



Universidad de Valladolid

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Matemáticas

Introducción a la integración numérica de ecuaciones diferenciales algebraicas de índice 2 con métodos de tipo Runge-Kutta.

Alumno: Diego Rojo García

Tutora: M.P. Calvo Cabrero

Índice general

Índice general	1
Introducción	3
1. Ecuaciones diferenciales algebraicas (EDAs)	5
1.1. Índice de una ecuación diferencial algebraica	6
1.1.1. Índice de diferenciación	6
1.1.2. Índice de perturbación	8
1.2. Un ejemplo: el péndulo matemático	9
1.2.1. El péndulo como EDO	11
1.2.2. El péndulo como EDA de índices 1, 2 y 3	12
2. Métodos RK para EDAs de índice 2	15
2.1. Métodos Runge-Kutta	15
2.2. Métodos Runge-Kutta semi-implícitos	17
2.3. Condiciones de orden	20
2.3.1. Desarrollo de Taylor de la solución exacta	20
2.3.2. Desarrollo de Taylor de la solución numérica generada con métodos RK semi-implícitos	29
2.3.3. Condiciones de orden	33
3. Una aplicación: sistemas mecánicos con restricciones	45
3.1. Descripción del problema	45
3.2. Algunos Ejemplos	48
3.2.1. Péndulo Matemático	49
3.2.2. Brazo plano articulado de dos grados de libertad	50
3.2.3. Andrews' Squeezing mechanism	52
3.3. Resultados numéricos	58
3.3.1. Péndulo matemático	59
3.3.2. Brazo plano articulado de dos grados de libertad	60

3.3.3. Andrews' Squeezing mechanism	62
Bibliografía	65
A. Código MATLAB	67
A.1. Programa principal: <code>hem4.m</code>	67
A.2. <code>Pendolo.m</code>	70
A.3. <code>Brazo.m</code>	71
A.4. <code>Andrews.m</code>	72

Introducción

El trabajo que presentamos tiene por objeto el estudio de la integración numérica de ecuaciones diferenciales algebraicas de índice 2, mediante métodos Runge-Kutta semi-implícitos. Las ecuaciones diferenciales algebraicas surgen de manera natural, por ejemplo, en el diseño de circuitos eléctricos al aplicar las leyes de Kirchhoff, cuando se consideran sistemas mecánicos formados por varias piezas simples (rígidas o flexibles) que están acopladas, o al estudiar problemas de control óptimo.

Para enmarcar el trabajo realizado en el contexto de los estudios del Grado en Matemáticas de la Universidad de Valladolid, hay que tener en cuenta que en los contenidos obligatorios del tercer curso se considera una muy breve introducción a la resolución numérica de ecuaciones diferenciales ordinarias (EDOs) y que he tenido la oportunidad de profundizar en el tratamiento numérico de dichas ecuaciones cursando la asignatura optativa de cuarto curso “Solución numérica de ecuaciones diferenciales”. Las ecuaciones diferenciales algebraicas (EDAs) eran, sin embargo, completamente nuevas para mí. Su tratamiento numérico plantea todas las dificultades que surgen al considerar sistemas de EDOs y, además, acarrea dificultades añadidas al tratar con las restricciones algebraicas que las complementan.

En esta memoria se ha considerado únicamente la integración de sistemas de ecuaciones diferenciales algebraicas de índice 2 en forma semi-implícita al no ser factible, debido a la amplia variedad de casos que existen, abordar la integración numérica de las ecuaciones diferenciales algebraicas en su totalidad. Los métodos elegidos para su integración numérica han sido los métodos Runge-Kutta (RK) semi-implícitos que, como se ve en la memoria, aportan algunas ventajas frente a los métodos Runge-Kutta generales.

En el Capítulo 1 se incluyen las definiciones básicas en relación con las EDAs y algunos ejemplos sencillos para ilustrar el concepto de índice.

El Capítulo 2, el más extenso, se ha dedicado a un estudio detallado de los métodos Runge-Kutta semi-implícitos surgidos en la literatura para integrar EDAs de índice 2 semi-implícitas. Especial atención se ha prestado a la obtención de las condiciones de orden que deben cumplir los coeficientes de los métodos considerados. Ello requiere la comparación de los desarrollos de Taylor de la solución exacta del problema y de la solución numérica calculada, para lo cual se ha desarrollado con detalle la *teoría de árboles* necesaria para este tipo de problemas y métodos.

En el Capítulo 3 se presenta una familia importante de problemas que conducen de manera natural a EDAs de índice 2: los sistemas mecánicos con restricciones holonómicas. Se implementa un método de orden 4 de la familia estudiada en el Capítulo 2 y se ilustra su comportamiento sobre tres sistemas modelo de la familia, que presentan distintos grados de dificultad.

En el Apéndice se incluyen los programas en MATLAB que implementan el método elegido y que se han utilizado para generar los resultados numéricos que se recogen en el Capítulo 3 para los tres ejemplos considerados.

Desde esta introducción agradezco el trabajo de dirección realizado por mi tutora, la Dra. Mari Paz Calvo Cabrero, sin olvidar la influencia del resto de profesores que he tenido a lo largo de estos años.

Valladolid a 1 de julio de 2015.

Capítulo 1

Ecuaciones diferenciales algebraicas

Las ecuaciones diferenciales algebraicas (EDAs) son, de manera genérica, ecuaciones diferenciales de la forma

$$\mathbf{F}(\mathbf{Y}'(t), \mathbf{Y}(t), t) = \mathbf{0}$$

donde $\mathbf{Y} = [Y_1, \dots, Y_d]^T : \mathcal{I} \subset \mathbb{R} \rightarrow \mathbb{R}^d$, $\mathbf{F} = [F_1, \dots, F_d]^T : \mathbb{R}^d \times \mathcal{D} \times \mathcal{I} \rightarrow \mathbb{R}^d$, \mathcal{I} es un intervalo abierto de \mathbb{R} , \mathcal{D} es un abierto de \mathbb{R}^d y la notación $'$ se refiere a la derivación respecto de la variable independiente t . Se omitirá la dependencia explícita de \mathbf{Y} e \mathbf{Y}' respecto de t y, a menos que se diga expresamente lo contrario, se supondrá que \mathbf{F} es tan regular como sea necesario.

