

Facultad de Ciencias  
Grado de Matemáticas



Universidad de Oviedo

*La Teoría de Butcher*  
*y*  
*los métodos de Runge-Kutta*

José García García

Trabajo Fin de Grado

Supervisado por Pablo Pérez Riera

Junio de 2020



# Índice general

<b>Resumen</b>	<b>v</b>
<b>Introducción</b>	<b>vii</b>
Introducción histórica . . . . .	vii
Principales autores . . . . .	ix
Motivación y objetivos . . . . .	x
<b>1. Preliminares</b>	<b>11</b>
1.1. Ecuaciones diferenciales . . . . .	11
1.2. Métodos de un paso . . . . .	15
<b>2. Métodos de Runge-Kutta explícitos</b>	<b>19</b>
2.1. Definición y primeras propiedades . . . . .	19
2.2. Estabilidad y consistencia de los métodos RK explícitos . . . . .	24
2.3. Estudio del orden mediante desarrollos de Taylor . . . . .	26
2.3.1. Métodos de Runge-Kutta explícitos de una etapa . . . . .	27
2.3.2. Métodos de Runge-Kutta explícitos de dos etapas . . . . .	28
2.3.3. Métodos de Runge-Kutta explícitos de tres etapas . . . . .	29
<b>3. Teoría de Butcher</b>	<b>31</b>
3.1. Árboles con raíz y árboles monótonamente ordenados . . . . .	31
3.2. Desarrollo de la solución exacta: derivadas de Frechet y diferenciales elementales	40
3.3. Desarrollo de la solución numérica: series de Butcher . . . . .	46
3.4. Condiciones de orden. . . . .	52

<b>4. Teoría de Butcher y métodos de Runge-Kutta explícitos</b>	<b>59</b>
4.1. Métodos de Runge-Kutta explícitos de una etapa. . . . .	60
4.2. Métodos de Runge-Kutta explícitos de dos etapas. . . . .	61
4.3. Métodos de Runge-Kutta explícitos de tres etapas. . . . .	61
4.4. Métodos de Runge-Kutta explícitos de cuatro etapas. . . . .	63
4.5. Métodos de Runge-Kutta explícitos de cinco etapas . . . . .	69
4.6. Conclusiones sobre métodos de Runge-Kutta explícitos . . . . .	72
<b>5. Métodos de Runge-Kutta implícitos</b>	<b>75</b>
5.1. Definición y primeras propiedades . . . . .	76
5.2. Orden de los métodos RK implícitos . . . . .	79
5.3. Otras familias de métodos RK implícitos . . . . .	90
5.4. Métodos DIRK, SDIRK y SIRK . . . . .	92
<b>A. Algunas propiedades de los métodos de Runge-Kutta implícitos</b>	<b>93</b>
A.1. Buena definición . . . . .	93
A.2. Estabilidad . . . . .	95
A.3. Regularidad . . . . .	97
<b>Bibliografía</b>	<b>99</b>

# Resumen

Los métodos de Runge-Kutta son ampliamente utilizados en la resolución numérica de las ecuaciones diferenciales ordinarias. Una técnica para analizar el orden de este tipo de fórmulas consiste en desarrollar el error local en serie de potencias del paso. Si bien estos desarrollos se pueden calcular fácilmente usando la fórmula de Taylor cuando la ecuación es escalar, el número de etapas es pequeño y el orden del método es bajo (hasta orden tres, por ejemplo), en otras situaciones este procedimiento se vuelve ciertamente complejo. En este punto, la Teoría de Butcher ofrece una buena alternativa. La innovadora y sorprendente idea fundamental en este nuevo marco consiste en relacionar los términos que aparecen en los desarrollos de Taylor con un tipo de grafos denominados árboles con raíz. La obra de Butcher no solo simplifica enormemente la obtención de las condiciones que debe satisfacer un método para alcanzar un orden determinado, sino que también facilita el cálculo del desarrollo del error en forma de series de potencias del paso.

**Palabras clave:** métodos de un paso, métodos explícitos, métodos implícitos, métodos de Runge-Kutta, error local, solución exacta, solución numérica, condición de orden, tabla de Butcher, árboles con raíz, derivada de Frechet, diferencial elemental, serie de Butcher.



# Introducción

Muchos de los fenómenos que ocurren a nuestro alrededor y a los que las Matemáticas tratan de dar una explicación encuentran su forma más natural en el lenguaje de las ecuaciones diferenciales. En efecto, buena parte de las leyes físicas y de los modelos matemáticos presentes en ámbitos tan diversos como la biología, la epidemiología, la química o la economía, involucran expresiones en las que, a menudo, intervienen las derivadas de una función con respecto a ciertas variables. El principal inconveniente que presentan tales fórmulas radica en la dificultad —cuando no en la imposibilidad— para obtener soluciones exactas. Se hace, por tanto, imprescindible recurrir a algún tipo de aproximación que nos permita tener al menos una idea del comportamiento de las soluciones de un problema dado. Es en este punto donde el Análisis Numérico juega un papel fundamental, pues no solo es importante conocer que un resultado obtenido numéricamente aproxima adecuadamente la solución analítica de un problema sino que es relevante también el modo en que lo hace (idea de convergencia).

Los métodos numéricos para la resolución de problemas de valor inicial se suelen clasificar en dos familias: los llamados métodos de un paso y los conocidos como métodos multipaso. En los primeros, para obtener una aproximación de la solución en un determinado instante se utiliza únicamente la información de la solución aproximada en el instante anterior; mientras que en los segundos la solución en un instante concreto se obtiene a partir de las aproximaciones realizadas en varios pasos previos. A su vez, para calcular la solución en cada instante se distinguen dos tipos de métodos: explícitos e implícitos, en función de si la aproximación numérica se define mediante una fórmula explícita o implícita. Los segundos suelen tener mejores propiedades numéricas pero requieren un mayor esfuerzo computacional. La elección, tanto entre una fórmula de un paso y una multipaso como entre un método explícito y uno implícito, dependerá, por tanto, del problema concreto que se plantee en cada caso.

## Introducción histórica

Durante el siglo XIX y buena parte del XX, la imposibilidad de realizar cálculos masivos de forma rápida y sistemática constituyó una barrera importante para la utilización de tales fórmulas. Sin embargo, a partir de los años cincuenta, con la aparición de los primeros ordenadores, esta situación comenzó a cambiar pues fue posible plantear y resolver multitud de problemas nuevos y dar respuesta a algunas cuestiones que habían quedado pendientes.

Como primer acontecimiento relevante en este campo, podemos destacar la publicación en 1768 del conocido como método de Euler explícito, en el cual la solución se aproxima a lo largo de la recta tangente a la curva solución en cada instante. Unos años más tarde, se probó que este método alcanza orden uno, lo cual significa que para obtener buenas aproximaciones de la solución es necesario tomar pasos muy pequeños por lo que la ejecución práctica del algoritmo es muy lenta. Este hecho provocó que los matemáticos buscaran desde entonces fórmulas que proporcionasen un orden mayor.

La siguiente gran aportación al estudio de la resolución numérica de las ecuaciones diferenciales fue obra de Adams y Bashforth, quienes describieron, en 1883, el primer método multipaso. Sus trabajos fueron revisados y ampliados por Moulton (1926), Milne (1926) y Dahlquist (1958).

Por otro lado, en el marco de los métodos de un paso, Runge propuso en 1895 generalizar el método de Euler mediante reevaluaciones sucesivas de la función en puntos adecuados para conseguir métodos de orden mayor que uno. Esta idea fue continuada unos años más tarde por Heun (1900), quien formuló el primer método de orden cuatro; y por Kutta (1901), quien describió todos los métodos explícitos de cuatro etapas y orden cuatro, mediante ocho ecuaciones que debían satisfacer sus coeficientes. Kutta aventuró también dieciséis condiciones para alcanzar orden cinco aunque posteriormente se probó que en realidad son necesarias diecisiete. Los primeros avances en el contexto de las fórmulas de Runge-Kutta culminaron con la publicación en 1925 de los trabajos de Nyström, que sirvieron para corregir algunas de las apreciaciones formuladas por Kutta unos años antes y para explicar cómo aplicar esta familia de métodos a los sistemas de ecuaciones diferenciales de segundo orden.

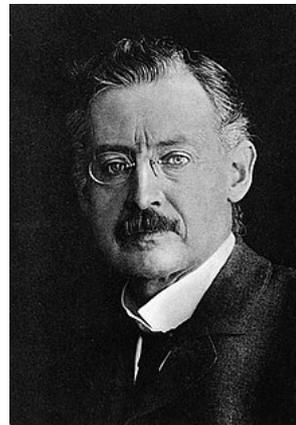
A partir de entonces el desarrollo de nuevos métodos quedó prácticamente paralizado hasta que, en 1956 y 1957, Hutta describió los métodos de orden seis y en 1963, Butcher, siguiendo los pasos de Gill (1951) y Merson (1956), publicó su innovadora teoría, en la que las soluciones analítica y numérica se expresan en forma de unas series, posteriormente llamadas series de Butcher o B-series en su honor, cuyos términos se corresponden con cierto tipo de árboles de la Teoría de Grafos. La obra de Butcher nos permite deducir las ecuaciones que deben satisfacer los coeficientes del método para alcanzar un orden determinado, al tiempo que nos proporciona una expresión del error en forma de series de potencias del paso. Estas ideas fueron revisadas y ampliadas años más tarde por los matemáticos austriacos Hairer y Wanner (1974), a quienes debemos la formalización y el desarrollo de la teoría de B-series.

Por último, cabe señalar que aunque en un inicio fueron los métodos de Runge-Kutta (RK) explícitos los que ocuparon todas las líneas de investigación en este campo, la aparición de un tipo de problemas —los llamados problemas *stiff*— ante los cuales las fórmulas explícitas se revelaron del todo ineficaces, ha motivado que, en los últimos tiempos, cobren importancia los métodos implícitos. Entre las aportaciones más recientes en el marco de las fórmulas implícitas pueden señalarse los métodos DIRK (*diagonally-implicit RK*) y SIRK (*single-implicit RK*).

## Principales autores

### Carl Runge

Carl David Tolmé Runge nació en Bremen (Alemania) el 30 de agosto de 1856 . Estudió Matemáticas y Física en la Universidad de Munich, donde coincidió con Max Planck, con quien poco después se trasladó a Berlín para asistir a las conferencias de Karl Weierstrass y Ernst Kummer. Tras unas primeras investigaciones en Álgebra y Geometría, en 1886 comenzó su labor docente en la Universidad de Hannover. En 1904, Klein consiguió que la Escuela de Matemáticas de Göttingen le ofreciera un puesto de profesor de *Matemática Aplicada* (el primero en Alemania), puesto que ocupó hasta que se retiró en 1925. Murió el 3 de enero de 1927 de un ataque al corazón.



### Martin Wilhem Kutta

Martin Wilhem Kutta nació en Pitschen (actual Polonia) el 3 de noviembre de 1867. Comenzó sus estudios en Breslau y continuó su formación en la Universidad de Munich, donde posteriormente ocupó un puesto en el departamento de Matemática Aplicada. Siguiendo los trabajos de Runge, abordó diferentes aspectos sobre la resolución numérica de ecuaciones diferenciales y publicó el conocido como *método de Runge-Kutta clásico* de cuatro etapas y orden cuatro. Destacables son también sus contribuciones en el campo de la aerodinámica. Kutta era un hombre solitario y decidió vivir apartado de la sociedad hasta su muerte el 25 de diciembre de 1944.



### John Charles Butcher

John Charles Butcher nació en Auckland (Nueva Zelanda) el 31 de marzo de 1933 . Acudió a las universidades de Auckland y Sidney. Posteriormente, se dedicó a labores de investigación en el campo del Análisis Numérico desde las universidades de Sidney y Canterbury, así como desde el Centro de Aceleración de Stamford. Influenciado por los trabajos previos de Gill y Merson, publicó en 1963 su innovadora teoría para el análisis del orden de los métodos de Runge-Kutta. En 1966 accedió al puesto de profesor de Matemática Aplicada en la Universidad de Auckland que ocuparía hasta su jubilación en 1998.



## Motivación y objetivos

Siguiendo la línea que nos marca la evolución histórica del estudio de los métodos numéricos para la resolución de ecuaciones diferenciales comenzaremos describiendo y analizando las fórmulas de Runge-Kutta explícitas para finalizar abordando el caso implícito. Se intentará en todo momento dejar constancia de todos aquellos resultados que, sin ser objeto de estudio propio de este trabajo, sí sean necesarios, no solo para facilitar el seguimiento y comprensión del tema, sino también para justificar, en la medida de lo posible, las hipótesis que en muchas ocasiones es necesario asumir a la hora de plantear un problema como los que aquí se van a tratar. Es por ello que comenzaremos presentando, sin entrar en detalle en las demostraciones, tanto el marco general de la Teoría de Ecuaciones Diferenciales como el contexto general de los métodos de un paso. Entraremos de lleno entonces en el tema principal del trabajo, que no es otro que el estudio de los métodos de Runge-Kutta, prestando especial atención al análisis del orden de dichas fórmulas. Así, podríamos señalar como objetivos fundamentales a abordar en el desarrollo de esta exposición los siguientes puntos:

- Presentar los métodos de Runge-Kutta en el marco general de los métodos de un paso.
- Obtener las condiciones que deben satisfacer los coeficientes de tales fórmulas para alcanzar un determinado orden desarrollando en serie de Taylor el error local.
- Mostrar la dificultad técnica que presenta dicho planteamiento, haciendo evidente la necesidad de buscar una alternativa que facilite ese estudio.
- Presentar, desarrollar y formalizar dicha alternativa, esto es, la Teoría de Butcher para el análisis del orden de los métodos de Runge-Kutta.
- Demostrar que efectivamente esta novedosa técnica simplifica el estudio del orden de los métodos de Runge-Kutta, presentando algunos ejemplos.

# Capítulo 1

## Preliminares

Presentamos en este primer capítulo un breve recordatorio de los conceptos y resultados fundamentales tanto en el marco de la Teoría General de Ecuaciones Diferenciales como en el contexto del Análisis Numérico de los llamados métodos de un paso. El objetivo no es otro que introducir todas aquellas herramientas necesarias para el posterior desarrollo de la Teoría de los métodos de Runge-Kutta y en particular de la Teoría de Butcher para el análisis del orden de estos. Al mismo tiempo, se pretende facilitar el seguimiento del trabajo, agilizando la lectura de esta parte, de manera que se ha optado por omitir las demostraciones de los resultados incluidos, prefiriendo, no obstante, recoger un buen número de ejemplos que los ilustren. Además, los ejemplos seleccionados estarán estrechamente relacionados con algunos de los elementos que irán apareciendo más adelante o que incluso ya han sido mencionados en la introducción, por lo que resultarán especialmente útiles.

### 1.1. Ecuaciones diferenciales

Comenzamos este repaso de la Teoría de Ecuaciones Diferenciales recordando una serie de definiciones básicas, la notación habitual y algunos resultados fundamentales.

**Definición 1.1.** Llamamos *ecuación diferencial* a cualquier ecuación que involucre una o varias funciones desconocidas (incógnitas) y algunas de sus derivadas respecto a ciertas variables independientes.

**Definición 1.2.** Una *solución* de una ecuación diferencial es una función que al ser sustituida en la ecuación la convierte en una identidad.

Si las funciones incógnitas de la ecuación diferencial dependen de una sola variable independiente se dirá que estamos ante una *ecuación diferencial ordinaria* (abreviadamente edo), mientras que si las funciones desconocidas dependen de más de una variable hablaremos de *ecuación en derivadas parciales* (abreviadamente edp).

Las ecuaciones diferenciales ordinarias, que son las que se van a tratar a lo largo de este trabajo, pueden venir escritas en *forma implícita*, esto es, mediante una expresión del tipo:

$$F(t, \mathbf{y}, \mathbf{y}', \mathbf{y}'', \dots, \mathbf{y}^{(n)}) = \mathbf{0}.$$

En ocasiones es posible despejar en la fórmula anterior la derivada de mayor orden obteniéndose lo que se conoce como *forma canónica, normal o explícita* de la ecuación:

$$\mathbf{y}^{(n)} = \mathbf{f}(t, \mathbf{y}, \mathbf{y}', \mathbf{y}'', \dots, \mathbf{y}^{(n-1)}).$$

En ambas formulaciones de la ecuación diferencial ordinaria, el número  $n$ , que denota el mayor orden de derivación que aparece, recibe el nombre de *orden* de la ecuación.

Cuando una ecuación diferencial no depende de la variable independiente se dice que es una *ecuación diferencial autónoma*. En caso contrario, esto es, cuando la variable independiente sí interviene en la ecuación, diremos que se trata de una *ecuación no autónoma*.

Una observación interesante a la que merece la pena dedicar unas pocas líneas es la siguiente: toda ecuación diferencial ordinaria de orden  $n > 1$  puede reescribirse de forma equivalente como un sistema de  $n$  ecuaciones diferenciales de orden uno.

En efecto, denotando:

$$\mathbf{y}_1 = \mathbf{y}; \quad \mathbf{y}_2 = \mathbf{y}'; \quad \mathbf{y}_3 = \mathbf{y}''; \quad \dots; \quad \mathbf{y}_n = \mathbf{y}^{(n-1)}$$

obtenemos el sistema siguiente:

$$\left\{ \begin{array}{l} \mathbf{y}'_1 = \mathbf{y}_2 \\ \mathbf{y}'_2 = \mathbf{y}_3 \\ \vdots \\ \mathbf{y}'_{n-1} = \mathbf{y}_n \\ \mathbf{y}'_n = \mathbf{f}(t, \mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-1}) \end{array} \right.$$

Es por ello que, en lo que sigue, nosotros haremos referencia únicamente a ecuaciones (o en su caso, sistemas de ecuaciones) diferenciales ordinarias de primer orden.

Sin embargo, encontrar las soluciones de una ecuación diferencial, como ya se ha señalado en el capítulo introductorio, no es en general una tarea sencilla hasta el punto de que en ocasiones no es ni tan siquiera posible. Si bien se conocen algunos métodos *clásicos* de resolución de este tipo de ecuaciones (variables separadas, factores de integración, ...) únicamente señalaremos aquí la necesidad de adjuntar condiciones que determinen una solución concreta dentro del conjunto de infinitas soluciones que puede tener una ecuación diferencial.

**Definición 1.3.** Llamamos *problema de valor inicial* (abreviadamente PVI) o *problema de Cauchy* de primer orden al sistema formado por la ecuación diferencial ordinaria  $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$  junto con la condición inicial  $\mathbf{y}(t_0) = \mathbf{y}_0$ .

Habitualmente, lo escribiremos de la siguiente manera:

$$(P) \quad \begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \mathbf{y}_0. \end{cases}$$

Un problema de valor inicial como el planteado en (P) admite una formulación equivalente mediante la siguiente expresión integral:

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{y}(s)) ds. \quad (1.1)$$

Como consecuencia, los métodos numéricos para resolver ecuaciones diferenciales ordinarias están estrechamente vinculados con las fórmulas de integración numérica y es por ello que en ocasiones haremos mención a la integración de un problema diferencial cuando, en verdad, se pretende hacer referencia a las técnicas de resolución del problema de Cauchy en cuestión.

**Ejemplo 1.1.** En la siguiente figura se recogen, en azul, varias de las infinitas soluciones de la ecuación diferencial  $y' = -y$ ; y en rojo, la única que satisface que  $y(0) = \frac{3}{8}$ .

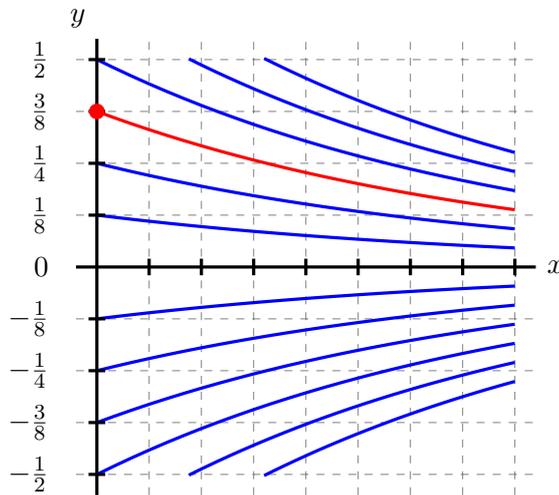


Figura 1.1: Soluciones de  $y' = y$  en el intervalo  $[0, 1]$ .

**Definición 1.4.** Una función  $\mathbf{f} : [a, b] \times \Omega \subseteq \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  es *lipschitziana* (o cumple la condición de Lipschitz) respecto de su segundo argumento en  $\Omega$  si:

$$\exists L > 0 \text{ tal que } \|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \Omega, \quad \forall t \in [a, b].$$

Se dice que  $L$  es una *constante de Lipschitz* de  $\mathbf{f}$  en  $\Omega$ .

**Teorema 1.1 (de existencia y unicidad de solución de Picard-Lipschitz).**

Sea  $f : [a, b] \times \Omega \subseteq \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  una función continua en  $[a, b] \times \Omega$  y lipschitziana respecto de su segundo argumento en  $\Omega$ . Entonces, dado  $(t_0, \mathbf{y}_0) \in [a, b] \times \Omega$  el problema de Cauchy

$$\begin{cases} \mathbf{y}' &= \mathbf{f}(t, \mathbf{y}), \\ \mathbf{y}(t_0) &= \mathbf{y}_0. \end{cases}$$

tiene una única solución definida en el intervalo  $[a, b]$ .

**Ejemplo 1.2.** La condición de continuidad por sí sola garantiza la existencia de solución (es el *Teorema de Peano o Cauchy-Peano*) pero para poder asegurar también la unicidad es necesario que la función  $f$  sea lipschitziana respecto de su segundo argumento. Para ilustrar estos resultados consideremos, sobre el cuadrado  $[-1, 1] \times [-1, 1]$ , el problema dado por:

$$\begin{cases} y' &= 3y^{\frac{2}{3}}, \\ y(0) &= 0. \end{cases}$$

En este caso, tenemos que  $f(t, y) = 3y^{\frac{2}{3}}$  que es una función continua pero no lipschitziana respecto de su segundo argumento en el dominio  $[-1, 1] \times [-1, 1]$  al cumplirse que:

$$\frac{|f(t, y) - f(t, 0)|}{|y - 0|} = \frac{3y^{\frac{2}{3}}}{y} = \frac{3}{y^{\frac{1}{3}}} \xrightarrow{y \rightarrow 0} \infty.$$

Además, es fácil comprobar que tanto  $y(t) = 0$  como  $y(t) = t^3$  son dos soluciones válidas del problema anterior. Luego tenemos un ejemplo en el que existe solución pero esta no es única.

**Observación 1.1.** Antes de entrar en el análisis numérico de las ecuaciones diferenciales ordinarias conviene recordar también el concepto de *problema bien planteado*. Así, diremos que un problema está bien planteado cuando presente las siguientes tres características:

- El problema tiene solución.
- Dicha solución es única.
- Además, la solución depende de los datos con continuidad.

Precisar la tercera condición requiere la utilización de normas adecuadas en los espacios de funciones que se consideran en cada problema concreto y resulta imprescindible para conseguir un “buen” resultado numérico, entendido este como una buena aproximación de la solución exacta. Dicha dependencia continua se traduce en que pequeños cambios en los datos del problema no conllevan grandes cambios en las soluciones obtenidas; mientras que si no disponemos de esa condición, puede ocurrir que pequeños errores, ya sean debidos a las limitaciones de cálculo del ordenador o directamente por el método considerado, podrían conducir a soluciones totalmente distintas, por lo que la aproximación numérica obtenida carecería de sentido y debería ser desechada.

## 1.2. Métodos de un paso

Consideraremos en lo que sigue el problema de valor inicial dado por:

$$(P) \quad \begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \boldsymbol{\eta}, \end{cases}$$

donde supondremos que  $\mathbf{f} : I \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  es una función continua y lipschitziana respecto de su segundo argumento (con constante de Lipschitz  $L > 0$  para alguna norma de  $\mathbb{R}^m$ ) con el objetivo de garantizar la existencia y unicidad de solución del problema (P) en base al Teorema 1.1 de Picard-Lipschitz.

Tomaremos además una partición  $\mathcal{P} = \{t_0 < t_1 < \dots < t_N = t_0 + T\}$  del intervalo  $I$  en  $N$  subintervalos de longitud  $h_n = t_{n+1} - t_n$  con  $n = 0, 1, \dots, N-1$ , y denotaremos por  $h$  el diámetro de dicha partición, esto es,

$$h := \|\mathcal{P}\| = \max\{h_n : n = 0, 1, \dots, N-1\}.$$

**Definición 1.5.** Un *método de un paso* para resolver numéricamente el problema (P) es un algoritmo que para cada partición  $\mathcal{P}$  permite calcular una aproximación  $\mathbf{y}_{n+1}$  del valor exacto  $\mathbf{y}(t_{n+1})$  utilizando únicamente la aproximación  $\mathbf{y}_n$  de  $\mathbf{y}(t_n)$ , el instante  $t_n$  y el paso  $h_n$ .

Escribiremos el método de la siguiente manera:

$$(M) \quad \begin{cases} \mathbf{y}_0 = \boldsymbol{\eta}_h, \\ \mathbf{y}_{n+1} = \mathbf{y}_n + h_n \Phi(t_n, \mathbf{y}_n, h_n), & n = 0, 1, \dots, N-1. \end{cases}$$

La función  $\Phi$ , que depende de la ecuación diferencial considerada a través de la función  $\mathbf{f}$ , recibe el nombre de *función incremento* del método y en lo que sigue supondremos que está bien definida y es continua en  $I \times \mathbb{R}^m \times [0, h^*]$ . Si la función  $\Phi$  se puede expresar en términos de  $\mathbf{f}$  el método se dirá *explícito*, mientras que si, por el contrario,  $\Phi$  está relacionada implícitamente con  $\mathbf{f}$ , el método se dirá *implícito*.

**Ejemplo 1.3.** Veamos algunos ejemplos muy conocidos de métodos explícitos e implícitos:

1. Método de Euler explícito:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \mathbf{f}(t_n, \mathbf{y}_n) \implies \Phi(t, \mathbf{y}, h) = \mathbf{f}(t, \mathbf{y}).$$

2. Método de Euler implícito:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}) \implies \Phi(t, \mathbf{y}, h) = \mathbf{f}(t + h, \mathbf{y} + h \Phi(t, \mathbf{y}, h)).$$

3. Método de Euler modificado (explícito):

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \mathbf{f}\left(t_n + \frac{h_n}{2}, \mathbf{y}_n + \frac{h_n}{2} \mathbf{f}(t_n, \mathbf{y}_n)\right) \implies \Phi(t, \mathbf{y}, h) = \mathbf{f}\left(t + \frac{h}{2}, \mathbf{y} + \frac{h}{2} \mathbf{f}(t, \mathbf{y})\right).$$

Como ya se señaló previamente, a la hora de analizar un resultado obtenido numéricamente no solo es interesante estudiar si este aproxima adecuadamente la solución analítica del problema considerado (idea de convergencia) sino que es relevante también el modo en que lo hace. Es por ello que prestaremos especial atención tanto a la rapidez con la que se alcanza esa buena aproximación (idea de orden de convergencia) como al comportamiento de esta frente a posibles cambios en los datos del problema (idea de estabilidad).

**Definición 1.6.** Si  $\mathbf{y}(\cdot)$  es la solución exacta del problema  $(P)$  e  $\{\mathbf{y}_n\}_{n=0}^N$  es una solución obtenida numéricamente, entonces se define el *error global* en el instante  $t_n$  como:

$$\mathbf{e}_n := \mathbf{y}(t_n) - \mathbf{y}_n.$$

**Definición 1.7.** Un método de un paso  $(M)$  se dice *convergente* con el problema  $(P)$  si:

$$\max\{\|\mathbf{e}_n\| : n = 1, 2, \dots, N\} \longrightarrow 0 \quad \text{cuando } h \rightarrow 0 \text{ y } \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\| \rightarrow 0.$$

Además, diremos que el método  $(M)$  es *convergente de orden (al menos)  $p > 0$*  si:

$$\max\{\|\mathbf{e}_n\| : n = 1, 2, \dots, N\} = \mathcal{O}(h^p) \quad \text{cuando } h \rightarrow 0 \text{ y } \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\| = \mathcal{O}(h^p).$$

**Definición 1.8.** Dada una colección de coeficientes reales,  $\{\mathbf{u}_n\}_{n=0}^N$ , podemos definir a partir de ella un operador en diferencias, que será denotado por  $L_h$ , mediante la siguiente expresión:

$$L_h(\mathbf{u}_n) := \mathbf{u}_{n+1} - \mathbf{u}_n - h_n \Phi(t_n, \mathbf{u}_n, h_n).$$

Entonces, diremos que un método de un paso  $(M)$  es *estable* cuando lo sea su operador asociado, esto es, cuando existan dos constantes  $M > 0$  y  $h^* > 0$  de forma tal que para cualquier partición  $\mathcal{P}$  del intervalo  $I$  con diámetro  $h \leq h^*$  y para cualesquiera tres colecciones de números reales,  $\{\mathbf{u}_n\}_{n=0}^N$ ,  $\{\mathbf{v}_n\}_{n=0}^N$  y  $\{\mathbf{w}_n\}_{n=0}^N$ , que verifiquen para  $n = 0, 1, \dots, N-1$ :

$$\begin{aligned} \triangleright L_h(\mathbf{u}_n) &= 0, \quad \text{es decir, } \mathbf{u}_{n+1} = \mathbf{u}_n + h_n \Phi(t_n, \mathbf{u}_n, h_n), \\ \triangleright L_h(\mathbf{v}_n) &= \mathbf{w}_n, \quad \text{es decir, } \mathbf{v}_{n+1} = \mathbf{v}_n + h_n \Phi(t_n, \mathbf{v}_n, h_n) + \mathbf{w}_n, \end{aligned}$$

se cumple

$$\|\mathbf{v}_n - \mathbf{u}_n\| \leq M \left( \|\mathbf{v}_0 - \mathbf{u}_0\| + \sum_{i=0}^{n-1} \|\mathbf{w}_i\| \right), \quad n = 1, 2, \dots, N.$$

El concepto de estabilidad se traduce en que pequeñas modificaciones en los datos del problema considerado (generalmente, errores de redondeo debidos a las limitaciones computacionales del ordenador) no causan grandes cambios en las soluciones numéricas obtenidas. Como puede comprobarse, esta noción está estrechamente relacionada con la condición de dependencia continua de la solución respecto a los datos del problema mencionada al hablar de la naturaleza de un problema bien planteado.

**Definición 1.9.** Para cada solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial del problema  $(P)$  se llama *error local* (o de truncamiento o de consistencia) en el instante  $t_n$  con paso  $h_n$  a:

$$\varepsilon(t_n, h_n; \mathbf{y}) := \mathbf{y}(t_n + h_n) - (\mathbf{y}(t_n) + h_n \Phi(t_n, \mathbf{y}(t_n), h_n)).$$

Es decir, el error local es la diferencia entre la solución exacta de la edo evaluada en  $t_n + h_n$  y la solución aproximada en  $t_n + h_n$  partiendo del dato exacto en  $t_n$ . En otras palabras, es el error inevitable aunque se dispusiese del dato exacto.

**Definición 1.10.** Un método de un paso  $(M)$  es *consistente* con el problema  $(P)$  si para cualquier solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial se cumple

$$\sum_{n=0}^{N-1} \|\varepsilon(t_n, h_n; \mathbf{y})\| \rightarrow 0, \quad \text{cuando } h \rightarrow 0.$$

Además, el método se dirá *de orden* (al menos)  $p > 0$  si para cualquier solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial del problema  $(P)$  se verifica

$$\sum_{n=0}^{N-1} \|\varepsilon(t_n, h_n; \mathbf{y})\| = \mathcal{O}(h^p), \quad \text{cuando } h \rightarrow 0.$$

Los conceptos de convergencia, estabilidad y consistencia que acabamos de introducir están relacionados entre sí de la siguiente manera.

**Proposición 1.1 (Estabilidad+Consistencia  $\implies$  Convergencia).**

*Supongamos que el método de un paso considerado  $(M)$  es estable y consistente (respectivamente, de orden  $p$ ) con el problema  $(P)$ . Entonces el método en cuestión es convergente (respectivamente, convergente de orden  $p$ ).*

Como consecuencia del resultado anterior, nuestros esfuerzos se centran ahora en presentar las condiciones que garantizan las propiedades de estabilidad y consistencia del método.

**Proposición 1.2.** *Un método de un paso  $(M)$  es estable si su función incremento  $\Phi$  es lipschitziana respecto de su segundo argumento.*

**Proposición 1.3.** *Un método de un paso  $(M)$  es consistente con el problema  $(P)$  si y solo si se cumple la siguiente condición:*

$$\Phi(t, \mathbf{y}, 0) = f(t, \mathbf{y}) \quad \forall (t, \mathbf{y}) \in I \times \mathbb{R}^m.$$

**Proposición 1.4.** *Un método de un paso  $(M)$  tal que su error local,  $\varepsilon(t_n, h_n; \mathbf{y})$ , se comporta como una  $\mathcal{O}(h^{p+1})$  alcanza orden  $p$ , esto es, se satisface que*

$$\sum_{n=0}^{N-1} \|\varepsilon(t_n, h_n; \mathbf{y})\| = \mathcal{O}(h^p),$$

para cualquier solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial del problema  $(P)$  de clase  $\mathcal{C}^{p+1}$ .



