gordo”}$
2. un vértice flaco tiene como mucho un hijo y este tiene que ser gordo.

Al conjunto de N -árboles ordenados monótonamente de orden q lo denotamos por LNT_q .

Al igual que los métodos RK los elementos $\{1, 2, 3, \dots\}$ se asocian a subíndices $\{j_1, j_2, j_3, \dots\}$ pero para simplificar presentación se reescribe como $\{i, j, k, \dots\}$.

Por tanto, los sumatorios de $y_i^{(3)}$ se asocian respectivamente a los árboles $\tau_1 = j \searrow_{\circ_i}$ y $\tau_2 = j \circlearrowleft_i$, y, se denotan las diferenciales elementales,

$$F_i \left(j \searrow_{\circ_i} \right) = \sum_j f_i^j y'_j \quad \text{y} \quad F_i \left(j \circlearrowleft_i \right) (y_0, y'_0) = \sum_j f_i^{j'} f_j$$

De la misma manera, los sumatorios de $y_i^{(4)}$ se asocian respectivamente a los árboles

¹ Traducción de “meagre vertex”

² Traducción de “fat vertex”

$$j \swarrow_i^k, j \searrow_i^k, j \circlearrowleft_i^k, j \circlearrowright_i^k, j \circlearrowleft_i^k, j \circlearrowright_i^k \text{ y } j \circlearrowleft_i^k$$

Y las deiferenciales elementales se denotan por

$$\begin{aligned} F_i \left(j \swarrow_i^k \right) &= \sum_{j,k}^m f_i^{jk} y_j' y_k', F_i \left(j \searrow_i^k \right) = \sum_{j,k}^m f_i^{jk'} y_j' f_k, F_i \left(j \circlearrowleft_i^k \right) = \sum_j^m f_i^j f_j, \\ F_i \left(j \circlearrowright_i^k \right) &= \sum_{j,k}^m f_i^{j'k} f_j y_k', F_i \left(j \circlearrowleft_i^k \right) = \sum_{jk}^m f_i^{j'k'} f_j f_k, \\ F_i \left(j \circlearrowright_i^k \right) &= \sum_{jk}^m f_i^{j'} f_j^k y_k' \text{ y } F_i \left(j \circlearrowleft_i^k \right) = \sum_{jk}^m f_i^{j'} f_j^{k'} f_k \end{aligned}$$

Cuando un vértice no tiene hijos decimos que es un **vértice extremo** (end-vertex). Además, si no hay confusión denotaremos τ en lugar de (τ, τ') . La definición de las diferenciales elementales para árboles de mayor orden se complica mucho por la notación. Por ello, en [3] la explicación de la siguiente forma.

Definición 4.2.5. Si $\tau \in LNT_q$ definimos

$$F_i(\tau)(y, y')$$

como una suma sobre los índices de todos los vértices gordos de τ y sobre los índices de todos los vértices extremos flacos. La expresión general de esta suma es un producto de expresiones del tipo:

$$\frac{\partial^r f_i}{\partial y_j \dots \partial y_k' \dots}(y, y') = f_i^{j \dots k' \dots}(y, y') \quad e \quad y_k'$$

Varios árboles pueden tener asociada la misma diferencial elemental. Se pueden identificar dichos árboles con la siguiente relación de equivalencia.

Definición 4.2.6. Dos N -árboles ordenados monótonamente τ y ν son equivalentes si tienen el mismo orden (q) y existe una biyección $\sigma : \{1, 2, \dots, q\} \rightarrow \{1, 2, \dots, q\}$ con $\sigma(1) = 1$ tal que $\tau \circ \sigma = \sigma \circ \nu$ en $\{2, 3, \dots, q\}$ y $\tau' \circ \sigma = \nu'$.

Ejemplo 4.2.7. El árbol asociado al segundo sumatorio de la derivada de orden 4 (4.2.15) y el asociado al cuarto sumatorio son equivalentes, pues existe la biyección σ tal que $\sigma(j) = k$ y $\sigma(k) = j$ verificando la Definición 4.2.6. Esto se traduce en que $F_i \left(j \searrow_i^k \right) = \sum_{j,k}^m f_i^{jk'} y_j' f_k = \sum_{j,k}^m f_i^{j'k} f_j y_k' = F_i \left(j \circlearrowright_i^k \right)$ pues

$$F_i \left(j \searrow_i^k \right) = \sum_{j,k}^m f_i^{jk'} y_j' f_k = \sum_{j,k}^m f_i^{j'k} f_j y_k' = F_i \left(j \circlearrowright_i^k \right)$$

lo único que cambian son los nombres de los contadores.

En cambio, obsérvese que el quinto sumatorio no es equivalente a los anteriores,

$$\text{ya que } F_i \left(j \circlearrowright_i^k \right) = \sum_{jk}^m f_i^{j'k'} f_j f_k \neq F_i \left(j \searrow_i^k \right), \text{ y no podemos encontrar}$$

una biyección tal que verifique las condiciones de la Definición 4.2.6.

Definición 4.2.8. *A la clase de equivalencia de los N -árboles ordenados monótonamente de orden q se le denomina N -árboles de orden q , que denotaremos por NT_q . Además, denotamos por $\alpha(\tau)$ al número de elementos de la clase de equivalencia de τ .*

Con estas definiciones podemos escribir $y^{(4)}$ como

$$y_i^{(4)} = \sum_{\tau \in LNT_3} F_i(\tau)(y, y') = \sum_{\tau \in NT_3} \alpha(\tau) F_i(\tau)(y, y')$$

Además igualando a (4.2.15) se obtiene que $\alpha(\tau) = 1$, $\forall \tau \neq j \searrow_i^o k$ y $\alpha\left(j \searrow_i^o k\right) = 2$. En general,

Teorema 4.2.9. *La solución exacta de (4.2.13) cumple que*

$$y^{(q)} = \sum_{\tau \in LNT_{q-1}} F(\tau)(y, y') = \sum_{\tau \in NT_{q-1}} \alpha(\tau) F(\tau)(y, y'), \quad q \geq 2 \quad (4.2.16)$$

4.3. Derivada de la solución numérica

Para desarrollar la solución numérica en serie de Butcher, reescribimos el método (4.0.5) como sigue

$$\begin{aligned} g_i(h) &= y_0 + c_i h y'_0 + \sum_{j=1}^s \bar{a}_{ij} (h^2 \varphi_j(h)), & g'_i(h) &= y'_0 + \sum_{j=1}^s a_{ij} (h \varphi_j(h)) \\ y_1 &= y_0 + h y'_0 + \sum_{i=1}^s \bar{b}_i (h^2 \varphi_i(h)), & y'_1 &= y'_0 + \sum_{i=1}^s b_i (h \varphi_i(h)) \end{aligned} \quad (4.3.17)$$

donde

$$\varphi_j(h) = f(g_j(h), g'_j(h)), \quad j = 1, \dots, s, \quad (4.3.18)$$

por lo que $\varphi_j(0) = f(g_j(0), g'_j(0)) = f(y_0, y'_0)$, $\forall j$.