Los sistemas que vamos a tratar en esta memoria serán sistemas autónomos de la forma

$$\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0} \tag{1.1}$$

puesto que un sistema no autónomo de la forma $\mathbf{F}(\mathbf{Y}', \mathbf{Y}, t) = \mathbf{0}$ puede transformarse en un sistema autónomo. Para ello, basta con introducir t en el vector \mathbf{Y} y añadir la ecuación $t' = 1$. Además, nos centraremos en los sistemas en los que la matriz $\partial\mathbf{F}/\partial\mathbf{Y}'$ es singular.

Si $\partial\mathbf{F}/\partial\mathbf{Y}'$ no es singular, podríamos, al menos formalmente, despejar \mathbf{Y}' de (1.1) en función de \mathbf{Y} para obtener una ecuación diferencial ordinaria (EDO).

En el caso en que $\mathbf{Y} = [\mathbf{X}, \mathbf{Z}]^T$ y (1.1) se pueda escribir como

$$\begin{aligned} \mathbf{X}' &= \mathbf{F}_1(\mathbf{X}, \mathbf{Z}), \\ \mathbf{0} &= \mathbf{F}_2(\mathbf{X}, \mathbf{Z}) \end{aligned}$$

diremos que la ecuación diferencial algebraica está en forma semi-explicita. Las derivadas de las componentes de \mathbf{Z} no aparecen explícitamente en la ecuación y, por ello, dichas componentes reciben el nombre de *variables algebraicas*. Por el contrario, las componentes de \mathbf{X} se llaman *variables diferenciales*.

1.1. Índice de una ecuación diferencial algebraica

1.1.1. Índice de diferenciación

El índice de diferenciación mide el número de veces que hay que derivar una ecuación diferencial algebraica para transformarla en una EDO.

Definición 1.1. La ecuación diferencial algebraica

$$\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0} \quad (1.2)$$

tiene **índice de diferenciación** m , si m es el menor entero tal que las ecuaciones

$$\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0}, \frac{d}{dt}\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0}, \dots, \frac{d^m}{dt^m}\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0}, \quad (1.3)$$

permiten obtener mediante manipulaciones algebraicas un sistema de EDOs explícitas $\mathbf{Y}' = \boldsymbol{\varphi}(\mathbf{Y})$, que llamaremos sistema de EDOs *subyacente*.

Veamos ahora algunas familias de sistemas de EDAs de índice de diferenciación 1, 2 y 3 que aparecen frecuentemente en las aplicaciones.

Sistemas de índice 1

Empezamos considerando problemas de la forma:

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad (1.4a)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{y}, \mathbf{z}). \quad (1.4b)$$

Derivando (1.4b) respecto de t y sustituyendo \mathbf{y}' por su expresión según (1.4a) tenemos

$$\mathbf{0} = \mathbf{g}_y(\mathbf{y}, \mathbf{z})\mathbf{f}(\mathbf{y}, \mathbf{z}) + \mathbf{g}_z(\mathbf{y}, \mathbf{z})\mathbf{z}'.$$

A lo largo de esta memoria indicaremos frecuentemente las derivadas parciales con la habitual notación de subíndices, por ejemplo, en la expresión anterior $\mathbf{g}_z = \partial \mathbf{g} / \partial \mathbf{z}$.

Si \mathbf{g}_z es invertible en un entorno de la solución tendremos

$$\mathbf{z}' = -\mathbf{g}_z^{-1}(\mathbf{y}, \mathbf{z})\mathbf{g}_y(\mathbf{y}, \mathbf{z})\mathbf{f}(\mathbf{y}, \mathbf{z})$$

y, por tanto, si \mathbf{g}_z es invertible, el problema (1.4) tiene índice de diferenciación 1, puesto que tras derivar una vez obtenemos el sistema de EDOs

$$\frac{d}{dt} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{f}(\mathbf{y}, \mathbf{z}) \\ -\mathbf{g}_z^{-1}(\mathbf{y}, \mathbf{z})\mathbf{g}_y(\mathbf{y}, \mathbf{z})\mathbf{f}(\mathbf{y}, \mathbf{z}) \end{pmatrix}.$$

Si los valores iniciales $(\mathbf{y}_0, \mathbf{z}_0)$ satisfacen la ecuación algebraica (1.4b), es decir, si $\mathbf{g}(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{0}$, diremos que son *consistentes*. Si los valores iniciales no son consistentes, no está garantizada la existencia de solución para (1.4).

Sistemas de índice 2

Consideramos ahora problemas de la forma

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad (1.5a)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{y}). \quad (1.5b)$$

Derivando (1.5b) respecto de t y sustituyendo \mathbf{y}' por su expresión según (1.5a) tenemos

$$\mathbf{0} = \mathbf{g}_y(\mathbf{y})\mathbf{f}(\mathbf{y}, \mathbf{z}). \quad (1.6)$$

Las ecuaciones (1.5a) y (1.6), junto con la condición de que $\mathbf{g}_y \mathbf{f}_z$ sea invertible, conforman un sistema de índice de diferenciación 1 de la forma (1.4). Si derivamos respecto de t (1.6), tenemos (omitiendo los argumentos)

$$\mathbf{0} = \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + \mathbf{g}_y \mathbf{f}_y \mathbf{f} + \mathbf{g}_y \mathbf{f}_z \mathbf{z}'$$

y, si $\mathbf{g}_y \mathbf{f}_z$ es invertible en un entorno de la solución, obtenemos la ecuación diferencial para \mathbf{z}

$$\mathbf{z}' = (-\mathbf{g}_y \mathbf{f}_z)^{-1}(\mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + \mathbf{g}_y \mathbf{f}_y \mathbf{f}).$$

Por tanto, el sistema (1.5) tiene índice de diferenciación 2.

Si los valores iniciales $(\mathbf{y}_0, \mathbf{z}_0)$ satisfacen las ecuaciones algebraicas (1.5b) y (1.6) diremos que son *consistentes*. Si los valores iniciales no son consistentes, no está garantizada la existencia de solución para (1.5).