## Capítulo 2

# Métodos de Runge-Kutta explícitos

En este capítulo se presentan los métodos de Runge-Kutta explícitos como un caso particular de los métodos de un paso introducidos en la sección anterior. Además, nos ocuparemos de estudiar sus propiedades de convergencia, estabilidad y consistencia, prestando especial atención a la obtención de las condiciones que deben satisfacer sus coeficientes para alcanzar un orden determinado. Se pretende poner de manifiesto la necesidad de introducir un nuevo procedimiento para analizar el orden del método, dejando al descubierto la inviabilidad de la técnica habitual, que consiste en desarrollar en serie de Taylor el error local.

### 2.1. Definición y primeras propiedades

**Definición 2.1.** Un *método de Runge-Kutta explícito de  $q$  etapas* (abreviado RKE) para resolver numéricamente el problema ( $P$ ) es un método de un paso

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \Phi(t_n, \mathbf{y}_n, h_n),$$

cuya función incremento es de la forma

$$\Phi(t, \mathbf{y}, h) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, h),$$

donde  $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_q$  vienen dados por:

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t, \mathbf{y}), \\ \mathbf{k}_2 &= \mathbf{f}(t + c_2 h, \mathbf{y} + h a_{21} \mathbf{k}_1), \\ \mathbf{k}_3 &= \mathbf{f}(t + c_3 h, \mathbf{y} + h a_{31} \mathbf{k}_1 + h a_{32} \mathbf{k}_2), \\ &\vdots \\ \mathbf{k}_q &= \mathbf{f}(t + c_q h, \mathbf{y} + h a_{q1} \mathbf{k}_1 + h a_{q2} \mathbf{k}_2 + \dots + h a_{qq-1} \mathbf{k}_{q-1}), \end{aligned}$$

con  $\{a_{ij}\}_{i,j=1}^q$ ,  $\{b_i\}_{i=1}^q$  y  $\{c_i\}_{i=1}^q$  tres colecciones de números reales que caracterizan el método.

Asumiremos que  $b_q \neq 0$  y  $c_2 \neq 0$ , pues, de lo contrario, el método sería equivalente a uno con menos etapas, esto es, sería un método reducible. En efecto, si la fórmula tiene  $q$  etapas y  $b_q = 0$ , la última etapa es innecesaria (porque  $\mathbf{k}_q$  no interviene) y por tanto el método equivale a otro con  $q - 1$  etapas. La razón por la que  $c_2 \neq 0$  se verá más adelante.

Los métodos de Runge-kutta explícitos son, efectivamente, métodos de un paso pues en la determinación de la solución en un instante determinado se utiliza únicamente la información relativa a la solución obtenida en el instante inmediatamente anterior. Por ejemplo, para calcular la aproximación  $\mathbf{y}_{n+1}$  del valor de la solución exacta en el instante  $t_{n+1}$  solo es necesario conocer la aproximación  $\mathbf{y}_n$  de la solución exacta en el instante  $t_n$ , el propio valor de  $t_n$  y el paso  $h_n$  puesto que con estos tres elementos podemos obtener los valores de  $\mathbf{k}_i$  con  $i = 1, 2, \dots, q$ ; utilizar estos, a su vez, para determinar  $\Phi$ ; y por último, calcular  $\mathbf{y}_{n+1}$  utilizando  $\mathbf{y}_n$ ,  $h_n$  y el valor de  $\Phi$ . Así, los resultados presentados en el Capítulo 1 para el caso más general de métodos de un paso pueden aplicarse a las fórmulas de Runge-Kutta.

Es habitual que los coeficientes del método aparezcan recogidos en forma de tabla:

$$\begin{array}{c|cccccc}
 0 & 0 & & & & \\
 c_2 & a_{21} & 0 & & & \\
 c_3 & a_{31} & a_{32} & 0 & & \\
 \vdots & \vdots & \vdots & & \ddots & \\
 c_q & a_{q1} & a_{q2} & \dots & a_{q,q-1} & 0 \\
 \hline
 & b_1 & b_2 & \dots & b_{q-1} & b_q
 \end{array}$$

Dicha tabla recibe el nombre de *tabla de Butcher* del método y algunos autores se refieren a la matriz estrictamente triangular inferior  $A = (a_{ij})_{i,j=1}^q \in \mathcal{M}_{q \times q}(\mathbb{R})$  como matriz de coeficientes, mientras que a los vectores de  $\mathbb{R}^q$ ,  $\mathbf{c} = (c_i)_{i=1}^q$  y  $\mathbf{b} = (b_i)_{i=1}^q$ , los llaman *vectores de nodos y pesos*, respectivamente. Esta terminología es debida a la estrecha relación que existe entre estos coeficientes y aquellos que aparecen en las fórmulas de integración numérica. Con esta notación, la tabla de Butcher anterior puede escribirse abreviadamente como sigue:

$$\begin{array}{c|c}
 \mathbf{c} & A \\
 \hline
 & \mathbf{b}^T
 \end{array}$$

**Ejemplo 2.1.** Métodos de Runge (1895) y Kutta (1905) de cuatro etapas.

$$\begin{array}{c|cccc}
 0 & 0 & & & \\
 \frac{1}{2} & \frac{1}{2} & 0 & & \\
 1 & 0 & 1 & 0 & \\
 1 & 0 & 0 & 1 & 0 \\
 \hline
 & \frac{1}{6} & \frac{2}{3} & 0 & \frac{1}{6}
 \end{array}
 \qquad
 \begin{array}{c|cccc}
 0 & 0 & & & \\
 \frac{1}{2} & \frac{1}{2} & 0 & & \\
 \frac{1}{2} & 0 & \frac{1}{2} & 0 & \\
 1 & 0 & 0 & 1 & 0 \\
 \hline
 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
 \end{array}$$

## Formulaciones alternativas de los métodos de Runge-Kutta explícitos

Existen varias formulaciones equivalentes de los métodos de Runge-Kutta explícitos que, en lo que sigue, utilizaremos indistintamente según nos convenga en cada momento.

Una formulación alternativa a la que aparece recogida en la propia definición de los métodos de Runge-Kutta explícitos y que nos facilitará algunas explicaciones posteriores es la siguiente:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}), \quad (2.1)$$

donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^{i-1} a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q.$$

En este caso, es fácil ver que ambas formulaciones son equivalentes si interpretamos:

$$\mathbf{k}_i = \mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}).$$

Otra opción, que cobrará especial relevancia a la hora de obtener el desarrollo de Taylor del error local en el marco de la Teoría de Butcher, para denotar una fórmula de Runge-Kutta explícita con  $q$  etapas es la siguiente:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \sum_{i=1}^q b_i \mathbf{g}_i, \quad (2.2)$$

con

$$\mathbf{g}_i = h_n \mathbf{f}(t_n + c_i h_n, \mathbf{y}_n + \sum_{j=1}^{i-1} a_{ij} \mathbf{g}_j), \quad i = 1, 2, \dots, q.$$

Resulta evidente que esta expresión es equivalente a la dada en la definición puesto que  $\mathbf{g}_i = h_n \mathbf{k}_i$  y en consecuencia tenemos que las tres formulaciones señaladas son equivalentes.

No obstante, cabe destacar que a la hora de implementar computacionalmente fórmulas de Runge-Kutta suele optarse por la primera de ellas, esto es, la de la propia definición.

## Problemas autónomos y no autónomos

En el marco general de la Teoría de Ecuaciones Diferenciales Ordinarias solamente se aborda la resolución de problemas de valor inicial autónomos. Recordemos por qué esto es así y veamos en qué condiciones esta reducción de casos se traslada también a la resolución numérica de problemas diferenciales mediante fórmulas de Runge-Kutta.

Es conocido que cualquier problema de valor inicial no autónomo puede transformarse en uno autónomo sin más que considerar la variable independiente como una nueva incógnita.

En efecto, si  $\mathbf{y}(t)$  es la solución del problema:

$$(P1) \quad \begin{cases} \mathbf{y}' &= \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) &= \boldsymbol{\eta} \in \mathbb{R}^m, \end{cases}$$

y denotamos  $\mathbf{Y}(t) = (t, \mathbf{y}(t))^T$ , se verifica que:

$$\begin{aligned} \triangleright \mathbf{Y}'(t) &= (1, \mathbf{y}'(t))^T = (1, \mathbf{f}(t, \mathbf{y}))^T. \\ \triangleright \mathbf{Y}(t_0) &= (t_0, \mathbf{y}(t_0))^T = (t_0, \boldsymbol{\eta})^T. \end{aligned}$$

Por tanto,  $\mathbf{Y}(t)$  es la solución del problema:

$$(P2) \quad \begin{cases} \mathbf{Y}' &= \mathbf{F}(\mathbf{Y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{Y}(t_0) &= (t_0, \boldsymbol{\eta})^T, \end{cases}$$

donde  $\mathbf{F}(\mathbf{Y}) = (1, \mathbf{f}(t, \mathbf{y}))^T$ . Y de la misma manera, si  $\mathbf{Y}(t)$  es la solución del problema (P2) y denotamos  $\mathbf{Y}(t) = (y_0(t), y_1(t), \dots, y_m(t))^T$ , entonces se tiene que  $y_0(t) = t$  y además  $\mathbf{y}(t) = (y_1(t), \dots, y_m(t))^T$  es la solución de (P1).

Esto quiere decir que, desde el punto de vista del Análisis Matemático, bastaría considerar problemas autónomos a costa, eso sí, de aumentar el número de ecuaciones en una unidad.

**Ejemplo 2.2.** El problema de valor inicial no autónomo dado por:

$$(P1) \quad \begin{cases} y' &= y \cos(t), & t \in I = [0, 2\pi], \\ y(0) &= 1, \end{cases}$$

puede convertirse en un sistema autónomo equivalente considerando:

$$(P2) \quad \begin{cases} y'_0 &= 1, \\ y'_1 &= y_1 \cos(t), & t \in I = [0, 2\pi], \\ y_0(0) &= 0, \\ y_1(0) &= 1. \end{cases}$$

Sus soluciones son, respectivamente,  $y(t) = e^{\text{sen}(t)}$  e  $\mathbf{Y}(t) = (t, e^{\text{sen}(t)})^T$ .

Parece razonable esperar que esta relación se mantenga para las soluciones numéricas de ambos problemas obtenidas mediante un método de Runge-Kutta explícito de  $q$  etapas.

Por consiguiente, si  $\{\mathbf{y}_n\}_{n=0}^N$  es una solución aproximada del problema no autónomo asociada a una partición  $\mathcal{P} = \{t_0, t_1, \dots, t_N\}$  y se cumple que la solución numérica del problema autónomo,  $\{\mathbf{Y}_n\}_{n=0}^N$ , verifica que  $\mathbf{Y}_n = (t_n, \mathbf{y}_n)^T$  para cierto  $n \in \{0, 1, \dots, N\}$ , entonces nos interesa esta relación se mantenga en el instante siguiente, esto es, buscamos que se cumpla que  $\mathbf{Y}_{n+1} = (t_{n+1}, \mathbf{y}_{n+1})^T$ . Veamos en qué se traduce esta afirmación.

Utilizando la notación de la formulación (2.1) de un método de Runge-Kutta explícito:

- Para el problema no autónomo ( $P1$ ) resulta:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}),$$

donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^{i-1} a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q.$$

- Análogamente, para el problema autónomo ( $P2$ ) obtenemos:

$$\mathbf{Y}_{n+1} = \mathbf{Y}_n + h_n \sum_{i=1}^q b_i \mathbf{F}(\mathbf{Y}_{n,i}),$$

donde

$$\mathbf{Y}_{n,i} = \mathbf{Y}_n + h_n \sum_{j=1}^{i-1} a_{ij} \mathbf{F}(\mathbf{Y}_{n,j}), \quad i = 1, 2, \dots, q.$$

Luego si tenemos que  $\mathbf{Y}_n = \begin{pmatrix} t_n \\ \mathbf{y}_n \end{pmatrix}$  para cierto  $n$  entonces se cumple que:

- $\mathbf{Y}_{n,1} = \begin{pmatrix} t_n \\ \mathbf{y}_n \end{pmatrix},$
- $\mathbf{Y}_{n,2} = \begin{pmatrix} t_n \\ \mathbf{y}_n \end{pmatrix} + h_n a_{21} \begin{pmatrix} 1 \\ \mathbf{f}(t_n, \mathbf{y}_n) \end{pmatrix} = \begin{pmatrix} t_n + h_n a_{21} \\ \mathbf{y}_n + h_n a_{21} \mathbf{f}(t_n, \mathbf{y}_n) \end{pmatrix},$
- $\mathbf{Y}_{n,3} = \begin{pmatrix} t_n + h_n a_{31} + h_n a_{32} \\ \mathbf{y}_n + h_n a_{31} \mathbf{f}(t_n, \mathbf{y}_n) + h_n a_{32} \mathbf{f}(t_n + h_n a_{21}, \mathbf{y}_n + h_n a_{21} \mathbf{f}(t_n, \mathbf{y}_n)) \end{pmatrix},$
- $\vdots$

lo que nos permite obtener lo siguiente:

$$\mathbf{Y}_{n+1} = \begin{pmatrix} t_n + h_n \sum_{i=1}^q b_i \\ \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}\left(t_n + h_n \sum_{j=1}^{i-1} a_{ij}, \mathbf{y}_{n,i}\right) \end{pmatrix}.$$

Así, deducimos que para que verifique que:

$$\mathbf{Y}_{n+1} = \begin{pmatrix} t_{n+1} \\ \mathbf{y}_{n+1} \end{pmatrix} = \begin{pmatrix} t_n + h_n \\ \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + h_n c_i, \mathbf{y}_{n,i}) \end{pmatrix},$$

deben cumplirse entonces las dos siguientes condiciones:

$$\sum_{i=1}^q b_i = 1,$$

$$\sum_{j=1}^{i-1} a_{ij} = c_i, \quad i = 1, 2, \dots, q.$$

Estas condiciones, que ya habían sido asumidas por Kutta en la primera formulación de estos métodos, se traducen en que los puntos en los que se evalúa la función  $\mathbf{f}$  son aproximaciones de primer orden de la solución. Además, simplifican enormemente la deducción de las condiciones que deben satisfacer los coeficientes de una fórmula de Runge-Kutta para alcanzar un determinado orden. En consecuencia, en todo lo que sigue, salvo indicación expresa en contra, se abordará únicamente el estudio de la resolución de problemas diferenciales mediante fórmulas de Runge-Kutta explícitas cuyos coeficientes verifican las dos condiciones anteriores.

Es habitual referirse a la segunda expresión con el nombre de *condición de suma de filas*, puesto que se traduce en que la suma de los elementos de cada fila de la matriz  $A$  coincide con el valor de la misma fila del vector  $\mathbf{c}$ . Además, cumpliéndose esta condición, si  $c_2 = 0$ , en el caso de métodos explícitos de  $q$  etapas, se deduce que  $a_{21} = c_2 = 0$ , luego las dos primeras filas de la tabla de Butcher son iguales y por tanto  $\mathbf{k}_1 = \mathbf{k}_2$ , por lo que el método equivale a otro con  $q - 1$  etapas. Por este motivo, asumíamos, en su momento, que  $c_2 \neq 0$ .

Por todo lo anterior, solo será necesario considerar la aplicabilidad de tales métodos a los siguientes tipos de problemas diferenciales:

- Caso escalar ( $m = 1$ ), que a su vez puede ser:
  - Escalar autónomo:  $y' = f(y)$  con  $f : \mathbb{R} \rightarrow \mathbb{R}$ .
  - Escalar no autónomo:  $y' = f(t, y)$  con  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ .  
Con el razonamiento anterior puede convertirse en un problema autónomo a costa de aumentar en una unidad su dimensión, abandonando, por tanto, el caso escalar.
- Caso vectorial ( $m > 1$ ):
  - Se reduce al problema autónomo:  $\mathbf{y}' = \mathbf{f}(\mathbf{y})$  con  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ .

## 2.2. Estabilidad y consistencia de los métodos RK explícitos

En la sección anterior hemos introducido las fórmulas de Runge-Kutta como un caso particular de los métodos de un paso y ahora nos proponemos utilizar esta condición para estudiar sus propiedades en base a los resultados establecidos anteriormente para este tipo de métodos. En particular, veremos en qué condiciones los métodos de Runge-Kutta son convergentes y para ello, por la Proposición 1.1, basta estudiar sus propiedades de estabilidad y consistencia.

**Proposición 2.1.** *La función incremento de un método de Runge-Kutta explícito de  $q$  etapas es lipschitziana respecto de su segundo argumento en su dominio de definición y continuidad, esto es, en el dominio  $I \times \mathbb{R}^m \times [0, h^*]$ .*

*Demostración.* Veamos que la función incremento del método,

$$\phi(t, \mathbf{y}, h) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, h),$$

es lipschitziana respecto de su segundo argumento en el dominio  $I \times \mathbb{R}^m \times [0, h^*]$ .

Bastará con probar que cada  $\mathbf{k}_i$  es lipschitziana respecto de su segundo argumento pues

$$\| \phi(t, \mathbf{y}, h) - \phi(t, \mathbf{z}, h) \| \leq \sum_{i=1}^q |b_i| \| \mathbf{k}_i(t, \mathbf{y}, h) - \mathbf{k}_i(t, \mathbf{z}, h) \|.$$

- Veamos que  $\mathbf{k}_1$  es lipschitziana con constante  $L_1$ :

$$\| \mathbf{k}_1(t, \mathbf{y}, h) - \mathbf{k}_1(t, \mathbf{z}, h) \| = \| \mathbf{f}(t, \mathbf{y}) - \mathbf{f}(t, \mathbf{z}) \| \leq L \| \mathbf{y} - \mathbf{z} \|.$$

Luego basta tomar  $L_1 = L$ .

- Hacemos lo mismo con  $\mathbf{k}_2$ :

$$\begin{aligned} & \| \mathbf{k}_2(t, \mathbf{y}, h) - \mathbf{k}_2(t, \mathbf{z}, h) \| \\ &= \| \mathbf{f}(t + c_2 h, \mathbf{y} + h a_{21} \mathbf{k}_1(t, \mathbf{y}, h)) - \mathbf{f}(t + c_2 h, \mathbf{z} + h a_{21} \mathbf{k}_1(t, \mathbf{z}, h)) \| \\ &\leq L (\| \mathbf{y} - \mathbf{z} \| + h |a_{21}| \| \mathbf{k}_1(t, \mathbf{y}, h) - \mathbf{k}_1(t, \mathbf{z}, h) \|) \\ &\leq L (\| \mathbf{y} - \mathbf{z} \| + h |a_{21}| L_1 \| \mathbf{y} - \mathbf{z} \|) \leq L (1 + h^* |a| L_1) \| \mathbf{y} - \mathbf{z} \| \end{aligned}$$

Luego  $\mathbf{k}_2$  es lipschitziana con constante  $L_2 = L(1 + h^* |a| L_1)$  donde  $|a| = \max_{i,j} |a_{ij}|$ .

- Reiterando el argumento deducimos que si  $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{i-1}$  son lipschitzianas respecto a su segundo argumento con constantes  $L_1, L_2, \dots, L_{i-1}$ , respectivamente, entonces  $\mathbf{k}_i$  también es lipschitziana respecto a su segundo argumento con contante  $L_i$ . En efecto:

$$\begin{aligned} & \| \mathbf{k}_i(t, \mathbf{y}, h) - \mathbf{k}_i(t, \mathbf{z}, h) \| \\ &\leq L \left( \| \mathbf{y} - \mathbf{z} \| + h \sum_{j=1}^{i-1} |a_{ij}| \| \mathbf{k}_j(t, \mathbf{y}, h) - \mathbf{k}_j(t, \mathbf{z}, h) \| \right) \\ &\leq L \left( 1 + h^* |a| \sum_{j=1}^{i-1} L_j \right) \| \mathbf{y} - \mathbf{z} \|. \end{aligned}$$

En este caso será suficiente tomar  $L_i = L \left( 1 + h^* |a| \sum_{j=1}^{i-1} L_j \right)$ . □

**Corolario 2.1.** *Todo método de Runge-Kutta explícito de  $q$  etapas es estable.*

*Demostración.* Basta tener en cuenta las proposiciones 1.2 y 2.1 anteriores.  $\square$

**Proposición 2.2.** *Un método Runge-Kutta explícito de  $q$  etapas es consistente si y solo si*

$$\sum_{i=1}^q b_i = 1.$$

*Demostración.* De acuerdo con la Proposición 1.3 el método será consistente si y solo si se cumple la condición de que  $\Phi(t, \mathbf{y}, 0) = f(t, \mathbf{y}) \quad \forall (t, \mathbf{y}) \in I \times \mathbb{R}^m$ . En este caso:

$$\Phi(t, \mathbf{y}, 0) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, 0) = \sum_{i=1}^q b_i \mathbf{f}(t, \mathbf{y}) = \mathbf{f}(t, \mathbf{y}) \sum_{i=1}^q b_i.$$

Por consiguiente,  $\Phi(t, \mathbf{y}, 0) = f(t, \mathbf{y})$  si y solo si  $\sum_{i=1}^q b_i = 1$ .  $\square$

En consecuencia, como ya habíamos señalado que asumíamos dicha condición, podemos afirmar que los métodos que nosotros vamos a considerar son consistentes. Más aún, en base a la Proposición 1.1, podemos concluir que las fórmulas que usaremos son, además, convergentes.

### 2.3. Estudio del orden mediante desarrollos de Taylor

Nuestro interés se centra ahora en el estudio de las condiciones que deben satisfacer los coeficientes de los métodos de Runge-Kutta explícitos para alcanzar el mayor orden posible. Con este fin obtendremos el desarrollo de Taylor del error local y deduciremos las condiciones que deben verificarse entonces para que tal desarrollo sea una  $\mathcal{O}(h^{p+1})$  con  $p$  lo más elevado posible, lo cual implicaría, según la Proposición 1.4, que el orden del método fuese  $p$ . Aunque esta forma de proceder no resulta especialmente complicada cuando la ecuación es escalar, el número de etapas es pequeño y el orden del método es “bajo” (hasta orden tres, por ejemplo), rápidamente, nos daremos cuenta que, a medida que el número de etapas de la fórmula considerada va en aumento, esta técnica se vuelve demasiado tediosa y hasta cierto punto compleja, por lo que nos va a interesar buscar un procedimiento alternativo que nos facilite este estudio. Seguramente, la dificultad de este primer método se deba a aquello que decía S.Gill, uno de los precursores de la investigación de Butcher, de que «*resulta muy complicado mantener la mente fría cuando se trata de calcular varias derivadas sucesivas*».

En esta sección se incluyen de la manera más detallada posible los cálculos necesarios para obtener las condiciones de orden asociadas a los métodos de Runge-Kutta explícitos de una, dos y tres etapas cuando se aplican a problemas escalares. Con ellos se pretende poner de manifiesto la dificultad técnica que conlleva este proceso, al tiempo que se intenta convencer al lector de la necesidad de desarrollar un procedimiento alternativo que facilite el análisis del orden de esta familia de métodos de resolución numérica de ecuaciones diferenciales.

### 2.3.1. Métodos de Runge-Kutta explícitos de una etapa

Debido a la simplicidad de este caso, no supondremos la condición de suma de filas y, en consecuencia, la tabla de Butcher asociada al método es de la forma:

$$\begin{array}{c|c} c & 0 \\ \hline & b \end{array}$$

El método se escribe como:

$$\begin{aligned} k_1 &= f(t_n + ch_n, y_n), \\ y_{n+1} &= y_n + h_n b k_1 = y_n + h_n b f(t_n + ch_n, y_n). \end{aligned}$$

y su error local es:

$$\varepsilon(t_n, h_n; y) = y(t_n + h_n) - y(t_n) - h_n b f(t_n + ch_n, y(t_n)).$$

Para escribir el error local en serie de potencias del paso, sustituiremos  $y(t_n + h_n)$  por su desarrollo de Taylor centrado en  $t_n$  y haremos lo propio con  $f(t_n + ch_n, y_n)$ , centrandlo en este caso el desarrollo en el punto  $(t_n, y(t_n))$  pero omitiremos los argumentos por brevedad y comodidad en la notación. En ambas expresiones aparecerán términos con  $y', y'', \dots$  que se relacionan con la función  $f$  y sus derivadas sin más que aplicar de forma sucesiva la Regla de la Cadena en la ecuación diferencial.

$$\begin{aligned} \triangleright y(t_n + h_n) &= y(t_n) + h_n y'(t_n) + \frac{h_n^2}{2!} y''(t_n) + \mathcal{O}(h_n^3) \\ &= y(t_n) + h_n f + \frac{h_n^2}{2!} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) + \mathcal{O}(h_n^3). \\ \triangleright f(t_n + ch_n, y(t_n)) &= f + \frac{\partial f}{\partial t} h_n + \mathcal{O}(h_n^2). \end{aligned}$$

Sustituyendo estos desarrollos en la expresión del error local obtenemos:

$$\varepsilon(t_n, h_n; y) = h_n(1 - b)f + h_n^2 \left( \left( \frac{1}{2} - bc \right) \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial f}{\partial y} f \right) + \mathcal{O}(h_n^3).$$

De esta manera, para que el método considerado tenga orden uno debe cumplirse lo siguiente:

$$\text{Condición de orden uno} \iff b = 1.$$

Nótese que para alcanzar orden dos es necesario elegir  $c = 1/2$  y además la función  $f$  debe cumplir  $\frac{\partial f}{\partial y} f = 0$ ; eso ocurre, por ejemplo, si  $f$  no depende de  $y$ , pero en tal caso la ecuación es puramente temporal. Por tanto, el orden máximo de una fórmula de Runge-Kutta explícita de una etapa, cualquiera que sea la función  $f$ , es uno. Además, como  $c$  no interviene en la condición de orden, es habitual tomar  $c = 0$  para que se cumpla la condición de suma de filas, obteniéndose la tabla de Butcher asociada al método de Euler explícito.

### 2.3.2. Métodos de Runge-Kutta explícitos de dos etapas

La tabla de Butcher de método será:

$$\begin{array}{c|cc} 0 & & 0 \\ c_2 & c_2 & 0 \\ \hline & b_1 & b_2 \end{array}$$

El método se escribe como:

$$\begin{cases} k_1 = f(t_n, y_n), \\ k_2 = f(t_n + c_2 h_n, y_n + c_2 h_n k_1), \end{cases}$$

$$y_{n+1} = y_n + h_n(b_1 k_1 + b_2 k_2).$$

y su error local es:

$$\varepsilon(t_n, h_n; y) = y(t_n + h_n) - y(t_n) - h_n(b_1 k_1(t_n, y(t_n), h_n) + b_2 k_2(t_n, y(t_n), h_n)).$$

En este caso necesitamos desarrollar:

$$\begin{aligned} \triangleright y(t_n + h_n) &= y(t_n) + h_n y'(t_n) + \frac{h_n^2}{2!} y''(t_n) + \frac{h_n^3}{3!} y'''(t_n) + \mathcal{O}(h_n^4) \\ &= y(t_n) + h_n f + \frac{h_n^2}{2!} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\ &\quad + \frac{h_n^3}{6} \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 + \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \right) + \mathcal{O}(h_n^4). \\ \triangleright k_1(t_n, y(t_n), h_n) &= f. \\ \triangleright k_2(t_n, y(t_n), h_n) &= f + \frac{\partial f}{\partial t} c_2 h_n + \frac{\partial f}{\partial y} c_2 h_n k_1 \\ &\quad + \frac{1}{2} \left( \frac{\partial^2 f}{\partial t^2} c_2^2 h_n^2 + 2 \frac{\partial^2 f}{\partial t \partial y} c_2^2 h_n^2 k_1 + \frac{\partial^2 f}{\partial y^2} f^2 c_2^2 h_n^2 k_1^2 \right) + \mathcal{O}(h_n^3). \end{aligned}$$

De nuevo en estos cálculos se han omitido los argumentos tanto de la función  $f$  como de sus derivadas parciales, que están siempre evaluadas en el punto  $(t_n, y(t_n))$ .

Sustituyendo estos desarrollos en la expresión del error local obtenemos:

$$\begin{aligned} \varepsilon(t_n, h_n; y) &= h_n(1 - b_1 - b_2)f \\ &\quad + h_n^2 \left( \frac{1}{2} - b_2 c_2 \right) \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\ &\quad + h_n^3 \left( \left( \frac{1}{6} - b_2 c_2^2 \right) \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 + \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \right) \right. \\ &\quad \left. + \frac{1}{6} \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \right) + \mathcal{O}(h_n^4). \end{aligned}$$

Así, obtenemos las siguientes condiciones de orden:

$$\text{Condición de orden uno} \iff b_1 + b_2 = 1,$$

$$\text{Condición de orden dos} \iff b_1 c_2 = \frac{1}{2}.$$

Evidentemente, para que el método alcance orden al menos dos, deben cumplirse de forma simultánea las dos condiciones anteriores. Por este motivo, en ocasiones nos referiremos a ambas expresiones de forma conjunta con el término “condiciones de orden dos”.

Además, como cumpliéndose estas condiciones no se puede anular el término en  $h_n^3$  no se podrá alcanzar orden tres. Por tanto, el orden máximo que se puede alcanzar con una fórmula de Runge-Kutta explícita de dos etapas es dos.

### 2.3.3. Métodos de Runge-Kutta explícitos de tres etapas

La tabla de Butcher de método será

0	0		
$c_2$	$c_2$	0	
$c_3$	$c_3 - a_{32}$	$a_{32}$	0
	$b_1$	$b_2$	$b_3$

El método se escribe como:

$$\begin{cases} k_1 = f(t_n, y_n), \\ k_2 = f(t_n + c_2 h_n, y_n + c_2 h_n k_1), \\ k_3 = f(t_n + c_3 h_n, y_n + h_n(c_3 - a_{32})k_1 + h_n a_{32} k_2), \end{cases}$$

$$y_{n+1} = y_n + h_n(b_1 k_1 + b_2 k_2 + b_3 k_3).$$

y su error local es:

$$\varepsilon(t_n, h_n; y) = y(t_n + h_n) - y(t_n) - h_n \sum_{i=1}^3 b_i k_i(t_n, y(t_n), h_n).$$

Desarrollamos primero:

$$\begin{aligned} \triangleright y(t_n + h_n) &= y(t_n) + h_n y'(t_n) + \frac{h_n^2}{2!} y''(t_n) + \frac{h_n^3}{3!} y'''(t_n) + \mathcal{O}(h_n^4) \\ &= y(t_n) + h_n f + \frac{h_n^2}{2!} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\ &\quad + \frac{h_n^3}{6} \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 + \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \right) + \mathcal{O}(h_n^4). \end{aligned}$$

$$\triangleright k_1(t_n, y(t_n), h_n) = f.$$

$$\begin{aligned}
\triangleright k_2(t_n, y(t_n), h_n) &= f + c_2 h_n \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\
&\quad + \frac{c_2^2 h_n^2}{2} \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 \right) + \mathcal{O}(h_n^3). \\
\triangleright k_3(t_n, y(t_n), h_n) &= f + h_n c_3 \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) + h_n^2 a_{32} c_2 \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\
&\quad + \frac{h_n^2 c_3^2}{2} \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 \right) + \mathcal{O}(h_n^3).
\end{aligned}$$

Otra vez más, con el fin de aligerar la notación, se han obviado los argumentos de la función  $f$  y de sus derivadas parciales puesto que siempre están evaluadas en el punto  $(t_n, y(t_n))$ .

Sustituyendo estos desarrollos en la expresión del error local obtenemos:

$$\begin{aligned}
\varepsilon(t_n, h_n; y) &= h_n(1 - b_1 - b_2 - b_3)f \\
&\quad + h_n^2 \left( \frac{1}{2} - b_2 c_2 - b_3 c_3 \right) \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \\
&\quad + h_n^3 \left( \left( \frac{1}{6} - \frac{1}{2} (b_2 c_2^2 + b_3 c_3^2) \right) \left( \frac{\partial^2 f}{\partial t^2} + 2 \frac{\partial^2 f}{\partial t \partial y} f + \frac{\partial^2 f}{\partial y^2} f^2 \right) \right. \\
&\quad \quad \left. + \left( \frac{1}{6} - a_{32} c_2 b_3 \right) \frac{\partial f}{\partial y} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} f \right) \right) \\
&\quad + \mathcal{O}(h_n^4).
\end{aligned}$$

Así, obtenemos las siguientes condiciones de orden tres:

$$\begin{aligned}
b_1 + b_2 + b_3 &= 1 \\
b_2 c_2 + b_3 c_3 &= \frac{1}{2} \\
b_2 c_2^2 + b_3 c_3^2 &= \frac{1}{3} \\
a_{32} c_2 b_3 &= \frac{1}{6}
\end{aligned}$$

Como vemos, los desarrollos son más laboriosos cuanto mayor es el número de etapas del método. A su vez, hay que tener en cuenta que en este caso hemos estudiado únicamente algunos casos de un problema escalar. En el caso de los sistemas diferenciales el proceso se vuelve más complicado si cabe, pues es necesario escribir las diferenciales de  $\mathbf{f}$  en términos de sus derivadas parciales respecto de sus argumentos. Con lo expuesto hasta el momento creemos que queda sobradamente probada la necesidad de desarrollar un procedimiento alternativo que facilite la obtención de las condiciones que deben satisfacer los coeficientes de un método de Runge-Kutta para alcanzar el mayor orden posible.