Nuestro objetivo es desarrollar $g_i(h)$, $g'_i(h)$, y_1 e y'_1 en serie de potencias de h , para lo que se necesita calcular sus derivadas respecto de h en $h = 0$. Obsérvese que y_1 tiene una expresión similar a $g_i(h)$ e y'_1 a $g'_i(h)$, por lo que si desarrollamos $g_i(h)$ y $g'_i(h)$, el desarrollo de y_1 e y'_1 será el mismo, salvo por los coeficientes $\{\bar{b}_i, b_i\}$.

En (4.3.17) aparecen expresiones del tipo $h\varphi(h)$ y $h^2\varphi(h)$, que tendremos que derivar. Denotando

$$(g(h))^{(k)} \Big|_{h=0} = \frac{d^k}{dh^k} g(h) \Big|_{h=0}, \quad k \geq 0$$

y aplicando la Regla de Leibniz para cualquier $\varphi(h)$, se tiene

$$\begin{aligned} (h\varphi(h))^{(q)}|_{h=0} &= \sum_{k=0}^q \binom{q}{k} [h^k(\varphi(h))^{(q-k)}]|_{h=0} = \binom{q}{0} [h(\varphi(h))^{(q)}]|_{h=0} \\ &+ \binom{q}{1} (\varphi(h))^{(q-1)}|_{h=0} = q \cdot (\varphi(h))^{(q-1)}|_{h=0} \end{aligned} \quad (4.3.19)$$

$$\begin{aligned} (h^2\varphi(h))^{(q)}|_{h=0} &= \sum_{k=0}^q \binom{q}{k} (h^2)^{(k)}(\varphi(h))^{(q-k)}|_{h=0} = \binom{q}{0} [h^2(\varphi(h))^{(q)}]|_{h=0} \\ &+ \binom{q}{1} [2h(\varphi(h))^{(q-1)}]|_{h=0} + \binom{q}{2} 2(\varphi(h))^{(q-1)}|_{h=0} \\ &= \frac{q!}{(q-2)!2!} 2(\varphi(h))^{(q-2)}|_{h=0} = q \cdot (q-1) \cdot (\varphi(h))^{(q-2)}|_{h=0} \end{aligned} \quad (4.3.20)$$

Ahora vamos a aplicar esto a $\varphi(h) = \varphi_j(h)$ para obtener las derivadas de g_i y g'_i en $h = 0$. Para cada etapa $i = 1, \dots, s$ y cada componente $l = 1, \dots, m$

$$((g_i(h))_l)^{(1)}|_{h=0} = c_i(y'_0)_l + \sum_{j=1}^s \bar{a}_{ij} (h^2((\varphi_j(h))_l)^{(1)}|_{h=0}) = c_i(y'_0)_l \quad (4.3.21)$$

$$((g'_i(h))_l)^{(1)}|_{h=0} = \sum_{j=1}^s a_{ij} (h(\varphi_j(h))_l)^{(1)}|_{h=0} = \sum_{j=1}^s a_{ij} (\varphi_j(0))_l = \left(\sum_{j=1}^s a_{ij} \right) f_l(y_0, y'_0) \quad (4.3.22)$$

$$\begin{aligned} ((g_i(h))_l)^{(2)} &= \sum_{j=1}^s \bar{a}_{ij} (h^2(\varphi_j(h))_l)^{(2)}|_{h=0} = \sum_{j=1}^s \bar{a}_{ij} 2(\varphi_j(h))_l|_{h=0} \\ &= 2 \left(\sum_{j=1}^s \bar{a}_{ij} \right) f_l(y_0, y'_0) \end{aligned} \quad (4.3.23)$$

Para la derivada segunda de g'_i y la derivada tercera de g_i tenemos que usar las fórmulas (4.3.19), (4.3.20), las derivadas anteriores y $(\varphi_j(h))^{(1)}$. Así, para $l = 1, \dots, m$,

$$(\varphi_j(h))_l^{(1)} = (f_l(g_j, g'_j))^{(1)} = \sum_{k=1}^m f_l^k(g_j, g'_j)(g_j(h))_k^{(1)} + \sum_{k=1}^m f_l^{k'}(g_j, g'_j)(g'_j(h))_k^{(1)} \quad (4.3.24)$$

Por tanto,

$$\begin{aligned}
(\varphi_j(h))_l^{(1)}|_{h=0} &= \sum_{k=1}^m f_l^k(y_0, y'_0) c_j(y'_0)_k + \sum_{k=1}^m f_l^{k'}(y_0, y'_0) \sum_{r=1}^s a_{jr} f_k(y_0, y'_0) \\
&= c_j \sum_{k=1}^m f_l^k y'_k |_{(y_0, y'_0)} + \left(\sum_{r=1}^s a_{jr} \right) \sum_{k=1}^m f_l^{k'} f_k |_{(y_0, y'_0)}
\end{aligned}$$

Para simplificar la notación, a partir de ahora entendemos que f_k , y_k e y'_k están evaluadas en (y_0, y'_0) . Una vez hallada la derivada de φ podemos deducir que

$$(g'_i(h))^{(2)}|_{h=0} = \sum_{j=1}^s a_{ij} (h\varphi_j(h))^{(2)}|_{h=0} = \sum_{j=1}^s a_{ij} 2(\varphi_j(h))^{(1)}|_{h=0} \quad (4.3.25)$$

por lo que, cada componente $l = 1, \dots, m$ verifica

$$\begin{aligned}
(g'_i(h))_l^{(2)}|_{h=0} &= \sum_{j=1}^s a_{ij} \left[2c_j \sum_{k=1}^m f_l^k y'_k + 2 \left(\sum_{r=1}^s a_{jr} \right) \sum_{k=1}^m f_l^{k'} f_k \right] \\
&2 \left[\sum_{j=1}^s a_{ij} c_j \right] \sum_{k=1}^m f_l^k y'_k + 2 \left[\sum_{j,r=1}^s a_{ij} a_{jr} \right] \sum_{k=1}^m f_l^{k'} f_k
\end{aligned} \quad (4.3.26)$$

Además,

$$\begin{aligned}
(g_i(h))^{(3)}|_{h=0} &= \sum_{j=1}^s \bar{a}_{ij} (h^2 \varphi_j(h))^{(3)}|_{h=0} = \sum_{j=1}^s \bar{a}_{ij} 3 \cdot 2(\varphi_j(h))^{(1)}|_{h=0} \Rightarrow \\
(g_i(h))_l^{(3)}|_{h=0} &= \sum_{j=1}^s \bar{a}_{ij} 3 \left[2c_j \sum_{k=1}^m f_l^k y'_k + 2 \sum_{r=1}^s a_{jr} \sum_{k=1}^m f_l^{k'} f_k \right] \\
&= 3 \cdot 2 \left[\sum_{j=1}^s \bar{a}_{ij} c_j \right] \sum_{k=1}^m f_l^k y'_k + 3 \cdot 2 \left[\sum_{j,r=1}^s \bar{a}_{ij} a_{jr} \right] \sum_{k=1}^m f_l^{k'} f_k
\end{aligned} \quad (4.3.27)$$