Sistemas de índice 3

Consideramos, por último, problemas de la forma

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad (1.7a)$$

$$\mathbf{z}' = \mathbf{k}(\mathbf{y}, \mathbf{z}, \mathbf{u}), \quad (1.7b)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{y}). \quad (1.7c)$$

Derivando (1.7c) respecto de t dos veces y sustituyendo \mathbf{y}' y \mathbf{z}' de (1.7a) y (1.7b) respectivamente tenemos que

$$\mathbf{0} = \mathbf{g}_y \mathbf{f}, \quad (1.8a)$$

$$\mathbf{0} = \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + \mathbf{g}_y \mathbf{f}_y \mathbf{f} + \mathbf{g}_y \mathbf{f}_z \mathbf{k}. \quad (1.8b)$$

Las ecuaciones (1.7a), (1.7b) y (1.8b), junto con la condición de que $\mathbf{g}_y \mathbf{f}_z \mathbf{k}_u$ sea invertible, conforman un sistema de índice 1 de la forma (1.4). Concluimos, por tanto, que (1.7) tiene índice de diferenciación 3 si $\mathbf{g}_y \mathbf{f}_z \mathbf{k}_u$ es invertible en un entorno de la solución.

Para que los valores iniciales $(\mathbf{y}_0, \mathbf{z}_0, \mathbf{u}_0)$ sean consistentes deberán verificar (1.7c), (1.8a) y (1.8b). Si los valores iniciales no son consistentes, no está garantizada la existencia de solución para (1.7).

1.1.2. Índice de perturbación

El índice de perturbación mide la sensibilidad de las soluciones de una EDA a perturbaciones en la ecuación.

Definición 1.2. La ecuación diferencial algebraica

$$\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0} \quad (1.9)$$

tiene **índice de perturbación** m a lo largo de una solución \mathbf{Y} en $[0, \bar{t}]$, si m es el menor entero tal que para toda función $\hat{\mathbf{Y}}$ que es solución de la ecuación perturbada

$$\mathbf{F}(\hat{\mathbf{Y}}', \hat{\mathbf{Y}}) = \boldsymbol{\delta}(t)$$

se verifica que

$$\|\hat{\mathbf{Y}}(t) - \mathbf{Y}(t)\| \leq C \left(\|\hat{\mathbf{Y}}(0) - \mathbf{Y}(0)\| + \max_{0 \leq \xi \leq t} \|\boldsymbol{\delta}(\xi)\| + \cdots + \max_{0 \leq \xi \leq t} \|\boldsymbol{\delta}^{(m-1)}(\xi)\| \right),$$

siempre que la expresión del segundo miembro sea suficientemente pequeña. Aquí, C denota una constante que depende únicamente de \mathbf{F} y de la longitud del intervalo.

Nota. No hemos escrito deliberadamente

$$\text{“Sea } \widehat{\mathbf{Y}}(t) \text{ la solución de } \mathbf{F}(\widehat{\mathbf{Y}}', \widehat{\mathbf{Y}}) = \boldsymbol{\delta}(t) \dots \text{”}$$

en la definición puesto que para un $\boldsymbol{\delta}(t)$ arbitrario no está asegurada la existencia de $\widehat{\mathbf{Y}}(t)$.

Nota. Con la definición dada, el índice de perturbación no será menor que 1.

Definición 1.3. Diremos que una EDA tiene **índice de perturbación** m si tiene índice de perturbación m a lo largo de cualquier solución.

El índice de diferenciación y el índice de perturbación de un sistema

$$\mathbf{F}(\mathbf{Y}', \mathbf{Y}) = \mathbf{0}$$

coincidirán si \mathbf{Y}' aparece multiplicada por una función constante, lo que ocurre en los sistemas semi-implícitos (1.4), (1.5) y (1.7), en los que aparece multiplicada por una matriz diagonal de ceros y unos.

Un estudio más detallado de la relación entre ambos índices se puede encontrar en [7], donde se demuestra

$$\begin{aligned} & \text{índice de diferenciación} \leq \text{índice de perturbación} \\ \text{e} \quad & \text{índice de perturbación} \leq \text{índice de diferenciación} + 1. \end{aligned}$$

Aunque, en lo sucesivo, cuando hablemos de índice nos estaremos refiriendo al índice de diferenciación, conviene mencionar aquí que además de los dos tipos de índice que se han definido en la memoria, en la literatura se han considerado otros tipos de índices (índice de nilpotencia, índice geométrico, índice estructural, índice de tractabilidad, ...) surgidos en su mayor parte en el contexto de sistemas de EDAs lineales [11, pp. 5–8].

1.2. Un ejemplo: el péndulo matemático

Veamos cuáles son las ecuaciones del movimiento de una masa puntual m suspendida al final de una varilla sin masa de longitud l bajo la influencia de una gravedad g , utilizando coordenadas cartesianas (x, y) . Un esquema del problema planteado puede verse en la Figura 1.1.

Si denotamos por θ el ángulo formado por la varilla con el eje vertical, tenemos que $x = l \sin \theta$ e $y = l \cos \theta$. Si denotamos por λ_x y λ_y a las componentes

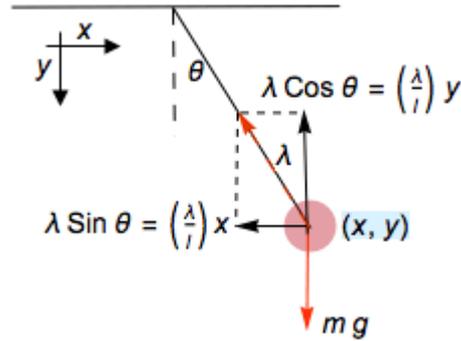


Figura 1.1: Péndulo (de <http://reference.wolfram.com> y elaboración propia)

horizontal y vertical de la tensión de la varilla y por $\lambda = \|\lambda_x, \lambda_y\|^T$ al módulo de dicha tensión, tendremos

$$\lambda_x = -\lambda \sin \theta = -\lambda \frac{x}{l},$$

$$\lambda_y = -\lambda \cos \theta = -\lambda \frac{y}{l}.$$

Aplicando las leyes de Newton tenemos entonces

$$mx'' = -\frac{\lambda}{l}x, \quad (1.10a)$$

$$my'' = -\frac{\lambda}{l}y + mg. \quad (1.10b)$$

Y reescribiendo cada una de las ecuaciones como un sistema diferencial de primer orden, tenemos

$$\begin{cases} x' = u, \\ u' = -\frac{\lambda}{ml}x, \\ y' = v, \\ v' = -\frac{\lambda}{ml}y + g. \end{cases}$$