## Capítulo 3

# Teoría de Butcher

En este capítulo se presentan las ideas fundamentales de la Teoría desarrollada a lo largo de los años sesenta del siglo pasado por el matemático neozelandés J.C. Butcher. Este novedoso planteamiento supuso un punto de inflexión muy importante en el estudio de los métodos numéricos utilizados para resolver ecuaciones diferenciales ordinarias. La obra de Butcher, que incorpora y formaliza algunos conceptos previamente señalados por Gill y Merson, fue revisada, refinada y ampliada a mediados de la década de los setenta y comienzos de los años ochenta por los austriacos Hairer y Wanner. En dichos trabajos se establece una correspondencia entre los términos que aparecen en los desarrollos de Taylor de las soluciones exacta y numérica y un tipo de grafos llamados árboles con raíz, facilitando enormemente la obtención de las ecuaciones que deben satisfacer los coeficientes de una fórmula de Runge-Kutta para alcanzar un determinado orden.

### 3.1. Árboles con raíz y árboles monótonamente ordenados

Aunque la Teoría de Árboles fue desarrollada originariamente por Arthur Cayley dentro del contexto más general de la Teoría de Grafos, en esta sección se formalizan sus conceptos fundamentales desde el punto de vista del Análisis Numérico (que es en definitiva el que nos interesa en este trabajo), procurando, en la medida de lo posible, establecer o, en su caso, mantener cierta ligadura entre su utilización en ambas áreas. Es probable, por tanto, que en dicha adaptación se produzca alguna ligera discrepancia con la terminología usual en uno y otro campo. En cualquier caso, siempre que se tenga constancia de la posible existencia de algún tipo de discordancia de este estilo se intentará señalarla convenientemente.

**Definición 3.1.** Un *grafo*  $G$  es un par  $(V, A)$  donde:

- $V$  es un conjunto de puntos.
- $A$  es un conjunto de pares de elementos de  $V$ .

Los elementos de  $V$  reciben el nombre de *vértices* o *nodos* del grafo  $G$ ; mientras que los elementos del conjunto  $A$  se llaman *arcos* o *aristas* y se denotan  $[x, y]$  con  $x, y \in V$ .

**Definición 3.2.** Un grafo  $G = (V, A)$  se dice *orientado* o *dirigido* cuando los arcos tienen un sentido definido, i.e., cuando dados  $x, y \in V$  se tiene que  $[x, y] \neq [y, x]$ . Cuando, por el contrario, no importa el orden de los pares, hablamos de grafos *no orientados* o *no dirigidos*. En este caso, se tiene que si  $x, y \in V$  entonces  $[x, y] = [y, x]$ .

En lo que sigue abordaremos única y exclusivamente el caso de grafos dirigidos.

**Definición 3.3.** Dado un grafo  $G = (V, A)$ :

- Un *bucle* o *lazo* de  $G$  es un arco que relaciona un nodo  $x \in V$  con sí mismo ( $[x, x]$ ).
- Se dice que  $G$  es un  $p$ -grafo si, dados  $x, y \in V$ ,  $A$  contiene a lo sumo  $p$  arcos  $[x, y]$ .
- Diremos que el grafo  $G$  es *simple* si es un 1-grafo sin bucles.

A continuación, haremos referencia solamente a grafos simples.

**Definición 3.4.** Dado un grafo  $G = (V, A)$ :

- Un *camino* de  $G$  es una sucesión de vértices  $x_1 x_2 \cdots x_n$  de manera que entre  $x_{i-1}$  y  $x_i$  existe un arco que los une ( $i = 2, 3, \dots, n$ ).
- La *longitud* de un camino es el número de arcos que lo forman.
- Un *camino simple* es aquel que no repite vértices en su recorrido.
- Un *ciclo* es un camino simple que empieza y termina en el mismo vértice.
- Se dice que  $G$  es *conexo* si todo par de vértices está unido por algún camino.

**Definición 3.5.** Un *árbol* es un grafo orientado, conexo y que no tiene bucles.

**Definición 3.6.** Un *árbol con raíz* (en inglés *rooted tree*) es un árbol en el que se puede distinguir un vértice, llamado *raíz*, de forma tal que existen caminos desde la raíz hasta los demás vértices del árbol.

- Llamamos *antecesores* o *predecesores* de un nodo  $x$  a todos aquellos vértices que forman parte del camino que une la raíz con dicho nodo  $x$ .
- Del mismo modo, llamamos *sucesores* o *descendientes* de un nodo  $x$  a aquellos vértices para los cuales existe un camino desde el vértice  $x$  hasta ellos mismos.

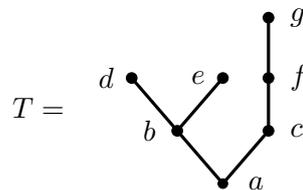
Reservaremos el nombre de *padres* e *hijos* para el caso de antecesores y sucesores inmediatos, esto es, cuando el camino que une los dos vértices en cuestión tiene longitud uno.

- Se llama *hoja* (o *vértice externo* o *colgante*) a todo vértice distinto de la raíz que no tiene sucesores.

**Nota 3.1.** Algunos autores, especialmente en el marco de la Teoría de Grafos, suelen referirse con el nombre de *arborescencias* (más concretamente *arborescencias de salida*) a lo que aquí hemos definido como árboles con raíz y reservan este término para el caso en el que el grafo es no dirigido. Otros, sin embargo, prefieren considerar que los árboles con raíz son grafos no dirigidos, evitando así la discrepancia con la terminología usual en la Teoría de Grafos, pero señalan que la raíz induce cierta orientación que les permite mantener los conceptos de antecesores y sucesores. Nosotros hemos optado por seguir la nomenclatura habitual en el campo del Análisis Numérico, por ser el objetivo fundamental del trabajo.

Gráficamente, los nodos de un árbol con raíz serán representados mediante puntos y los arcos, mediante segmentos que unen tales puntos según las relaciones establecidas entre ellos. Además, optaremos por situar el vértice raíz en la parte inferior del diagrama, de manera que este se “despliegue” hacia arriba y por comodidad y al no existir riesgo de ambigüedad omitiremos la flecha que indica el sentido de los arcos. También es conveniente señalar en este punto que es habitual etiquetar los puntos del grafo mediante los elementos que forman parte del conjunto de vértices y en nuestro caso serán o bien números  $\{1, 2, 3, 4, 5, \dots\}$  o bien letras  $\{a, b, c, d, e, \dots\}$  o  $\{i, j, k, l, m, \dots\}$ . En ese caso particular hablaremos de *árboles con raíz y etiquetados* (en inglés *labelled rooted trees*).

**Ejemplo 3.1.** Consideremos el árbol con raíz  $T = (V, A)$  cuya representación gráfica es:



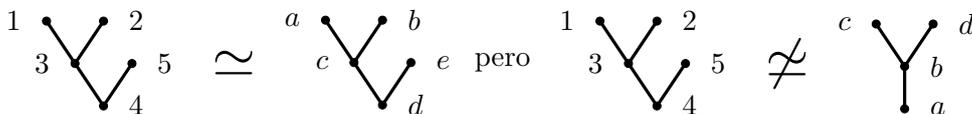
Entonces, se tiene que:

- El conjunto de vértices (y de etiquetas) es  $V = \{a, b, c, d, e, f, g\}$ .
- El conjunto de aristas es  $A = \{[a, b], [a, c], [b, d], [c, e], [c, f], [f, g]\}$ .
- El vértice  $a$  es la raíz del árbol  $T$ .
- Los antecesores del nodo  $g$  son  $\{a, c, f\}$  y su padre es el vértice  $f$ .
- Los sucesores del nodo  $b$  son  $\{d, e\}$ , que son, al mismo tiempo, sus hijos.
- Los vértices  $\{d, e, g\}$  son las hojas del árbol con raíz  $T$ .

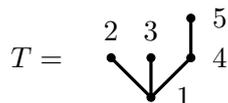
**Definición 3.7.** Dos árboles etiquetados con raíz  $T = (V, A)$  y  $T' = (V', A')$  se dicen *isomorfos* y se denota  $T \simeq T'$  si existe una aplicación biyectiva  $\varphi : V \rightarrow V'$  (que recibe el nombre de *isomorfismo* entre  $T$  y  $T'$ ) de modo tal que  $[x, y] \in A$  si y solo si  $[\varphi(x), \varphi(y)] \in A'$ .

Cuando  $T = T'$  el isomorfismo  $\varphi$  se dice *automorfismo*. Además, el conjunto de todos los automorfismos de un árbol con raíz  $T$ , que se denota por  $\text{Aut}(T)$ , tiene estructura de grupo con la composición de aplicaciones. El orden de dicho grupo es el cardinal del conjunto  $\text{Aut}(T)$ .

**Ejemplo 3.2.** Gráficamente, el hecho de que dos árboles sean isomorfos se traduce en que sus árboles sin etiquetas son idénticos como se aprecia en los siguientes diagramas:



**Ejemplo 3.3.** Consideremos el árbol con raíz  $T = (V, A)$  cuya representación gráfica es:



En este caso, tenemos que  $V = \{1, 2, 3, 4, 5\}$  y  $A = \{[1, 2], [1, 3], [1, 4], [4, 5]\}$ .

La aplicación que permuta los vértices 2 y 3, más concretamente, la aplicación  $\varphi$  dada por:

$$\begin{aligned} \varphi: V &\longrightarrow V \\ 1 &\mapsto 1 \\ 2 &\mapsto 3 \\ 3 &\mapsto 2 \\ 4 &\mapsto 4 \\ 5 &\mapsto 5 \end{aligned}$$

es un automorfismo del árbol con raíz  $T$ . De hecho,  $\text{Aut}(T) = \{Id, \varphi\}$  donde  $Id$  denota la aplicación identidad de  $V$  y, en consecuencia, el orden del grupo de automorfismos es 2.

**Nota 3.2.** En relación con los ejemplos anteriores conviene señalar lo siguiente:

- Aunque se presenten sin etiquetas, no debe olvidarse que, en este punto, los árboles se definen a partir de los conjuntos de nodos y aristas. La importancia a la hora de determinar los isomorfismos o automorfismos de un árbol recae, precisamente, sobre estos dos conjuntos, independientemente de su representación gráfica. Así pues, el Ejemplo 3.2 podría llevarnos a pensar que los dos árboles siguientes no son isomorfos, puesto que sus árboles sin etiquetas no son exactamente iguales. Sin embargo, los nodos y las aristas del segundo árbol son los mismos que aparecían en el mencionado ejemplo, por lo que sí son isomorfos. No obstante, lo más habitual es dejar fijo el diagrama sin etiquetas y establecer una correspondencia entre los nodos y las aristas de uno y otro árbol.



- El hecho de que los árboles sin etiquetas coincidan no significa que los árboles en cuestión vayan a ser isomorfos, como se aprecia con los siguientes diagramas, puesto que, entre otros, el arco  $[1, 2]$  aparece en el primer árbol pero no en el segundo.



**Proposición 3.1.** *La relación binaria «ser árboles isomorfos» es de equivalencia.*

*Demostración.* Veamos que se cumplen las propiedades reflexiva, simétrica y transitiva.

- Reflexiva, i.e.,  $T \simeq T$ : basta considerar  $\varphi = Id$ .
- Simétrica, i.e.,  $T \simeq T' \Leftrightarrow T' \simeq T$ : basta considerar  $\varphi^{-1}$  (que existe por ser  $\varphi$  biyectiva).
- Transitiva, i.e., si  $T \simeq T'$  y  $T' \simeq T''$  entonces  $T \simeq T''$ : basta considerar  $\varphi'' = \varphi' \circ \varphi$ .  $\square$

Podemos entender entonces los árboles con raíz como las clases de equivalencia bajo este tipo de isomorfismo. De esta manera la clase de equivalencia de un árbol  $T$ ,  $[T]$ , que en lo que sigue será denotada por  $[\tau]$ , no es más que la colección de todos los árboles con raíz y con etiquetas que resultan de renombrar los vértices de  $T$  de todas las formas posibles (con las etiquetas del árbol dado). Identificaremos, por tanto, el representante de dicha clase de equivalencia,  $\tau$ , con el árbol sin etiquetas común a todos los elementos de la familia.

**Notación:** Para referirnos a un árbol  $\tau$  tal que al eliminar su raíz se obtienen  $k$  árboles no conexos  $\tau_1, \tau_2, \dots, \tau_k$  adoptaremos la notación  $\tau = \{\tau_1, \tau_2, \dots, \tau_k\}$  donde, a su vez, cada  $\tau_i$  con  $i = 1, 2, \dots, k$  será denotado siguiendo estas mismas pautas.

Añadimos las siguientes puntualizaciones:

- El árbol con un único nodo será denotado por  $\tau_0$ .
- En consecuencia todos los árboles se expresan en función de  $\tau_0$ .

**Ejemplo 3.4.** Veamos cómo usar esta nomenclatura en un caso concreto:

$$\tau = \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \Rightarrow \tau = \left\{ \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array}, \bullet \right\} = \left\{ \{ \bullet, \bullet \}, \bullet \right\} = \left\{ \{ \tau_0, \tau_0 \}, \tau_0 \right\}.$$

**Definición 3.8.** Dado un árbol con raíz  $\tau = \{\tau_1, \tau_2, \dots, \tau_k\}$ :

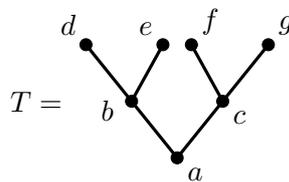
- Llamamos *orden* de  $\tau$  y denotamos  $\rho(\tau)$  al número de vértices de  $\tau$ .  
Se calcula fácilmente de la siguiente manera:
  - $\rho(\tau_0) = 1$ .
  - $\rho(\tau) = 1 + \sum_{i=1}^k \rho(\tau_i)$ .
- Se define la *simetría* de  $\tau$ , que será denotada por  $\sigma(\tau)$ , como el orden del grupo de automorfismos de  $\tau$ . Para obtener su valor es necesario que en caso de que algunos árboles  $\tau_i$  sean iguales expresemos estas repeticiones en términos de potencias. Así:
  - $\sigma(\tau_0) = 1$ .
  - $\sigma(\tau) = \prod_{i=1}^r n_i! \sigma(s_i)^{n_i}$  donde  $\tau = \{s_1^{n_1}, \dots, s_r^{n_r}\}$ .

- Se llama *densidad* o *producto de inundación* de  $\tau$  y se denota  $\gamma(\tau)$  al producto a lo largo de todos los vértices del orden del subárbol con raíz en ese mismo vértice. Para conocer su valor se utilizan las siguientes fórmulas:

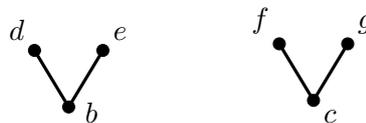
- $\gamma(\tau_0) = 1.$
- $\gamma(\tau) = \rho(\tau) \prod_{i=1}^k \gamma(\tau_i).$

**Observación 3.1.** Por extensión, y con cierto abuso de notación, dado un árbol con raíz y etiquetado,  $T$ , perteneciente a la clase de equivalencia de  $\tau$  (esto es,  $\tau$  es el árbol  $T$  sin etiquetas) entenderemos que  $\rho(T) = \rho(\tau)$ ,  $\sigma(T) = \sigma(\tau)$  y  $\gamma(T) = \gamma(\tau)$ .

**Ejemplo 3.5.** Para ilustrar el cálculo de estas magnitudes consideraremos el siguiente árbol:

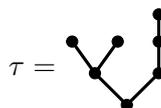


Claramente  $\rho(T) = 7$  y  $\gamma(T) = 7 \cdot 3 \cdot 3 = 63$ . Ahora bien, el cálculo de  $\sigma(T)$  requiere un poco más de explicación. Notemos que el grafo es isomorfo a sí mismo (automorfo) si permutamos los pares  $d$  y  $e$  o  $f$  y  $g$  y también si intercambiamos los dos subgrafos siguientes:

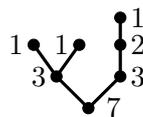


Entonces  $\sigma(T) = 2! \cdot 2! \cdot 2! = 2 \cdot 2 \cdot 2 = 8$ .

**Ejemplo 3.6.** Con el fin de afianzar el manejo de estos tres conceptos, calculamos nuevamente sus valores para el árbol con raíz y sin etiquetas  $\tau$  siguiente:

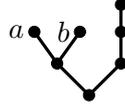


Por un lado, es claro que  $\rho(\tau) = 7$ , pues en el diagrama anterior se aprecian 7 nodos o vértices. Por otra parte, para facilitar el estudio de su densidad consideramos el diagrama siguiente, en el cual al lado de cada nodo se incluye el orden del subárbol con raíz en dicho nodo.



Por tanto, a partir de la propia definición de la densidad de un árbol, resulta evidente que el valor de  $\gamma(\tau)$  puede obtenerse multiplicando los números anotados anteriormente, esto es, en este caso,  $\gamma(\tau) = 7 \cdot 3 \cdot 3 \cdot 2 = 126$ .

Y finalmente, para calcular la simetría de  $\tau$  solamente necesitamos tener en cuenta que la única transformación que mantiene las relaciones que existen entre los vértices es aquella que consiste en intercambiar los nodos  $a$  y  $b$  que aparecen en la representación siguiente:



Por consiguiente, en este caso, se tiene que  $\sigma(\tau) = 2! = 2$ .

**Nota 3.3.** Podemos dar, además, una interpretación del significado de la densidad de un árbol con raíz, esto es, de la idea intuitiva que hay detrás de su definición. Nos fijamos para ello en las densidades de los cuatro árboles de orden cuatro siguientes.

$\tau$				
$\gamma(\tau)$	4	8	12	24

Tabla 3.1: Densidad de los cuatro árboles de orden cuatro.

Comparando los valores de las densidades y las formas de estos cuatro árboles con raíz, se puede concluir entonces que la densidad de un árbol es mayor cuanto más tarda en ramificarse, entendiendo esto, a su vez, como la altura a la que el árbol abre o despliega sus ramas. Así, el primer árbol, que se despliega desde el inicio, es el menos denso de todos; mientras que el último, que no se ramifica en ningún punto, es el más denso. Además, cuando dos árboles se ramifican en el mismo punto, la densidad es mayor en aquel que tiene una altura global mayor, como puede apreciarse en los dos primeros árboles de la tabla anterior, que se abren desde el principio pero el segundo tiene tres niveles de altura y el primero, solo dos.

Nuestro interés por la teoría de árboles se debe a que permite identificar cada término del desarrollo de Taylor de una función vectorial con un árbol con raíz. A su vez, esto nos servirá para desarrollar el error local de un método de forma sistemática.

Un árbol  $T = (V, A)$  con raíz en un vértice  $r$  puede entenderse como una aplicación que asocia a cada elemento con su antecesor, esto es:

$$T: V \setminus \{r\} \longrightarrow V \\ x \longmapsto T(x) := \text{padre}(x)$$

La Teoría de Butcher para el análisis del orden de consistencia de un método de Runge-Kutta utiliza en su desarrollo un tipo concreto de árboles con raíz, los llamados árboles con raíz monótonamente ordenados, cuya formalización introducimos a continuación.

**Definición 3.9.** Sea  $\mathcal{C} = \{i < j < k < l < m < \dots\}$  una cadena ordenada de índices. Denotemos por  $\mathcal{C}_q$  el subconjunto de  $\mathcal{C}$  formado por sus primeros  $q$  elementos. Entonces, llamamos *árbol con raíz monótonamente ordenado de orden  $q \geq 2$*  a cualquier árbol con raíz  $T: \mathcal{C}_q \setminus \{i\} \longrightarrow \mathcal{C}_q$  tal que  $T(x) < x, \forall x \in \mathcal{C}_q \setminus \{i\}$ .

Denotamos por  $\mathcal{LT}$  el conjunto de todos los árboles con raíz monótonamente ordenados y por  $\mathcal{LT}_q$  la colección de todos los árboles con raíz monótonamente ordenados de orden  $q$ .

Finalmente, restringiendo la noción de isomorfismo al caso de árboles monótonamente ordenados (que además, en este caso concreto, recibe el nombre de equivalencia de árboles y se denota por  $\sim$ ), podemos introducir los siguientes conceptos:

- La colección de árboles con raíz será el conjunto cociente  $\mathcal{T} = \mathcal{LT}/\sim$ .
- La colección de los árboles con raíz de orden  $q$  será el conjunto cociente  $\mathcal{T}_q = \mathcal{LT}_q/\sim$ .

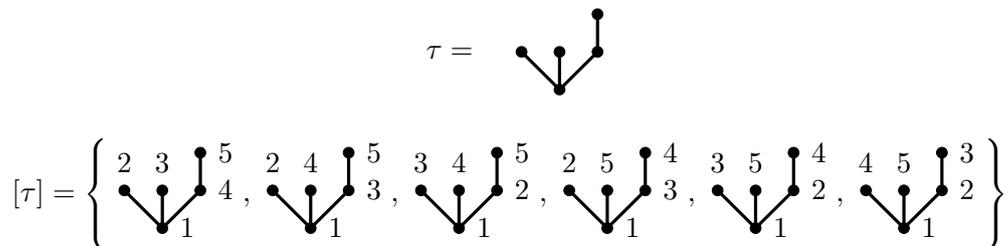
$q$	1	2	3	4	5	6	7	8	9	10
$\text{Card}(\mathcal{T}_q)$	1	1	2	4	9	20	48	115	286	719

Tabla 3.2: Número de árboles con raíz hasta orden diez.

**Definición 3.10.** Dado un árbol con raíz  $\tau$  denotaremos por  $\alpha(\tau)$  el número de árboles con raíz monótonamente ordenados que son equivalentes (isomorfos) a  $\tau$ , esto es, el número de elementos que forman la clase de equivalencia de  $\tau$  bajo la relación  $\sim$ . En otras palabras,  $\alpha(\tau)$  es el número de formas distintas de etiquetar monótonamente el árbol  $\tau$ .

**Nota 3.4.** En lo que sigue, adoptaremos  $\{1, 2, \dots, \rho(\tau)\}$  como conjunto de vértices.

**Ejemplo 3.7.** Veamos cómo y cuántos son los elementos de la clase de equivalencia bajo la relación definida por  $\sim$  cuyo representante es el árbol con raíz  $\tau$  donde:



Por tanto  $\alpha(\tau) = 6$ .

**Proposición 3.2 (Cálculo práctico de  $\alpha(\tau)$ ).**

Dado un árbol  $\tau \in \mathcal{T}$ , se cumple que:

$$\alpha(\tau) = \frac{\rho(\tau)!}{\sigma(\tau)\gamma(\tau)}.$$

*Demostración.* Con  $\rho(\tau)!$  calculamos todas las posibles formas de etiquetar los árboles de ese orden; al dividir por la simetría  $\sigma(\tau)$  contamos una sola vez aquellos que son equivalentes; y finalmente, dividiendo de nuevo por la densidad  $\gamma(\tau)$  nos quedamos únicamente con aquellos que son monótonamente ordenados □

**Nota 3.5.** Por completitud suele denotarse por  $\emptyset$  el único árbol vacío de orden 0.

Tabla 3.3: Árboles hasta orden cinco y sus propiedades

Árbol	Nombre	$\rho(\tau)$	$\sigma(\tau)$	$\gamma(\tau)$	$\alpha(\tau)$
	$\tau_0$	1	1	1	1
	$\{\tau_0\}$	2	1	2	1
	$\{\tau_0, \tau_0\}$	3	2	3	1
	$\{\{\tau_0\}\}$	3	1	6	1
	$\{\tau_0, \tau_0, \tau_0\}$	4	6	4	1
	$\{\{\tau_0\}, \tau_0\}$	4	1	8	3
	$\{\{\tau_0, \tau_0\}\}$	4	2	12	1
	$\{\{\{\tau_0\}\}\}$	4	1	24	1
	$\{\tau_0, \tau_0, \tau_0, \tau_0\}$	5	24	5	1
	$\{\{\tau_0\}, \tau_0, \tau_0\}$	5	2	10	6
	$\{\{\tau_0, \tau_0\}, \tau_0\}$	5	2	15	4
	$\{\{\{\tau_0\}\}, \tau_0\}$	5	1	30	4
	$\{\{\tau_0\}, \{\tau_0\}\}$	5	2	20	3
	$\{\{\tau_0, \tau_0, \tau_0\}\}$	5	6	20	1
	$\{\{\{\tau_0\}, \tau_0\}\}$	5	1	40	3
	$\{\{\{\tau_0, \tau_0\}\}\}$	5	2	60	1
	$\{\{\{\{\tau_0\}\}\}\}$	5	1	120	1

### 3.2. Desarrollo de la solución exacta: derivadas de Frechet y diferenciales elementales

La idea de Butcher consiste en relacionar los términos del desarrollo de Taylor del error local,

$$\varepsilon(t_n, h; \mathbf{y}) = \mathbf{y}(t_n + h) - (\mathbf{y}(t_n) + h\Phi(t_n, \mathbf{y}(t_n), h)),$$

con los árboles con raíz introducidos en el apartado anterior, analizando, por separado, los desarrollos de las soluciones exacta,  $\mathbf{y}(t_n + h)$ , y aproximada,  $\mathbf{y}(t_n) + h\Phi(t_n, \mathbf{y}(t_n), h)$ .

En esta sección nos ocuparemos de lo primero y para ello consideraremos el sistema autónomo

$$(P) \quad \begin{cases} \mathbf{y}' &= \mathbf{f}(\mathbf{y}(t)), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) &= \boldsymbol{\eta}, \end{cases}$$

donde  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  con  $m > 1$  y denotaremos las derivadas parciales de cada componente  $f_i$  de la función  $\mathbf{f}$  de la siguiente manera:

$$\begin{aligned} f_i^j &= \frac{\partial f_i}{\partial x_j}, \\ f_i^{jk} &= \frac{\partial^2 f_i}{\partial x_j \partial x_k}, \\ &\vdots \\ f_i^{j_1, \dots, j_n} &= \frac{\partial^n f_i}{\partial x_{j_1} \dots \partial x_{j_n}}. \end{aligned}$$

Así, los subíndices indican la componente que estamos derivando y los superíndices denotan las variables respecto a las cuales derivamos. Además, asumiremos la conocida como *notación indicial o de Einstein*, en la cual se suprimen el símbolo de sumatorio y los índices de sumación de los pares de índices repetidos que van desde 1 hasta la dimensión  $m$  del problema (P).

Por ejemplo:  $f_i^j f_j = \sum_{j=1}^m f_i^j f_j$ .

**Definición 3.11.** Sea  $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  con  $m > 1$  una función analítica en  $\mathcal{U}$  y sea  $\mathbf{z} \in \mathcal{U}$ . Llamamos *derivada M-ésima (o de orden M) de Frechet de  $\mathbf{f}$  en  $\mathbf{z}$*  a la aplicación

$$\mathbf{f}^{(M)}(\mathbf{z}) : \overbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}^{M \text{ veces}} \rightarrow \mathbb{R}^m \\ (\mathbf{U}_1, \dots, \mathbf{U}_M) \mapsto \mathbf{f}^{(M)}(\mathbf{z})(\mathbf{U}_1, \dots, \mathbf{U}_M) := \left( f_i^{(M)}(\mathbf{z})(\mathbf{U}_1, \dots, \mathbf{U}_M) \right)_{i=1}^m$$

donde

$$f_i^{(M)}(\mathbf{z})(\mathbf{U}_1, \dots, \mathbf{U}_M) := \sum_{j_1=1}^m \dots \sum_{j_M=1}^m f_i^{j_1, \dots, j_M}(\mathbf{z}) u_{1, j_1} \dots u_{M, j_M}$$

siendo  $\mathbf{U}_k = (u_{k,1}, \dots, u_{k,m})^T \in \mathbb{R}^m$  para  $k = 1, \dots, M$ .

**Ejemplo 3.8.** Derivadas de Frechet de órdenes 1 y 2 de  $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  con  $m > 1$ .

- La derivada de Frechet de orden 1 de  $\mathbf{f}$  en un punto  $\mathbf{z} \in \mathbb{R}^m$  se define como la aplicación

$$\begin{aligned} \mathbf{f}^{(1)}(\mathbf{z}) : \mathbb{R}^m &\longrightarrow \mathbb{R}^m \\ \mathbf{U}_1 &\mapsto \mathbf{f}^{(1)}(\mathbf{z})(\mathbf{U}_1) := \left( f_i^{(1)}(\mathbf{z})(\mathbf{U}_1) \right)_{i=1}^m \end{aligned}$$

donde

$$f_i^{(1)}(\mathbf{z})(\mathbf{U}_1) := \sum_{j=1}^m f_i^j(\mathbf{z})u_{1,j}, \quad \text{siendo } \mathbf{U}_1^T = (u_{1,r})_{r=1}^m.$$

- Por su parte, la derivada de Frechet de orden 2 de  $\mathbf{f}$  en  $\mathbf{z} \in \mathbb{R}^m$  viene dada por

$$\begin{aligned} \mathbf{f}^{(2)}(\mathbf{z}) : \mathbb{R}^m \times \mathbb{R}^m &\longrightarrow \mathbb{R}^m \\ (\mathbf{U}_1, \mathbf{U}_2) &\mapsto \mathbf{f}^{(2)}(\mathbf{z})(\mathbf{U}_1, \mathbf{U}_2) := \left( f_i^{(2)}(\mathbf{z})(\mathbf{U}_1, \mathbf{U}_2) \right)_{i=1}^m \end{aligned}$$

donde

$$f_i^{(2)}(\mathbf{z})(\mathbf{U}_1, \mathbf{U}_2) := \sum_{j=1}^m \sum_{k=1}^m f_i^{jk}(\mathbf{z})u_{1,j}u_{2,k}, \quad \text{siendo } \mathbf{U}_s^T = (u_{s,r})_{r=1}^m, \quad s = 1, 2.$$

Entonces, como la componente  $i$ -ésima de la solución del problema (P) viene dada por:

$$y_i'(t) = f_i(\mathbf{y}(t)),$$

aplicando la Regla de la Cadena, obtenemos:

$$y_i''(t) = \sum_{j=1}^m \frac{\partial f_i}{\partial y_j}(\mathbf{y}(t))y_j'(t) = \sum_{j=1}^m f_i^j(\mathbf{y}(t))f_j(\mathbf{y}(t)). \quad (3.1)$$

Ahora, si comparamos esta expresión con la obtenida en el Ejemplo 3.8 para la derivada de Frechet de orden 1, podemos deducir que:

$$y_i''(t) = f_i^{(1)}(\mathbf{y}(t))\left(\mathbf{f}(\mathbf{y}(t))\right).$$

Derivando en la ecuación (3.1), obtenemos, por la Regla de la cadena:

$$\begin{aligned} y_i'''(t) &= \sum_{j=1}^m \sum_{k=1}^m \frac{\partial f_i^j}{\partial y_k}(\mathbf{y}(t))y_k'(t)f_j(\mathbf{y}(t)) + \sum_{j=1}^m f_i^j(\mathbf{y}(t)) \sum_{k=1}^m \frac{\partial f_j}{\partial y_k}(\mathbf{y}(t))y_k'(t) \\ &= \sum_{j=1}^m \sum_{k=1}^m f_i^{jk}(\mathbf{y}(t))f_j(\mathbf{y}(t))f_k(\mathbf{y}(t)) + \sum_{j=1}^m f_i^j(\mathbf{y}(t)) \sum_{k=1}^m f_j^k(\mathbf{y}(t))f_k(\mathbf{y}(t)). \end{aligned}$$

Y comparando estas expresiones con las obtenidas en el Ejemplo 3.8, podemos concluir:

$$y_i'''(t) = f_i^{(2)}(\mathbf{y}(t))\left(\mathbf{f}(\mathbf{y}(t)), \mathbf{f}(\mathbf{y}(t))\right) + f_i^{(1)}(\mathbf{y}(t))\left(\mathbf{f}^{(1)}(\mathbf{y}(t))\left(\mathbf{f}(\mathbf{y}(t))\right)\right).$$

En definitiva, aplicando de manera reiterada la Regla de la Cadena, deducimos que las derivadas sucesivas de la componente  $i$ -ésima de la solución del problema  $(P)$  vienen dadas por las siguientes expresiones, donde tanto  $y_i$  como sus derivadas están evaluadas en  $t$  mientras que la función  $f_i$  y sus derivadas se evalúan en  $\mathbf{y}(t)$ . No obstante, omitimos estos argumentos para no sobrecargar en exceso la notación.

- $y_i' = f_i.$
- $y_i'' = f_i^j f_j.$
- $y_i''' = f_i^{jk} f_j f_k + f_i^j f_j^k f_k.$
- $y_i^{(4)} = f_i^{jkl} f_j f_k f_l + f_i^{jk} f_j^l f_k f_l + f_i^{jk} f_j f_k^l f_l + f_i^{jl} f_l f_j^k f_k + f_i^j f_j^{kl} f_k f_l + f_i^j f_j^k f_k^l f_l.$
- $\vdots$

Y por tanto podemos escribir las derivadas sucesivas de  $\mathbf{y}$  en términos de las derivadas de Frechet como sigue, donde de nuevo omitimos, por brevedad en la notación, el argumento  $t$  de  $\mathbf{y}$  y sus derivadas y el argumento  $\mathbf{y}(t)$  de las funciones  $\mathbf{f}$  y  $\mathbf{f}^{(M)}$  con  $M = 1, 2, \dots$

- $\mathbf{y}' = \mathbf{f}.$
- $\mathbf{y}'' = \mathbf{f}^{(1)}(\mathbf{f}).$
- $\mathbf{y}''' = \mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}) + \mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f})).$
- $\mathbf{y}^{(4)} = \mathbf{f}^{(3)}(\mathbf{f}, \mathbf{f}, \mathbf{f}) + \mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}) + \mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}^{(1)}(\mathbf{f})) + \mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}) + \mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})) + \mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))).$
- $\vdots$

Observamos además que, en general,  $\mathbf{y}^{(n)}$  es una combinación lineal de varias derivadas de Frechet de orden menor o igual que  $n - 1$ .