Para la derivada tercera de g'_i y la derivada cuarta de g_i necesitaremos, además de todo lo anterior, para cada $l = 1, \dots, m$, derivando respecto h en (4.3.24),

$$\begin{aligned}
(\varphi_j(h))_l^{(2)} &= \sum_{k,q=1}^m f_l^{kq} (g_j(h))_q^{(1)} (g_j(h))_k^{(1)} + \sum_{k,q=1}^m f_l^{kq'} (g'_j(h))_q^{(1)} (g_j(h))_k^{(1)} \\
&+ \sum_{k=1}^m f_l^k (g_j(h))_k^{(2)} + \sum_{k,q=1}^m f_l^{k'q} (g_j(h))_q^{(1)} (g'_j(h))_k^{(1)} \\
&+ \sum_{k,q=1}^m f_l^{k'q'} (g'_j(h))_q^{(1)} (g'_j(h))_k^{(1)} + \sum_{k=1}^m f_l^{k'} (g'_j(h))_k^{(2)}
\end{aligned}$$

Por tanto,

$$\begin{aligned}
(\varphi_j(h))_l^{(2)}|_{h=0} &= c_j^2 \sum_{k,q=1}^m f_l^{kq} y'_k y'_q + \sum_{r=1}^s a_{jr} c_j \sum_{k,q=1}^m f_l^{kq'} f_q y'_k + 2 \sum_{r=1}^s \bar{a}_{jr} \sum_{k=1}^m f_l^k f_k \\
&+ \left(\sum_{r=1}^s a_{jr} \right)^2 \sum_{k,q=1}^m f_l^{k'q'} f_k f_q + \sum_{r=1}^s a_{jr} c_j \sum_{k,q=1}^m f_l^{k'q} f_k y'_q \\
&+ 2 \left(\sum_{r=1}^s a_{jr} c_r \right) \sum_{k,q=1}^m f_l^{k'} f_k^q y'_q + 2 \left(\sum_{r,p=1}^s a_{jr} a_{rp} \right) \sum_{k,q=1}^m f_l^{k'} f_k^q f_q
\end{aligned}$$

Así tenemos que

$$(g'_i(h))^{(3)}|_{h=0} = \sum_{j=1}^s a_{ij} (h\varphi_j(h))^{(3)}|_{h=0} = \sum_{j=1}^s a_{ij} 3(\varphi_j(h))^{(2)}|_{h=0} \quad (4.3.28)$$

Por componentes,

$$\begin{aligned}
((g'_i(h))_l)^{(3)}|_{h=0} &= \sum_{j=1}^s a_{ij} \left[3c_j^2 \sum_{k,q=1}^m f_l^{kq} y'_k y'_q + 3 \sum_{r=1}^s a_{jr} c_j \sum_{k,q=1}^m f_l^{kq'} f_q y'_k \right. \\
&+ 3 \cdot 2 \sum_{r=1}^s \bar{a}_{jr} \sum_{k=1}^m f_l^k f_k + 3 \left(\sum_{r=1}^s a_{jr} \right)^2 \sum_{k,q=1}^m f_l^{k'q'} f_k f_q \\
&+ 3 \sum_{r=1}^s a_{jr} c_j \sum_{k,q=1}^m f_l^{k'q} f_k y'_q + 3 \cdot 2 \left(\sum_{r=1}^s a_{jr} c_r \right) \sum_{k,q=1}^m f_l^{k'} f_k^q y'_q \\
&\left. + 3 \cdot 2 \left(\sum_{r,p=1}^s a_{jr} a_{rp} \right) \sum_{k,q=1}^m f_l^{k'} f_k^q f_q \right] \quad (4.3.29)
\end{aligned}$$

Analizando los resultados obtenidos observamos que, usando la misma aplicación densidad (2.1.6) que para los árboles de LT (sin distinguir si los nodos son gordos o flacos)

$$(g'_i(h))_l^{(2)}|_{h=0} = \gamma(\mathfrak{l}) \sum_{j=1}^s a_{ij} c_j F_l(\mathfrak{l})(y_0, y'_0) + \gamma(\mathfrak{g}) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s a_{jr} \right) F_l(\mathfrak{g})(y_0, y'_0),$$

$$\begin{aligned}
g_i(h)_l^{(3)}|_{h=0} &= 3 \cdot \gamma(\mathfrak{l}) \left(\sum_{j=1}^s \bar{a}_{ij} \right) c_j F_l(\mathfrak{l})(y_0, y'_0) \\
&+ 3 \cdot \gamma(\mathfrak{g})(y_0, y'_0) \sum_{j=1}^s \bar{a}_{ij} \left(\sum_{r=1}^s a_{jr} \right) F_l(\mathfrak{g})(y_0, y'_0)
\end{aligned}$$

y que

$$\begin{aligned}
 (g'_i(h))_l^{(3)}|_{h=0} &= \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} c_j^2 F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s a_{jr} c_j \right) F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s \bar{a}_{jr} \right) F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s a_{jr} \right)^2 F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s a_{jr} c_j \right) F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r=1}^s a_{jr} c_r \right) F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0) \\
 &+ \gamma \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) \sum_{j=1}^s a_{ij} \left(\sum_{r,p=1}^s a_{jr} a_{rp} \right) F_l \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) (y_0, y'_0)
 \end{aligned}$$

Así vemos que en el caso de las derivadas $(g_i(h))^{(q)}|_{h=0}$ aparecen unos coeficientes acompañando al $\sum_{j=1}^s \bar{a}_{ij}$ y en las derivadas $(g'_i(h))^{(q)}|_{h=0}$ aparecen los mismos coeficientes pero acompañando a $\sum_{j=1}^s a_{ij}$. Esto motiva la siguiente definición:

Definición 4.3.10. Para un N -árbol ordenado monótonamente τ denotamos por $\Phi_i(\tau)$ a una suma sobre los índices de todos los vértices gordos de τ , cuyo término general viene dado por productos de expresiones del tipo:

a_{kl} si el vértice gordo “ k ” tiene un hijo gordo “ l ”,
 \bar{a}_{kl} si el vértice gordo “ k ” está conectado a través de un hijo flaco con “ l ” y
 c_k^m si el vértice gordo “ k ” está conectado con m vértices extremos flacos.