Como la longitud de la varilla es fija, la restricción de la posición de la masa viene dada por la ecuación algebraica

$$x^2 + y^2 = l^2.$$

Las ecuaciones que controlan el movimiento del péndulo descrito son entonces

$$\begin{cases} x' = u, \\ y' = v, \end{cases} \quad (1.11a)$$

$$\begin{cases} u' = -\frac{\lambda}{ml}x, \\ v' = -\frac{\lambda}{ml}y + g, \end{cases} \quad (1.11b)$$

$$0 = x^2 + y^2 - l^2. \quad (1.11c)$$

1.2.1. El péndulo como EDO

Si realizamos el cambio de variables

$$\begin{aligned} x &= l \sin \theta, \\ y &= l \cos \theta, \end{aligned}$$

que garantiza que la restricción (1.11c) se satisface automáticamente, tendremos que las derivadas de primer y segundo orden de x e y son

$$\begin{aligned} x' &= l\theta' \cos \theta, & x'' &= l\theta'' \cos \theta - l(\theta')^2 \sin \theta, \\ y' &= -l\theta' \sin \theta, & y'' &= -l\theta'' \sin \theta - l(\theta')^2 \cos \theta. \end{aligned}$$

Sustituimos ahora estos valores en (1.10a) y (1.10b) obteniendo

$$\left(\frac{\lambda}{l} - m(\theta')^2\right) l \sin \theta + ml\theta'' \cos \theta = 0, \quad (1.12a)$$

$$-ml\theta'' \sin \theta + \left(\frac{\lambda}{l} - m(\theta')^2\right) l \cos \theta - mg = 0. \quad (1.12b)$$

Multiplicando (1.12a) por $\cos \theta$, (1.12b) por $-\sin \theta$ y sumando ambas ecuaciones obtenemos

$$ml\theta'' + mg \sin \theta = 0,$$

y despejando θ'' obtenemos la conocida EDO de segundo orden para el péndulo

$$\theta'' = -\frac{g}{l} \sin \theta.$$

Este tipo de simplificaciones sencillas que pueden realizarse en el péndulo por ser un caso muy simple de sistema mecánico multicuerpo no se podrán realizar, generalmente, en sistemas más complejos.

1.2.2. El péndulo como EDA de índices 1, 2 y 3

Índice 3

Con la formulación (1.11) el péndulo es un sistema de ecuaciones diferenciales algebraicas de índice 3 de la forma (1.7), donde

$$\begin{aligned}\mathbf{y} &= [x, y]^T, \\ \mathbf{z} &= [u, v]^T, \\ \mathbf{u} &= \lambda,\end{aligned}$$

y las ecuaciones del sistema son

$$\begin{aligned}\mathbf{y}' &= \mathbf{z} = \mathbf{f}(\mathbf{y}, \mathbf{z}), \\ \mathbf{z}' &= \left[-\frac{\lambda}{ml}x, -\frac{\lambda}{ml}y + g \right]^T = \mathbf{k}(\mathbf{y}, \mathbf{z}, \mathbf{u}), \\ 0 &= x^2 + y^2 - l^2 = \mathbf{g}(\mathbf{y}).\end{aligned}$$

Se tiene entonces

$$\mathbf{g}_y = [2x, 2y], \quad \mathbf{f}_z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \mathbf{k}_u = \left[-\frac{x}{ml}, -\frac{y}{ml} \right]^T$$

y, por tanto,

$$\mathbf{g}_y \mathbf{f}_z \mathbf{k}_u = -\frac{2}{ml}(x^2 + y^2),$$

que es invertible en un entorno de la solución.

Índice 2

Si derivamos (1.11c) respecto de t , tras dividir entre 2 y sustituir x' e y' por sus expresiones según (1.11a) obtenemos la ecuación

$$0 = xu + yv. \tag{1.13}$$

La formulación (1.11a), (1.11b), (1.13) es la de un sistema de EDAs de índice 2 de la forma (1.5), donde $\mathbf{y} = [x, y, u, v]^T$ y $\mathbf{z} = \lambda$. Las ecuaciones del sistema son

$$\begin{aligned}\mathbf{y}' &= \left[u, v, -\frac{\lambda}{ml}x, -\frac{\lambda}{ml}y + g \right]^T = \mathbf{f}(\mathbf{y}, \mathbf{z}), \\ 0 &= xu + yv = \mathbf{g}(\mathbf{y}).\end{aligned}$$

En este caso

$$\mathbf{g}_y = [u, v, x, y] \quad \text{y} \quad \mathbf{f}_z = \left[0, 0, -\frac{x}{ml}, -\frac{y}{ml}\right]^T,$$

y, por tanto,

$$\mathbf{g}_y \mathbf{f}_z = -\frac{1}{ml}(x^2 + y^2),$$

que es invertible en un entorno de la solución.

Índice 1

Si derivamos (1.13) respecto de t y sustituimos de nuevo x' e y' por sus expresiones según (1.11a) tenemos que

$$u^2 + v^2 + xu' + yv' = 0.$$

Sustituyendo ahora u' y v' por sus expresiones según (1.11b), tenemos que

$$u^2 + v^2 - \frac{\lambda}{ml}(x^2 + y^2) + gy = 0.$$

Sabemos por (1.11c) que $x^2 + y^2 = l^2$ y, por tanto, tras multiplicar por m obtenemos la ecuación

$$m(u^2 + v^2) - \lambda l + mgy = 0. \quad (1.14)$$

Con la formulación (1.11a), (1.11b), (1.14) el péndulo es un sistema de EDAs de índice 1 de la forma (1.4), donde $\mathbf{y} = [x, y, u, v]^T$ y $\mathbf{z} = \lambda$. Las ecuaciones del sistema son

$$\begin{aligned} \mathbf{y}' &= \left[u, v, -\frac{\lambda}{ml}x, -\frac{\lambda}{ml}y + g \right]^T = \mathbf{f}(\mathbf{y}, \mathbf{z}), \\ 0 &= m(u^2 + v^2) - \lambda l + mgy = \mathbf{g}(\mathbf{y}, \mathbf{z}) \end{aligned}$$

y $\mathbf{g}_z = -l$ es invertible.