Continuando los trabajos de Merson, el matemático neozelandés John Charles Butcher, en la década de los años sesenta del siglo pasado, fue capaz de establecer una relación entre estos conceptos y los árboles introducidos en la sección anterior. Con lo expuesto hasta el momento, podemos afirmar que, en general, cada derivada de la componente  $i$ -ésima de la solución exacta del problema  $(P)$  es una combinación lineal de varias derivadas de Frechet de la componente  $i$ -ésima de la función  $\mathbf{f}$  que define el sistema diferencial considerado. La idea de Butcher consiste en asociar a cada uno de esos sumandos un árbol con raíz cuyos vértices serán los índices de tales fórmulas y los arcos vendrán determinados por los pares [subíndice,superíndice] que aparecen.

En la tabla siguiente se incluyen algunos ejemplos.

Orden	Derivada	Grafos
1	$y'_i = f_i$	
2	$y''_i = f_i^j f_j$	
3	$y'''_i = f_i^{jk} f_j f_k + f_i^j f_j^k f_k$	
4	$y_i^{(4)} = f_i^{jkl} f_j f_k f_l + f_i^{jk} f_j^l f_k f_l + f_i^{jk} f_j f_k^l f_l + f_i^j f_j^k f_l f_k + f_i^j f_j f_l f_k + f_i^j f_j^k f_l f_l$	

Tabla 3.4: Correspondencia entre derivadas de la componente i-ésima y árboles etiquedados

**Definición 3.12.** Sea  $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  una función analítica en  $\mathcal{U}$  y sea  $\mathbf{y} \in \mathcal{U}$ . Dado un árbol  $\tau \in \mathcal{T}_q$  definimos la *diferencial elemental de orden q de f asociada a tau en el punto y* de la siguiente manera:

$$\begin{aligned} \mathbf{F}(\emptyset)(\mathbf{y}) &= \mathbf{y}, \\ \mathbf{F}(\tau_0)(\mathbf{y}) &= \mathbf{f}(\mathbf{y}), \\ \mathbf{F}(\tau)(\mathbf{y}) &= \mathbf{f}^{(M)}(\mathbf{y})(\mathbf{F}(\tau_1)(\mathbf{y}), \dots, \mathbf{F}(\tau_M)(\mathbf{y})), \text{ cuando } \tau = \{\tau_1, \dots, \tau_M\}. \end{aligned}$$

**Nota 3.6.** Por extensión, y con cierto abuso de notación, dado un árbol con raíz monótonamente ordenado y etiquetado de orden  $q$ ,  $T \in \mathcal{LT}_q$ , cuyo árbol sin etiquetas subyacente es  $\tau$  (y por tanto,  $\tau \in \mathcal{T}_q$ ), o lo que es lo mismo, si  $T$  pertenece a la clase de equivalencia de representante  $\tau$ , entenderemos que la diferencial elemental de orden  $q$  de  $\mathbf{f}$  asociada a  $T$  en el punto  $\mathbf{y}$  no es más que la correspondiente diferencial elemental de orden  $q$  de  $\mathbf{f}$  asociada a  $\tau$  en dicho punto, esto es,  $\mathbf{F}(T)(\mathbf{y}) = \mathbf{F}(\tau)(\mathbf{y})$ .

**Ejemplo 3.9.** Veamos cómo se calcula la diferencial elemental de una función analítica  $\mathbf{f} : \mathcal{U} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  en un punto  $\mathbf{y} \in \mathcal{U}$  asociada al árbol  $\tau$  que aparece en el Ejemplo 3.4:

$$\tau = \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \quad \tau = \left\{ \underbrace{\{\tau_0, \tau_0\}}_{\tau_1}, \underbrace{\tau_0}_{\tau_2} \right\}$$

Como se cumple que:

- $\mathbf{F}(\tau_1)(\mathbf{y}) = \mathbf{f}^{(2)}(\mathbf{y})(\mathbf{F}(\tau_0)(\mathbf{y}), \mathbf{F}(\tau_0)(\mathbf{y})) = \mathbf{f}^{(2)}(\mathbf{y})(\mathbf{f}(\mathbf{y}), \mathbf{f}(\mathbf{y}))$ .
- $\mathbf{F}(\tau_2)(\mathbf{y}) = \mathbf{F}(\tau_0)(\mathbf{y}) = \mathbf{f}(\mathbf{y})$ .

entonces, obtenemos:

$$\mathbf{F}(\tau)(\mathbf{y}) = \mathbf{f}^{(2)}(\mathbf{y})(\mathbf{F}(\tau_1)(\mathbf{y}), \mathbf{F}(\tau_2)(\mathbf{y})) = \mathbf{f}^{(2)}(\mathbf{y})\left(\mathbf{f}^{(2)}(\mathbf{y})(\mathbf{f}(\mathbf{y}), \mathbf{f}(\mathbf{y})), \mathbf{f}(\mathbf{y})\right).$$

En la Tabla 3.5 se incluyen las diferenciales elementales asociadas a los árboles con raíz hasta orden cinco, donde se omite, por brevedad en la notación, su argumento.

A partir de todo lo anterior podemos deducir los siguientes resultados que establecen una relación entre la solución exacta o teórica del problema diferencial considerado y las diferenciales elementales asociadas a los árboles con raíz.

**Teorema 3.1.** *La derivada  $k$ -ésima de cualquier solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial del problema (P) puede escribirse en términos de las diferenciales elementales como sigue:*

$$\mathbf{y}^{(k)}(t_n) = \sum_{T \in \mathcal{LT}_k} \mathbf{F}(T)(\mathbf{y}(t_n)) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)), \quad k = 0, 1, 2, \dots$$

*Demostración.* Con ayuda de las Tablas 3.4 y 3.5, es fácil ver que dicho resultado es cierto para  $k = 0, 1, 2, 3$  y 4 e incluso para  $k = 5$ . Para tratar de generalizarlo, fijémonos en el razonamiento que hay detrás de esos cálculos, dando una interpretación gráfica de los mismos. Para calcular  $\mathbf{y}^{(i+1)}$  derivamos  $\mathbf{y}^{(i)}$  y esto, a su vez, consiste en derivar  $\mathbf{F}(T)(\mathbf{y})$  para cada uno de los árboles con raíz monótonamente ordenados de orden  $i$ . Este proceso se traduce, gráficamente, en la adición, en cada caso, de una nueva rama con una hoja a cada nodo de cada uno de los árboles anteriores. Luego, para cada árbol considerado obtenemos  $i$  árboles de orden  $i + 1$ , por lo que en el cálculo de  $\mathbf{y}^{(i+1)}$  aparecen todos los árboles monótonamente ordenados de orden  $i + 1$  y, además, cada uno de ellos aparece exactamente una única vez. Finalmente, agrupando aquellos que dan lugar a la misma diferencial elemental obtenemos la segunda expresión que aparece en el enunciado del Teorema.  $\square$

**Corolario 3.1.** *El desarrollo de Taylor de cualquier solución  $\mathbf{y}(\cdot)$  de la ecuación diferencial del problema (P):*

$$\mathbf{y}(t_n + h) = \mathbf{y}(t_n) + h\mathbf{y}'(t_n) + \frac{h^2}{2!}\mathbf{y}''(t_n) + \dots$$

*puede escribirse formalmente en términos de las diferenciales elementales asociadas a árboles con raíz de la siguiente manera:*

$$\mathbf{y}(t_n + h) = \sum_{T \in \mathcal{LT}} \mathbf{F}(T)(\mathbf{y}(t_n)) \frac{h^{\rho(T)}}{\rho(T)!} = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}.$$

**Nota 3.7.** Conviene destacar que el concepto de diferencial elemental (ver Definición 3.12) se aplica sobre funciones  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  analíticas en  $\mathcal{U} \subseteq \mathbb{R}^m$ , esto es, para funciones que, localmente, es decir, que en un entorno de cada punto  $\mathbf{y} \in \mathcal{U}$ , pueden expresarse como una serie de potencias. Como consecuencia de esta definición se tiene que si  $\mathbf{f}$  es una función analítica en  $\mathcal{U} \subseteq \mathbb{R}^m$ , entonces  $\mathbf{f}$  es infinitamente derivable en  $\mathcal{U}$  y, además, su serie de Taylor converge y coincide con  $\mathbf{f}$  en  $\mathcal{U}$ .

Tabla 3.5: Correspondencia entre diferenciales elementales y árboles

Orden	Árbol	Diferencial elemental
0	$\emptyset$	$\mathbf{y}$
1	$\bullet$	$\mathbf{f}$
2	$\begin{array}{c}   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f})$
3	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})$
	$\begin{array}{c} \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))$
4	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(3)}(\mathbf{f}, \mathbf{f}, \mathbf{f})$
	$\begin{array}{c} \bullet \\   \\ \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f})$
	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}))$
	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f})))$
5	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \quad \diagup \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(4)}(\mathbf{f}, \mathbf{f}, \mathbf{f}, \mathbf{f})$
	$\begin{array}{c} \bullet \\   \\ \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(3)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}, \mathbf{f})$
	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(2)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}), \mathbf{f})$
	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f})), \mathbf{f})$
	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}^{(1)}(\mathbf{f}))$
	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(3)}(\mathbf{f}, \mathbf{f}, \mathbf{f}))$
	$\begin{array}{c} \bullet \\   \\ \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}))$
	$\begin{array}{c} \bullet \\ \diagdown \quad \diagup \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})))$
	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))))$

### 3.3. Desarrollo de la solución numérica: series de Butcher

Nuestro objetivo es desarrollar, en potencias crecientes del paso  $h$ , el error local en el instante  $t_n$ , cuya expresión venía dada, según la Definición 1.9, por la diferencia:

$$\varepsilon(t_n, h; \mathbf{y}) = \mathbf{y}(t_n + h) - (\mathbf{y}(t_n) + h\Phi(t_n, \mathbf{y}(t_n), h)).$$

En la sección anterior hemos conseguido escribir el desarrollo de Taylor de  $\mathbf{y}(t_n + h)$  mediante las diferenciales elementales asociadas a árboles con raíz. En lo que sigue trataremos de obtener un resultado semejante para el segundo término del error local, que se corresponde con la solución aproximada del problema  $(P)$  en el instante  $t_{n+1}$  partiendo del dato exacto  $\mathbf{y}(t_n)$  y que en adelante se denotará, por comodidad,  $\mathbf{y}_{n+1}(t_n)$ , es decir,

$$\mathbf{y}_{n+1}(t_n) = \mathbf{y}(t_n) + h\Phi(t_n, \mathbf{y}(t_n), h).$$

Para ello, introduciremos el concepto de B-series o series de Butcher formalizado a mediados de los años setenta del siglo pasado por los matemáticos austriacos Hairer y Wanner.



Figura 3.1: Ernst Hairer (1949) y Gerhard Wanner (1942)

**Definición 3.13.** Dada una aplicación  $a : \mathcal{T} \rightarrow \mathbb{R}$ , llamamos *serie de Butcher* o *B-serie asociada a la aplicación  $a$  en punto  $\mathbf{y} \in \mathbb{R}^m$*  a la serie formal dada por:

$$B(a, \mathbf{y})(h) := \sum_{\tau \in \mathcal{T}} \alpha(\tau) a(\tau) \mathbf{F}(\tau)(\mathbf{y}) \frac{h^{\rho(\tau)}}{\rho(\tau)!}.$$

**Nota 3.8.** En las condiciones de la definición anterior, dado un árbol con raíz monótonamente ordenado y etiquetado,  $T \in \mathcal{LT}$ , cuyo árbol sin etiquetas subyacente es  $\tau$  (y por tanto,  $\tau \in \mathcal{T}$ ), o lo que es lo mismo, si  $T$  pertenece a la clase de equivalencia de representante  $\tau$ , entenderemos que, con cierto abuso de notación,  $a(T) = a(\tau)$ . Así, en ese caso, la serie de Butcher o B-serie asociada a la aplicación  $a$  en punto  $\mathbf{y} \in \mathbb{R}^m$  quedaría definida como

$$B(a, \mathbf{y})(h) := \sum_{T \in \mathcal{LT}} a(T) \mathbf{F}(T)(\mathbf{y}) \frac{h^{\rho(T)}}{\rho(T)!}.$$

Usaremos indistintamente ambas notaciones, aunque la primera es más habitual.

**Definición 3.14.** Dada una aplicación  $a : \mathcal{T} \rightarrow \mathbb{R}$  en las condiciones de la definición anterior, una función  $\mathbf{g} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$  de clase  $\mathcal{C}^\infty$  en un entorno del origen se dice que es *representable en serie de Butcher mediante  $a$* , y se denota  $\mathbf{g}(h) = B(a, \mathbf{y})(h)$ , si:

$$\mathbf{g}^{(k)}(0) = \sum_{T \in \mathcal{LT}_k} a(T) \mathbf{F}(T)(\mathbf{y}) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) a(\tau) \mathbf{F}(\tau)(\mathbf{y}), \quad k = 0, 1, 2, \dots$$

**Teorema 3.2 (Fundamental de B-series).**

Sea  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  una función analítica en un entorno abierto del punto  $\mathbf{y} \in \mathbb{R}^m$  y sea  $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^m$  una función representable en serie de Butcher mediante cierta aplicación  $a : \mathcal{T} \rightarrow \mathbb{R}$  verificando que  $a(\emptyset) = 1$ , esto es,

$$\mathbf{g}^{(k)}(0) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) a(\tau) \mathbf{F}(\tau)(\mathbf{y}), \quad k = 0, 1, 2, \dots$$

Entonces,  $h\mathbf{f}(\mathbf{g}(h))$  es representable en serie de Butcher mediante la aplicación  $a' : \mathcal{T} \rightarrow \mathbb{R}$  definida por:

$$\begin{cases} a'(\emptyset) &= 0, \\ a'(\tau_0) &= 1, \\ a'(\tau) &= \rho(\tau) a(\tau_1) \cdots a(\tau_M), \text{ cuando } \tau = \{\tau_1, \dots, \tau_M\}. \end{cases}$$

*Demostración.* Denotemos  $\mathbf{c}(h) = h\mathbf{f}(\mathbf{g}(h))$ . Tenemos que probar que:

$$\mathbf{c}^{(k)}(0) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) a'(\tau) \mathbf{F}(\tau)(\mathbf{y}), \quad k = 0, 1, 2, \dots$$

- Para  $k = 0$  se tiene:

$$\mathbf{c}^{(0)}(h) = \mathbf{c}(h) \implies \mathbf{c}^{(0)}(0) = \mathbf{c}(0) = \mathbf{0}.$$

Como  $\emptyset$  es el único árbol de orden 0, entonces, efectivamente,  $a'(\emptyset) = 0$ .

- Para  $k = 1$  se cumple que:

$$\mathbf{c}^{(1)}(h) = \mathbf{c}'(h) = (\mathbf{f} \circ \mathbf{g})(h) + h(\mathbf{f} \circ \mathbf{g})'(h) \implies \mathbf{c}^{(1)}(0) = (\mathbf{f} \circ \mathbf{g})(0).$$

Como  $a(\emptyset) = 1$  entonces  $\mathbf{g}(0) = \mathbf{y}$ , y por tanto  $\mathbf{c}^{(1)}(0) = \mathbf{f}(\mathbf{y})$ . Así,  $a'(\tau_0) = 1$ .

- Para  $k \geq 2$ , aplicando la Regla de Leibniz:

$$\begin{aligned} \mathbf{c}^{(k)}(h) &= \sum_{i=0}^k \binom{k}{i} Id^{(i)}(h) (\mathbf{f} \circ \mathbf{g})^{(k-i)}(h) = \binom{k}{0} h (\mathbf{f} \circ \mathbf{g})^{(k)}(h) + \binom{k}{1} (\mathbf{f} \circ \mathbf{g})^{(k-1)}(h) \\ &\implies \mathbf{c}^{(k)}(0) = \binom{k}{1} (\mathbf{f} \circ \mathbf{g})^{(k-1)}(0) = k (\mathbf{f} \circ \mathbf{g})^{(k-1)}(0). \end{aligned}$$

Siguiendo un razonamiento similar al utilizado para relacionar las derivadas de la solución exacta del problema  $(P)$  con las derivadas de Frechet y posteriormente para establecer una correspondencia entre estas y los árboles con raíz, obtenemos que las derivadas sucesivas de  $(\mathbf{f} \circ \mathbf{g})(h)$  pueden expresarse en términos de las derivadas de Frechet de  $\mathbf{f}$  en  $\mathbf{g}(h)$  al tiempo que podemos relacionar biunívocamente cada una de estas con un tipo concreto de árboles de la manera siguiente:

$$\begin{aligned}
 (\mathbf{f} \circ \mathbf{g})' &= \mathbf{f}^{(1)}(\mathbf{g})(\mathbf{g}') && \rightsquigarrow && \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \\
 (\mathbf{f} \circ \mathbf{g})'' &= \mathbf{f}^{(2)}(\mathbf{g})(\mathbf{g}', \mathbf{g}') + \mathbf{f}^{(1)}(\mathbf{g})(\mathbf{g}'') && \rightsquigarrow && \begin{array}{cc} \begin{array}{c} \bullet & \bullet \\ \diagdown & / \\ \bullet \end{array} & \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \end{array} \\
 (\mathbf{f} \circ \mathbf{g})''' &= \mathbf{f}^{(3)}(\mathbf{g})(\mathbf{g}', \mathbf{g}', \mathbf{g}') + \mathbf{f}^{(2)}(\mathbf{g})(\mathbf{g}'', \mathbf{g}') && \rightsquigarrow && \begin{array}{ccc} \begin{array}{c} \bullet & \bullet & \bullet \\ \diagdown & | & / \\ \bullet \end{array} & \begin{array}{c} \bullet & \bullet \\ \diagdown & / \\ \bullet \end{array} & \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \\
 + \mathbf{f}^{(2)}(\mathbf{g})(\mathbf{g}', \mathbf{g}'') + \mathbf{f}^{(2)}(\mathbf{g})(\mathbf{g}', \mathbf{g}'') + \mathbf{f}^{(1)}(\mathbf{g})(\mathbf{g}''') &&& && \begin{array}{ccc} \begin{array}{c} \bullet & \bullet \\ \diagdown & / \\ \bullet \end{array} & \begin{array}{c} \bullet & \bullet \\ \diagdown & / \\ \bullet \end{array} & \begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \bullet \end{array} \end{array}
 \end{aligned}$$

donde se ha omitido el argumento  $h$  por brevedad en la notación.

Observamos que, en este caso, los árboles que aparecen tienen en común que las ramas que parten de los hijos de la raíz son de tipo cadena, esto es, de la forma:



Además, podemos deducir que la raíz tiene tantos hijos como indica el orden de la derivada de Frechet en cuestión y cada hijo forma con sus descendientes una cadena cuya longitud coincide con el orden de derivación de cada  $\mathbf{g}$ .

Introducimos la siguiente notación en relación con esta nueva variedad de árboles:

- Denotaremos por  $\mathcal{ST}$  al conjunto de árboles para los que las ramas que parten de los hijos de raíz son de tipo cadena y si queremos hacer referencia únicamente a aquellos árboles de esta familia con un determinado orden  $k$  emplearemos  $\mathcal{ST}_k$ .
- Ahora denotaremos por  $u_i$  al árbol de tipo cadena con  $\rho(u_i) = i$ .

Entonces un  $\tau \in \mathcal{ST}_k$  cualquiera se expresará como  $\tau = \{\tau_1, \dots, \tau_M\} = \{u_{i_1}, \dots, u_{i_M}\}$ .

Nótese que, además, al árbol  $\tau = \{u_{i_1}, \dots, u_{i_M}\}$  le corresponde la derivada de Frechet  $M$ -ésima de  $\mathbf{f}$  en  $\mathbf{g}(h)$  evaluada en  $(\mathbf{g}^{(i_1)}, \dots, \mathbf{g}^{(i_M)})$ , que abreviadamente denotaremos  $\mathbf{G}(\tau)(h)$ . Esto es,  $\mathbf{G}(\tau)(h) = \mathbf{f}^{(M)}(\mathbf{g}(h))(\mathbf{g}^{(i_1)}(h), \dots, \mathbf{g}^{(i_M)}(h))$ .

Entonces:

$$(\mathbf{f} \circ \mathbf{g})^{(k-1)}(h) = \sum_{\tau \in \mathcal{ST}_k} \mathbf{G}(\tau)(h),$$

y en consecuencia:

$$\begin{aligned}
(\mathbf{f} \circ \mathbf{g})^{(k-1)}(0) &= \sum_{\tau \in \mathcal{ST}_k} \mathbf{G}(\tau)(0) = \sum_{\tau \in \mathcal{ST}_k} \mathbf{f}^{(M)}(\mathbf{y}) \left( \mathbf{g}^{(i_1)}(0), \dots, \mathbf{g}^{(i_M)}(0) \right) \\
&= \sum_{\tau \in \mathcal{ST}_k} \mathbf{f}^{(M)}(\mathbf{y}) \left( \sum_{\tau_1 \in \mathcal{T}_{i_1}} \alpha(\tau_1) a(\tau_1) \mathbf{F}(\tau_1)(\mathbf{y}), \dots, \sum_{\tau_M \in \mathcal{T}_{i_M}} \alpha(\tau_M) a(\tau_M) \mathbf{F}(\tau_M)(\mathbf{y}) \right) \\
&= \sum_{\tau \in \mathcal{ST}_k} \sum_{\tau_1 \in \mathcal{T}_{i_1}} \cdots \sum_{\tau_M \in \mathcal{T}_{i_M}} \alpha(\tau_1) \cdots \alpha(\tau_M) a(\tau_1) \cdots a(\tau_M) \mathbf{f}^{(M)}(\mathbf{y}) \left( \mathbf{F}(\tau_1)(\mathbf{y}), \dots, \mathbf{F}(\tau_M)(\mathbf{y}) \right) \\
&= \sum_{\tau \in \mathcal{ST}_k} \sum_{\tau_1 \in \mathcal{T}_{i_1}} \cdots \sum_{\tau_M \in \mathcal{T}_{i_M}} \alpha(\{\tau_1, \dots, \tau_M\}) \frac{a'(\{\tau_1, \dots, \tau_M\})}{\rho(\{\tau_1, \dots, \tau_M\})} \mathbf{F}(\{\tau_1, \dots, \tau_M\})(\mathbf{y}) \\
&= \frac{1}{k} \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) a'(\tau) \mathbf{F}(\tau)(\mathbf{y}).
\end{aligned}$$

Por consiguiente,

$$\mathbf{c}^{(k)}(0) = k(\mathbf{f} \circ \mathbf{g})^{(k-1)}(0) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) a'(\tau) \mathbf{F}(\tau)(\mathbf{y}) \iff h\mathbf{f}(\mathbf{g}(h)) = B(a', \mathbf{y})(h). \quad \square$$

Adaptando la formulación (2.2) de un método de Runge-Kutta explícito de  $q$  etapas al caso autónomo obtenemos la siguiente expresión:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \sum_{i=1}^q b_i \mathbf{g}_i, \tag{3.2}$$

donde

$$\mathbf{g}_i = h_n \mathbf{f} \left( \mathbf{y}_n + \sum_{j=1}^{i-1} a_{ij} \mathbf{g}_j \right), \quad i = 1, 2, \dots, q.$$

Esta formulación es más conveniente ahora pues nos permite representar fácilmente la solución obtenida numéricamente en serie de Butcher, en base a los siguientes resultados.

**Lema 3.1.** *Cada  $\mathbf{g}_i$  es representable en serie de Butcher en la forma*

$$\mathbf{g}_i = B(\varphi_i, \mathbf{y}_n), \quad i = 1, 2, \dots, q,$$

donde las aplicaciones  $\varphi_i : \mathcal{T} \rightarrow \mathbb{R}$  están definidas por:

$$\begin{cases} \varphi_i(\emptyset) = 0, \\ \varphi_i(\tau_0) = 1, \\ \varphi_i(\tau) = \rho(\tau) \sum_{j_1, \dots, j_M}^{i-1} a_{ij_1} \cdots a_{ij_M} \varphi_{j_1}(\tau_1) \cdots \varphi_{j_M}(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\}. \end{cases}$$

*Demostración.* Denotemos por  $\varphi_0 : \mathcal{T} \rightarrow \mathbb{R}$  la aplicación definida por:

$$\varphi_0(\tau) = \begin{cases} 1 & \text{si } \tau = \emptyset, \\ 0 & \text{en otro caso.} \end{cases}$$

Procedemos por inducción:

- Para  $i = 1$ , consideramos  $\varphi_1(\tau) = \begin{cases} 1 & \text{si } \tau = \tau_0, \\ 0 & \text{en otro caso.} \end{cases}$

Entonces,

$$B(\varphi_1, \mathbf{y}_n)(h) = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \varphi_1(\tau) \mathbf{F}(\tau)(\mathbf{y}_n) \frac{h^{\rho(\tau)}}{\rho(\tau)!} = \alpha(\tau_0) \varphi_1(\tau_0) \mathbf{F}(\tau_0)(\mathbf{y}_n) h = \mathbf{f}(\mathbf{y}_n) h = \mathbf{g}_1.$$

Luego para  $i = 1$  se cumple.

- Supongamos el resultado cierto hasta  $i - 1$  y veamos que se cumple para  $i$ .

Por un lado, por construcción, sabemos que

$$\mathbf{g}_i = h \mathbf{f} \left( \mathbf{y}_n + \sum_{j=1}^{i-1} a_{ij} \mathbf{g}_j \right),$$

y por otra parte, por hipótesis de inducción, tenemos que:

$$\begin{aligned} \mathbf{y}_n + \sum_{j=1}^{i-1} a_{ij} \mathbf{g}_j &= B(\varphi_0, \mathbf{y}_n)(h) + \sum_{j=1}^{i-1} a_{ij} B(\varphi_j, \mathbf{y}_n) \\ &= B(\xi_i, \mathbf{y}_n)(h), \quad \text{donde } \xi_i = \varphi_0 + \sum_{j=1}^{i-1} a_{ij} \varphi_j. \end{aligned}$$

Aplicando el Teorema Fundamental de B-series (Teorema 3.2) concluimos que:

$$\mathbf{g}_i = h \mathbf{f} \left( \mathbf{y}_n + \sum_{j=1}^{i-1} a_{ij} \mathbf{g}_j \right) = B(\varphi_i, \mathbf{y}_n)(h),$$

donde la función  $\varphi_i$  se obtiene de la siguiente manera:

$$\begin{aligned} \varphi_i(\emptyset) &= 0, \\ \varphi_i(\tau_0) &= 1, \\ \varphi_i(\tau) &= \rho(\tau) \xi_i(\tau_1) \cdots \xi_i(\tau_M) \\ &= \rho(\tau) \left( \varphi_0(\tau_1) + \sum_{j=1}^{i-1} a_{ij} \varphi_j(\tau_1) \right) \cdots \left( \varphi_0(\tau_M) + \sum_{j=1}^{i-1} a_{ij} \varphi_j(\tau_M) \right) \\ &= \rho(\tau) \left( \sum_{j=1}^{i-1} a_{ij} \varphi_j(\tau_1) \right) \cdots \left( \sum_{j=1}^{i-1} a_{ij} \varphi_j(\tau_M) \right) \\ &= \rho(\tau) \sum_{j_1, \dots, j_M}^{i-1} a_{ij_1} \cdots a_{ij_M} \varphi_{j_1}(\tau_1) \cdots \varphi_{j_M}(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\}. \quad \square \end{aligned}$$

**Observación 3.2.** Podemos considerar la aplicación:

$$\begin{aligned}\varphi: \mathcal{T} &\longrightarrow \mathbb{R}^q \\ \tau &\mapsto \varphi(\tau) = (\varphi_i(\tau))_{i=1}^q\end{aligned}$$

De acuerdo con el Lema anterior  $\varphi$  puede computarse de la siguiente manera:

$$\begin{aligned}\varphi(\emptyset) &= (0, 0, \dots, 0)^T = \mathbf{0}, \\ \varphi(\tau_0) &= (1, 1, \dots, 1)^T = \mathbf{e}, \\ \varphi(\tau) &= \rho(\tau) \cdot A\varphi(\tau_1) \cdot \dots \cdot A\varphi(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\},\end{aligned}$$

donde “ $\cdot$ ” denota el producto de vectores componente a componente, esto es, dados dos vectores,  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ , se tiene que  $\mathbf{u} \cdot \mathbf{v} = (u_1v_1, u_2v_2, \dots, u_mv_m)^T$ .

**Definición 3.15.** Se define la *función de pesos internos* como la aplicación dada por:

$$\begin{aligned}\pi: \mathcal{T} &\longrightarrow \mathbb{R}^q \\ \tau &\mapsto \pi(\tau) = \frac{1}{\gamma(\tau)}\varphi(\tau)\end{aligned}$$

**Lema 3.2.** La función de pesos internos puede calcularse más rápidamente como:

$$\begin{aligned}\pi(\emptyset) &= (0, 0, \dots, 0)^T = \mathbf{0}, \\ \pi(\tau_0) &= (1, 1, \dots, 1)^T = \mathbf{e}, \\ \pi(\tau) &= A\pi(\tau_1) \cdot \dots \cdot A\pi(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\},\end{aligned}$$

donde, de nuevo, “ $\cdot$ ” denota el producto de vectores componente a componente.

**Definición 3.16.** Llamamos *función peso elemental* a la aplicación definida como:

$$\begin{aligned}\Psi: \mathcal{T} &\longrightarrow \mathbb{R} \\ \tau &\mapsto \Psi(\tau) = \mathbf{b}^T \pi(\tau)\end{aligned}$$

**Observación 3.3** (Regla práctica para obtener  $\Psi(\tau)$ ).

Dado un método de  $q$  etapas, la función peso elemental asociada a un árbol con raíz  $\tau$  cualquiera puede calcularse fácilmente siguiendo los pasos que se indican a continuación.

- En primer lugar, asignamos etiquetas  $(i, j, k, l, m, \dots)$  a todos los vértices del árbol dado salvo sus hojas, asumiendo que  $i$  es la etiqueta asociada a la raíz.
- A continuación escribimos una secuencia de factores donde el primer elemento es  $b_i$ .
- Luego, para cada arco  $[j, k]$  que no acabe en una hoja añadimos el término  $a_{jk}$ ; mientras que si, por el contrario, si acaba en una hoja, añadimos el factor  $c_j$ .
- Sumamos para todas las posibles elecciones de etiquetas en el conjunto  $\{1, 2, \dots, q\}$ .

La expresión resultante coincide con el valor de  $\Psi(\tau)$ .



**Teorema 3.4 (Condiciones de orden).**

Un método de Runge-Kutta explícito con  $q$  etapas alcanza orden  $p$  si y solo si se verifica que:

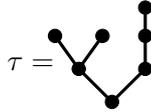
$$\Psi(\tau) = \frac{1}{\gamma(\tau)}, \quad \forall \tau \in \mathcal{T}_k \text{ con } k \leq p.$$

*Demostración.* Comparando los desarrollos que hemos obtenido para las soluciones exacta (Corolario 3.1) y numérica (Corolario 3.3), cuyas formulaciones vienen das por:

$$\begin{aligned} \triangleright \mathbf{y}(t_n + h) &= \sum_{\tau \in \mathcal{T}} \alpha(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}, \\ \triangleright \mathbf{y}_{n+1}(t_n) &= \sum_{\tau \in \mathcal{T}} \alpha(\tau) \omega(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}, \end{aligned}$$

deducimos que para que la diferencia entre ambos desarrollos o lo que es lo mismo, el desarrollo del error local, se comporte como una  $\mathcal{O}(h^{p+1})$  y en consecuencia, el método alcance orden  $p$  (ver Proposición 1.4), debe cumplirse que  $\omega(\tau) = 1$  para todo árbol  $\tau$  de orden menor o igual que  $p$ . Así, como  $\omega(\tau) = \gamma(\tau)\Psi(\tau)$ , despejando obtenemos la condición de orden dada.  $\square$

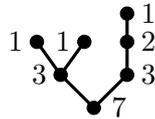
**Ejemplo 3.11.** Veamos cómo obtener la condición de orden asociada al siguiente árbol:



En el Ejemplo 3.10 ya habíamos explicado cómo se calcula la función peso elemental relativa al árbol  $\tau$  dado, cuya formulación para un método de  $q$  etapas resultaba ser la siguiente:

$$\Psi(\tau) = \sum_{i=1}^q \sum_{j=1}^q \sum_{k=1}^q \sum_{l=1}^q b_i a_{ij} a_{ik} c_j^2 a_{kl} c_l.$$

Además teniendo en cuenta el siguiente diagrama es fácil deducir que el valor de la función de densidad (o producto de inundación) de  $\tau$  es  $\gamma(\tau) = 7 \cdot 3 \cdot 3 \cdot 2 \cdot 1 \cdot 1 \cdot 1 = 126$ .



Así, la condición de orden asociada al árbol dado, según el Teorema 3.4, es:

$$\sum_{i=1}^q \sum_{j=1}^q \sum_{k=1}^q \sum_{l=1}^q b_i a_{ij} a_{ik} c_j^2 a_{kl} c_l = \frac{1}{126}.$$

Incluimos a continuación en la Tabla 3.6 un resumen de todas aquellas funciones que intervinen en las condiciones de orden de todos y cada uno de los árboles hasta orden cinco. Dicha tabla cobrará especial relevancia en el capítulo siguiente, en el que haremos un estudio pormenorizado de las familias de fórmulas de Runge-Kutta explícitas de hasta cinco etapas.