Por ejemplo,

$$\begin{aligned}
 \Phi_j(\circ) &= c_j & \Phi_j(\circ) &= \sum_{r=1}^s a_{jr} \\
 \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= c_j^2, & \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= \sum_{r=1}^s a_{jr} c_j = \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right), & \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= \left(\sum_{r=1}^s a_{jr} \right)^2, \\
 \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= \sum_{r=1}^s \bar{a}_{jr}, & \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= \sum_{r=1}^s a_{jr} c_r, & \Phi_j \left(\begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \circ \end{array} \right) &= \sum_{r,p=1}^s a_{jr} a_{rp}
 \end{aligned}$$

Esto se generaliza para cualquier derivada:

Teorema 4.3.11. g_i y g'_i de (4.3.17) satisfacen

$$(g_i)^{(q+1)}|_{h=0} = (q+1) \sum_{\tau \in LNT_q} \gamma(\tau) \sum_{j=1}^s \bar{a}_{ij} \Phi_j(\tau) F(\tau)(y_0, y'_0) \quad (4.3.30)$$

$$(g'_i)^{(q)}|_{h=0} = \sum_{\tau \in LNT_q} \gamma(\tau) \sum_{j=1}^s a_{ij} \Phi_j(\tau) F(\tau)(y_0, y'_0) \quad (4.3.31)$$

donde $\gamma(\tau)$ es la aplicación definida en el Lema 2.1.7.

Por la similitud de g_i y y_1 , g'_i y y'_1 se deduce que:

Teorema 4.3.12. La solución numérica de un método RKN, (y_1, y'_1) , satisface

$$(y_1)^{(q)}|_{h=0} = q \sum_{\tau \in LNT_{q-1}} \gamma(\tau) \sum_{i=1}^s \bar{b}_i \Phi_i(\tau) F(\tau)(y_0, y'_0), \quad q \geq 2 \quad (4.3.32)$$

$$(y'_1)^{(q-1)}|_{h=0} = \sum_{\tau \in LNT_{q-1}} \gamma(\tau) \sum_{i=1}^s b_i \Phi_i(\tau) F(\tau)(y_0, y'_0), \quad q \geq 2 \quad (4.3.33)$$

donde $\gamma(\tau)$ es la aplicación definida en el Lema 2.1.7.

En la Tabla 4.1 se detallan los árboles y los correspondientes derivadas sucesivas.

Por lo tanto tenemos una fórmula general para hallar las derivadas de y_1 e y'_1 , y así poder calcular el desarrollo en serie de potencias de h .

4.4. Condiciones de orden

Para realizar el estudio del orden de un método RKN tenemos que comparar las series de Maclaurin en potencias de h de la solución del método,

$$\begin{aligned} y_1 &= y_0 + y'_0 h + \sum_{q=2}^{\infty} (y_1)^{(q)}|_{h=0} \frac{h^q}{q!} = y_0 + y'_0 h \\ &\quad + \sum_{\tau \in NT} \alpha(\tau) \gamma(\tau) \left(\sum_{i=1}^s \bar{b}_i \Phi_i(\tau) \right) F(\tau)(y_0, y'_0) \frac{h^{\rho(\tau)+1}}{\rho(\tau)!} \\ y'_1 &= y'_0 + \sum_{q=1}^{\infty} (y'_1)^{(q)}|_{h=0} \frac{h^q}{q!} = y'_0 \\ &\quad + \sum_{\tau \in NT} \alpha(\tau) \gamma(\tau) \left(\sum_{i=1}^s b_i \Phi_i(\tau) \right) F(\tau)(y_0, y'_0) \frac{h^{\rho(\tau)}}{\rho(\tau)!} \end{aligned}$$

con las series de la solución exacta y su derivada,

$$y(t_0 + h) = y_0 + y_0' h + \sum_{q=2}^{\infty} y^{(q)}(t_0) \frac{h^q}{q!} = y_0 + y_0' h + \sum_{\tau \in NT} \alpha(\tau) F(\tau)(y_0, y_0') \frac{h^{\rho(\tau)+1}}{(\rho(\tau) + 1)!}$$

$$y'(t_0 + h) = y_0' + \sum_{q=1}^{\infty} y^{(q+1)}(t_0) \frac{h^q}{q!} = y_0' + \sum_{\tau \in NT} \alpha(\tau) F(\tau)(y_0, y_0') \frac{h^{\rho(\tau)}}{\rho(\tau)!}$$

Así obtenemos el siguiente Teorema:

Teorema 4.4.13. *Un método RKN es de orden p si y solo si*

$$\sum_{i=1}^s \bar{b}_i \Phi_i(\tau) = \frac{1}{(\rho(\tau) + 1) \cdot \gamma(\tau)} \quad \text{para los } N\text{-árboles } \tau \text{ con } \rho(\tau) \leq p - 1, \quad (4.4.34)$$

$$\sum_{i=1}^s b_i \Phi_i(\tau) = \frac{1}{\gamma(\tau)} \quad \text{para los } N\text{-árboles } \tau \text{ con } \rho(\tau) \leq p. \quad (4.4.35)$$

Ejemplo 4.4.14. Para que un método RKN tenga orden 3 debe cumplir,

$$\begin{aligned} \sum_{i=1}^s b_i &= 1, \quad \sum_{i=1}^s \bar{b}_i = \frac{1}{2}, \quad \sum_{i=1}^s b_i c_i = \frac{1}{2}, \quad \sum_{i,j=1}^s b_i a_{ij} = \frac{1}{2}, \\ \sum_{i,j=1}^s \bar{b}_i a_{ij} &= \frac{1}{6}, \quad \sum_{i=1}^s \bar{b}_i c_i = \frac{1}{6}, \quad \sum_{i=1}^s b_i c_i^2 = \frac{1}{3}, \quad \sum_{i,j=1}^s b_i c_i a_{ij} = \frac{1}{3}, \\ \sum_{i=1}^s b_i \left(\sum_{j=1}^s a_{ij} \right)^2 &= \frac{1}{3}, \quad \sum_{i,j=1}^s b_i \bar{a}_{ij} = \frac{1}{6}, \quad \sum_{i,j=1}^s b_i a_{ij} c_j = \frac{1}{6}, \quad \sum_{i,j,r=1}^s b_i a_{ij} a_{jr} = \frac{1}{6} \end{aligned}$$

Sin embargo, estas condiciones de orden se reducen cuando aplicamos el RKN (4.0.5) a problemas del tipo:

$$y'' = f(y) \quad (4.4.36)$$

En estos problemas desaparecen las diferenciales elementales que contienen derivadas con respecto de y' . Por consiguiente, se considera solo el siguiente subconjunto de N -árboles:

Definición 4.4.15. *Un N -árbol τ se llama N -árbol especial o **SN-árbol**, si los vértices gordos sólo tienen hijos flacos.*

En consecuencia, podemos enunciar un Teorema análogo al 4.4.13 pero para SN-árbol:

Orden	τ	$\rho(\tau)$	$\alpha(\tau)$	$\gamma(\tau)$	$\Phi_i(\tau)$	$F_i(\tau)(y, y')$
1		1	1	1	1	f_i
2		2	1	2	c_i	$\sum_j f_i^j y_j'$
		2	1	2	$\sum_j a_{ij}$	$\sum_{j,k} f_i^{j'} f_j$
3		3	1	3	c_i^2	$\sum_{j,k} f_i^{jk} y_j' y_k'$
		3	2	3	$\sum_j a_{ij} c_i$	$\sum_{j,k} f_i^{jk'} f_k y_j'$
		3	1	3	$\left(\sum_j a_{ij}\right)^2$	$\sum_{j,k} f_i^{j'k'} f_j f_k$
		3	1	6	$\sum_j \bar{a}_{ij}$	$\sum_j f_i^j f_j$
		3	1	6	$\sum_j a_{ij} c_j$	$\sum_{j,k} f_i^{j'} f_j^k y_k'$
		3	1	6	$\sum_{j,r} a_{ij} a_{jr}$	$\sum_{j,k} f_i^{j'} f_j^{k'} f_k$

Tabla 4.1. N-árboles y diferenciales elementales asociados

Teorema 4.4.16. *Un método RKN (4.0.5) para la ecuación diferencial especial (4.4.36) es de orden p , si*

$$\sum_{i=1}^s \bar{b}_i \Phi_i(\tau) = \frac{1}{(\rho(\tau) + 1) \cdot \gamma(\tau)} \quad \text{para los SN-árboles } \tau \text{ con } \rho(\tau) \leq p - 1, \tag{4.4.37}$$

$$\sum_{i=1}^s b_i \Phi_i(\tau) \quad \text{para los SN-árboles } \tau \text{ con } \rho(\tau) \leq p. \tag{4.4.38}$$

Por tanto, las condiciones de orden 3 para PVI's del tipo (4.4.36) son solamente:

$$\sum_{i=1}^s \bar{b}_i = \frac{1}{2}, \quad \sum_{i=1}^s \bar{b}_i c_i = \frac{1}{6}, \quad \sum_{i=1}^s b_i = 1, \quad \sum_{i=1}^s b_i c_i = \frac{1}{2}, \quad \sum_{i=1}^s b_i c_i^2 = \frac{1}{3}, \quad \sum_{i,j=1}^s b_i \bar{a}_{ij} = \frac{1}{6}$$

Aplicación a la ecuación de ondas

Supongamos que tenemos una cuerda de longitud L sujeta por sus dos extremos en la cual se producen pequeñas vibraciones. Además asumimos que la cuerda está totalmente estirada, elaborada de un material homogéneo y no hay gravedad. La vibración a lo largo del tiempo se puede describir mediante la ecuación [5]

$$u_{tt}(x, t) = \alpha^2 u_{xx}(x, t), \quad x \in [0, L], t \geq 0, \quad (5.0.1)$$

donde $u(x, t)$ es el desplazamiento vertical desde la posición de equilibrio del punto x de la cuerda en el instante t , por lo que u_{tt} representa la aceleración vertical de la cuerda en el punto x .

De esta manera la ecuación (5.0.1), que se denomina **ecuación de ondas**, puede interpretarse como que la aceleración en cada punto de la cuerda es proporcional a la tensión de la misma y que será mayor cuanto mayor sea la convexidad de u .

Se puede obtener una aproximación a la solución de (5.0.1), mediante el uso del **Método de Líneas**. Consideramos el ejemplo

$$\begin{aligned} u_{tt} &= \alpha^2 u_{xx}, & t \geq 0, & \quad x \in [0, 1] \\ u(x, 0) &= \sin(2\pi x), & \forall x \in [0, 1] \\ u_t(x, 0) &= \frac{1}{2} \sin(\pi x) & \forall x \in [0, 1] \\ u(0, t) &= u(1, t) = 0, & \forall t \geq 0 \end{aligned} \quad (5.0.2)$$

cuya solución exacta es conocida:

$$u(x, t) = \frac{1}{2\pi\alpha} \sin(\pi x) \sin(\pi\alpha t) + \sin(2\pi x) \cos(2\pi\alpha t) \quad (5.0.3)$$

por lo que podremos comprobar numéricamente la eficiencia y la precisión de los métodos que vamos a testar.

Para aplicar el método de líneas primero se realiza una partición equiespaciada del intervalo $[0, 1]$: $x_i = i\Delta_x$, $\Delta_x = \frac{1}{M+1}$, $0 \leq i \leq M + 1$. Se aplica la ecuación (5.0.2) sobre los puntos interiores de dicha partición

$$u_{tt}(x_i, t) = \alpha^2 u_{xx}(x_i, t), \quad i = 1, \dots, M$$

ya que los valores para $x_0 = 0$ y $x_{M+1} = 1$ vienen dadas por las condiciones de frontera de Dirichlet homogéneas. Si aproximamos el segundo miembro de la ecuación mediante diferencias finitas centrales de segundo orden, se tiene cuando $\Delta_x \rightarrow 0$,

$$u_{xx}(x_i, t) \simeq \frac{u(x_i + \Delta_x, t) - 2u(x_i, t) + u(x_i - \Delta_x, t)}{\Delta_x^2} - \frac{\Delta_x^2}{12} u_{xxxx}(x_i, t) + \mathcal{O}(\Delta_x^4) \quad (5.0.4)$$

Suponiendo entonces que Δ_x es suficientemente pequeño, aproximaremos la solución de (5.0.2) por la del sistema de ecuaciones diferenciales ordinarias

$$u_{tt}(x_i, t) = \alpha^2 \frac{u(x_{i+1}, t) - 2u(x_i, t) + u(x_{i-1}, t)}{\Delta_x^2} \quad 1 \leq i \leq M, \quad t \geq 0 \quad (5.0.5)$$

O, denotando,

$$w_i(t) = u(x_i, t), \quad w_i''(t) = u_{tt}(x_i, t)$$

se reescribe,

$$\begin{cases} w_i'' = \alpha^2 \frac{w_{i+1} - 2w_i + w_{i-1}}{\Delta_x^2} & 1 \leq i \leq M \\ w_i(0) = \sin(2\pi x_i) \\ w_i'(0) = \frac{1}{2} \sin(\pi x_i) \end{cases}$$

teniendo en cuenta que $w_0 = u(0, t) = 0$ y $w_{M+1} = u(1, t) = 0$.

Se agrupan todas las componentes vectorialmente

$$U = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix}, \quad U_0 = \begin{pmatrix} w_1(0) \\ w_2(0) \\ \vdots \\ w_M(0) \end{pmatrix} = \begin{pmatrix} \sin(2\pi x_1) \\ \sin(2\pi x_2) \\ \vdots \\ \sin(2\pi x_M) \end{pmatrix}, \quad U_0' = \begin{pmatrix} w_1'(0) \\ w_2'(0) \\ \vdots \\ w_M'(0) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \sin(\pi x_1) \\ \frac{1}{2} \sin(\pi x_2) \\ \vdots \\ \frac{1}{2} \sin(\pi x_M) \end{pmatrix}$$

y definiendo la matriz,

$$\mathcal{A} = \frac{\alpha^2}{\Delta_x^2} \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & -2 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 1 & -2 \end{pmatrix} \in \mathcal{M}_{M \times M}(\mathbb{R})$$

el PVI de segundo orden se escribe

$$U'' = \mathcal{A}U = f(U), \quad U(0) = U_0, \quad U'(0) = U'_0 \quad t \geq 0 \quad (5.0.6)$$

Para la integración temporal de este sistema aplicaremos los métodos RKN vistos en el capítulo anterior.