Capítulo 2

Métodos RK semi-explícitos para EDAs de índice 2

Consideramos en este capítulo ecuaciones diferenciales algebraicas de la forma

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad (2.1a)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{y}), \quad (2.1b)$$

donde $\mathbf{g}_y \mathbf{f}_z$ es invertible en un entorno de la solución, es decir, ecuaciones diferenciales algebraicas semi-implícitas de índice 2. En todo el capítulo supondremos, salvo que se diga lo contrario, que los valores iniciales $\mathbf{y}_0, \mathbf{z}_0$ son consistentes con (2.1), i.e.

$$\mathbf{g}(\mathbf{y}_0) = \mathbf{0} \quad \text{y} \quad \mathbf{g}_y(\mathbf{y}_0) \mathbf{f}(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{0}.$$

El objetivo de este capítulo es el estudio de métodos Runge-Kutta semi-implícitos para la integración numérica de (2.1). Sin embargo, conviene comenzar con algunos aspectos generales del uso de los métodos Runge-Kutta aplicados a sistemas de la forma (2.1), que nos servirán de apoyo para el desarrollo de ciertos resultados relativos a los métodos semi-implícitos.

2.1. Métodos Runge-Kutta

En esta sección utilizaremos para la integración numérica de (2.1) métodos Runge-Kutta con matriz $\mathcal{A} = (a_{ij})$ invertible y que, por tanto, serán implícitos. Dada una aproximación \mathbf{y}_n a la solución de (2.1) en tiempo t_n

satisfaciendo $\mathbf{g}(\mathbf{y}_n) = \mathbf{0}$, el método genera una aproximación a la solución de (2.1) en $t_{n+1} = t_n + h$ mediante las relaciones

$$\begin{aligned}\mathbf{y}_{n+1} &= \mathbf{y}_n + h \sum_{i=1}^s b_i \mathbf{k}_{ni}, \\ \mathbf{z}_{n+1} &= \mathbf{z}_n + h \sum_{i=1}^s b_i \boldsymbol{\ell}_{ni},\end{aligned}\tag{2.2a}$$

donde

$$\mathbf{k}_{ni} = \mathbf{f}(\mathbf{Y}_{ni}, \mathbf{Z}_{ni}),\tag{2.2b}$$

$$\mathbf{0} = \mathbf{g}(\mathbf{Y}_{ni}),\tag{2.2c}$$

y las etapas internas vienen dadas por

$$\mathbf{Y}_{ni} = \mathbf{y}_n + h \sum_{i=1}^s a_{ij} \mathbf{k}_{nj},\tag{2.2d}$$

$$\mathbf{Z}_{ni} = \mathbf{z}_n + h \sum_{i=1}^s a_{ij} \boldsymbol{\ell}_{nj}.\tag{2.2e}$$

El parámetro h es la longitud del paso de integración, s denota el número de etapas del método y los coeficientes a_{ij} y b_i definen el método.

Una vez resuelto el sistema no lineal (2.2b)–(2.2d) que proporciona los valores de $(\mathbf{Y}_{ni}, \mathbf{Z}_{ni})$ y teniendo en cuenta que la matriz \mathcal{A} de los coeficientes del método Runge-Kutta es invertible, los vectores $\boldsymbol{\ell}_{ni}$ pueden escribirse como

$$\boldsymbol{\ell}_{ni} = \frac{1}{h} \sum_{j=1}^s w_{ij} (\mathbf{Z}_{nj} - \mathbf{z}_n),\tag{2.3}$$

donde w_{ij} denota el elemento (i, j) de la matriz inversa de \mathcal{A} . Una vez calculados $\boldsymbol{\ell}_{ni}$, $1 \leq i \leq s$, obtenemos \mathbf{y}_{n+1} y \mathbf{z}_{n+1} de (2.2a) de forma explícita.

A continuación citamos un resultado sobre la existencia y unicidad de la solución del método Runge-Kutta, donde para simplificar la notación, escribiremos (η, ζ) en lugar de $(\mathbf{y}_n, \mathbf{z}_n)$, que supondremos dependientes de h y suprimiremos el índice n en \mathbf{Y}_{ni} , \mathbf{Z}_{ni} , \mathbf{k}_{ni} y $\boldsymbol{\ell}_{ni}$. Una prueba del teorema puede encontrarse en [4, pp. 31–32].

Teorema 2.1. Sean (η, ζ) tales que

$$\mathbf{g}(\eta) = O(h^2) \quad y \quad \mathbf{g}_y(\eta)\mathbf{f}(\eta, \zeta) = O(h), \quad (2.4)$$

y supongamos que $\mathbf{g}_y\mathbf{f}_z$ es invertible en un entorno de la solución de (2.1). Si la matriz de coeficientes del método Runge-Kutta $\mathcal{A} = (a_{ij})$ es invertible entonces el sistema no lineal

$$\begin{cases} \mathbf{Y}_i = \eta + h \sum_{j=1}^s a_{ij}\mathbf{f}(\mathbf{Y}_j, \mathbf{Z}_j), & i = 1, \dots, s, \\ \mathbf{0} = \mathbf{g}(\mathbf{Y}_i), & i = 1, \dots, s. \end{cases} \quad (2.5)$$

tiene solución para $h \leq h_0$. Esta solución es localmente única y verifica

$$\mathbf{Y}_i - \eta = O(h) \quad y \quad \mathbf{Z}_i - \zeta = O(h).$$

Nota. Aunque comencemos con valores iniciales consistentes, para $n \geq 1$ la solución numérica $(\mathbf{y}_n, \mathbf{z}_n)$ sólo verificará aproximadamente las condiciones de consistencia (1.5b) y (1.6). Esto es lo que expresa la condición (2.4) del teorema.