Tabla 3.6: Árboles y funciones que intervienen en las condiciones de orden

Orden	Árbol	$F(\tau)(\mathbf{y})$	$\gamma(\tau)$	$\pi(\tau)$
0	$\emptyset$	$\mathbf{y}$	1	
1	$\bullet$	$\mathbf{f}$	1	$\mathbf{e}$
2	$\vdots$	$\mathbf{f}^{(1)}(\mathbf{f})$	2	$\mathbf{c}$
3	$\vee$	$\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})$	3	$\mathbf{c}^2$
	$\vdots$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))$	6	$A\mathbf{c}$
4	$\vee$	$\mathbf{f}^{(3)}(\mathbf{f}, \mathbf{f}, \mathbf{f})$	4	$\mathbf{c}^3$
	$\bullet$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f})$	8	$A\mathbf{c} \cdot \mathbf{c}$
	$\vee$	$\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}))$	12	$A\mathbf{c}^2$
	$\vdots$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f})))$	24	$A^2\mathbf{c}$
5	$\vee$	$\mathbf{f}^{(4)}(\mathbf{f}, \mathbf{f}, \mathbf{f}, \mathbf{f})$	5	$\mathbf{c}^4$
	$\bullet$	$\mathbf{f}^{(3)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}, \mathbf{f})$	10	$A\mathbf{c} \cdot \mathbf{c}^2$
	$\vee$	$\mathbf{f}^{(2)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f}), \mathbf{f})$	15	$A\mathbf{c}^2 \cdot \mathbf{c}$
	$\bullet$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f})), \mathbf{f})$	30	$A^2\mathbf{c} \cdot \mathbf{c}$
	$\vee$	$\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}^{(1)}(\mathbf{f}))$	20	$A\mathbf{c} \cdot A\mathbf{c}$
	$\vee$	$\mathbf{f}^{(1)}(\mathbf{f}^{(3)}(\mathbf{f}, \mathbf{f}, \mathbf{f}))$	20	$A\mathbf{c}^3$
	$\vee$	$\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{f}), \mathbf{f}))$	40	$A(A\mathbf{c} \cdot \mathbf{c})$
	$\vee$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})))$	60	$A^2\mathbf{c}^2$
	$\vdots$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))))$	120	$A^3\mathbf{c}$

### Condiciones de orden para problemas escalares

Si bien en el Capítulo 2 habíamos diferenciado tres tipos de problemas diferenciales, a saber, escalar autónomo (*I*), escalar no autónomo (*II*) o vectorial -que sabemos que se reduce al caso autónomo- (*III*), hasta el momento solo hemos abordado el estudio de las condiciones de orden para el último de estos tres casos ( $m > 1$ ). Veamos lo que sucede con los otros dos.

Podría pensarse (y de hecho en un inicio se pensó) que el orden que alcanza una fórmula de Runge-Kutta es independiente del tipo de problema al que se aplica. Sin embargo, de la Teoría de Butcher se deduce que esto no siempre es así, pues según el tipo de problema en cuestión pueden coincidir ciertas diferenciales elementales y esto nos permite, a su vez, simplificar las condiciones de orden que se derivan del Teorema 3.4. En definitiva, los postulados de Butcher permiten demostrar que el orden de un método de Runge-Kutta dado puede depender del tipo de problema (*I*, *II* o *III*) que se considere.

También señalamos en el Capítulo 2 que un problema escalar no autónomo puede convertirse al caso autónomo considerando la variable independiente como una nueva incógnita, lo que supone aumentar su dimensión en una unidad, perdiendo, por tanto, su condición de escalar. Incluso se llegó a probar que cumpliéndose la condición de suma de filas la resolución de ambos problemas mediante una fórmula de Runge-Kutta es equivalente. Como además, el problema escalar autónomo puede interpretarse como un caso particular del escalar no autónomo, podemos afirmar lo siguiente sobre el orden de un método dado:

$$\text{Orden } p \text{ en (I)} \Leftrightarrow \text{Orden } p \text{ en (II)} \Leftrightarrow \text{Orden } p \text{ en (III)}.$$

Para analizar la validez de los recíprocos de estas dos implicaciones necesitamos adaptar la Definición 3.11 de la  $M$ -ésima derivada de Frechet, que en su momento se construyó para sistemas autónomos, al caso también autónomo pero escalar.

Así, dados  $z \in \mathbb{R}$  y  $f : \mathbb{R} \rightarrow \mathbb{R}$ , la  $M$ -ésima derivada de Frechet vendrá dada por:

$$f^{(M)}(z) : \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{M \text{ veces}} \rightarrow \mathbb{R} \\ (u_1, \dots, u_M) \mapsto f^{(M)}(z)(u_1, \dots, u_M) := f^{(M)}(z)u_1 \cdots u_M$$

Recordemos también que, según el Corolario 3.1, el desarrollo de Taylor de la solución exacta se expresa en términos de las diferenciales elementales asociadas a árboles con raíz,

$$\mathbf{y}(t_n + h) = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}, \quad (3.3)$$

y, de la misma manera, según el Corolario 3.3, la solución numérica puede desarrollarse como:

$$\mathbf{y}_{n+1} = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \omega(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}. \quad (3.4)$$

Un análisis riguroso de las condiciones que deben satisfacer los coeficientes de un método para alcanzar orden  $p$  permite demostrar que, salvo para  $1 \leq p \leq 3$ , estas dependen del tipo de problema que se plantee en cada caso.

En concreto se verifica lo siguiente:

- Orden  $p \leq 3$ :

Orden	Árbol	$F(\tau)(\mathbf{y})$ en (I)	$F(\tau)(\mathbf{y})$ en (III)
0	$\emptyset$	$\mathbf{y}$	$\mathbf{y}$
1	$\bullet$	$f$	$\mathbf{f}$
2	$\begin{array}{c} \bullet \\   \\ \bullet \end{array}$	$f' f$	$\mathbf{f}^{(1)}(\mathbf{f})$
3	$\begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad / \\ \bullet \end{array}$	$f'' f^2$	$\mathbf{f}^{(2)}(\mathbf{f}, \mathbf{f})$
	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	$(f')^2 f$	$\mathbf{f}^{(1)}(\mathbf{f}^{(1)}(\mathbf{f}))$

Tabla 3.7: Diferenciales elementales hasta orden tres en (I) y (III)

Como puede apreciarse en la tabla anterior, todas las diferenciales elementales de orden menor o igual que tres en (I) y en (III) son distintas y, por tanto, para que un método alcance orden  $p$  es necesario que se satisfagan todas las condiciones que se derivan del Teorema 3.4. En consecuencia, podemos afirmar que:

$$\text{Orden } p \text{ en (I)} \Leftrightarrow \text{Orden } p \text{ en (III)}.$$

Y por consiguiente, tenemos que:

$$\text{Orden } p \text{ en (I)} \Leftrightarrow \text{Orden } p \text{ en (II)} \Leftrightarrow \text{Orden } p \text{ en (III)}.$$

- Orden  $p = 4$ :

- En el caso vectorial, todas las diferenciales elementales de orden cuatro son distintas (ver Tabla 3.5) y, de nuevo, deben cumplirse las condiciones del Teorema 3.4.
- Sin embargo, en el problema escalar autónomo coinciden las diferenciales elementales asociadas a los árboles  $\tau_1$  y  $\tau_2$  siguientes:

$$\tau_1 = \begin{array}{c} \bullet \\ | \\ \bullet \quad \bullet \end{array} \quad \tau_2 = \begin{array}{c} \bullet \quad \bullet \\ \diagdown \quad / \\ \bullet \\ | \\ \bullet \end{array}$$

En efecto, se verifica que

$$F(\tau_1) = f^{(2)}(f^{(1)}(f), f) = f'' f' f f,$$

$$F(\tau_2) = f^{(1)}(f^{(2)}(f, f)) = f' f'' f f.$$

Entonces, teniendo en cuenta que según el Corolario 3.2 para cualquier árbol  $\tau \neq \emptyset$   $\omega(\tau) = \Psi(\tau)\gamma(\tau)$ , las partes de los desarrollos de las soluciones exacta y numérica correspondientes a los árboles  $\tau_1$  y  $\tau_2$ , que por las ecuaciones (3.3) y (3.4) sabemos que, en el caso escalar, son respectivamente,

$$\begin{aligned} \triangleright & \alpha(\tau_1)F(\tau_1)\frac{h^{\rho(\tau_1)}}{\rho(\tau_1)!} + \alpha(\tau_2)F(\tau_2)\frac{h^{\rho(\tau_2)}}{\rho(\tau_2)!}, \\ \triangleright & \alpha(\tau_1)\Psi(\tau_1)\gamma(\tau_1)F(\tau_1)\frac{h^{\rho(\tau_1)}}{\rho(\tau_1)!} + \alpha(\tau_2)\Psi(\tau_2)\gamma(\tau_2)F(\tau_2)\frac{h^{\rho(\tau_2)}}{\rho(\tau_2)!}, \end{aligned}$$

nos quedan, en dicho caso, de la siguiente manera:

$$\begin{aligned} \triangleright & (\alpha(\tau_1) + \alpha(\tau_2))f''f'ff\frac{h^4}{4!}, \\ \triangleright & (\alpha(\tau_1)\Psi(\tau_1)\gamma(\tau_1) + \alpha(\tau_2)\Psi(\tau_2)\gamma(\tau_2))f''f'ff\frac{h^4}{4!}. \end{aligned}$$

Por tanto, las dos condiciones de orden asociadas a los árboles  $\tau_1$  y  $\tau_2$ ,

$$\Psi(\tau_1) = \frac{1}{\gamma(\tau_1)} \quad \text{y} \quad \Psi(\tau_2) = \frac{1}{\gamma(\tau_2)},$$

pueden reemplazarse por la condición simplificada siguiente:

$$\alpha(\tau_1)\Psi(\tau_1)\gamma(\tau_1) + \alpha(\tau_2)\Psi(\tau_2)\gamma(\tau_2) = \alpha(\tau_1) + \alpha(\tau_2).$$

Es evidente que cumpliéndose las dos primeras ecuaciones, la última también se cumple pero el recíproco no tiene por qué ser cierto. Así, en el caso escalar autónomo basta con asumir la condición simplificada, pero en el caso vectorial, como las diferenciales elementales asociadas a  $\tau_1$  y  $\tau_2$  son distintas, esta no nos garantiza que se cumplan las dos condiciones de orden generales. Concluimos, por tanto, que:

$$\text{Orden 4 en (I)} \Leftarrow \text{Orden 4 en (III)}.$$

Por consiguiente, podemos afirmar que para orden  $p \geq 4$ :

$$\text{Orden 4 en (I)} \Leftarrow \text{Orden 4 en (III)}.$$

Por último, dado un problema escalar no autónomo

$$\begin{cases} y' = f(t, y), & f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \\ y(t_0) = \eta, \end{cases}$$

pasamos a considerar el sistema autónomo equivalente

$$\begin{cases} \mathbf{Y}' = \boldsymbol{\varphi}(\mathbf{Y}), \\ \mathbf{Y}(t_0) = \boldsymbol{\mathcal{H}}, \end{cases}$$

donde los vectores de  $\mathbb{R}^2$ ,  $\mathbf{Y}$ ,  $\mathcal{H}$  y  $\varphi$ , vienen dados por:

$$\mathbf{Y} = \begin{pmatrix} t \\ y \end{pmatrix}, \quad \mathcal{H} = \begin{pmatrix} t_0 \\ \eta \end{pmatrix}, \quad \varphi(\mathbf{Y}) = \begin{pmatrix} 1 \\ f(t, y) \end{pmatrix}.$$

En este caso las diferenciales elementales asociadas a los árboles  $\tau_1$  y  $\tau_2$ ,

$$\mathbf{F}(\tau_1) = \varphi^{(2)}(\varphi^{(1)}(\varphi), \varphi) = \begin{pmatrix} 0 \\ (f_t + f_y f)(f_{yt} + f_{yy} f) \end{pmatrix},$$

$$\mathbf{F}(\tau_2) = \varphi^{(1)}(\varphi^{(2)}(\varphi, \varphi)) = \begin{pmatrix} 0 \\ f_y (f_{tt} + 2f_{ty} f + f_{yy} f^2) \end{pmatrix},$$

son distintas y por tanto no es posible asumir la condición simplificada, lo que nos obliga a que se satisfagan las condiciones originales y en consecuencia:

$$\text{Orden 4 en (I)} \Leftrightarrow \text{Orden 4 en (II)} \Leftrightarrow \text{Orden 4 en (III)}.$$

- Por último, en el problema escalar no autónomo coinciden las diferenciales elementales de orden cinco asociadas a los dos árboles siguientes:

$$\tau_3 = \begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \bullet \end{array} \quad \tau_4 = \begin{array}{c} \bullet \\ | \\ \bullet \\ / \quad \backslash \\ \bullet \quad \bullet \end{array}$$

En efecto:

$$\mathbf{F}(\tau_3) = \varphi^{(2)}(\varphi^{(1)}(\varphi^{(1)}(\varphi)), \varphi) = \begin{pmatrix} 0 \\ f_{yt} f_y (f_t + f_y f) + f_{yy} f_y (f_t + f_y f) f \end{pmatrix},$$

$$\mathbf{F}(\tau_4) = \varphi^{(1)}(\varphi^{(2)}(\varphi^{(1)}(\varphi), \varphi)) = \begin{pmatrix} 0 \\ f_y (f_{yt} (f_t + f_y f) + f_{yy} (f_t + f_y f) f) \end{pmatrix}.$$

Por consiguiente, para  $p \geq 5$  concluimos que:

$$\text{Orden } p \text{ en (I)} \Leftrightarrow \text{Orden } p \text{ en (II)} \Leftrightarrow \text{Orden } p \text{ en (III)}.$$

**Observación 3.4.** Como ya se señaló anteriormente, en el inicio de los tiempos esta distinción de órdenes en función del tipo de problema diferencial en cuestión no se tuvo en cuenta y este hecho fue lo que motivó que Kutta tan solo llegara a aventurar dieciséis condiciones para alcanzar orden cinco, que son las que serían necesarias en problemas escalares no autónomos pero no son suficientes para sistemas, ya que en ese caso son necesarias diecisiete.

## Capítulo 4

# Teoría de Butcher y métodos de Runge-Kutta explícitos

Llegados a este punto, una vez explicada la Teoría de Butcher dentro del contexto general de los métodos de Runge-Kutta para resolver numéricamente ecuaciones diferenciales ordinarias, resulta interesante volver la vista atrás y recordar lo expuesto hasta el momento. En el Capítulo 2, además de presentar las fórmulas de Runge-Kutta explícitas y sus distintas formulaciones, nos interesamos por estudiar de manera pormenorizada algunas de sus propiedades como la convergencia, la estabilidad o la consistencia. Introdujimos también una técnica para obtener las ecuaciones que deben satisfacer los coeficientes del método para alcanzar un determinado orden de consistencia. Esta consistía en desarrollar en serie de Taylor el error local, pero, rápidamente comprobamos, tras ponerla en práctica para fórmulas de hasta tres etapas, que al incrementar el número de etapas dicho procedimiento se volvía difícilmente manejable y nos convencimos de la necesidad de poner en marcha una forma de proceder alternativa. Fue en ese momento cuando, en el Capítulo 3, expusimos la innovadora idea propuesta inicialmente por Butcher y ampliada y refinada posteriormente por Hairer y Wanner y constatamos de primera mano cómo esta teoría simplificaba de forma considerable el cálculo de las condiciones de orden asociadas a una fórmula de Runge-Kutta explícita. Veamos ahora como la Teoría de Butcher ayuda también a realizar una discusión detallada y razonada de las distintas familias de métodos explícitos hasta orden cinco.

En un método de Runge-Kutta explícito de  $q$  etapas cumpliendo la condición de suma de filas disponemos de  $\frac{q(q+1)}{2}$  coeficientes por determinar. Aunque en un inicio se pensó que tal fórmula podría alcanzar orden  $p$  si y solamente si el número de coeficientes era mayor o igual que el número de condiciones, posteriormente se descubrió que esto no era así.

Etapas	1	2	3	4	5	6	7	8	9	10
Coefficientes	1	3	6	10	15	21	28	36	45	55

Tabla 4.1: Relación entre el número de etapas y el número de coeficientes disponibles.

Orden	1	2	3	4	5	6	7	8	9	10
Condiciones	1	2	4	8	17	37	85	200	486	1205

Tabla 4.2: Relación entre el orden y el número de condiciones.

**Proposición 4.1.** *Un método de Runge-Kutta explícito de  $q$  etapas alcanza a lo sumo orden  $q$ .*

*Demostración.* Sea  $\tau$  el árbol de tipo cadena de orden  $q + 1$ , esto es,

$$\tau = \begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \vdots \\ | \\ \bullet \end{array}$$

Entonces:

- $\gamma(\tau) = (q + 1)!$
- $\Psi(\tau) = \mathbf{b}^T \boldsymbol{\pi}(\tau) = \mathbf{b}^T \underbrace{A\boldsymbol{\pi}(\tau_0) \cdot \dots \cdot A\boldsymbol{\pi}(\tau_0)}_{q \text{ veces}} = \mathbf{b}^T A^q \mathbf{e} = 0.$

Por consiguiente, no se verifica la condición de orden  $q + 1$  para el árbol  $\tau$ . □

Incluimos a continuación una exposición detallada y razonada de todas las distintas familias de métodos de Runge-Kutta explícitos que podemos encontrarnos con hasta cinco etapas, analizando detenidamente la resolución de las ecuaciones que plantean las condiciones de orden que se derivan, en cada caso, de los postulados de Butcher.

## 4.1. Métodos de Runge-Kutta explícitos de una etapa.

Consideremos un método de Runge-Kutta explícito cuya tabla de Butcher tenga la forma:

$$\begin{array}{c|c} 0 & 0 \\ \hline & b \end{array}$$

La condición de orden uno que se obtiene según la Teoría de Butcher se traduce en:

$$b = 1.$$

Por tanto, la tabla de Butcher resultante es:

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

que se corresponde con aquella que está asociada al método de Euler explícito.

## 4.2. Métodos de Runge-Kutta explícitos de dos etapas.

Consideremos una fórmula de Runge-Kutta explícita de dos etapas cuya tabla de Butcher sea:

$$\begin{array}{c|cc} 0 & & 0 \\ c_2 & c_2 & 0 \\ \hline & b_1 & b_2 \end{array}$$

En este caso, las condiciones de orden dos que se derivan de la Teoría de Butcher son:

$$\begin{aligned} \mathbf{b}^T \mathbf{e} = 1 &\Leftrightarrow b_1 + b_2 = 1 \\ \mathbf{b}^T \mathbf{c} = \frac{1}{2} &\Leftrightarrow b_2 c_2 = \frac{1}{2} \end{aligned}$$

Dejando  $c_2$  como parámetro y despejando obtenemos que:

$$c_2 \neq 0, \quad b_2 = \frac{1}{2c_2} \quad \text{y} \quad b_1 = 1 - \frac{1}{2c_2}.$$

Por tanto, la tabla de Butcher resultante es:

$$\begin{array}{c|cc} 0 & & 0 \\ c_2 & c_2 & 0 \\ \hline & 1 - \frac{1}{2c_2} & \frac{1}{2c_2} \end{array}$$

**Ejemplo 4.1.** Tablas de Butcher asociadas a los métodos de Euler modificado ( $c_2 = \frac{1}{2}$ ) y de Euler mejorado —también conocido como método de Heun— ( $c_2 = 1$ ), respectivamente.

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

## 4.3. Métodos de Runge-Kutta explícitos de tres etapas.

Consideremos ahora un método de Runge-Kutta explícito de tres etapas verificando la condición de suma de filas cuya tabla de Butcher viene dada por:

$$\begin{array}{c|ccc} 0 & & & 0 \\ c_2 & & c_2 & 0 \\ c_3 & c_3 - a_{32} & a_{32} & 0 \\ \hline & b_1 & b_2 & b_3 \end{array}$$

En este caso, las condiciones de orden tres que se siguen de los postulados de Butcher son:

$$\mathbf{b}^T \mathbf{e} = 1 \quad \Leftrightarrow \quad b_1 + b_2 + b_3 = 1$$

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

$$\mathbf{b}^T \mathbf{c}^2 = \frac{1}{3} \quad \Leftrightarrow \quad b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3}$$

$$\mathbf{b}^T \mathbf{A} \mathbf{c} = \frac{1}{6} \quad \Leftrightarrow \quad b_2 a_{32} c_2 = \frac{1}{6}$$

La segunda y la tercera condición se pueden disponer en forma matricial como sigue:

$$\begin{pmatrix} c_2 & c_3 \\ c_2^2 & c_3^2 \end{pmatrix} \begin{pmatrix} b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \end{pmatrix}.$$

Para obtener las soluciones de dicho sistema calculamos el determinante de la matriz anterior:

$$\det \begin{pmatrix} c_2 & c_3 \\ c_2^2 & c_3^2 \end{pmatrix} = c_2 c_3 (c_3 - c_2).$$

Como de la cuarta condición se deduce que  $c_2 \neq 0$ , podemos distinguir dos casos:

- Si  $c_2 c_3 (c_3 - c_2) \neq 0$  y  $c_2 \neq \frac{2}{3}$  la tabla de Butcher quedaría:

$$\begin{array}{c|ccc} 0 & & 0 & \\ c_2 & & c_2 & 0 \\ c_3 & \frac{c_3(3c_2-3c_2^2-c_3)}{c_2(2-3c_2)} & \frac{c_3(c_3-c_2)}{c_2(2-3c_2)} & 0 \\ \hline & \frac{2-3c_2-3c_3+6c_2c_3}{6c_2c_3} & \frac{2-3c_3}{6c_2(c_2-c_3)} & \frac{2-3c_2}{6c_3(c_3-c_2)} \end{array}$$

**Ejemplo 4.2.** Tablas de Butcher de los métodos de orden tres de Heun y Kutta.

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

- Si  $c_2 c_3 (c_3 - c_2) = 0$ , como  $c_2 \neq 0$  entonces  $c_3 (c_3 - c_2) = 0$  y distinguimos dos casos:

- $c_3 = 0$ :

$$\begin{array}{c|ccc} 0 & & 0 & \\ \frac{2}{3} & & \frac{2}{3} & 0 \\ 0 & -\frac{1}{4b_3} & \frac{1}{4b_3} & 0 \\ \hline & \frac{1}{4} - b_3 & \frac{3}{4} & b_3 \end{array}$$

- $c_2 = c_3$ :

0	0			
$\frac{2}{3}$	$\frac{2}{3}$	0		
$\frac{2}{3}$	$\frac{2}{3} - \frac{1}{4b_3}$	$\frac{1}{4b_3}$	0	
	$\frac{1}{4}$	$\frac{3}{4} - b_3$	$b_3$	

#### 4.4. Métodos de Runge-Kutta explícitos de cuatro etapas.

Consideremos una fórmula Runge-Kutta explícita de cuatro etapas cuya tabla de Butcher es:

0	0			
$c_2$	$a_{21}$	0		
$c_3$	$a_{31}$	$a_{32}$	0	
$c_4$	$a_{41}$	$a_{42}$	$a_{43}$	0
	$b_1$	$b_2$	$b_3$	$b_4$

Supondremos además que verifica la condición de suma de filas, lo que se traduce en que:

$$\begin{aligned} c_2 &= a_{21} \\ c_3 &= a_{31} + a_{32} \\ c_4 &= a_{41} + a_{42} + a_{43} \end{aligned}$$

En base a la Tabla 3.6 obtenemos las ocho condiciones de orden cuatro siguientes:

$$\begin{aligned} \mathbf{b}^T \mathbf{e} &= 1 \\ \mathbf{b}^T \mathbf{c} &= \frac{1}{2} \\ \mathbf{b}^T \mathbf{c}^2 &= \frac{1}{3} \\ \mathbf{b}^T \mathbf{c}^3 &= \frac{1}{4} \\ \mathbf{b}^T \mathbf{A}^2 \mathbf{c} &= \frac{1}{24} \\ \mathbf{b}^T \mathbf{A} \mathbf{c} &= \frac{1}{6} \\ \mathbf{b}^T \mathbf{A} \mathbf{c}^2 &= \frac{1}{12} \\ \mathbf{b}^T (\mathbf{A} \mathbf{c} \cdot \mathbf{c}) &= \frac{1}{8} \end{aligned} \tag{4.1}$$

Continuando con el razonamiento seguido en el caso de las fórmulas de tres etapas, para obtener una clasificación de las distintas familias de métodos de este tipo necesitamos resolver el sistema de ecuaciones que plantean las ocho condiciones de orden anteriores. Esta tarea no resulta para nada sencilla, aunque ya Kutta fue capaz de aventurar los cinco casos en los que dicho sistema algebraico tiene solución. Nosotros haremos uso de los siguientes resultados.

**Lema 4.1.** *Dado un método Runge-Kutta explícito de orden cuatro, se tiene que los vectores,*

$$\mathbf{e}, \mathbf{c}, \mathbf{c}^2 \text{ y } A\mathbf{c},$$

*son linealmente independientes (es decir, forman una base de  $\mathbb{R}^4$ ).*

*Demostración.* Tomamos  $\lambda_1, \lambda_2, \lambda_3$  y  $\lambda_4$  tales que  $\lambda_1\mathbf{e} + \lambda_2\mathbf{c} + \lambda_3\mathbf{c}^2 + \lambda_4A\mathbf{c} = \mathbf{0}$ . Como  $c_1 = 0$  entonces  $\lambda_1 = 0$  y en consecuencia,  $\lambda_2\mathbf{c} + \lambda_3\mathbf{c}^2 + \lambda_4A\mathbf{c} = \mathbf{0}$ . Si multiplicamos, ahora, esta expresión por la izquierda por  $\mathbf{b}^T$  y  $\mathbf{b} \cdot \mathbf{c}$ , respectivamente, obtenemos:

$$\triangleright \mathbf{b}^T (\lambda_2\mathbf{c} + \lambda_3\mathbf{c}^2 + \lambda_4A\mathbf{c}) = 0 \quad \Leftrightarrow \quad \frac{1}{2}\lambda_2 + \frac{1}{3}\lambda_3 + \frac{1}{6}\lambda_4 = 0.$$

$$\triangleright \mathbf{b} \cdot \mathbf{c} (\lambda_2\mathbf{c} + \lambda_3\mathbf{c}^2 + \lambda_4A\mathbf{c}) = 0 \quad \Leftrightarrow \quad \frac{1}{3}\lambda_2 + \frac{1}{4}\lambda_3 + \frac{1}{8}\lambda_4 = 0.$$

Luego resolviendo el sistema anterior, obtenemos que  $\lambda_2 = 0$  y  $\lambda_4 = -2\lambda_3$ , de donde se deduce que  $\lambda_3\mathbf{c}^2 - 2\lambda_3A\mathbf{c} = \mathbf{0}$ . Finalmente, como el método no es reducible, tenemos que  $c_2 \neq 0$  y por tanto,  $\lambda_3 = 0$ , y, a su vez, también  $\lambda_4 = 0$ .  $\square$

**Corolario 4.1.** *Un método de Runge-Kutta explícito de cuatro etapas y orden cuatro verifica:*

$$\mathbf{b}^T A = \mathbf{b}^T - (\mathbf{b} \cdot \mathbf{c})^T$$

*Demostración.* Sea  $\mathbf{v}^T = \mathbf{b}^T A - \mathbf{b}^T + (\mathbf{b} \cdot \mathbf{c})^T$ .

Entonces, por las condiciones de orden cuatro, se tiene que:

- $\mathbf{v}^T \mathbf{e} = \mathbf{b}^T A\mathbf{e} - \mathbf{b}^T \mathbf{e} + \mathbf{b}^T \mathbf{c} = \frac{1}{2} - 1 + \frac{1}{2} = 0.$
- $\mathbf{v}^T \mathbf{c} = \mathbf{b}^T A\mathbf{c} - \mathbf{b}^T \mathbf{c} + \mathbf{b}^T \mathbf{c}^2 = \frac{1}{6} - \frac{1}{2} + \frac{1}{3} = 0.$
- $\mathbf{v}^T \mathbf{c}^2 = \mathbf{b}^T A\mathbf{c}^2 - \mathbf{b}^T \mathbf{c}^2 + \mathbf{b}^T \mathbf{c}^3 = \frac{1}{12} - \frac{1}{3} + \frac{1}{4} = 0.$
- $\mathbf{v}^T A\mathbf{c} = \mathbf{b}^T A^2\mathbf{c} - \mathbf{b}^T A\mathbf{c} + \mathbf{b}^T (A\mathbf{c} \cdot \mathbf{c}) = \frac{1}{24} - \frac{1}{6} + \frac{1}{8} = 0.$

La única posibilidad que hay es que  $\mathbf{v}^T = \mathbf{0}$  y por tanto  $\mathbf{b}^T A = \mathbf{b}^T - (\mathbf{b} \cdot \mathbf{c})^T$ .  $\square$

**Lema 4.2.** *Un método de Runge-Kutta explícito de cuatro etapas alcanza orden cuatro si y solamente si se verifican las siguientes ecuaciones:*

$$\begin{aligned} \mathbf{b}^T \mathbf{e} &= 1 \\ \mathbf{b}^T \mathbf{c} &= \frac{1}{2} \\ \mathbf{b}^T \mathbf{c}^2 &= \frac{1}{3} \\ \mathbf{b}^T \mathbf{c}^3 &= \frac{1}{4} \\ \mathbf{b}^T A^2\mathbf{c} &= \frac{1}{24} \\ \mathbf{b}^T A &= \mathbf{b}^T - (\mathbf{b} \cdot \mathbf{c})^T \end{aligned} \tag{4.2}$$

*Demostración.* Sabemos que un método de Runge-kutta explícito de cuatro etapas alcanza orden cuatro si y solo si se verifican las ocho ecuaciones de (4.1). Ahora bien, si comparamos estas con las ecuaciones de (4.2) observamos que las cinco primeras ecuaciones de cada caso coinciden, por lo que, cumpliéndose estas, para demostrar el resultado es bastará probar que las ecuaciones restantes de ambos casos también son equivalentes, esto es,

$$\left. \begin{aligned} \mathbf{b}^T A \mathbf{c} &= \frac{1}{6} \\ \mathbf{b}^T A \mathbf{c}^2 &= \frac{1}{12} \\ \mathbf{b}^T (A \mathbf{c} \cdot \mathbf{c}) &= \frac{1}{8} \end{aligned} \right\} \iff \mathbf{b}^T A = \mathbf{b}^T - (\mathbf{b} \cdot \mathbf{c})^T.$$

Asumimos que las cinco igualdades que coinciden en (4.1) y (4.2) son ciertas y si, además, se verifican las tres ecuaciones de la izquierda, entonces se cumplen todas las ecuaciones de (4.1) y, en consecuencia, el método alcanza orden cuatro. En ese caso, por el Corolario 4.1 se tiene la condición de la derecha. Recíprocamente, multiplicando la igualdad de la derecha por  $\mathbf{c}$ ,  $\mathbf{c}^2$  y  $A\mathbf{c}$ , respectivamente, y teniendo en cuenta las cinco expresiones anteriores podemos comprobar que se satisfacen las tres ecuaciones de la izquierda. En efecto:

- $\mathbf{b}^T A \mathbf{c} = \mathbf{b}^T \mathbf{c} - \mathbf{b}^T \mathbf{c}^2 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ .
- $\mathbf{b}^T A \mathbf{c}^2 = \mathbf{b}^T \mathbf{c}^2 - \mathbf{b}^T \mathbf{c}^3 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$ .
- $\mathbf{b}^T A^2 \mathbf{c} = \mathbf{b}^T A \mathbf{c} - \mathbf{b}^T (A \mathbf{c} \cdot \mathbf{c}) \Rightarrow \mathbf{b}^T (A \mathbf{c} \cdot \mathbf{c}) = \frac{1}{6} - \frac{1}{24} = \frac{1}{8}$ . □

Por consiguiente, desarrollando las seis igualdades de (4.2), concluimos que las ecuaciones algebraicas que deben cumplir los coeficientes de una fórmula de Runge-Kutta explícita de cuatro etapas para alcanzar orden cuatro son las siguientes, aunque es necesario tener en cuenta algunas consideraciones adicionales (véase la Nota 4.1):

$$\begin{aligned} (1) \quad b_1 + b_2 + b_3 + b_4 &= 1 \\ (2) \quad b_2 c_2 + b_3 c_3 + b_4 c_4 &= \frac{1}{2} \\ (3) \quad b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2 &= \frac{1}{3} \\ (4) \quad b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 &= \frac{1}{4} \\ (5) \quad b_4 a_{43} a_{32} c_2 &= \frac{1}{24} \\ (6) \quad b_4 a_{42} + b_3 a_{32} &= b_2(1 - c_2) \\ (7) \quad b_4 a_{43} &= b_3(1 - c_3) \\ (8) \quad c_4 &= 1 \end{aligned}$$

a las que hay que añadir las tres ecuaciones que se derivan de la condición de suma de filas:

$$\begin{aligned} (9) \quad c_2 &= a_{21} \\ (10) \quad c_3 &= a_{31} + a_{32} \\ (11) \quad c_4 &= a_{41} + a_{42} + a_{43} \end{aligned}$$

**Nota 4.1.** Las ecuaciones (6), (7) y (8) anteriores se deducen a partir de la última igualdad de la colección (4.2), esto es, se obtienen desarrollando la siguiente expresión:

$$\mathbf{b}^T A = \mathbf{b}^T - (\mathbf{b} \cdot \mathbf{c})^T.$$

Sin embargo, si escribimos en extenso la transpuesta de dicha igualdad observamos que no se ha tenido en cuenta la primera ecuación que aparece. En efecto,

$$\begin{pmatrix} b_2 a_{21} + b_3 a_{31} + b_4 a_{41} \\ b_3 a_{32} + b_4 a_{42} \\ b_4 a_{43} \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2(1 - c_2) \\ b_3(1 - c_3) \\ b_4(1 - c_4) \end{pmatrix} \quad (4.3)$$

La razón por la que la primera expresión no se incluye en el listado (1 – 11) anterior es que es redundante, es decir, se puede obtener a partir de las demás como se prueba a continuación:

$$\begin{aligned} b_1 &= b_2 a_{21} + b_3 a_{31} + b_4 a_{41} \stackrel{(9-11)}{\iff} b_1 = b_2 c_2 + b_3(c_3 - a_{32}) + b_4(c_4 - a_{42} - a_{43}) \\ \stackrel{(2,6,7)}{\iff} b_1 &= \frac{1}{2} - b_2(1 - c_2) - b_3(1 - c_3) \stackrel{(2)}{\iff} b_1 = 1 - b_2 - b_3 - b_4 c_4 \\ \stackrel{(8)}{\iff} b_1 &= 1 - b_2 - b_3 - b_4 \end{aligned}$$

Así, la primera igualdad de (4.3) es equivalente a la ecuación (1), y, por tanto, solo es necesario añadir una de ellas, de ahí que únicamente se consideren las expresiones (1 – 11) dadas.