En particular, los métodos que usaremos serán:

$\begin{array}{c ccc} \frac{1}{2} & 0 & & \\ \frac{1}{2} & 0 & \frac{1}{2} & \\ \hline & 0 & 0 & \frac{1}{2} \\ & 0 & 0 & 1 \end{array}$	$\begin{array}{c ccc} \frac{1}{2} & 0 & & \\ \frac{1}{2} & 0 & \frac{1}{30} & \\ \hline & 0 & 0 & 0 & \frac{1}{12} \\ & 0 & 0 & 0 & \frac{1}{2} \\ & 0 & 0 & 0 & 1 \end{array}$	$\begin{array}{c ccc} \frac{1}{2} & \frac{1}{8} & & \\ 1 & 0 & \frac{1}{2} & \\ \hline & \frac{1}{6} & \frac{1}{3} & 0 \\ & \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{array}$	$\begin{array}{c ccc} \frac{1}{5} & \frac{1}{50} & & \\ \frac{2}{3} & \frac{-1}{27} & \frac{7}{27} & \\ \hline 2 & \frac{3}{10} & \frac{-2}{35} & \frac{9}{35} \\ \hline & \frac{14}{336} & \frac{100}{336} & \frac{54}{336} & 0 \\ & \frac{14}{336} & \frac{125}{336} & \frac{162}{336} & \frac{35}{336} \end{array}$
M1	M2	M3	M4

Los métodos **M1** y **M2** son de orden 2 y han sido propuestos en [6] como apropiados para resolver problemas del tipo (5.0.6) por tener un buen orden de *dispersión* y tener *disipación nula* (*zero-dissipation*). Estas propiedades de los métodos se saben que son buenas cuando la matriz \mathcal{A} de (5.0.6) tiene autovalores reales negativos de módulo grande, como es el caso de nuestra matriz \mathcal{A} , cuyos autovalores son

$$\lambda_k = -4 \frac{\alpha^2}{\Delta_x^2} \sin^2 \left(\frac{k\pi}{2(M+1)} \right), \quad k = 1, \dots, M, \quad (5.0.7)$$

por ser una matriz Toeplitz tridiagonal, por lo que el módulo de los autovalores aumenta cuando $\Delta_x \rightarrow 0$. Los métodos **M3** y **M4** lo hemos extraído de [3] y tienen órdenes 4 y 5 respectivamente.

Al resolver el problema mediante este proceso se producen dos tipos de errores, uno debido a la discretización espacial (5.0.4) y el otro debido al método RKN empleado. A los primeros lo llamaremos **errores espaciales** y a los segundos **errores temporales**. Además llamaremos **errores globales** a la diferencia entre la solución numérica y la solución de la EDP (5.0.2).

Para estudiar los errores debidos a la integración temporal mediante los métodos RKN, hemos creado un código en Octave en el que realizamos dicha integración en el intervalo $t \in [0, T]$ a paso fijo $h = \frac{1}{N}$, dado un número de pasos N , y lo comparamos con la aproximación de la rutina ode45 de Octave¹ con tolerancia del error relativo y absoluto de 10^{-13} . En las Tablas 5.1 y 5.2 presentamos los errores temporales, en norma uniforme, que hemos obtenido para cada método, tomando una malla espacial de $M = 40$ puntos en el caso $\alpha = 1$, tanto para la aproximación de la solución exacta del PVI (5.0.6) como para su derivada.

¹ <https://octave.sourceforge.io/octave/function/ode45.html>

M1

N	y_t	y'_t
20	2.119710e+25	8.344529e+26
40	4.593201e+15	2.678235e+17
80	2.100631e-04	5.051369e-02
160	4.739199e-05	1.261993e-02
320	1.156237e-05	3.154445e-03
640	2.872754e-06	7.885774e-04
1280	7.170738e-07	1.971422e-04
2560	1.791988e-07	4.928543e-05
5120	4.479530e-08	1.232135e-05
10240	1.119852e-08	3.080337e-06

M2

N	y_t	y'_t
20	2.101617e-04	1.409401e-02
40	6.282098e-08	7.929858e-05
80	1.570310e-08	1.979143e-05
160	3.925687e-09	4.945627e-06
320	9.814196e-10	1.236265e-06
640	2.453579e-10	3.090574e-07
1280	6.134146e-11	7.726384e-08
2560	1.533930e-11	1.931601e-08
5120	3.840123e-12	4.829125e-09
10240	9.670875e-13	1.207352e-09

Tabla 5.1. Errores temporales para $M = 40$ y $\alpha = 1$ **M3**

N	y_t	$y_t p$
20	3.491711e+03	2.262343e+05
40	1.150342e-06	7.448185e-05
80	3.951982e-08	4.666419e-06
160	1.462807e-09	2.918446e-07
320	5.993528e-11	1.824389e-08
640	2.708611e-12	1.140325e-09

M4

N	y_t	y'_t
20	5.232102e-06	6.123710e-06
40	1.650590e-07	9.467875e-08
80	5.171276e-09	1.594315e-09
160	1.617696e-10	2.851186e-11
320	5.112799e-12	5.857537e-13
640	2.163825e-13	1.380562e-13

Tabla 5.2. Errores temporales para $M = 40$ y $\alpha = 1$

Debemos observar que, por el alto orden de los métodos **M3** y **M4**, rápidamente alcanzan la precisión del ode45 de Octave, por lo que los resultados para tamaños de paso h menores que $1/640$ no son relevantes. En las Tablas 5.1 y 5.2 se observa claramente la diferencia de aplicar un método de orden 2 o superior. También hay que observar que, salvo el **M4**, ninguno de los otros tres métodos es capaz de dar la solución aceptable con solo 20 pasos, lo que revela que estos métodos tienen algunos problemas de estabilidad. Si queremos comprobar que un RKN de orden $p \geq 1$ está alcanzando dicho orden sobre un problema concreto, podemos hacerlo tomando el error temporal obtenido integrando a paso fijo h y a paso $h/2$. Como sabemos que, para h suficientemente pequeño, y C independiente de h , si $\text{error}(h)$ denota la norma uniforme del error temporal

integrando a paso fijo h ,

$$\text{error}(h) \simeq Ch^p \Rightarrow \text{error}(h/2) \simeq C \frac{h^p}{2^p} \simeq \frac{\text{error}(h)}{2^p} \Rightarrow 2^p \simeq \frac{\text{error}(h)}{\text{error}(h/2)} \Rightarrow$$