2.2. Métodos Runge-Kutta semi-implícitos

Abordamos ahora la integración numérica de (2.1) utilizando métodos Runge-Kutta semi-implícitos que, dada una aproximación \mathbf{y}_n a la solución de (2.1) en tiempo t_n satisfaciendo $\mathbf{g}(\mathbf{y}_n) = \mathbf{0}$ generan una aproximación a la solución de (2.1) en $t_{n+1} = t_n + h$ mediante las relaciones

$$\mathbf{Y}_{n1} = \mathbf{y}_n, \quad (2.6a)$$

$$\mathbf{Y}_{ni} = \mathbf{y}_n + h \sum_{j=1}^{i-1} a_{ij}\mathbf{f}(\mathbf{Y}_{nj}, \mathbf{Z}_{nj}), \quad i = 2, \dots, s, \quad (2.6b)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{Y}_{ni}), \quad i = 1, \dots, s, \quad (2.6c)$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{i=1}^s b_i\mathbf{f}(\mathbf{Y}_{ni}, \mathbf{Z}_{ni}), \quad (2.6d)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{y}_{n+1}). \quad (2.6e)$$

Igual que en la sección anterior, el parámetro h es la longitud del paso de integración, s denota el número de etapas del método y los coeficientes a_{ij} y b_i

definen el método. Como viene siendo habitual se introducen los coeficientes

$$c_1 = 0, \quad c_i = \sum_{j=1}^{i-1} a_{ij}, \quad 2 \leq i \leq s.$$

Veamos cómo se aplica el algoritmo. Como $\mathbf{Y}_{n1} = \mathbf{y}_n$ la igualdad (2.6c) para $i = 1$ se satisface por hipótesis. Si ahora sustituímos en (2.6c) la expresión de \mathbf{Y}_{n2} dada por (2.6b) tenemos que resolver el sistema no lineal en \mathbf{Z}_{n1}

$$0 = \mathbf{g}(\mathbf{y}_n + ha_{21}\mathbf{f}(\mathbf{Y}_{n1}, \mathbf{Z}_{n1})).$$

Una vez hemos calculado \mathbf{Z}_{n1} , obtenemos \mathbf{Y}_{n2} de (2.6b) de forma explícita.

Podemos repetir este proceso en las sucesivas etapas; en la etapa $(i+1)$ -ésima deberemos resolver el sistema no lineal en \mathbf{Z}_{ni}

$$0 = \mathbf{g}\left(\mathbf{y}_n + h \sum_{j=1}^i a_{i+1,j}\mathbf{f}(\mathbf{Y}_{nj}, \mathbf{Z}_{nj})\right), \quad (2.7)$$

para lo que es necesario que $a_{i+1,i} \neq 0$. A continuación, obtenemos \mathbf{Z}_{ns} resolviendo el sistema resultante de sustituir en (2.6e) la expresión de \mathbf{y}_{n+1} dada por (2.6d), es decir,

$$0 = \mathbf{g}\left(\mathbf{y}_n + h \sum_{i=1}^s b_i\mathbf{f}(\mathbf{Y}_{ni}, \mathbf{Z}_{ni})\right), \quad (2.8)$$

para lo que es necesario que $b_s \neq 0$. Finalmente, calculamos de forma explícita \mathbf{y}_{n+1} de (2.6d). Se puede probar (ver [3, LEMMA 1]) que para h suficientemente pequeño la solución de los sistemas (2.7) y (2.8) existe y es localmente única.

En general, no podemos determinar \mathbf{z}_{n+1} . Para obtener una aproximación a $\mathbf{z}(t_{n+1})$ de las ecuaciones (2.6) consideraremos métodos con $c_s = 1$ y tomaremos $\mathbf{z}_{n+1} = \mathbf{Z}_{ns}$.

Los métodos Runge-Kutta semi-implícitos como (2.6) se pueden reescribir en el formato general (2.2) de la siguiente manera. La ecuación de la primera etapa viene dada por $\mathbf{Y}_{n1} = \mathbf{y}_n$ y si utilizamos la notación

$$\tilde{\mathbf{Y}}_{ni} := \mathbf{Y}_{n,i+1} \quad \text{si } i = 0, \dots, s-1, \quad \tilde{\mathbf{Y}}_{ns} := \mathbf{y}_{n+1} \quad (2.9)$$

y

$$\tilde{\mathcal{A}} = (\tilde{a}_{ij})_{i,j=1}^s = \begin{bmatrix} a_{21} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{s1} & \cdots & a_{s,s-1} & 0 \\ b_1 & \cdots & b_{s-1} & b_s \end{bmatrix}, \quad (2.10)$$

el resto de ecuaciones se transforman en

$$\begin{aligned}\tilde{\mathbf{Y}}_{ni} &= \mathbf{y}_n + h \sum_{j=1}^i \tilde{a}_{ij} \mathbf{f}(\tilde{\mathbf{Y}}_{n,j-1}, \mathbf{Z}_{nj}), \quad i = 1, \dots, s, \\ \mathbf{0} &= \mathbf{g}(\tilde{\mathbf{Y}}_{ni}), \quad i = 1, \dots, s.\end{aligned}\tag{2.11}$$

Excepto por el desplazamiento del segundo índice de $\tilde{\mathbf{Y}}_{n,j-1}$, (2.11) es equivalente a (2.5) y por tanto a (2.2) con matriz de coeficientes invertible $\tilde{\mathbf{A}}$, pues la solución numérica del método Runge-Kutta semi-implícito es

$$\begin{aligned}\mathbf{y}_{n+1} &= \tilde{\mathbf{Y}}_{ns}, \\ \mathbf{z}_{n+1} &= \mathbf{Z}_{ns},\end{aligned}$$

y dado que $\tilde{a}_{sj} = b_j$, $1 \leq j \leq s$, esto coincide con (2.2a).

Si consideramos valores iniciales consistentes \mathbf{y}_0 , \mathbf{z}_0 y aplicamos un paso de longitud h de un método Runge-Kutta semi-implícito tenemos que

$$\begin{aligned}\mathbf{Y}_1 &= \mathbf{y}_0, \\ \mathbf{Y}_i &= \mathbf{y}_0 + h \sum_{j=1}^{i-1} a_{ij} \mathbf{f}(\mathbf{Y}_j, \mathbf{Z}_j), \quad i = 2, \dots, s, \\ \mathbf{0} &= \mathbf{g}(\mathbf{Y}_i), \quad i = 1, \dots, s, \\ \mathbf{y}_1 &= \mathbf{y}_0 + h \sum_{i=1}^s b_i \mathbf{f}(\mathbf{Y}_i, \mathbf{Z}_i), \\ \mathbf{0} &= \mathbf{g}(\mathbf{y}_1), \\ \mathbf{z}_1 &= \mathbf{Z}_s.\end{aligned}$$

En esta situación, la diferencia entre la solución numérica $(\mathbf{y}_1, \mathbf{z}_1)$ tras un paso de longitud h y la solución exacta en $t + h$

$$\delta \mathbf{y}_h(t) = \mathbf{y}_1 - \mathbf{y}(t + h), \quad \delta \mathbf{z}_h(t) = \mathbf{z}_1 - \mathbf{z}(t + h)$$

recibe el nombre de *error local*.