La resolución del sistema que plantean estas once ecuaciones no es una tarea sencilla. Kutta fue capaz de encontrar los cinco casos para los que existe una solución pero no ofreció ninguna explicación de como había llegado a ellos.

Exponemos a continuación un posible razonamiento para resolver dicho sistema algebraico y como consecuencia mostramos las distintas familias de métodos de Runge-Kutta explícitos de cuatro etapas y orden cuatro que resultan.

En primer lugar, observamos que todos los métodos que se obtienen no son reducibles, pues  $b_4 \neq 0$  y  $c_2 \neq 0$ , puesto que la ecuación (5), cuya formulación es:

$$b_4 a_{43} a_{32} c_2 = \frac{1}{24},$$

implica que todos los factores de la izquierda,  $b_4$  y  $c_2$ , entre ellos, son no nulos. Ahora, sutituyendo la expresión (7) en la condición (5), podemos deducir la siguiente igualdad,

$$b_3(1 - c_3) a_{32} c_2 = \frac{1}{24},$$

que, como antes, nos permite concluir inmediatamente que  $b_3 \neq 0$  y  $c_3 \neq 1$ . De esta manera, tomaremos como parámetros los coeficientes  $c_2$  y  $c_3$ , que satisfacen que  $c_2 \neq 0$  y  $c_3 \neq 1$ .

A continuación, calculamos los valores de  $b_2$ ,  $b_3$  y  $b_4$  resolviendo el sistema lineal planteado por las ecuaciones (2), (3) y (4), es decir,

$$\begin{pmatrix} c_2 & c_3 & 1 \\ c_2^2 & c_3^2 & 1 \\ c_2^3 & c_3^3 & 1 \end{pmatrix} \begin{pmatrix} b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{4} \end{pmatrix}$$

y acto seguido usamos (1) para obtener  $b_1$ . La resolución de dicho sistema de ecuaciones nos lleva a distinguir los siguientes cinco casos:

- Si  $c_3(1 - c_2)(c_2 - c_3) \neq 0$  obtenemos los siguientes resultados:

$$\begin{aligned} b_4 &= \frac{3 - 4(c_2 + c_3) + 6c_2c_3}{12(1 - c_2)(1 - c_3)}, \\ b_3 &= \frac{1 - 2c_2}{12c_3(1 - c_3)(c_3 - c_2)}, \\ b_2 &= \frac{2c_3 - 1}{12c_2(1 - c_2)(c_3 - c_2)}, \\ b_1 &= \frac{1 - 2(c_2 + c_3) + 6c_2c_3}{12c_2c_3}. \end{aligned}$$

Recordemos que  $b_3 \neq 0$  y  $b_4 \neq 0$ . Sobre esto cabe puntualizar lo siguiente:

- Si  $c_2 \neq \frac{1}{2}$  es claro que  $b_3 \neq 0$  y en ese caso  $b_4 \neq 0$  si y solo si  $3 - 4(c_2 + c_3) + 6c_2c_3 \neq 0$ . Los valores de  $b_1$ ,  $b_2$ ,  $b_3$  y  $b_4$  serían, por tanto, los que acabamos de señalar.
- Si  $c_3 = \frac{1}{2}$  entonces  $b_3 = \frac{2}{3}$  independientemente del valor de  $c_2$ . En ese caso, además, tendríamos que  $b_1 = \frac{1}{6}$ ,  $b_2 = 0$  y  $b_4 = \frac{1}{6}$ .
- Si  $c_3 = 0$  entonces el sistema tiene solución solo cuando  $c_2 = \frac{1}{2}$ . En ese caso, los coeficientes anteriores resultan ser:  $b_3 \neq 0$ ,  $b_2 = \frac{2}{3}$ ,  $b_4 = \frac{1}{6}$  y  $b_1 = \frac{1}{6} - b_3$ .
- Si  $c_2 = 1$  solo existirá solución si  $c_3 = \frac{1}{2}$ . Obtenemos entonces que los parámetros anteriores toman los siguientes valores:  $b_4 \neq 0$ ,  $b_2 = \frac{1}{6} - b_4$ ,  $b_3 = \frac{2}{3}$  y  $b_1 = \frac{1}{6}$ .
- Si  $c_2 = c_3$  deducimos que la única posibilidad para que exista solución es que  $c_2 = c_3 = \frac{1}{2}$  y los resultados obtenidos en ese caso son:  $b_3 \neq 0$ ,  $b_2 = \frac{2}{3} - b_3$ ,  $b_4 = \frac{1}{6}$  y  $b_1 = \frac{1}{6}$ .

Una vez calculados los valores de  $b_1$ ,  $b_2$ ,  $b_3$  y  $b_4$  y teniendo en cuenta los parámetros  $c_2$  y  $c_3$ , podemos determinar los demás coeficientes. El valor de  $a_{32}$  se deduce de la ecuación

$$b_3(1 - c_3)a_{32}c_2 = \frac{1}{24},$$

y, finalmente, los coeficientes  $a_{31}$ ,  $a_{43}$ ,  $a_{42}$ ,  $a_{41}$  pueden obtenerse fácilmente a partir de las condiciones (10), (7), (6) y (11), respectivamente.

Conseguimos clasificar, de esta manera, los métodos de Runge-Kutta explícitos de cuatro etapas y orden cuatro en cinco familias paramétricas distintas. Aunque los cinco casos que resultan recogen todas las situaciones que pueden darse, no son excluyentes pues, por ejemplo, el caso (II) se solapa con los casos (I), (IV) y (V). En efecto, si en (II) elegimos  $c_2 = \frac{1}{2}$  y en (V),  $b_3 = \frac{2}{3}$ , obtenemos el mismo método.

Las tablas de Butcher resultantes en cada caso son las siguientes:

(I)  $c_2 \notin \{0, \frac{1}{2}, 1\}$ ,  $c_3 \notin \{0, 1, c_2\}$  y  $3 - 4(c_2 + c_3) + 6c_2c_3 \neq 0$ :

0	0			
$c_2$	$c_2$	0		
$c_3$	$\frac{c_3(3c_2 - c_3 - 4c_2^2)}{2c_2(1 - 2c_2)}$	$\frac{c_3(c_3 - c_2)}{2c_2(1 - 2c_2)}$	0	
1	$a_{41}$	$a_{42}$	$a_{43}$	0
	$\frac{1 - 2(c_2 + c_3) + 6c_2c_3}{12c_2c_3}$	$\frac{2c_3 - 1}{12c_2(1 - c_2)(c_3 - c_2)}$	$\frac{1 - 2c_2}{12c_3(1 - c_3)(c_3 - c_2)}$	$\frac{3 - 4(c_2 + c_3) + 6c_2c_3}{12(1 - c_2)(1 - c_3)}$

donde los valores de  $a_{41}$ ,  $a_{42}$  y  $a_{43}$  vienen dados por:

$$a_{41} = \frac{c_3^2(12c_2^2 - 12c_2 + 4) - c_3(12c_2^2 - 15c_2 + 5) + 4c_2^2 - 6c_2 + 2}{2c_2c_3(3 - 4(c_2 + c_3) + 6c_2c_3)},$$

$$a_{42} = \frac{(1 - c_2)(c_2 + 5c_3 - 4c_3^2 - 2)}{2c_2(c_3 - c_2)(3 - 4(c_2 + c_3) + 6c_2c_3)},$$

$$a_{43} = \frac{(1 - c_2)(1 - c_3)(1 - 2c_2)}{c_3(c_3 - c_2)(3 - 4(c_2 + c_3) + 6c_2c_3)}.$$

(II)  $c_2 \neq 0$ ,  $c_3 = \frac{1}{2}$ :

0	0			
$c_2$	$c_2$	0		
$\frac{1}{2}$	$\frac{1}{2} - \frac{1}{8c_2}$	$\frac{1}{8c_2}$	0	
1	$\frac{1}{2c_2} - 1$	$-\frac{1}{2c_2}$	2	0
	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$

(III)  $c_2 = \frac{1}{2}$ ,  $c_3 = 0$ :

0	0			
$\frac{1}{2}$	$\frac{1}{2}$	0		
0	$-\frac{1}{12b_3}$	$\frac{1}{12b_3}$	0	
1	$-\frac{1}{2} - 6b_3$	$\frac{3}{2}$	$6b_3$	0
	$\frac{1}{6} - b_3$	$\frac{2}{3}$	$b_3$	$\frac{1}{6}$

(IV)  $c_2 = 1, c_3 = \frac{1}{2}$ :

0	0				
1	1	0			
$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{8}$	0		
1	$1 - \frac{1}{4b_4}$	$-\frac{1}{12b_4}$	$\frac{1}{3b_4}$	0	
	$\frac{1}{6}$	$\frac{1}{6} - b_4$	$\frac{2}{3}$	$b_4$	

(V)  $c_2 = c_3 = \frac{1}{2}$ :

0	0				
$\frac{1}{2}$	$\frac{1}{2}$	0			
$\frac{1}{2}$	$\frac{1}{2} - \frac{1}{6b_3}$	$\frac{1}{6b_3}$	0		
1	0	$1 - 3b_3$	$3b_3$	0	
	$\frac{1}{6}$	$\frac{2}{3} - b_3$	$b_3$	$\frac{1}{6}$	

**Ejemplo 4.3.** Si elegimos  $b_3 = \frac{1}{3}$  en este caso, obtenemos la tabla de Butcher asociada al conocido como método de Runge-Kutta clásico de cuatro etapas y orden cuatro, formulado por el propio Kutta en el año 1901.

0	0				
$\frac{1}{2}$	$\frac{1}{2}$	0			
$\frac{1}{2}$	0	$\frac{1}{2}$	0		
1	0	0	1	0	
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	

## 4.5. Métodos de Runge-Kutta explícitos de cinco etapas

En una fórmula de Runge-Kutta explícita que consta de cinco etapas disponemos, a la vista de lo que aparece recogido en la Tabla 4.1, de 15 coeficientes por determinar. Si siguiendo la línea que parecen indicarnos los desarrollos realizados hasta el momento quisiéramos que tal fórmula alcanzase orden cinco, de acuerdo con la Tabla 4.2 tendrían que verificarse 17 condiciones. Esto supone un cambio con lo planteado hasta ahora pues, por primera vez, nos encontramos con que el número de parámetros es inferior al número de ecuaciones. El primero en abordar el estudio de esta familia de métodos fue, una vez más, Kutta, quien ya a comienzos del siglo pasado sugirió la imposibilidad de alcanzar orden cinco con “solo” cinco etapas, sin embargo, no fue hasta mediados de los años sesenta cuando, de forma totalmente independiente, varios autores (entre ellos Butcher) consiguieron demostrarlo. Para ello vamos a seguir un razonamiento similar al que se hizo en el caso de cuatro etapas.

**Lema 4.3.** *Si existiese algún método de Runge-Kutta explícito de cinco etapas y orden cinco, debería cumplirse que los cinco vectores,*

$$\mathbf{e}, \mathbf{c}, \mathbf{c}^2, \mathbf{Ac} \text{ y } \mathbf{e}_5 = (0, 0, 0, 0, 1)^T,$$

*fuesen linealmente independientes (y por tanto formasen una base de  $\mathbb{R}^5$ ).*

*Demostración.* Tomamos  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  y  $\lambda_5$  tales que  $\lambda_1 \mathbf{e} + \lambda_2 \mathbf{c} + \lambda_3 \mathbf{c}^2 + \lambda_4 \mathbf{Ac} + \lambda_5 \mathbf{e}_5 = \mathbf{0}$ . Como  $c_1 = 0$  entonces  $\lambda_1 = 0$  y en consecuencia,  $\lambda_2 \mathbf{c} + \lambda_3 \mathbf{c}^2 + \lambda_4 \mathbf{Ac} + \lambda_5 \mathbf{e}_5 = \mathbf{0}$ . Ahora, si multiplicamos esta expresión por la izquierda por los vectores  $\mathbf{b}^T A$  y  $\mathbf{b}^T A^2$ , respectivamente, obtenemos las dos ecuaciones siguientes:

$$\triangleright \mathbf{b}^T A (\lambda_2 \mathbf{c} + \lambda_3 \mathbf{c}^2 + \lambda_4 \mathbf{Ac} + \lambda_5 \mathbf{e}_5) = 0 \quad \Leftrightarrow \quad \frac{1}{6} \lambda_2 + \frac{1}{12} \lambda_3 + \frac{1}{24} \lambda_4 = 0,$$

$$\triangleright \mathbf{b}^T A^2 (\lambda_2 \mathbf{c} + \lambda_3 \mathbf{c}^2 + \lambda_4 \mathbf{Ac} + \lambda_5 \mathbf{e}_5) = 0 \quad \Leftrightarrow \quad \frac{1}{24} \lambda_2 + \frac{1}{60} \lambda_3 + \frac{1}{120} \lambda_4 = 0,$$

Luego resolviendo el sistema anterior, obtenemos que  $\lambda_2 = 0$  y  $\lambda_4 = -2\lambda_3$ , de donde se deduce que  $\lambda_3 \mathbf{c}^2 - 2\lambda_3 \mathbf{Ac} + \lambda_5 \mathbf{e}_5 = \mathbf{0}$ . Finalmente, como el método no es reducible, tenemos que  $c_2 \neq 0$  y por tanto,  $\lambda_3 = 0$ , y, a su vez, también  $\lambda_4 = 0$ , y, así,  $\lambda_5 = 0$ .  $\square$

**Lema 4.4.** *Si existiese algún método de Runge-Kutta explícito de cinco etapas y orden cinco, entonces se verificarían las dos siguientes igualdades:*

$$\mathbf{u}^T = \left( \mathbf{b}^T A - \mathbf{b}^T + (\mathbf{b} \cdot \mathbf{c})^T \right) A = 0,$$

$$\mathbf{v}^T = 2\mathbf{b}^T A^2 - \mathbf{b}^T A - \mathbf{b}^T A \cdot \mathbf{c}^T = 0.$$

*Demostración.* Por las condiciones de orden cinco (vease la Tabla 3.6) tenemos que:

- $\mathbf{u}^T \mathbf{e} = \mathbf{b}^T \mathbf{Ac} - \mathbf{b}^T \mathbf{c} + \mathbf{b}^T \mathbf{c}^2 = \frac{1}{6} - \frac{1}{2} + \frac{1}{3} = 0.$
- $\mathbf{u}^T \mathbf{c} = \mathbf{b}^T A^2 \mathbf{c} - \mathbf{b}^T \mathbf{Ac} + \mathbf{b}^T (\mathbf{Ac} \cdot \mathbf{c}) = \frac{1}{24} - \frac{1}{6} + \frac{1}{8} = 0.$
- $\mathbf{u}^T \mathbf{c}^2 = \mathbf{b}^T A^2 \mathbf{c}^2 - \mathbf{b}^T \mathbf{Ac}^2 + \mathbf{b}^T (\mathbf{Ac}^2 \cdot \mathbf{c}) = \frac{1}{60} - \frac{1}{12} + \frac{1}{15} = 0.$
- $\mathbf{u}^T \mathbf{Ac} = \mathbf{b}^T A^3 \mathbf{c} - \mathbf{b}^T A^2 \mathbf{c} + \mathbf{b}^T (A^2 \mathbf{c} \cdot \mathbf{c}) = \frac{1}{120} - \frac{1}{24} + \frac{1}{30} = 0.$
- $\mathbf{u}^T \mathbf{e}_5 = \left( \mathbf{b}^T A - \mathbf{b}^T + (\mathbf{b} \cdot \mathbf{c})^T \right) A \mathbf{e}_5 = 0$  por ser la última columna de A toda cero.

Entonces, como por el Lema 4.3 anterior, los vectores  $\mathbf{e}, \mathbf{c}, \mathbf{c}^2, \mathbf{Ac}$  y  $\mathbf{e}_5$  son linealmente independientes, es decir, forman una base de  $\mathbb{R}^5$ , la única posibilidad es que  $\mathbf{u}^T = \mathbf{0}$ .

De forma completamente análoga se prueba que  $\mathbf{v}^T = \mathbf{0}$ . En este caso, habría que calcular:

$$\mathbf{v}^T \mathbf{e} = 2\mathbf{b}^T A^2 \mathbf{c} - \mathbf{b}^T \mathbf{c} + \mathbf{b}^T \mathbf{Ac} = 2\frac{1}{6} - \frac{1}{2} + \frac{1}{6} = 0.$$

$$\mathbf{v}^T \mathbf{c} = 2\mathbf{b}^T A^2 \mathbf{c} - \mathbf{b}^T \mathbf{Ac} + \mathbf{b}^T \mathbf{Ac}^2 = 2\frac{1}{24} - \frac{1}{6} + \frac{1}{12} = 0.$$

- $\mathbf{v}^T \mathbf{c}^2 = 2\mathbf{b}^T A^2 \mathbf{c}^2 - \mathbf{b}^T A \mathbf{c}^2 + \mathbf{b}^T A \mathbf{c}^3 = 2\frac{1}{60} - \frac{1}{12} + \frac{1}{20} = 0.$
- $\mathbf{v}^T A \mathbf{c} = 2\mathbf{b}^T A^3 \mathbf{c} - \mathbf{b}^T A^2 \mathbf{c} + \mathbf{b}^T A(A \mathbf{c} \cdot \mathbf{c}) = 2\frac{1}{120} - \frac{1}{24} + \frac{1}{40} = 0.$
- $\mathbf{v}^T \mathbf{e}_5 = (2\mathbf{b}^T A^2 - \mathbf{b}^T A - \mathbf{b}^T A \cdot \mathbf{c}^T) \mathbf{e}_5 = 0$  por ser la última columna de A toda cero.

Y, de nuevo, razonando como antes, la única posibilidad sería que  $\mathbf{v}^T = \mathbf{0}$ . □

**Corolario 4.2.** *Si existiese algún método de Runge-Kutta explícito de cinco etapas y orden cinco, entonces debería cumplirse la siguiente condición:*

$$c_4 = c_5 = 1.$$

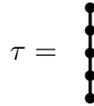
*Demostración.* Sabemos, por Lema 4.4, que en las condiciones del enunciado del Corolario, se cumple que los vectores definidos como:

$$\begin{aligned} \mathbf{u}^T &= (\mathbf{b}^T A - \mathbf{b}^T + (\mathbf{b} \cdot \mathbf{c})^T) A, \\ \mathbf{v}^T &= 2\mathbf{b}^T A^2 - \mathbf{b}^T A - \mathbf{b}^T A \cdot \mathbf{c}^T, \end{aligned}$$

son idénticamente nulos y, por tanto, si nos fijamos en sus cuartas componentes, resulta que:

$$\begin{aligned} u_4 &= b_5(c_5 - 1)a_{54} = 0, \\ v_4 &= b_5(c_4 - 1)a_{54} = 0. \end{aligned}$$

Además, teniendo en cuenta la condición de orden asociada al árbol  $\tau$  siguiente:



cuya formulación, según el Teorema 3.4, es:

$$\Psi(\tau) = \frac{1}{\gamma(\tau)}, \text{ y que, en este caso, equivale a que } b_5 a_{54} a_{43} a_{32} c_2 = \frac{1}{120},$$

podemos deducir que  $b_5 \neq 0$  y  $a_{54} \neq 0$  y, entonces,  $u_4 = 0$  si y solo si  $c_5 = 1$  y, de la misma manera, obtendríamos que  $c_4 = 1$  es condición necesaria y suficiente para que  $v_4 = 0$ . □

**Teorema 4.1.** *No es posible alcanzar orden cinco con un método de Runge-Kutta explícito de cinco etapas.*

*Demostración.* Razonamos por reducción al absurdo, suponiendo que existen fórmulas de Runge-Kutta explícitas de cinco etapas que sí alcanzan orden cinco. En ese caso, se cumplirían las condiciones de orden asociadas a los dos árboles  $\tau_1$  y  $\tau_2$  siguientes,



cuyas expresiones, según el Teorema 3.4 y la Tabla 3.6, son las siguientes:

$$\Psi(\tau_1) = \frac{1}{\gamma(\tau_1)}, \text{ que equivale a que } \mathbf{b}^T A^2 \mathbf{c} = \frac{1}{24},$$

$$\Psi(\tau_2) = \frac{1}{\gamma(\tau_2)}, \text{ que equivale a que } \mathbf{b}^T (A^2 \mathbf{c} \cdot \mathbf{c}) = \frac{1}{30}.$$

Luego, restando ambas ecuaciones concluimos que:

$$\mathbf{b}^T ((\mathbf{e} - \mathbf{c}) \cdot A^2 \mathbf{c}) = \mathbf{b}^T A^2 \mathbf{c} - \mathbf{b}^T (A^2 \mathbf{c} \cdot \mathbf{c}) = \frac{1}{24} - \frac{1}{30} \neq 0. \quad (4.4)$$

Además, el Corolario 4.2 nos permite deducir que el vector,

$$(\mathbf{e} - \mathbf{c})^T = (1, 1, 1, 1, 1)^T - (0, c_2, c_3, 1, 1)^T = (1, 1 - c_2, 1 - c_3, 0, 0)^T,$$

y como también sabemos que,

$$A^2 \mathbf{c} = (0, 0, 0, x, y)^T, \quad \text{con } x, y \in \mathbb{R},$$

entonces podemos concluir que:

$$(\mathbf{e} - \mathbf{c}) \cdot A^2 \mathbf{c} = 0,$$

lo que contradice el resultado obtenido en (4.4) y nos permite afirmar que un método de Runge-Kutta explícito de cinco etapas no puede alcanzar orden cinco.  $\square$

## 4.6. Conclusiones sobre métodos de Runge-Kutta explícitos

Podemos generalizar el Teorema 4.1 para concluir lo siguiente:

**Teorema 4.2.** *Si  $q \geq 5$  no es posible alcanzar orden  $q$  con un método de Runge-Kutta explícito de  $q$  etapas.*

*Demostración.* El razonamiento a seguir es similar al que se ha explicado en la sección anterior para el caso  $q = 5$ . En primer lugar, se prueba que, si existiese alguna fórmula con  $q$  etapas que alcanzase orden  $q$ , entonces los vectores

$$\mathbf{e}, \mathbf{c}, \mathbf{c}^2, A\mathbf{c}, \mathbf{e}_5, \dots, \mathbf{e}_q$$

serían linealmente independientes, es decir, formarían una base de  $\mathbb{R}^q$ . A continuación, se demuestra que, en caso de existir algún método con  $q$  etapas y orden  $q$ , las dos últimas componentes del vector de nodos  $\mathbf{c}$  deberían cumplir la condición

$$c_q = c_{q-1} = 1.$$

Finalmente, a partir de estos dos resultados, buscaríamos una contradicción.  $\square$

Para finalizar esta sección y este capítulo, conviene resaltar que tras la introducción de los primeros métodos de Runge-Kutta explícitos hacia finales del siglo XIX por Runge y las posteriores contribuciones de Kutta y Heun a comienzos del siglo XX, el interés se centró entonces en encontrar fórmulas de orden cada vez mayor con el menor número de etapas posible. Muchos fueron los autores que se ocuparon de este asunto, algunos de los cuales aparecen señalados en la tabla siguiente, en la que se recoge la evolución histórica del estudio del orden en función del número de etapas de las fórmulas de Runge-Kutta explícitas.

Orden	Etapas	Autor	Año
2	2	Coriolis	1837
2	2	Runge	1895
3	4	Runge	1895
3	3	Heun	1900
4	8	Heun	1900
4	4	Kutta	1901
5	6	Kutta	1901
5	6	Nyström	1925
6	8	Hutta	1956
6	7	Butcher	1964
7	9	Butcher	1968
8	11	Curtis	1970
8	11	Cooper y Verner	1972
10	18	Curtis	1975
10	17	Hairer	1978

Tabla 4.3: Evolución estudio orden y número de etapas en los métodos de RK explícitos.

La cuestión de qué orden puede alcanzarse con un método explícito de  $q$  etapas es un problema que sigue abierto. En la tabla siguiente se recoge la información disponible hasta la fecha sobre el número mínimo de etapas que debe tener un método para alcanzar un cierto orden.

Orden	1	2	3	4	5	6	7	8	9	10
Etapas	1	2	3	4	6	7	9	11	$12 \leq q \leq 17$	$13 \leq q \leq 17$

Tabla 4.4: Número mínimo de etapas para alcanzar un orden determinado.



## Capítulo 5

# Métodos de Runge-Kutta implícitos

Los métodos de Runge-Kutta que hemos introducido en capítulos anteriores son fórmulas de un paso de tipo explícito: la aproximación en un instante concreto depende de manera explícita de la aproximación en el instante anterior; en la práctica, esto se traduce en que la función incremento  $\Phi$  tiene una expresión explícita. En un inicio, fueron precisamente este tipo de métodos los que ocuparon las principales líneas de investigación en el campo del Análisis Numérico de las ecuaciones diferenciales ordinarias. Se trataba entonces de encontrar fórmulas explícitas de orden cada vez más elevado, intentando incurrir en el menor gasto computacional posible, lo que se traducía en la utilización del menor número de etapas necesarias para alcanzar el orden deseado. Sin embargo, la “aparición”, a mediados del siglo XX de cierto tipo de problemas diferenciales, los llamados *problemas stiff* o *problemas rígidos*, que aparecen frecuentemente en la práctica (reacciones químicas, circuitos eléctricos, etc.) y cuya resolución mediante fórmulas explícitas resulta casi imposible, motivó la consideración de otro tipo de métodos, más generales, que son de tipo implícito. En palabras del propio Dahlquist, uno de los principales estudiosos de esta clase de problemas, «*a partir de los años sesenta las tornas cambiaron radicalmente y todo el mundo se dio cuenta de que el universo a nuestro alrededor estaba lleno de problemas stiff*». Aunque la definición formal no está nada clara, sí podemos dar una idea general de lo que significa este fenómeno. Cuando se resuelve numéricamente una ecuación diferencial, uno espera que el tamaño del paso sea relativamente pequeño en aquellas regiones donde la curva solución experimenta mucha variación (fase transitoria) y que, por el contrario, sea relativamente grande donde la curva solución se mantiene prácticamente constante (fase estacionaria). Sin embargo, en algunas ocasiones el tamaño del paso propuesto es tan pequeño que el número de iteraciones necesarias para completar el proceso es tan grande que resulta inviable. En los métodos implícitos, la aproximación en un instante se define mediante una función implícita en la que interviene la aproximación en el instante anterior. La función incremento  $\Phi$  también se conoce de forma implícita, lógicamente. En general, el coste computacional es mayor que en el caso de los métodos explícitos, pero presentan buenas propiedades de estabilidad, alcanzan mayor orden a igual número de etapas y en ocasiones son la única alternativa eficaz para resolver problemas *stiff*.

## 5.1. Definición y primeras propiedades

Recordemos, en primer lugar, que a lo largo de todo este trabajo estamos estudiando la resolución numérica de un problema de valor inicial cuya formulación es la siguiente:

$$(P) \quad \begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \boldsymbol{\eta}, \end{cases}$$

donde  $\mathbf{f} : I \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  es una función continua y lipschitziana respecto de su segundo argumento (con constante de Lipschitz  $L > 0$  para alguna norma de  $\mathbb{R}^m$ ).

A diferencia de lo que ocurre con las fórmulas de Runge-Kutta explícitas, en las que la matriz de coeficientes  $A = (a_{i,j})_{i,j=1}^q \in \mathcal{M}_{q \times q}(\mathbb{R})$  es estrictamente triangular inferior, en el caso implícito dicha matriz es llena, por lo que la tabla de Butcher asociada toma la forma:

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1q} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_q & a_{q1} & a_{q2} & \cdots & a_{qq} \\ \hline & b_1 & b_2 & \cdots & b_q \end{array}$$

Nótese también que, ahora, el valor del nodo  $c_1$  no es necesariamente nulo como sucede en el caso de los métodos explícitos que satisfacen la condición de suma de filas y, por tanto, podemos formalizar las fórmulas de Runge-Kutta implícitas como sigue:

**Definición 5.1.** Un método de Runge-Kutta implícito de  $q$  etapas (abreviado RKI) para resolver numéricamente el problema (P) es un algoritmo de un paso

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \Phi(t_n, \mathbf{y}_n, h_n),$$

cuya función incremento se define como

$$\Phi(t, \mathbf{y}, h) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, h),$$

donde  $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_q$  resultan de calcular:

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t + c_1 h, \mathbf{y} + h a_{11} \mathbf{k}_1 + h a_{12} \mathbf{k}_2 + \cdots + h a_{1q} \mathbf{k}_q), \\ \mathbf{k}_2 &= \mathbf{f}(t + c_2 h, \mathbf{y} + h a_{21} \mathbf{k}_1 + h a_{22} \mathbf{k}_2 + \cdots + h a_{2q} \mathbf{k}_q), \\ &\vdots \\ \mathbf{k}_q &= \mathbf{f}(t + c_q h, \mathbf{y} + h a_{q1} \mathbf{k}_1 + h a_{q2} \mathbf{k}_2 + \cdots + h a_{qq} \mathbf{k}_q), \end{aligned} \tag{5.1}$$

con  $\{a_{ij}\}_{i,j=1}^q$ ,  $\{b_i\}_{i=1}^q$  y  $\{c_i\}_{i=1}^q$  tres colecciones de números reales que caracterizan el método.

Como sucede también en el caso explícito, las fórmulas de Runge-Kutta implícitas son, en efecto, métodos de un paso pues para obtener la solución aproximada en un concreto instante solo se necesita la información relativa a la solución numérica en el instante anterior. Así, una vez conocida la solución en un instante  $t_n$ , que sabemos viene determinada por la terna  $\{t_n, \mathbf{y}_n, h_n\}$ , podemos calcular la solución el instante siguiente,  $t_{n+1}$ , de la siguiente manera:

- En primer lugar, resolvemos el sistema de ecuaciones (posiblemente) no lineales (5.1).
- Una vez conocidos los valores de  $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_q$ , calculamos  $\Phi$ .
- Y finalmente, obtenemos  $\mathbf{y}_{n+1}$  a partir de  $\mathbf{y}_n, h_n$  y  $\Phi$ .

Por tanto, los resultados que se han presentado en el Capítulo 1 para el caso más general de métodos de un paso pueden aplicarse en particular a las fórmulas de Runge-Kutta implícitas.

Por otro lado, dependiendo de la forma de la matriz  $A$  distinguimos tres tipos de métodos:

- Si la matriz  $A$  es estrictamente triangular inferior, es decir, si  $a_{ij} = 0$  para  $i \leq j$ , entonces estamos en el caso ya conocido de los métodos de Runge-Kutta explícitos.
- Si la matriz  $A$  es triangular inferior, esto es, si  $a_{ij} = 0$  para  $i < j$  pero  $a_{kk} \neq 0$  para ciertos  $k \in \{1, 2, \dots, q\}$ , el método se dice *semi-implícito*, *semi-explícito* o *diagonalmente implícito* y se denota habitualmente por sus siglas en inglés, DIRK.

Cuando los elementos de la diagonal principal de la matriz  $A$  toman el mismo valor se habla de *métodos simplemente diagonalmente implícitos*, que también se denotan por sus siglas en inglés SDIRK (traducción literal de “*Singly Diagonally Implicit RK*”).

- Si la matriz  $A$  es llena, esto es, si  $a_{ij} \neq 0$  para algún  $i \leq j$ , el método se dice implícito.

También podemos señalar los métodos simplemente implícitos (abreviadamente, SIRK) que son aquellos métodos implícitos en los que la matriz  $A$  tiene un único valor propio.