$$p \simeq \frac{\log(\text{error}(h)) - \log(\text{error}(h/2))}{\log 2}$$

Como $\log(h) - \log(h/2) = \log 2$, $\forall h > 0$, si dado un $N_0 \geq 1$ consideramos los errores obtenidos con las sucesiones anteriores con paso fijo h para $h = h_k = \frac{h_0}{2^k}$, $k = 0, 1, \dots$ siendo $h_0 = \frac{1}{N_0}$. Se tendrá

$$p \simeq \frac{\log\left(\text{error}\left(\frac{h_0}{2^k}\right)\right) - \log\left(\text{error}\left(\frac{h_0}{2^{k+1}}\right)\right)}{\log 2}$$

por lo que si dibujamos en una gráfica en escala logarítmica los pasos $(h_k, \text{error}(h_k))$, $k = 0, 1, 2, \dots$ y el método está integrando el problema con el orden p previsto, esta gráfica será aproximadamente un recta de pendiente p . Podemos ver esto en las Figuras 5.1 y 5.2 en escala logarítmica para $M = 80$, $\alpha = 1$ y $\alpha = 10$, con los cuatro métodos considerados.

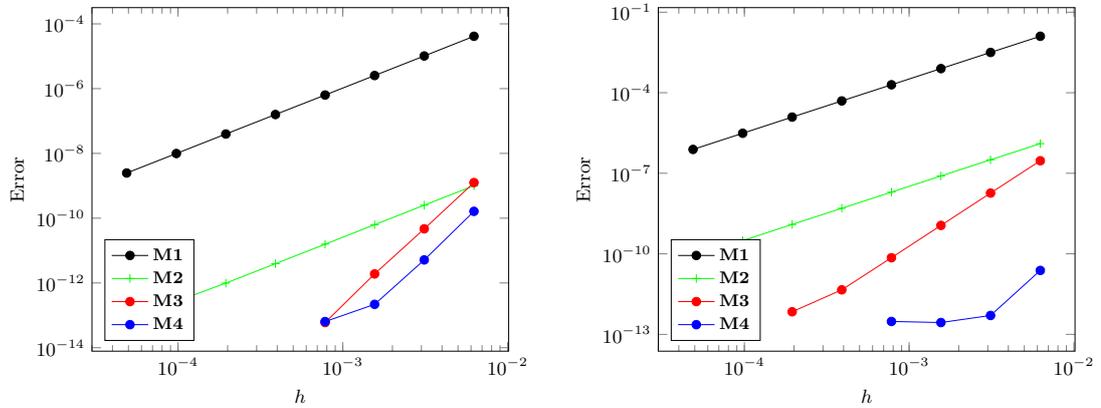


Figura 5.1. Errores de la solución exacta y de su derivada cuando $\alpha = 1$ y $M = 80$

En estas gráficas podemos ver claramente que los errores de cada método generan rectas de pendiente 2 (**M1** y **M2**), 4 (**M3**) y 5 (**M4**) respectivamente, aunque en este último caso la recta se convierte rápidamente en horizontal ya que se alcanza la precisión del ode45 de Octave con muy pocos pasos. También se puede observar que aunque el **M1** y el **M2** tienen el mismo orden, **M2** da mucha más precisión que el otro. Además, se observa que todos los métodos empeoran su precisión cuando $\alpha = 10$ respecto al caso $\alpha = 1$, lo que se debe al aumento del módulo de los autovalores (5.0.7) de la matriz \mathcal{A} de (5.0.6).

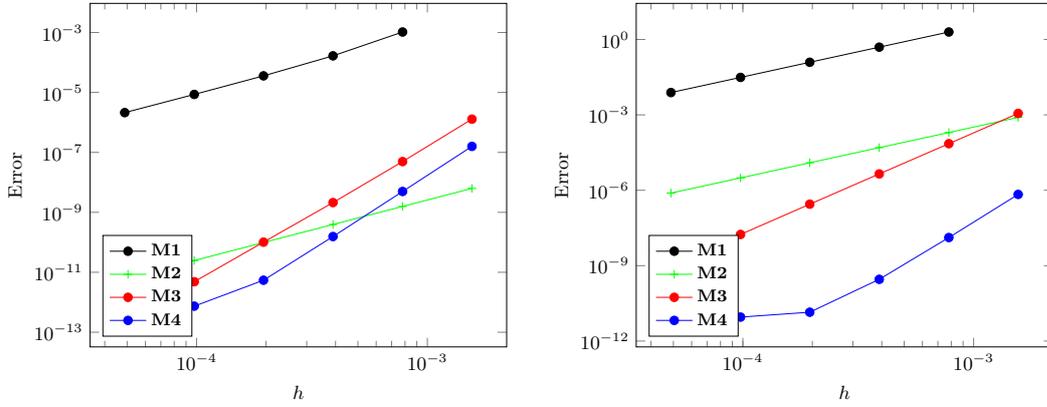


Figura 5.2. Errores de la solución exacta y de su derivada cuando $\alpha = 10$ y $M = 80$

Finalmente, en la Tabla 5.3 hemos comparado la aproximación con la solución exacta de la EDP (5.0.3) obteniendo los errores globales de cada método. Vemos que todos los métodos alcanzan el mismo error global porque los errores temporales son muy inferiores a los errores espaciales. Además, podemos ver que cuando duplicamos el número de puntos de la malla espacial se obtiene un error global, que se debe solo al error espacial, que es casi 4 veces inferior, lo que corrobora que la discretización espacial está aproximando correctamente con orden 2 a la solución exacta.

M	10	20	40
M1	6.60e-02	1.82e-02	4.78e-03
M2	6.60e-02	1.82e-02	4.78e-03
M3	6.60e-02	1.82e-02	4.78e-03
M4	6.60e-02	1.82e-02	4.78e-03

M	10	20	40
M1	8.11e-01	2.30e-01	6.05e-02
M2	8.04e-01	2.29e-01	6.00e-02
M3	8.04e-01	2.28e-01	6.00e-02
M4	8.04e-01	2.28e-01	6.00e-02

Tabla 5.3. Errores globales para $\alpha = 1$, $\alpha = 10$, respectivamente, y $N = 2^9 N_0$ donde $N_0 = \frac{M}{2}$

Bibliografía

- [1] CALVO, M. MONTIJANO, J.I. y RÁNDEZ, L. *Métodos de Runge-Kutta para la resolución de ecuaciones diferenciales ordinarias*. Secretariado de Publicaciones, Universidad de Zaragoza, 1988.
- [2] LAMBERT, J.D. *Numerical methods for ordinary differential systems (the initial value problem)*. John Wiley, 1991.
- [3] HAIRER, E., NORSETT, S.P., y WANNER, G. *Solving ordinary differential equations. I: nonstiff problems*. Springer 2nd Revised Edition, Berlin, 1993.
- [4] BUTCHER, J.C. *Numerical methods for ordinary differential equations*. John Wiley and Sons, 2003.
- [5] FARLOW, STANLEY J. *Partial differential equations for scientists and engineers*. Dover 2nd Edition, New York, 1993, pp. 123–160.
- [6] VAN DER HOUWEN, P.J. y SOMMEIJER, B.P. Explicit Runge-Kutta(-Nyström) Methods with Reduced Phase Errors for Computing Oscillating Solutions. *SIAM Journal on Numerical Analysis*, Jun, 1987, vol. 24, no. 3 (1987) pp. 595–617.