Teorema 2.2. *Supongamos que $\mathbf{g}_y \mathbf{f}_z$ es invertible en un entorno de la solución $(\mathbf{y}(t), \mathbf{z}(t))$ de (2.1) y que los valores iniciales son consistentes. Si los coeficientes del método Runge-Kutta verifican que $a_{i,i-1} \neq 0$ para $i = 2, \dots, s$, $b_s \neq 0$ y el error local verifica*

$$\delta \mathbf{y}_h(t) = O(h^r), \quad P(t) \delta \mathbf{y}_h(t) = O(h^{r+1}), \tag{2.12}$$

donde $P(t)$ es la proyección dada por

$$P(t) = I - Q(t), \quad Q(t) = \mathbf{f}_z(\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y(\mathbf{y}(t), \mathbf{z}(t)), \quad (2.13)$$

entonces el método (2.6) es convergente de orden r , i.e.,

$$\mathbf{y}_n - \mathbf{y}(t_n) = O(h^r) \quad \text{para} \quad t_n - t_0 = nh \leq \text{Const.} \quad (2.14)$$

Una prueba de este resultado se puede encontrar en [4, p. 49], donde se utiliza la notación (2.9)–(2.10), la formulación (2.11) y los resultados existentes para los métodos Runge-Kutta de la forma (2.2).

2.3. Condiciones de orden de métodos RK semi-implícitos para EDAs de índice 2

En esta sección obtendremos condiciones sobre los coeficientes del método Runge-Kutta aplicado al sistema (2.1) que aseguran que el error local es de un orden dado.

Obtendremos dichas condiciones comparando los desarrollos de Taylor de la solución exacta de (2.1) y de la solución numérica calculada tras un paso con (2.6).

2.3.1. Desarrollo de Taylor de la solución exacta

Si los valores iniciales son consistentes se puede probar [4, p. 4] que el sistema (2.1) tiene, localmente, solución única. En esta subsección vamos a describir el desarrollo de Taylor de las componentes \mathbf{z} e \mathbf{y} de la solución exacta del sistema (2.1). Para ello calcularemos en primer lugar las sucesivas derivadas respecto de t de \mathbf{z} e \mathbf{y} y las escribiremos como combinación lineal de las llamadas *diferenciales elementales*, i.e. expresiones que contienen únicamente derivadas de \mathbf{f} y \mathbf{g} .

La derivada respecto de t de la primera componente de la solución, $\mathbf{y}(t)$, viene dada por (2.1a). Si derivamos esta ecuación respecto de t tenemos

$$\mathbf{y}'' = \mathbf{f}_y(\mathbf{y}, \mathbf{z})\mathbf{y}' + \mathbf{f}_z(\mathbf{y}, \mathbf{z})\mathbf{z}', \quad (2.15)$$

donde se ha omitido la dependencia respecto de t de \mathbf{y} , \mathbf{z} y sus derivadas. De ahora en adelante se omitirán, generalmente, tanto esta dependencia como los argumentos de \mathbf{f} , \mathbf{g} y sus derivadas, siempre que estos sean obvios.

Para calcular \mathbf{z}' primero derivamos dos veces respecto de t la ecuación (2.1b)

$$\begin{aligned} 0 &= \mathbf{g}_y \mathbf{y}', \\ 0 &= \mathbf{g}_{yy}(\mathbf{y}', \mathbf{y}') + \mathbf{g}_y \mathbf{y}'', \end{aligned} \quad (2.16)$$

y a continuación sustituimos en (2.16) \mathbf{y}' e \mathbf{y}'' por sus expresiones según (2.1a) y (2.15), respectivamente. Obtenemos entonces

$$0 = \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + \mathbf{g}_y \mathbf{f}_y \mathbf{f} + \mathbf{g}_y \mathbf{f}_z \mathbf{z}' \quad (2.17)$$

o, equivalentemente,

$$\mathbf{z}' = (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}, \quad (2.18)$$

donde hemos utilizado la hipótesis de que $\mathbf{g}_y \mathbf{f}_z$ es invertible en un entorno de la solución.

Sustituyendo en (2.15) \mathbf{z}' e \mathbf{y}' por sus expresiones según (2.18) y (2.1a) obtenemos la expresión de \mathbf{y}'' como combinación lineal de diferenciales elementales

$$\mathbf{y}'' = \mathbf{f}_y \mathbf{f} + \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}. \quad (2.19)$$

Para calcular \mathbf{z}'' derivamos (2.18) respecto de t ayudándonos de que para un vector constante \mathbf{u}

$$\begin{aligned} \frac{d}{dt} (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u} &= \\ &= (-\mathbf{g}_y \mathbf{f}_z)^{-1} \left(\mathbf{g}_{yy}(\mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, \mathbf{f}) + \mathbf{g}_y \mathbf{f}_{zy}((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, \mathbf{f}) \right. \\ &\quad \left. + \mathbf{g}_y \mathbf{f}_{zz}((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}) \right) \end{aligned} \quad (2.20)$$

y, por tanto,

$$\begin{aligned} \mathbf{z}'' &= (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yyy}(\mathbf{f}, \mathbf{f}, \mathbf{f}) + 3(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}_y \mathbf{f}) \\ &\quad + 3(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f})) \\ &\quad + 3(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}) + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{yy}(\mathbf{f}, \mathbf{f}) \\ &\quad + 2(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{yz}(\mathbf{f}, (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f})) \\ &\quad + 2(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{yz}(\mathbf{f}, \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}) + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}_y \mathbf{f} \\ &\quad + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) \\ &\quad + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f} \\ &\quad + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zz}((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}), (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f})) \\ &\quad + 2(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zz}((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}), (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}) \\ &\quad + (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zz}((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}, (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f}). \end{aligned} \quad (2.21)$$

Derivando sucesivamente (2.19) y (2.21) iremos obteniendo las expresiones de las derivadas sucesivas, aunque la larga expresión de \mathbf{z}'' como combinación lineal de diferenciales elementales muestra la gran ayuda que supondría tener una representación gráfica más compacta de las mismas.