**Ejemplo 5.1.** Métodos de Euler implícito y del punto medio.

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

**Ejemplo 5.2.** Algunos métodos de Runge-Kutta implícitos de dos etapas.

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 1 & \frac{1}{2} & \frac{1}{2} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \quad \begin{array}{c|cc} 3 - 2\sqrt{2} & \frac{5-3\sqrt{2}}{4} & \frac{7-5\sqrt{2}}{4} \\ \hline 1 & \frac{1+\sqrt{2}}{4} & \frac{3-\sqrt{2}}{4} \\ \hline & \frac{1+\sqrt{2}}{4} & \frac{3-\sqrt{2}}{4} \end{array}$$

Una posibilidad alternativa para definir un método de Runge-Kutta implícito es:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}), \quad (5.2)$$

donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q.$$

Otra formulación equivalente a las ya señaladas es también la que sigue:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \sum_{i=1}^q b_i \mathbf{g}_i, \quad (5.3)$$

con

$$\mathbf{g}_i = h_n \mathbf{f}(t_n + c_i h_n, \mathbf{y}_n + \sum_{j=1}^q a_{ij} \mathbf{g}_j), \quad i = 1, 2, \dots, q.$$

El principal inconveniente de las fórmulas de Runge-Kutta implícitas radica en la dificultad de resolver el sistema algebraico que plantean las ecuaciones de las etapas puesto que, a diferencia del caso explícito, ahora no se pueden resolver recursivamente una tras otra al intervenir cada una de ellas en todas las demás ecuaciones. Por consiguiente, nuestro primer objetivo va a ser garantizar “la buena definición” del método, esto es, conseguir resultados que permitan asegurar la existencia y la unicidad de solución de dicho sistema no lineal. Además, bajo determinadas condiciones, las fórmulas implícitas son estables y satisfacen ciertas propiedades de regularidad, pero como las demostraciones de todos estos resultados son excesivamente técnicas y se desvían notablemente del objetivo del trabajo, que es la Teoría de Butcher, se ha optado por omitirlas en esta sección, aunque se incluyen en el Apéndice A.

**Teorema 5.1.** *El sistema de ecuaciones (5.1) que definen las etapas de una fórmula de Runge-Kutta implícita tiene solución única y en consecuencia el método está bien definido si*

$$h_n L \|A\|_\infty < 1,$$

donde  $L$  es una constante de Lipschitz de la función  $\mathbf{f}$  respecto a cierta norma  $\|\cdot\|$  de  $\mathbb{R}^m$  y  $\|A\|_\infty$  denota la norma del supremo de la matriz  $A$ , es decir,

$$\|A\|_\infty = \max_{i=1, \dots, q} \sum_{j=1}^q |a_{ij}|.$$

El resultado también es cierto si se cumple que

$$h_n L \rho(|A|) < 1,$$

donde  $|A|$  denota la matriz cuyos elementos son los valores absolutos de los elementos de  $A$  y  $\rho(|A|)$  es el radio espectral de la matriz  $|A|$ , es decir,

$$\rho(|A|) = \max\{|\lambda| : \lambda \text{ es autovalor de } |A|\}.$$

**Proposición 5.1.** *La función incremento de un método de Runge-Kutta implícito de  $q$  etapas,*

$$\Phi(t, \mathbf{y}, h) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, h),$$

*es lipschitziana respecto de su segundo argumento en el dominio  $I \times \mathbb{R}^m \times [0, h^*]$ , y por tanto el método es estable, si se verifica que*

$$h^* L\rho(|A|) < 1.$$

**Proposición 5.2.** *Si la función  $\mathbf{f} : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  que define el problema diferencial (P) es de clase  $C^p$  entonces las funciones de las etapas,  $\mathbf{k}_i$  o  $\mathbf{g}_i$ , según la formulación utilizada, (vistas como funciones de  $h$ ) son también de clase  $C^p$  en un entorno de  $h = 0$ .*

## 5.2. Orden de los métodos RK implícitos

Dedicamos esta sección a presentar y desarrollar la Teoría de Butcher para el análisis del orden de consistencia en el marco más general de las fórmulas de Runge-Kutta implícitas. La línea argumental no supone un gran cambio con lo expuesto hasta el momento, pues no es más que una mera generalización de lo que se hizo en el Capítulo 3 para el caso explícito. Es por ello que, en ocasiones, para no sobrecargar artificialmente o en exceso esta parte, prescindiremos de algunas demostraciones. No obstante, sí aparecen algunos resultados notablemente diferentes a los que nos hemos encontrado anteriormente y, además, surgirán otras nociones interesantes al tiempo que quedará patente la conexión que existe entre estas formas de resolver ecuaciones diferenciales ordinarias y otras partes del Análisis Numérico, a las que ya hemos venido haciendo referencia, como es el caso de las fórmulas de integración numérica, y que merece la pena, por tanto, abordar más detenidamente.

Supondremos que los métodos de Runge-Kutta que trataremos en esta sección, al igual que hacíamos en el caso explícito, verifican la condición de suma de filas, esto es,

$$c_i = \sum_{j=1}^q a_{ij}, \quad i = 1, 2, \dots, q,$$

pues, en ese caso, sabemos que la resolución de un problema no autónomo es equivalente a la resolución de un problema autónomo (con una dimensión más, pues la variable independiente pasa a considerarse como una nueva incógnita). Por tanto, consideraremos únicamente, en un principio, la aplicación de las fórmulas implícitas a sistemas autónomos de la forma siguiente:

$$(P) \quad \begin{cases} \mathbf{y}' = \mathbf{f}(\mathbf{y}(t)), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \boldsymbol{\eta}, \end{cases}$$

Como la solución exacta del problema  $(P)$  anterior es la misma, independientemente de si el método utilizado en la resolución numérica es explícito o implícito, podemos usar de nuevo el Corolario 3.1 para escribirla en serie de potencias del paso  $h$  en términos de las diferenciales elementales asociadas a los árboles con raíz de la siguiente manera:

$$\mathbf{y}(t_n + h) = \sum_{T \in \mathcal{LT}} \mathbf{F}(T)(\mathbf{y}(t_n)) \frac{h^{\rho(T)}}{\rho(T)!} = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}.$$

Replicando lo hecho en el caso explícito, trataremos de obtener, en lo que sigue, un desarrollo en serie de potencias del paso  $h$  de la solución numérica obtenida al aplicar un método de Runge-Kutta implícito al problema  $(P)$ , por lo que necesitamos adaptar la formulación (5.3) al caso autónomo. Resulta entonces que la solución aproximada viene dada por:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \sum_{i=1}^q b_i \mathbf{g}_i, \quad (5.4)$$

donde

$$\mathbf{g}_i = h_n \mathbf{f} \left( \mathbf{y}_n + \sum_{j=1}^q a_{ij} \mathbf{g}_j \right), \quad i = 1, 2, \dots, q.$$

Nótese que en dicha formulación la única diferencia con respecto al caso explícito (véase la expresión (3.2)) es que en el sumatorio que aparece en la expresión de cada  $\mathbf{g}_i$  intervienen los  $q$  coeficientes  $a_{ij}$  (con  $j = 1, 2, \dots, q$ ), pues, al tratarse de una fórmula implícita, no son necesariamente nulos. En los métodos explícitos es suficiente con considerar los primeros  $i - 1$  coeficientes  $a_{ij}$  (con  $j = 1, 2, \dots, i - 1$ ) pues, como  $a_{ij} = 0$  para  $j = i, i + 1, \dots, q$ , estos no hacen variar el valor de la suma.

Este hecho nos permitirá, además, generalizar los resultados obtenidos en la Sección 3.3 para la solución numérica obtenida con una fórmula explícita al caso implícito, por lo que en algunos casos en vez de reproducir una demostración completamente análoga a la ya realizada anteriormente se optará por hacer referencia únicamente al resultado correspondiente.

**Lema 5.1.** *Cada  $\mathbf{g}_i$  es representable en serie de Butcher en la forma*

$$\mathbf{g}_i = B(\varphi_i, \mathbf{y}_n)(h), \quad i = 1, 2, \dots, q,$$

donde las aplicaciones  $\varphi_i : \mathcal{T} \rightarrow \mathbb{R}$  están definidos por:

$$\begin{cases} \varphi_i(\emptyset) = 0, \\ \varphi_i(\tau_0) = 1, \\ \varphi_i(\tau) = \rho(\tau) \sum_{j_1, \dots, j_M}^q a_{ij_1} \cdots a_{ij_M} \varphi_{j_1}(\tau_1) \cdots \varphi_{j_M}(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\}. \end{cases}$$

*Demostración.* Véase el Lema 3.1 y téngase en cuenta el comentario del párrafo anterior.  $\square$

A partir de este resultado construimos la aplicación

$$\begin{aligned}\varphi: \mathcal{T} &\longrightarrow \mathbb{R}^q \\ \tau &\mapsto \varphi(\tau) = (\varphi_i(\tau))_{i=1}^q\end{aligned}$$

que se obtiene abreviadamente como:

$$\begin{aligned}\varphi(\emptyset) &= (0, 0, \dots, 0)^T = \mathbf{0}, \\ \varphi(\tau_0) &= (1, 1, \dots, 1)^T = \mathbf{e}, \\ \varphi(\tau) &= \rho(\tau) \cdot A\varphi(\tau_1) \cdot \dots \cdot A\varphi(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\}.\end{aligned}$$

Construimos entonces la función de pesos internos:

$$\begin{aligned}\pi: \mathcal{T} &\longrightarrow \mathbb{R}^q \\ \tau &\mapsto \pi(\tau) = \frac{1}{\gamma(\tau)} \varphi(\tau)\end{aligned}$$

que resulta de calcular

$$\begin{aligned}\pi(\emptyset) &= (0, 0, \dots, 0)^T = \mathbf{0}, \\ \pi(\tau_0) &= (1, 1, \dots, 1)^T = \mathbf{e}, \\ \pi(\tau) &= A\pi(\tau_1) \cdot \dots \cdot A\pi(\tau_M), \quad \text{cuando } \tau = \{\tau_1, \dots, \tau_M\}.\end{aligned}$$

Por último obtenemos la función peso elemental:

$$\begin{aligned}\Psi: \mathcal{T} &\longrightarrow \mathbb{R} \\ \tau &\mapsto \Psi(\tau) = \mathbf{b}^T \pi(\tau)\end{aligned}$$

**Corolario 5.1.** *La solución aproximada del problema (P) obtenida numéricamente mediante una fórmula de Runge-Kutta implícita de  $q$  etapas se puede representar en serie de Butcher mediante la aplicación  $\omega: \mathcal{T} \longrightarrow \mathbb{R}$  definida por:*

$$\omega(\tau) = \begin{cases} 1 & \text{si } \tau = \emptyset, \\ \gamma(\tau)\Psi(\tau) & \text{si en otro caso.} \end{cases}$$

*Demostración.* La formulación (5.4) de un método de Runge-Kutta implícito de  $q$  etapas en el caso autónomo y el Lema 5.1 nos permiten expresar la solución numérica del problema (P) en serie de Butcher de la siguiente manera:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \sum_{i=1}^q b_i \mathbf{g}_i = \mathbf{y}_n + \sum_{i=1}^q b_i B(\varphi_i, \mathbf{y}_n)(h) = B(\omega, \mathbf{y}_n)(h). \quad \square$$

Finalmente, sin más que tener en cuenta los conceptos de serie de Butcher (ver Definición 3.13) y de función representable en serie de Butcher (ver Definición 3.14), podemos obtener el desarrollo de la solución numérica como se indica en los siguientes resultados.

**Teorema 5.2.** *La derivada  $k$ -ésima de la solución aproximada del problema (P) obtenida numéricamente mediante una fórmula de Runge-Kutta implícita de  $q$  etapas viene dada por:*

$$\mathbf{y}_{n+1}^{(k)}(t_n) = \sum_{T \in \mathcal{L}\mathcal{T}_k} \omega(T) \mathbf{F}(T)(\mathbf{y}(t_n)) = \sum_{\tau \in \mathcal{T}_k} \alpha(\tau) \omega(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)), \quad k = 0, 1, 2, \dots$$

**Corolario 5.2.** *El desarrollo de Taylor de la solución aproximada del problema (P) obtenida numéricamente mediante una fórmula de Runge-Kutta implícita de  $q$  etapas puede escribirse en términos de las diferenciales elementales asociadas a árboles con raíz como sigue:*

$$\mathbf{y}_{n+1}(t_n) = \sum_{T \in \mathcal{L}\mathcal{T}} \omega(T) \mathbf{F}(T)(\mathbf{y}(t_n)) \frac{h^{\rho(T)}}{\rho(T)!} = \sum_{\tau \in \mathcal{T}} \alpha(\tau) \omega(\tau) \mathbf{F}(\tau)(\mathbf{y}(t_n)) \frac{h^{\rho(\tau)}}{\rho(\tau)!}.$$

Por último, comparando los desarrollos obtenidos para las soluciones exacta y aproximada podemos deducir una condición necesaria y suficiente para que un método de Runge-Kutta de tipo implícito alcance un determinado orden de consistencia.

**Teorema 5.3 (Condiciones de orden).**

*Un método de Runge-Kutta implícito con  $q$  etapas alcanza orden  $p$  si y solo si se verifica que:*

$$\Psi(\tau) = \frac{1}{\gamma(\tau)}, \quad \forall \tau \in \mathcal{T}_k \text{ con } k \leq p$$

*Demostración.* Véase la prueba del Teorema 3.4. □

Como se señaló en el Capítulo 1, un problema de valor inicial (P) puede formalizarse de manera equivalente como una ecuación integral (véase la ecuación (1.1)) y, en consecuencia, las fórmulas de Runge-Kutta y de integración numérica están estrechamente relacionadas. Dicha conexión va a ser utilizada para obtener métodos implícitos a partir de ciertas familias conocidas de fórmulas de cuadratura y, además, el grado de precisión de dichas fórmulas determina el orden de consistencia que alcanzan tales métodos.

Recordemos que una *fórmula de cuadratura* o *fórmula de integración numérica* es una regla que aproxima el valor de una integral definida evaluando el integrando en una serie de puntos determinados, como se indica en la siguiente expresión:

$$\int_{\alpha}^{\beta} f(s) ds \simeq \sum_{i=1}^n \lambda_i f(\mu_i),$$

donde  $\{\lambda_i\}_{i=1}^n$  y  $\{\mu_i\}_{i=1}^n$  reciben el nombre de *pesos* y *nodos*, respectivamente.

Por otro lado, se dice que una fórmula de cuadratura tiene *grado de precisión al menos  $N$*  si es exacta (es decir, el valor de la solución exacta coincide con el de la aproximación numérica) para polinomios de grado menor o igual que  $N$ . Del mismo modo, se dice que tiene *grado de precisión justamente  $N$*  si la fórmula tiene grado de precisión al menos  $N$  pero no es exacta para algún polinomio de grado de  $N + 1$ .

Además, respecto al grado de precisión máximo que puede alcanzar una determinada fórmula de integración numérica podemos citar el siguiente resultado:

**Lema 5.2.** *Una fórmula de cuadratura con  $q$  nodos tiene a lo sumo grado de precisión  $2q - 1$ .*

*Demostración.* Véase, por ejemplo, [10]. □

Veamos ahora cómo definir un método de Runge-Kutta a partir de una fórmula de cuadratura.

En primer lugar, si  $\mathbf{y}(\cdot)$  es la solución del problema de valor inicial dado por:

$$(P) \quad \begin{cases} \mathbf{y}' &= \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) &= \boldsymbol{\eta}, \end{cases}$$

podemos escribir la solución en  $t_n + h_n$  como sigue:

$$\mathbf{y}(t_n + h_n) = \mathbf{y}(t_n) + \int_{t_n}^{t_n + h_n} \mathbf{f}(t, \mathbf{y}(t)) dt. \quad (5.5)$$

Consideramos entonces una fórmula de integración numérica definida en el intervalo  $[0, 1]$  que utiliza  $q$  pesos  $\{b_1, b_2, \dots, b_q\}$  y  $q$  nodos  $\{c_1, c_2, \dots, c_q\}$  y cuya expresión es, por tanto,

$$\int_0^1 \psi(s) ds \simeq \sum_{i=1}^q b_i \psi(c_i). \quad (5.6)$$

Si modificamos el intervalo de integración mediante el cambio de variable  $t = t_n + h_n s$ , resulta:

$$\int_{t_n}^{t_n + h_n} \xi(t) dt \simeq h_n \sum_{i=1}^q b_i \xi(t_n + c_i h_n), \quad \text{con } \xi(t) = \psi\left(\frac{t - t_n}{h_n}\right),$$

y teniendo en cuenta esta expresión y la ecuación (5.5) aproximamos la solución en  $t_n + h_n$ :

$$\mathbf{y}(t_n + h_n) \simeq \mathbf{y}(t_n) + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + c_i h_n, \mathbf{y}(t_n + c_i h_n)). \quad (5.7)$$

Con el mismo razonamiento pero utilizando las fórmulas de cuadratura siguientes,

$$\int_0^{c_i} \psi(s) ds \simeq \sum_{j=1}^q a_{ij} \psi(c_j), \quad i = 1, 2, \dots, q, \quad (5.8)$$

podemos aproximar el valor de la solución en  $t_n + c_i h_n$ . Nos quedaría entonces,

$$\begin{aligned} \mathbf{y}(t_n + c_i h_n) &= \mathbf{y}(t_n) + \int_{t_n}^{t_n + c_i h_n} \mathbf{f}(t, \mathbf{y}(t)) dt \\ &\simeq \mathbf{y}(t_n) + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}(t_n + c_j h_n)). \end{aligned} \quad (5.9)$$

Así, a partir de las fórmulas de cuadratura que acabamos de mencionar y considerando las expresiones (5.7) y (5.9) podemos obtener el valor de la aproximación de la solución exacta en un determinado instante de la siguiente manera:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \sum_{i=1}^q b_i \mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}),$$

donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q,$$

que coincide con la formulación (5.2) de un método de Runge-Kutta implícito de  $q$  etapas.

Para completar este apartado recordamos también un resultado que relaciona el grado de precisión de las fórmulas de cuadratura y el orden de los métodos de Runge-Kutta asociados.

**Lema 5.3.** *Si la fórmula de cuadratura en el intervalo  $[0, 1]$  definida en (5.6) tiene grado de precisión al menos  $p - 1$  y para cada  $i = 1, 2, \dots, q$  la fórmula de cuadratura en el intervalo  $[0, c_i]$  definida en (5.8) tiene grado de precisión al menos  $p - 2$ , el método de Runge-Kutta asociado tiene orden al menos  $p$ .*

*Demostración.* Véase, por ejemplo, [3]. □

Al disponer de más parámetros libres que en el caso explícito, cabe esperar entonces que las fórmulas implícitas alcancen un orden más elevado. Veamos que efectivamente esto es así.

**Teorema 5.4.** *Un método Runge-Kutta implícito de  $q$  etapas puede alcanzar hasta orden  $2q$ .*

*Demostración.* Supongamos que el método dado alcanza orden  $p$ . Entonces, se cumplen las condiciones de orden asociadas a los árboles  $\tau_i$  con  $i$  nodos y cuyas formas son:

$$\tau_i = \begin{array}{c} \bullet \quad \bullet \quad \dots \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \end{array} \quad i = 1, 2, \dots, p.$$

Asumiendo que  $\mathbf{c}^k = (c_1^k, c_2^k, \dots, c_q^k)^T$ , según el Teorema 5.3, tales condiciones se traducen en:

$$\Psi(\tau_i) = \frac{1}{\gamma(\tau_i)}, \text{ o equivalentemente, } \mathbf{b}^T \mathbf{c}^{i-1} = \frac{1}{i}, \text{ para } i = 1, 2, \dots, p,$$

Además, en ese caso la fórmula de cuadratura (5.6) con vector de pesos  $\mathbf{b}$  y de nodos  $\mathbf{c}$ :

$$\int_0^1 \psi(s) ds \simeq \sum_{i=1}^q b_i \psi(c_i),$$

tiene grado de precisión al menos  $p - 1$  y como por Lema 5.2 sabemos que el grado de precisión de la fórmula de integración numérica anterior es a lo sumo  $2q - 1$ , entonces:

$$p - 1 \leq 2q - 1 \Leftrightarrow p \leq 2q. \quad \square$$

De la demostración del Teorema anterior se deduce que en el caso de que finalmente existan métodos de Runge-Kutta implícitos de  $q$  etapas y orden  $2q$  entonces sus coeficientes  $\{b_i\}_{i=1}^q$  y  $\{c_i\}_{i=1}^q$  son, respectivamente, los pesos y los nodos de la única fórmula de cuadratura de Gauss con  $q$  nodos en el intervalo  $[0, 1]$ , esto es,

$$\int_0^1 \psi(x) dx \simeq \sum_{i=1}^q b_i \psi(c_i),$$

que alcanza grado de precisión  $2q - 1$ .

Para demostrar que efectivamente existen métodos con esas características es conveniente manejar las llamadas *hipótesis simplificadoras*. Así, dada una fórmula de Runge-Kutta implícita de  $q$  etapas y asumiendo, de nuevo, la notación  $\mathbf{c}^k = (c_1^k, c_2^k, \dots, c_q^k)^T$ :

- Se dice que dicha fórmula satisface la propiedad  $B(n)$  si sus coeficientes cumplen que

$$\mathbf{b}^T \mathbf{c}^{k-1} = \frac{1}{k}, \quad k = 1, 2, \dots, n. \quad (5.10)$$

- Del mismo modo, diremos que verifica la propiedad  $C(n)$  si se tiene que

$$A \mathbf{c}^{k-1} = \frac{\mathbf{c}^k}{k}, \quad k = 1, 2, \dots, n. \quad (5.11)$$

- Finalmente, se dice que satisface la propiedad  $D(n)$  si sus coeficientes verifican que

$$\left( \mathbf{b} \cdot \mathbf{c}^{k-1} \right)^T A = \frac{1}{k} \left( \mathbf{b} - \left( \mathbf{b} \cdot \mathbf{c}^k \right) \right)^T, \quad k = 1, 2, \dots, n. \quad (5.12)$$

Estas propiedades permiten “simplificar”, las condiciones que, según el Teorema 5.3, deben verificarse para alcanzar un determinado orden de la siguiente manera:

- La propiedad  $B(n)$  se traduce en que los árboles de la forma  $\tau = \{\tau_1, \tau_2, \dots, \tau_{k-1}\}$  con  $\tau_i = \tau_0$  para  $i = 1, 2, \dots, k-1$  con  $k \leq n$ , esto es, los árboles:



verifican la condición de orden que les corresponde. En efecto, a un árbol  $\tau$  de esa forma le corresponde la condición de orden

$$\Psi(\tau) = \frac{1}{\gamma(\tau)}, \quad \text{o equivalentemente, } \mathbf{b}^T \boldsymbol{\pi}(\tau) = \frac{1}{\gamma(\tau)}. \quad (5.13)$$

Además, se cumple que:

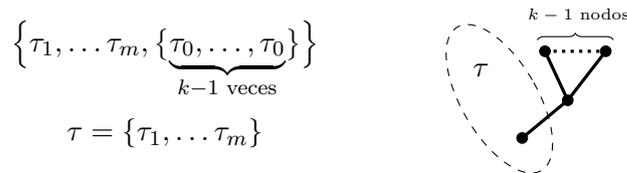
$$\begin{aligned} \triangleright \quad \gamma(\tau) &= \rho(\tau) = k. \\ \triangleright \quad \boldsymbol{\pi}(\tau) &= A \boldsymbol{\pi}(\tau_1) \cdot \dots \cdot A \boldsymbol{\pi}(\tau_{k-1}) = \underbrace{A \mathbf{e} \cdot \dots \cdot A \mathbf{e}}_{k-1 \text{ veces}} = \underbrace{\mathbf{c} \cdot \dots \cdot \mathbf{c}}_{k-1 \text{ veces}} = \mathbf{c}^{k-1}. \end{aligned}$$

Entonces, sustituyendo estos valores en la expresión (5.13) obtenemos,

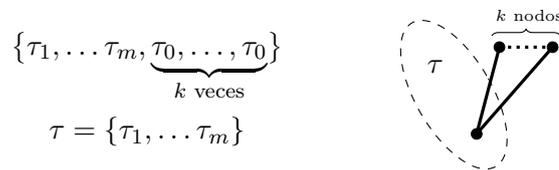
$$\mathbf{b}^T \mathbf{c}^{k-1} = \frac{1}{k},$$

que es justamente la condición (5.10) que garantiza la propiedad  $B(n)$

- La propiedad  $C(n)$  significa que para todo  $k \leq n$ , un árbol de la forma



da lugar a la misma condición que otro árbol de su mismo orden que tiene la forma:

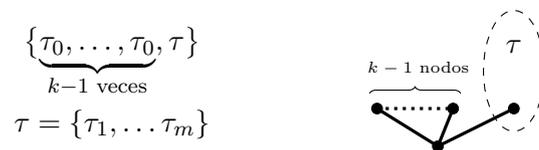


En efecto, razonando como antes, es fácil comprobar que las condiciones de orden asociadas a estos dos tipos de árboles resultan ser,

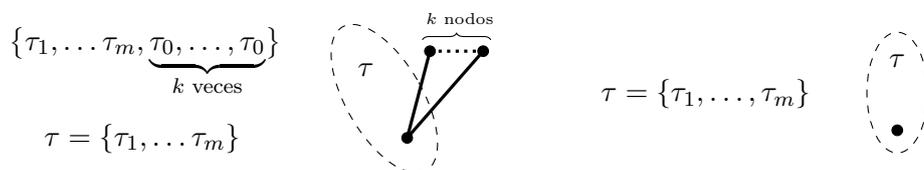
$$\begin{aligned} \triangleright \mathbf{b}^T \boldsymbol{\pi}(\tau) \cdot A \mathbf{c}^{k-1} &= \frac{\rho(\tau)}{k\gamma(\tau)(\rho(\tau) + k)}, \\ \triangleright \mathbf{b}^T \boldsymbol{\pi}(\tau) \cdot \mathbf{c}^k &= \frac{\rho(\tau)}{\gamma(\tau)(\rho(\tau) + k)}. \end{aligned}$$

Entonces, es evidente que, cumpliéndose la condición (5.11) que garantiza la propiedad  $C(n)$ , las dos expresiones anteriores son son claramente equivalentes, por lo que, a la hora de estudiar el orden de un método, será suficiente con comprobar únicamente las condiciones asociadas a una de las dos familias de árboles señaladas.

- La propiedad  $D(n)$  significa que la condición de orden de un árbol de la forma



se puede obtener a partir de las condiciones de orden asociadas a los árboles:



En efecto, las condiciones de orden asociadas a estos tres tipos de árboles son:

$$\begin{aligned} \triangleright \mathbf{b}^T \mathbf{c}^{k-1} \cdot A\boldsymbol{\pi}(\tau) &= \frac{1}{\gamma(\tau)(\rho(\tau) + k)}, \\ \triangleright \mathbf{b}^T \boldsymbol{\pi}(\tau) \cdot \mathbf{c}^k &= \frac{\rho(\tau)}{\gamma(\tau)(\rho(\tau) + k)}, \\ \triangleright \mathbf{b}^T \boldsymbol{\pi}(\tau) &= \frac{1}{\gamma(\tau)}, \end{aligned}$$

y cumpliéndose la condición (5.12) que garantiza la propiedad  $D(n)$ , es fácil comprobar que si se satisfacen las dos últimas, se verifica entonces la primera.

Además, cada una de las condiciones  $B(n)$ ,  $C(n)$  y  $D(n)$  anteriores se traduce, respectivamente, en que las siguientes fórmulas de cuadratura tienen grado de precisión  $n - 1$ :

$$\begin{aligned} \int_0^1 \psi(x) dx &\simeq \sum_{i=1}^q b_i \psi(c_i), \\ \int_0^{c_i} \psi(x) dx &\simeq \sum_{j=1}^q a_{ij} \psi(c_j), \quad i = 1, 2, \dots, q, \\ \int_{c_j}^1 b_j \psi(x) dx &\simeq \sum_{i=1}^q b_i a_{ij} \psi(c_i), \quad j = 1, 2, \dots, q. \end{aligned}$$

**Proposición 5.3.** *Si un método de Runge-Kutta implícito satisface las propiedades  $B(p)$ ,  $C(k)$  y  $D(r)$  con  $p \leq k + r + 1$  y  $p \leq 2(k + 1)$  entonces tiene orden al menos  $p$ .*

*Demostración.* Como se verifica la condición  $C(k)$  entonces basta comprobar que se cumplen las condiciones de orden asociadas a los árboles de la forma  $\tau = \{\tau_1, \dots, \tau_n\}$  con  $\rho(\tau) \leq p$  para los cuales  $\tau_i$  o bien es  $\tau_0$  o bien tiene orden al menos  $k + 1$ . Como, además, se tiene que  $p \leq 2(k + 1)$  entonces o bien todos los  $\tau_i$  son  $\tau_0$  o bien solo hay uno de orden al menos  $k + 1$ . En el primer supuesto, las condiciones de orden asociadas a ese tipo de árboles son ciertas puesto que se cumple la condición  $B(p)$ ; mientras que en el segundo caso, como  $p \leq k + r + 1$  se tiene que el número de árboles de la forma  $\tau_0$  es a lo sumo  $r - 1$  y al satisfacerse la propiedad  $D(r)$ , se verifican las condiciones de orden asociadas a estos árboles.  $\square$

**Teorema 5.5.** *Un método de Runge-Kutta implícito de  $q$  etapas tiene orden  $2q$  si y solamente si se satisfacen las propiedades  $B(2q)$  y  $C(q)$ . Además, en ese caso, los coeficientes quedan unívocamente determinados de la siguiente manera:*

- Al verificarse la propiedad  $B(2q)$ ,  $\{b_i\}_{i=1}^q$  y  $\{c_i\}_{i=1}^q$  serían, respectivamente, los pesos y los nodos de la única fórmula de cuadratura de Gauss con  $q$  nodos en el intervalo  $[0, 1]$ , esto es,

$$\int_0^1 \psi(x) dx \simeq \sum_{i=1}^q b_i \psi(c_i)$$

que alcanza grado de precisión  $2q - 1$ .

- Como se satisface la propiedad  $C(q)$  entonces para cada  $i = 1, 2, \dots, q$  se tiene que  $\{a_{ij}\}_{j=1}^q$  son los pesos de la fórmula de cuadratura de tipo interpolatorio en el intervalo  $[0, c_i]$  respecto a los nodos  $\{c_i\}_{i=1}^q$ , esto es,

$$\int_0^{c_i} \psi(x) dx \simeq \sum_{j=1}^q a_{ij} \psi(c_j), \quad i = 1, 2, \dots, q.$$

*Demostración.* Veamos que una fórmula de Runge-Kutta implícita de  $q$  etapas que verifica las propiedades  $B(2q)$  y  $C(q)$  satisface también la propiedad  $D(q)$ , pues, en ese caso, por la Proposición 5.3 sabríamos que el orden es al menos  $2q$  y, por el Teorema 5.4, que es a lo sumo  $2q$ , de manera que, en definitiva, el orden de dicha fórmula sería exactamente  $2q$ .

El hecho de que la propiedad  $D(q)$  se verifique cuando se cumple que:

$$\left(\mathbf{b} \cdot \mathbf{c}^{k-1}\right)^T A = \frac{1}{k} \left(\mathbf{b} - \left(\mathbf{b} \cdot \mathbf{c}^k\right)\right)^T, \quad k = 1, 2, \dots, q,$$

nos induce a considerar los vectores de  $\mathbb{R}^q$  dados por:

$$\mathbf{d}_k^T = \left(\mathbf{b} \cdot \mathbf{c}^{k-1}\right)^T A - \frac{1}{k} \left(\mathbf{b} - \left(\mathbf{b} \cdot \mathbf{c}^k\right)\right)^T, \quad k = 1, 2, \dots, q.$$

Entonces, para  $m = 1, 2, \dots, q$  se tiene que:

$$\begin{aligned} \mathbf{d}_k^T \mathbf{c}^{m-1} &= \left(\mathbf{b} \cdot \mathbf{c}^{k-1}\right)^T A \mathbf{c}^{m-1} - \frac{1}{k} \mathbf{b}^T \mathbf{c}^{m-1} + \frac{1}{k} \left(\mathbf{b} \cdot \mathbf{c}^k\right)^T \mathbf{c}^{m-1} \\ &= \left(\mathbf{b} \cdot \mathbf{c}^{k-1}\right)^T \frac{\mathbf{c}^m}{m} - \frac{1}{k} \frac{1}{m} + \frac{1}{k} \mathbf{b}^T \mathbf{c}^{k+m-1} \\ &= \frac{1}{m} \mathbf{b}^T \mathbf{c}^{k+m-1} - \frac{1}{k} \frac{1}{m} + \frac{1}{k} \frac{1}{k+m} \\ &= \frac{1}{m} \frac{1}{k+m} - \frac{1}{k} \frac{1}{m} + \frac{1}{k} \frac{1}{k+m} = 0. \end{aligned}$$

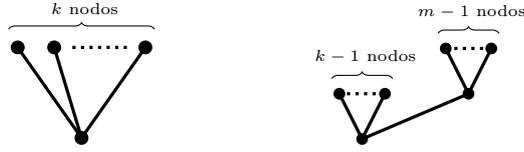
Deducimos pues que  $\mathbf{d}_k^T$  es ortogonal a  $\mathbf{c}^{m-1}$  para  $m = 1, 2, \dots, q$ .

Además, como se verifica la propiedad  $B(2q)$  sabemos que  $\{c_i\}_{i=1}^q$  son los nodos de la única fórmula de cuadratura de Gauss con  $q$  nodos en el intervalo  $[0, 1]$  que alcanza grado de precisión  $2q - 1$ , por lo que los coeficientes  $c_i$  son todos reales y distintos. De esta manera, concluimos que  $\{\mathbf{c}^{m-1}\}_{m=1}^q$  es una base de  $\mathbb{R}^q$  y, en consecuencia, tenemos que  $\mathbf{d}_k^T = \mathbf{0}$ , lo que se traduce en el cumplimiento de la propiedad  $D(q)$ .