Numerical methods to approximate the solution of differential equations. Second order problems.



Alba del Pino Sosa Déniz

Facultad de Ciencias · Sección de Matemáticas
Universidad de La Laguna
alu0101160375@ull.edu.es

Abstract

The aim of this work is to study Runge-Kutta-Nyström (RKN) methods for solving initial value problems (IVPs) in second-order ordinary differential equations (ODEs).

1. Runge-Kutta methods

A Runge-Kutta (RK) method of s stages given an approximation y_n to the solution of the IVPs

$$y' = f(t, y), \quad t \in [0, T], \quad y, f \in \mathbb{R}^m, \quad y(0) = y_0$$

at the point t_n , with f satisfying certain conditions, gives an approximation to such a solution at the point $t_n + h$, that we denote by y_{n+1} , obtained by the formulation

$$K_i = f(t_n + c_i h, y_n + h \sum_{j=1}^s a_{ij} K_j), \quad 1 \leq i \leq s, \quad (1)$$

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i K_i \quad (2)$$

Coefficients of the RK: $A = (a_{ij})_{i,j=1}^s$, $b = (b_i)_{i=1}^s$, $c = (c_i)_{i=1}^s$

2. Order of Runge-Kutta methods

The RK has order p if and only if

$$l(t_0, h) = \mathcal{O}(h^{p+1}), \quad h \rightarrow 0.$$

An application

$$\tau : \{2, 3, \dots, q\} \rightarrow \{1, 2, \dots, q-1\} \text{ con } \tau(i) < i, \forall i = 2, 3, \dots, q$$

is called **labeled trees** of order q . LT = set of all labeled trees.

Theorem. A RK method has order $p \geq 1$ if and only if

$$b^T \Phi(\tau) = \frac{1}{\gamma(\tau)}, \quad \forall \tau \in \text{LT with } \rho(\tau) \leq p$$

where $\rho(\tau)$ = order, $\Phi(\tau)$ and $\gamma(\tau)$ are functions that only depend on the RK coefficients.

Example: conditions for order 3: $b^T c = 1$, $b^T c^2 = \frac{1}{2}$, $b^T c^3 = \frac{1}{3}$, $b^T A c = \frac{1}{6}$

3. Convergence of the Runge-Kutta methods

Theorem. If a RK is of order p and $f \in C^p(\mathcal{T}_\delta)$, there is $\bar{h} > 0$ such that if we consider any partition $\mathcal{P} = \{t_0 < t_1 < t_2 < \dots < t_N = t_f\}$ of $[t_0, t_f]$ with $h_{\max} := \max_j h_j \leq \bar{h} = \min\{h^*, h'\}$ then

$$\|y(t_n) - y_n\| \leq h_{\max}^p \frac{C}{\Lambda} [e^{\Lambda(t_n - t_0)} - 1], \quad n = 0, 1, \dots, N,$$

where C is independent of h and Λ is the Lipschitz constant of the function $\Phi(t_n, y_n, h) = \sum_{i=1}^s b_i K_i$.

4. Runge-Kutta-Nyström methods

A Runge-Kutta-Nyström (RKN(\bar{A}, A, \bar{b}, b)) method gives an approximation to the solution of the second order IPVs

$$y'' = f(t, y, y') \quad y(t_0) = y_0, \quad y'(t_0) = y'_0 \quad y, f \in \mathbb{R}^m, \quad t \in [t_0, T]$$

by the following formulation

$$\begin{aligned} k'_i &= f\left(t_0 + c_i h, y_0 + c_i h y'_0 + h^2 \sum_{j=1}^s \bar{a}_{ij} k'_j, y_0 + h \sum_{j=1}^s a_{ij} k'_j\right) \\ y_1 &= y_0 + h y'_0 + h^2 \sum_{i=1}^s \bar{b}_i k'_i \quad y'_1 = y'_0 + h \sum_{i=1}^s b_i k'_i \end{aligned} \quad (3)$$

A **labeled N-tree** of order q is a labeled tree τ together with the mapping $\tau' : \{1, 2, \dots, q\} \rightarrow \{\text{"meagre"}, \text{"fat"}\}$.

Theorem. A RKN method is of order p if and only if

$$\sum_{i=1}^s \bar{b}_i \Phi_i(\tau) = \frac{1}{(\rho(\tau) + 1) \cdot \gamma(\tau)} \quad \text{for the N-trees } \tau \text{ with } \rho(\tau) \leq p-1,$$

$$\sum_{i=1}^s b_i \Phi_i(\tau) = \frac{1}{\gamma(\tau)} \quad \text{for the N-trees } \tau \text{ with } \rho(\tau) \leq p$$

where $\rho(\tau)$ = τ order, $\Phi(\tau)$ and $\gamma(\tau)$ are functions that only depend on the RKN coefficients, and in that case, it is convergent of order p .

5. Application to the wave equation

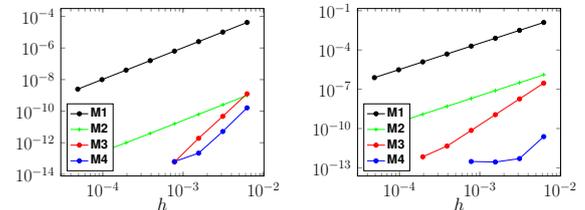
Wave equation

$$\begin{cases} u_{tt} = \alpha^2 u_{xx}, & u(0, t) = u(1, t) = 0 \quad t \geq 0, \quad x \in [0, 1] \\ u(x, 0) = \sin(2\pi x), & u_t(x, 0) = \frac{1}{2} \sin(\pi x), \quad \forall x \in [0, 1] \end{cases} \quad (4)$$

Method of lines: the partial derivate u_{xx} is discretized with second-order finite differences, obtaining the WP

$$U'' = AU = f(U), \quad U(0) = U_0, \quad U'(0) = U'_0 \quad t \geq 0.$$

Test with methods **M1**, **M2** of order 2, **M3** order 4, **M4** order 5. Temporal errors for y and y' with $\alpha = 1$, $M = 80$:



References

[1] HAIRER, E. NORSETT, S.P. and WANNER, G. *Solving ordinary differential equations. I: nonstiff problems*, Springer 2nd Revised Edition, Berlin, 1993.