Recíprocamente, veamos que si un método de Runge-Kutta implícito de  $q$  etapas tiene orden  $2q$  entonces se verifican las hipótesis simplificadoras  $B(2q)$  y  $C(q)$ .

Como la fórmula en cuestión alcanza orden  $2q$ , según el Teorema 5.3, se satisfacen todas las condiciones de orden asociadas a todos los árboles de orden  $p \leq 2q$ .

En particular, se cumplen aquellas que están asociadas a los árboles de las formas:



- En el primer caso,  $\tau = \{\tau_1, \tau_2, \dots, \tau_k\}$  con  $\tau_i = \tau_0$  para  $i = 1, 2, \dots, k$  con  $k \leq 2q - 1$ . Esto significa que se cumple la propiedad  $B(2q)$ .
- Por su parte, en el segundo caso,  $\tau = \{\tau_1, \dots, \tau_{k-1}, \{\tilde{\tau}_1, \dots, \tilde{\tau}_{m-1}\}\}$  con  $\tau_i = \tilde{\tau}_j = \tau_0$  para  $i = 1, 2, \dots, k - 1$  y para  $j = 1, 2, \dots, m - 1$  con  $k + m \leq 2q$ . La condición de orden asociada a los árboles de esta familia se traduce en:

$$\mathbf{b}^T (\mathbf{c}^{k-1} \cdot A \mathbf{c}^{m-1}) = \frac{1}{m(k+m)}, \quad (5.14)$$

puesto que, realizando los correspondientes cálculos, se tiene que:

$$\begin{aligned} \triangleright \gamma(\tau) &= m(k+m). \\ \triangleright \pi(\tau) &= \underbrace{A \varphi(\tau_0) \cdots A \varphi(\tau_0)}_{k-1 \text{ veces}} \cdot A \left( \underbrace{A \varphi(\tau_0) \cdots A \varphi(\tau_0)}_{m-1 \text{ veces}} \right) \\ &= \underbrace{\mathbf{c} \cdots \mathbf{c}}_{k-1 \text{ veces}} \cdot A \left( \underbrace{\mathbf{c} \cdots \mathbf{c}}_{m-1 \text{ veces}} \right) = \mathbf{c}^{k-1} \cdot A \mathbf{c}^{m-1}. \end{aligned}$$

Al cumplirse la propiedad  $B(2q)$  tenemos que

$$\mathbf{b}^T \mathbf{c}^{k+m-1} = \frac{1}{k+m},$$

y entonces, la condición de orden (5.14) nos quedaría de la siguiente manera:

$$\mathbf{b}^T (\mathbf{c}^{k-1} \cdot A \mathbf{c}^{m-1}) = \frac{1}{m} \mathbf{b}^T \mathbf{c}^{k+m-1} \Leftrightarrow (\mathbf{b} \cdot \mathbf{c}^{k-1})^T \left( A \mathbf{c}^{m-1} - \frac{1}{m} \mathbf{c}^m \right) = 0.$$

Finalmente, por ser  $\{b_i\}_{i=1}^q$  y  $\{c_i\}_{i=1}^q$  los pesos y los nodos de una fórmula de cuadratura de Gauss tenemos que  $\{\mathbf{b} \cdot \mathbf{c}^{k-1}\}_{k=1}^q$  es una base de  $\mathbb{R}^q$ , de donde deducimos que:

$$A \mathbf{c}^{m-1} - \frac{1}{m} \mathbf{c}^m = \mathbf{0}, \quad m = 1, 2, \dots, q.$$

En consecuencia, se satisface la propiedad  $C(q)$ . □

**Ejemplo 5.3.** Método de Gauss de una etapa y orden dos.

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

**Ejemplo 5.4.** Método de Gauss de dos etapas y orden cuatro.

$$\begin{array}{c|cc} \frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

**Ejemplo 5.5.** Método de Gauss de tres etapas y orden seis.

$$\begin{array}{c|ccc} \frac{5-\sqrt{15}}{10} & \frac{5}{36} & \frac{10-3\sqrt{15}}{45} & \frac{25-6\sqrt{15}}{180} \\ \frac{1}{2} & \frac{10+3\sqrt{15}}{72} & \frac{2}{9} & \frac{10-3\sqrt{15}}{72} \\ \frac{5+\sqrt{15}}{10} & \frac{25+6\sqrt{15}}{180} & \frac{10+3\sqrt{15}}{45} & \frac{5}{36} \\ \hline & \frac{5}{18} & \frac{4}{9} & \frac{5}{18} \end{array}$$

### 5.3. Otras familias de métodos RK implícitos

Hasta el momento hemos presentado únicamente métodos de Runge-Kutta implícitos con  $q$  etapas cuyo orden es el máximo posible, esto es,  $2q$ ; y hemos visto también como este tipo de métodos está estrechamente vinculado con las fórmulas de integración numérica de Gauss. Precisamente, esa relación existente entre los métodos de Runge-Kutta implícitos y las fórmulas de cuadratura nos permite construir sistemáticamente nuevas familias de métodos.

Tabla 5.1: Métodos Runge-Kutta implícitos a partir de fórmulas de cuadratura

Método	Orden	Hipótesis simplificadoras	$c_i$	$a_{ij}$
Gauss	$2q$	$B(2q)$ , $C(q)$ y $D(q)$	---	---
Radau I	$2q - 1$	$B(2q - 1)$ , $C(q)$ y $D(q - 1)$	$c_1 = 0$	$a_{1j} = 0$
Radau IA	$2q - 1$	$B(2q - 1)$ , $C(q - 1)$ y $D(q)$	$c_1 = 0$	---
Radau II	$2q - 1$	$B(2q - 1)$ , $C(q - 1)$ y $D(q)$	$c_q = 1$	---
Radau IIA	$2q - 1$	$B(2q - 1)$ , $C(q)$ y $D(q - 1)$	$c_q = 1$	$a_{qj} = b_j$
Lobatto IIIA	$2q - 2$	$B(2q - 2)$ , $C(q)$ y $D(q - 2)$	$c_1 = 0$ $c_q = 1$	$a_{1j} = 0$ $a_{qj} = b_j$
Lobatto IIIB	$2q - 2$	$B(2q - 2)$ , $C(q - 2)$ y $D(q)$	$c_1 = 0$ $c_q = 1$	$a_{iq} = 0$
Lobatto IIIC	$2q - 2$	$B(2q - 2)$ , $C(q - 1)$ y $D(q - 1)$	$c_1 = 0$ $c_q = 1$	$a_{qj} = b_j$

Especialmente destacable es el caso de las fórmulas de Radau y de Lobatto, que dan lugar a métodos de Runge-Kutta implícitos con  $q$  etapas y con órdenes  $2q - 1$  y  $2q - 2$ , respectivamente. Esta reducción del orden puede ser entendida como consecuencia de que en las fórmulas de Radau y Lobatto se impone que el primer y/o el último nodo coincidan con los extremos del intervalo de integración, mientras que en las fórmulas de Gauss los nodos se eligen con total libertad con el único propósito de alcanzar el máximo grado de precisión posible.

Podemos señalar también la familia de los métodos de colocación, que consiste en encontrar un polinomio de grado  $q$  cuya derivada coincida en  $q$  puntos con la dirección marcada por el campo vectorial que determina la ecuación diferencial. De los métodos recogidos en la tabla anterior solo las fórmulas de Gauss, Radau I, Radau IIA y Lobatto IIIA son este tipo. Salta a la vista que todos ellos verifican la propiedad  $C(q)$  mientras que esta no se mantiene para los que por el contrario no son métodos de colocación. De hecho, puede probarse que tal propiedad es condición necesaria y suficiente para que un método de Runge-Kutta implícito con todos los nodos  $c_i$  distintos y con orden al menos  $q$  sea un método de colocación. No obstante, la demostración de este resultado se vuelve demasiado técnica y se aleja en exceso de los objetivos fundamentales de la memoria, por lo que se ha preferido no incluirla. Para un análisis más exhaustivo de esta familia de métodos véase, por ejemplo, [4].

**Ejemplo 5.6.** Métodos de Radau (I, IA, II y IIA) de dos etapas y orden tres.

0	0	0	0	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{5}{12}$	$-\frac{1}{12}$
$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{5}{12}$	1	1	0	1	$\frac{3}{4}$	$\frac{1}{4}$
	$\frac{1}{4}$	$\frac{3}{4}$		$\frac{1}{4}$	$\frac{3}{4}$		$\frac{3}{4}$	$\frac{1}{4}$		$\frac{3}{4}$	$\frac{1}{4}$

**Ejemplo 5.7.** Métodos de Radau (I, IA, II y IIA) de tres etapas y orden cinco.

0	0	0	0	0	$\frac{1}{9}$	$\frac{-1-\sqrt{6}}{18}$	$\frac{-1+\sqrt{6}}{18}$
$\frac{6-\sqrt{6}}{10}$	$\frac{9+\sqrt{6}}{75}$	$\frac{24+\sqrt{6}}{120}$	$\frac{168-73\sqrt{6}}{600}$	$\frac{6-\sqrt{6}}{10}$	$\frac{1}{9}$	$\frac{88+7\sqrt{6}}{360}$	$\frac{88-43\sqrt{6}}{360}$
$\frac{6+\sqrt{6}}{10}$	$\frac{9-\sqrt{6}}{75}$	$\frac{168+73\sqrt{6}}{600}$	$\frac{24-\sqrt{6}}{120}$	$\frac{6+\sqrt{6}}{10}$	$\frac{88+43\sqrt{6}}{360}$	$\frac{88-7\sqrt{6}}{360}$	$\frac{88-7\sqrt{6}}{360}$
	$\frac{1}{9}$	$\frac{16+\sqrt{6}}{36}$	$\frac{16-\sqrt{6}}{36}$		$\frac{1}{9}$	$\frac{16+\sqrt{6}}{36}$	$\frac{16-\sqrt{6}}{36}$
$\frac{4-\sqrt{6}}{10}$	$\frac{24-\sqrt{6}}{120}$	$\frac{24-11\sqrt{6}}{120}$	0	$\frac{4-\sqrt{6}}{10}$	$\frac{88-7\sqrt{6}}{360}$	$\frac{296-169\sqrt{6}}{1800}$	$\frac{-2+3\sqrt{6}}{225}$
$\frac{4+\sqrt{6}}{10}$	$\frac{24+11\sqrt{6}}{120}$	$\frac{24+\sqrt{6}}{120}$	0	$\frac{4+\sqrt{6}}{10}$	$\frac{296+169\sqrt{6}}{1800}$	$\frac{88+7\sqrt{6}}{360}$	$\frac{-2-3\sqrt{6}}{225}$
1	$\frac{6-\sqrt{6}}{12}$	$\frac{6+\sqrt{6}}{12}$	0	1	$\frac{16-\sqrt{6}}{36}$	$\frac{16+\sqrt{6}}{36}$	$\frac{1}{9}$
	$\frac{16-\sqrt{6}}{36}$	$\frac{16+\sqrt{6}}{36}$	$\frac{1}{9}$		$\frac{16-\sqrt{6}}{36}$	$\frac{16+\sqrt{6}}{36}$	$\frac{1}{9}$

**Ejemplo 5.8.** Métodos de Lobatto (IIIA, IIIB, IIIC) de dos etapas y orden dos.

0	0	0	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$-\frac{1}{2}$
1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	0	1	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$		$\frac{1}{2}$	$\frac{1}{2}$		$\frac{1}{2}$	$\frac{1}{2}$

**Ejemplo 5.9.** Métodos de Lobatto (IIIA, IIIB, IIIC) de tres etapas y orden cuatro.

$$\begin{array}{c|ccc}
 0 & 0 & 0 & 0 \\
 \frac{1}{2} & \frac{5}{24} & \frac{1}{3} & -\frac{1}{24} \\
 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\
 \hline
 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6}
 \end{array}
 \quad
 \begin{array}{c|ccc}
 0 & \frac{1}{6} & -\frac{1}{6} & 0 \\
 \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\
 1 & \frac{1}{6} & \frac{5}{6} & 0 \\
 \hline
 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6}
 \end{array}
 \quad
 \begin{array}{c|ccc}
 0 & \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\
 \frac{1}{2} & \frac{1}{6} & \frac{5}{12} & -\frac{1}{12} \\
 1 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\
 \hline
 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6}
 \end{array}$$

## 5.4. Métodos DIRK, SDIRK y SIRK

Para completar este capítulo hacemos referencia ahora a los métodos semi-implícitos (también llamados semi-implícitos, diagonalmente implícitos o más brevemente, por sus siglas en inglés, DIRK), a los métodos simplemente diagonalmente implícitos (también conocidos por sus siglas en inglés, SDIRK) y a los métodos simplemente implícitos (SIRK).

En cuanto a las fórmulas DIRK y SDIRK, simplemente señalaremos que fueron impulsadas por Norsett (1974), Crouziex (1976) y Alexander (1977) e incluiremos algunos ejemplos.

**Ejemplo 5.10.** Métodos de dos etapas y orden tres (máximo en general).

$$\begin{array}{c|cc}
 \frac{3\alpha-1}{6\alpha} & \frac{3\alpha-1}{6\alpha} & 0 \\
 \frac{1+\alpha}{2} & \alpha & \frac{1-\alpha}{2} \\
 \hline
 & \frac{3\alpha^2}{3\alpha^2+1} & \frac{1}{3\alpha^2+1}
 \end{array}
 \quad
 \begin{array}{c|cc}
 \frac{3\pm\sqrt{3}}{6} & \frac{3\pm\sqrt{3}}{6\alpha} & 0 \\
 \frac{3\mp\sqrt{3}}{6} & \mp\frac{\sqrt{3}}{3} & \frac{3\pm\sqrt{3}}{6} \\
 \hline
 & \frac{1}{2} & \frac{1}{2}
 \end{array}$$

Por su parte los métodos SIRK, desarrollados principalmente por Burrage y Butcher, son aquellos que verifican las propiedades  $B(q)$  y  $C(q-1)$ , siendo  $q$  el número de etapas del método en cuestión y cuyos coeficientes cumplen, además, las dos condiciones siguientes:

- $c_i \neq c_j$  para todo  $i \neq j$  en  $\{1, 2, \dots, q\}$ .
- La matriz  $A$  posee un único valor propio.

**Ejemplo 5.11.** Familia paramétrica de métodos SIRK de dos etapas y orden dos (máximo en general salvo cuando  $\alpha = \frac{3\pm\sqrt{3}}{6}$  que alcanza orden tres).

$$\begin{array}{c|cc}
 (2-\sqrt{2})\alpha & \frac{4-\sqrt{2}}{4}\alpha & \frac{4-3\sqrt{2}}{4}\alpha \\
 (2+\sqrt{2})\alpha & \frac{4+3\sqrt{2}}{4}\alpha & \frac{4+\sqrt{2}}{4}\alpha \\
 \hline
 & \frac{4(1+\sqrt{2})\alpha-\sqrt{2}}{8\alpha} & \frac{4(1-\sqrt{2})\alpha+\sqrt{2}}{8\alpha}
 \end{array}$$

## Apéndice A

# Algunas propiedades de los métodos de Runge-Kutta implícitos

En el Capítulo 5 enunciamos los resultados que garantizan la buena definición, la estabilidad y la regularidad de las fórmulas de Runge-Kutta implícitas. Sin embargo, se omiten sus demostraciones por ser demasiado técnicas y por averse en exceso del objetivo fundamental del trabajo, que es presentar la Teoría de Butcher para el análisis del orden de esta familia de métodos. Para explicar adecuadamente tales pruebas es necesario recordar diversas nociones del Análisis Matemático y del Análisis Numérico Matricial, y es por ello que se ha optado por incluirlas a modo de apéndice.

### A.1. Buena definición

Comenzamos demostrando la buena definición de los métodos de Runge-Kutta implícitos, pues al no poder resolver sucesivamente las ecuaciones de las etapas no está garantizada la existencia y unicidad de solución del sistema algebraico no lineal (5.1). Es necesario recordar en este punto algunas nociones fundamentales sobre espacios normados y algunos resultados básicos en ese campo, como el Teorema del punto fijo de Banach.

**Definición A.1.** Sean  $(E, \|\cdot\|)$  un espacio normado y  $f : E \rightarrow E$  una función:

- Decimos que  $f$  es *contractiva* si existe  $0 < \lambda < 1$  de manera que se verifica que:

$$\|f(x) - f(y)\| \leq \lambda \|x - y\|, \quad \forall x, y \in E.$$

- Llamamos *punto fijo* de  $f$  a un  $x_0 \in E$  tal que  $f(x_0) = x_0$ .

**Teorema A.1 (del punto fijo de Banach).**

*Sea  $E$  un espacio de Banach y sea  $f : E \rightarrow E$  una función contractiva. Entonces  $f$  tiene un único punto fijo en  $E$ .*

**Teorema A.2.** *El sistema de ecuaciones (5.1) que definen las etapas de una fórmula de Runge-Kutta implícita tiene solución única y en consecuencia el método está bien definido si*

$$h_n L \|A\|_\infty < 1,$$

donde  $L$  es una constante de Lipschitz de la función  $\mathbf{f}$  respecto a cierta norma  $\|\cdot\|$  de  $\mathbb{R}^m$  y  $\|A\|_\infty$  denota la norma del supremo de la matriz  $A$ , es decir,

$$\|A\|_\infty = \max_{i=1,\dots,q} \sum_{j=1}^q |a_{ij}|.$$

El resultado también es cierto si se cumple que

$$h_n L \rho(|A|) < 1,$$

donde  $|A|$  denota la matriz cuyos elementos son los valores absolutos de los elementos de  $A$  y  $\rho(|A|)$  es el radio espectral de la matriz  $|A|$ , es decir,

$$\rho(|A|) = \max\{|\lambda| : \lambda \text{ es autovalor de } |A|\}.$$

*Demostración.* Sabemos que el sistema de ecuaciones (5.1) que definen las etapas de una fórmula de Runge-Kutta implícita es equivalente a aquel que plantean las ecuaciones de la formulación (5.2) y, entonces, podemos considerar el vector  $\mathbf{Y} = (\mathbf{y}_{n,i})_{i=1}^q \in \mathbb{R}^{mq}$ , donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q.$$

Esto nos lleva a definir la aplicación  $\mathbf{G}$  de la siguiente manera:

$$\begin{aligned} \mathbf{G} : \quad \mathbb{R}^{mq} &\longrightarrow \mathbb{R}^{mq} \\ \mathbf{V} = (\mathbf{v}_i)_{i=1}^q &\mapsto \mathbf{G}(\mathbf{V}) = (\mathbf{G}_i(\mathbf{V}))_{i=1}^q \end{aligned}$$

donde

$$\mathbf{G}_i(\mathbf{V}) := \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{v}_j).$$

Entonces el sistema no lineal (5.1) que definen las etapas del método considerado puede formularse de manera equivalente mediante la ecuación de tipo punto fijo siguiente:

$$\mathbf{Y} = \mathbf{G}(\mathbf{Y}),$$

y, por tanto, dicho sistema tiene solución única si y solamente si el vector  $\mathbf{Y} = (\mathbf{y}_{n,i})_{i=1}^q$  es el único punto fijo de la aplicación  $\mathbf{G}$ . Para probar esto último es suficiente con demostrar que  $\mathbf{G}$  es contractiva, pues, en ese caso, por el Teorema del punto fijo de Banach (véase Teorema A.1)  $\mathbf{G}$  tiene un solo punto fijo y este tendría que ser, en consecuencia,  $\mathbf{Y}$ .

No obstante, antes de probar que la aplicación  $\mathbf{G}$  es efectivamente contractiva conviene recordar y precisar lo siguiente respecto a las distintas normas que se consideran. En primer lugar, nótese que  $\|\cdot\|$  denota, según se establece en el enunciado del Teorema, cierta norma en  $\mathbb{R}^m$ . Además, dado un vector de  $\mathbb{R}^{mq}$ ,  $\mathbf{V} = (\mathbf{v}_i)_{i=1}^q$  con  $\mathbf{v}_i \in \mathbb{R}^m$  para  $i = 1, 2, \dots, q$ , denotamos por  $\widehat{\mathbf{V}}$  al vector de  $\mathbb{R}^q$  dado por  $\widehat{\mathbf{V}} = (\|\mathbf{v}_i\|)_{i=1}^q$  y definimos la siguiente norma en  $\mathbb{R}^{mq}$ :

$$\|\mathbf{V}\|_{\mathbb{R}^{mq}} = \max_{i=1, \dots, q} \|\mathbf{v}_i\| = \|\widehat{\mathbf{V}}\|_{\infty}.$$

Entonces, dados dos vectores cualesquiera de  $\mathbb{R}^{mq}$ ,  $\mathbf{U} = (\mathbf{u}_i)_{i=1}^q$  y  $\mathbf{V} = (\mathbf{v}_i)_{i=1}^q$ , se verifica que  $\mathbf{G}$  es una aplicación contractiva puesto que, a su vez, se cumple, por la hipótesis, que:

$$\begin{aligned} \|\mathbf{G}(\mathbf{U}) - \mathbf{G}(\mathbf{V})\|_{\mathbb{R}^{mq}} &= \max_{i=1, \dots, q} \|\mathbf{G}_i(\mathbf{U}) - \mathbf{G}_i(\mathbf{V})\| \leq \max_{i=1, \dots, q} h_n L \sum_{j=1}^q |a_{ij}| \|\mathbf{u}_j - \mathbf{v}_j\| \\ &\leq h_n L \|A\|_{\infty} \|\mathbf{U} - \mathbf{V}\|_{\mathbb{R}^{mq}} < \|\mathbf{U} - \mathbf{V}\|_{\mathbb{R}^{mq}}. \end{aligned}$$

Con ello queda probada la validez de la primera afirmación del enunciado. Veamos la segunda.

Por inducción se obtiene que para las composiciones sucesivas de la función  $\mathbf{G}$  se cumple que:

$$\|\mathbf{G}^k(\mathbf{U}) - \mathbf{G}^k(\mathbf{V})\|_{\mathbb{R}^{mq}} \leq h_n^k L^k \| |A|^k \|_{\infty} \|\mathbf{U} - \mathbf{V}\|_{\mathbb{R}^{mq}}.$$

Como  $\rho(|A|) = \lim_{k \rightarrow \infty} \| |A|^k \|_{\infty}^{\frac{1}{k}}$  y por hipótesis se tiene que  $h_n L \rho(|A|) < 1$  entonces

$$\exists k_0 \in \mathbb{N} \text{ tal que } \forall k \geq k_0 \quad h_n^k L^k \| |A|^k \|_{\infty} < 1.$$

Luego en particular, concluimos que  $\mathbf{G}^{k_0}$  es contractiva y por el Teorema A.1 del punto fijo de Banach tiene un único punto fijo, que será también el único punto fijo de  $\mathbf{G}$ .  $\square$

## A.2. Estabilidad

Veamos ahora bajo qué condiciones una fórmula de Runge-Kutta implícita es estable.

**Proposición A.1.** *La función incremento de un método de Runge-Kutta implícito de  $q$  etapas,*

$$\Phi(t, \mathbf{y}, h) = \sum_{i=1}^q b_i \mathbf{k}_i(t, \mathbf{y}, h),$$

*es lipschitziana respecto de su segundo argumento en el dominio  $I \times \mathbb{R}^m \times [0, h^*]$ , y por tanto el método es estable, si se verifica que*

$$h^* L \rho(|A|) < 1.$$

*Demostración.* En base a la Proposición 1.2 basta con probar que la función incremento,

$$\Phi(t_n, \mathbf{y}_n, h_n) = \sum_{i=1}^q b_i \mathbf{k}_i = \sum_{i=1}^q b_i \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,i}),$$

donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q,$$

es lipschitziana respecto de su segundo argumento.

Entonces, sustituyendo y teniendo en cuenta que la función  $\mathbf{f}$  es lipschitziana respecto de su segundo argumento, con constante de Lipschitz  $L$  respecto a la norma  $\|\cdot\|$  de  $\mathbb{R}^m$ ,

$$\begin{aligned} \|\Phi(t_n, \mathbf{y}_n, h_n) - \Phi(t_n, \mathbf{z}_n, h_n)\| &\leq \sum_{i=1}^q |b_i| \|\mathbf{f}(t_n + c_i h_n, \mathbf{y}_{n,i}) - \mathbf{f}(t_n + c_i h_n, \mathbf{z}_{n,i})\| \\ &\leq \sum_{i=1}^q |b_i| L \|\mathbf{y}_{n,i} - \mathbf{z}_{n,i}\|. \end{aligned} \quad (\text{A.1})$$

Como, además se cumple que,

$$\begin{aligned} \|\mathbf{y}_{n,i} - \mathbf{z}_{n,i}\| &\leq \|\mathbf{y}_n - \mathbf{z}_n\| + h_n \sum_{j=1}^q |a_{ij}| \|\mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}) - \mathbf{f}(t_n + c_j h_n, \mathbf{z}_{n,j})\| \\ &\leq \|\mathbf{y}_n - \mathbf{z}_n\| + h_n \sum_{j=1}^q |a_{ij}| L \|\mathbf{y}_{n,j} - \mathbf{z}_{n,j}\|. \end{aligned}$$

Además, dado un vector cualquiera de  $\mathbb{R}^{mq}$ ,  $\mathbf{U} = (\mathbf{u}_i)_{i=1}^q$  con  $\mathbf{u}_i \in \mathbb{R}^m$  para  $i = 1, 2, \dots, q$ , denotaremos por  $\widehat{\mathbf{U}}$  al vector de  $\mathbb{R}^q$  dado por  $\widehat{\mathbf{U}} = (\|\mathbf{u}_i\|)_{i=1}^q$  y dado otro vector cualquiera de  $\mathbb{R}^{mq}$ ,  $\mathbf{V} = (\mathbf{v}_i)_{i=1}^q$ , usaremos la notación  $\mathbf{U} \preceq \mathbf{V}$  para indicar que

$$\mathbf{u}_i \leq \mathbf{v}_i, \quad i = 1, 2, \dots, q.$$

De esta manera, considerando los vectores de  $\mathbb{R}^{mq}$   $\mathbf{Y} = (\mathbf{y}_{n,i})_{i=1}^q$  y  $\mathbf{Z} = (\mathbf{z}_{n,i})_{i=1}^q$  se verifica:

$$\widehat{\mathbf{Y} - \mathbf{Z}} \preceq \|\mathbf{y}_n - \mathbf{z}_n\| \mathbf{e} + h^* L |A| \widehat{\mathbf{Y} - \mathbf{Z}},$$

donde  $\mathbf{e} = (1, 1, \dots, 1)^T \in \mathbb{R}^q$  y entonces, denotando  $B = h^* L |A|$ , también se tiene que:

$$\|\widehat{\mathbf{Y} - \mathbf{Z}}\|_\infty \leq \|\mathbf{y}_n - \mathbf{z}_n\| \|I + B + B^2 + \dots + B^k\|_\infty + \|B^{k+1}\|_\infty \|\widehat{\mathbf{Y} - \mathbf{Z}}\|_\infty. \quad (\text{A.2})$$

Como, por hipótesis,  $\rho(B) < 1$ , se cumplen las siguientes dos propiedades,

$$\lim_{k \rightarrow \infty} B^k = 0 \quad \text{y} \quad (I - B)^{-1} = \sum_{k=0}^{\infty} B^k.$$

Así, podemos reducir la ecuación (A.2) a la siguiente expresión:

$$\| \widehat{\mathbf{Y} - \mathbf{Z}} \|_{\infty} \leq \| \mathbf{y}_n - \mathbf{z}_n \| \| (I - B)^{-1} \|_{\infty},$$

y, sustituyendo finalmente en la expresión (A.1), obtenemos:

$$\| \Phi(t_n, \mathbf{y}_n, h_n) - \Phi(t_n, \mathbf{z}_n, h_n) \| \leq L \| (I - B)^{-1} \|_{\infty} \sum_{i=1}^q |b_i| \| \mathbf{y}_n - \mathbf{z}_n \|,$$

de donde se deduce que la función  $\phi$  es lipschitziana respecto de su segundo argumento.  $\square$

### A.3. Regularidad

Por último, abordaremos el estudio de la regularidad de las fórmulas de Runge-Kutta de tipo implícito. Para ello necesitamos recordar otro resultado muy importante y conocido del Análisis Matemático: el Teorema de la Función Implícita.

#### Teorema A.3 (de la Función Implícita).

Sean  $\mathbf{f} : \Omega \subseteq \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{R}^M$  una función de clase  $C^p$  en el abierto  $\Omega$  y  $(\mathbf{x}_0, \mathbf{y}_0) \in \Omega$  tal que

$$\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = 0 \quad \text{y} \quad \det \left( \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) \right) \neq 0.$$

Entonces existen dos abiertos  $\mathcal{U} \subseteq \mathbb{R}^N$  con  $\mathbf{x}_0 \in \mathcal{U}$  y  $\mathcal{W} \subseteq \mathbb{R}^M$  con  $\mathbf{y}_0 \in \mathcal{W}$  y una función  $\mathbf{g} : \mathcal{U} \rightarrow \mathcal{W}$  de clase  $C^p$  tal que  $\mathbf{g}(\mathbf{x}_0) = \mathbf{y}_0$  y  $\mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x})) = 0 \quad \forall \mathbf{x} \in \mathcal{U}$ .

**Corolario A.1.** Si la función  $\mathbf{f} : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  que define el problema diferencial (P) es de clase  $C^p$  entonces las funciones de las etapas,  $\mathbf{k}_i$  o  $\mathbf{g}_i$ , según la formulación utilizada en cada momento, (vistas como funciones de  $h$ ) son también de clase  $C^p$  en un entorno de  $h = 0$ .

*Demostración.* Al igual que en el Teorema A.2, usando la formulación (5.2) para definir las funciones de las etapas, podemos considerar el vector  $\mathbf{Y} = (\mathbf{y}_{n,i})_{i=1}^q \in \mathbb{R}^{mq}$ , donde

$$\mathbf{y}_{n,i} = \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{y}_{n,j}), \quad i = 1, 2, \dots, q,$$

y, esto nos lleva, a su vez, a definir la aplicación  $\mathbf{G}$  de la siguiente manera:

$$\begin{aligned} \mathbf{G} : \quad \mathbb{R}^{mq} &\longrightarrow \mathbb{R}^{mq} \\ \mathbf{V} = (\mathbf{v}_i)_{i=1}^q &\mapsto \mathbf{G}(\mathbf{V}) = (\mathbf{G}_i(\mathbf{V}))_{i=1}^q \end{aligned}$$

donde

$$\mathbf{G}_i(\mathbf{V}) := \mathbf{y}_n + h_n \sum_{j=1}^q a_{ij} \mathbf{f}(t_n + c_j h_n, \mathbf{v}_j).$$

Buscamos entonces un punto fijo de  $\mathbf{G}$ , es decir, una solución de la ecuación siguiente:

$$\mathbf{F}(t_n, \mathbf{y}_n, h_n, \mathbf{V}) = \mathbf{Y} - \mathbf{G}(\mathbf{V}) = \mathbf{0}.$$

Claramente, el vector  $\mathbf{Y} = (\mathbf{y}_{n,i})_{i=1}^q$  es una solución del sistema anterior puesto que se verifica:

$$\mathbf{F}(t_n, \mathbf{y}_n, h_n, \mathbf{Y}) = \mathbf{0}.$$

Además, se cumple también que:

$$\det \left( \frac{\partial \mathbf{F}}{\partial \mathbf{V}} (t_n, \mathbf{y}_n, h_n, \mathbf{Y}) \right) = I_q - \det \left( \frac{\partial \mathbf{G}}{\partial \mathbf{V}} (t_n, \mathbf{y}_n, h_n, \mathbf{Y}) \right) \neq 0.$$

Luego, por el Teorema A.3 de la Función Implícita concluimos que:

$$\mathbf{Y} = \mathbf{g}(t_n, \mathbf{y}_n, h_n) \text{ con } \mathbf{g} \text{ de clase } \mathcal{C}^p,$$

y por consiguiente, tanto  $\mathbf{k}_i$  como  $\mathbf{g}_i$  (y por extensión  $\Phi$ ) son de clase  $\mathcal{C}^p$ . □

# Bibliografía

- [1] J.C. Butcher, *Numerical methods for ordinary differential equations in the 20th century*, Elsevier, 2000.
- [2] J.C. Butcher, *Numerical Methods for Ordinary Differential Equations*, John Wiley & Sons, 2008 (Segunda edición).
- [3] M. Calvo, J.I. Montijano, L. Rández, *Curso de Análisis Numérico. Métodos de Runge-Kutta para la resolución numérica de ecuaciones diferenciales ordinarias*, Departamento de Matemática Aplicada de la Universidad de Zaragoza, 1998.
- [4] E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer, 1993 (Segunda edición).
- [5] E. Hairer, G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential Algebraic Problems*, Springer, 1996 (Segunda edición).
- [6] J.D. Lambert, *Numerical Methods for Ordinary Differential Systems. The Initial Value Problem*, John Wiley & Sons, 1991.
- [7] D.F. Griffiths, D.J. Higham, *Numerical Methods for Ordinary Differential Equations, Initial Value Problems*, Springer, 2010.
- [8] J. Sotomayor, *Lições de equações diferenciais ordinarias*, I.M.P.A., 1979.
- [9] G.F. Simmons, *Ecuaciones diferenciales con aplicaciones y notas históricas*, McGraw-Hill, 1993.
- [10] E. Isaacson, H.P. Keller, *Analysis of Numerical Methods*, Dover, 1994.