Árboles y diferenciales elementales

Para lograr una representación más compacta de las expresiones de las derivadas respecto de t de \mathbf{z} e \mathbf{y} como combinación lineal de diferenciales elementales, introduciremos como herramienta auxiliar los árboles con raíz para representar a estas últimas.

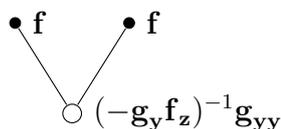
Los *árboles con raíz* son grafos conexos que no tienen ciclos y en los que existe un nodo o vértice destacado que denominaremos *raíz* del árbol. Las aristas del grafo suelen denominarse *ramas* en el contexto de los árboles. A lo largo de la memoria la raíz será el nodo inferior de cada grafo y las ramas que salgan de ella y el resto de nodos serán ascendentes.

A diferencia de los árboles utilizados en el estudio de las condiciones de orden de las EDOs, en este caso tendremos dos tipos de nodos que diferenciaremos gráficamente como vértices negros y blancos.

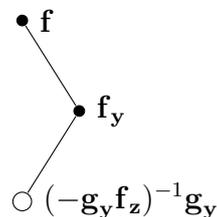
Para la asociación entre diferenciales elementales y árboles con raíz identificamos \mathbf{f} con un vértice negro, una derivada parcial $\frac{\partial^{m+n}\mathbf{f}}{\partial\mathbf{y}^m\partial\mathbf{z}^n}$ con un vértice negro del que salen m ramas ascendentes que lo conectan con m vértices negros y n ramas ascendentes que lo conectan con n vértices blancos. También identificamos $(-\mathbf{g}_y\mathbf{f}_z)^{-1}\frac{\partial^q\mathbf{g}}{\partial\mathbf{y}^q}$ con un vértice blanco del que salen q ramas ascendentes que lo conectan con q vértices negros.

\mathbf{y}' : • \mathbf{f}

\mathbf{z}' :



$$(-\mathbf{g}_y\mathbf{f}_z)^{-1}\mathbf{g}_{yy}(\mathbf{f}, \mathbf{f})$$



$$(-\mathbf{g}_y\mathbf{f}_z)^{-1}\mathbf{g}_y\mathbf{f}_y\mathbf{f}$$

Figura 2.1: Representación gráfica de los diferenciales elementales de \mathbf{y}' y \mathbf{z}' .

Los árboles con raíz asociados a las diferenciales elementales de las expresiones de \mathbf{y}' y \mathbf{z}' dadas por (2.1a) y (2.18) pueden verse en la Figura 2.1.

Las derivadas de \mathbf{y} están asociadas a árboles cuyas raíces son negras. Denotaremos estos árboles por t o t_i ; al árbol que está formado por un único vértice negro, el asociado a \mathbf{y}' , lo denotaremos por τ . Los árboles asociados a las derivadas de \mathbf{z} tienen, por el contrario, raíces blancas. Denotaremos estos árboles por u o u_i .

Definición 2.1. Denotamos por $T = T_{\mathbf{y}} \cup T_{\mathbf{z}}$ al conjunto de árboles (diferenciales algebraicos de índice 2) definidos de manera recursiva por

- $\tau \in T_{\mathbf{y}}$,
- $[t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}}$ si $t_1, \dots, t_m \in T_{\mathbf{y}}$ y $u_1, \dots, u_n \in T_{\mathbf{z}}$,
- $[t_1, \dots, t_m]_{\mathbf{z}} \in T_{\mathbf{z}}$ si $t_1, \dots, t_m \in T_{\mathbf{y}}$ y o bien $m > 1$ o bien $m = 1$ y $t_1 \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

En la definición anterior $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}}$ es el árbol que se obtiene al conectar las raíces de $t_1, \dots, t_m, u_1, \dots, u_n$ mediante $m + n$ ramas a un nuevo vértice negro que se convierte en la raíz del árbol t y $u = [t_1, \dots, t_m]_{\mathbf{z}}$ es el árbol que se obtiene de igual modo al conectar las raíces de t_1, \dots, t_m a un nuevo vértice blanco que se convierte en la raíz de u .

Definición 2.2. El **orden** de un árbol $t \in T_{\mathbf{y}}$ o $u \in T_{\mathbf{z}}$, que denotaremos por $\rho(t)$ o $\rho(u)$, es la diferencia entre el número de vértices negros y el número de vértices blancos del árbol.

Podemos ver en la Figura (2.1) como el orden de los árboles asociados a \mathbf{y}' y \mathbf{z}' es 1.

Definición 2.3. Las **diferenciales elementales** $F(t)$ o $F(u)$ asociadas a los árboles de T quedan definidas de la manera siguiente:

- $F(\tau) = \mathbf{f}$,
- $F(t) = \frac{\partial^{m+n} \mathbf{f}}{\partial \mathbf{y}^m \partial \mathbf{z}^n} \left(F(t_1), \dots, F(t_m), F(u_1), \dots, F(u_n) \right)$, si
 $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}}$,
- $F(u) = (-\mathbf{g}_y \mathbf{f}_z)^{-1} \frac{\partial^q \mathbf{g}}{\partial \mathbf{y}^q} \left(F(t_1), \dots, F(t_q) \right)$, si $u = [t_1, \dots, t_q]_{\mathbf{z}} \in T_{\mathbf{z}}$.

Ahora que tenemos definidas las diferenciales elementales asociadas a los árboles con raíz, veamos cómo obtener, con la ayuda de estos últimos, la

expresión de las sucesivas derivadas de \mathbf{z} e \mathbf{y} como combinación lineal de diferenciales elementales.

Para ello debemos derivar, con la ayuda de los árboles asociados, las diferenciales elementales. Por la regla de Leibniz o del producto, la derivación respecto de t de una diferencial elemental da lugar a una suma de diferenciales elementales. Dado un árbol $\sigma \in T$ con diferencial elemental asociada $F(\sigma)$, obtenemos los árboles asociados a las diferenciales elementales de la derivada respecto de t de la diferencial elemental $F(\sigma)$ siguiendo las siguientes reglas:

1. En el caso de que en la aplicación de la regla de Leibniz se derive un término $\frac{\partial^{m+n}\mathbf{f}}{\partial\mathbf{y}^m\partial\mathbf{z}^n}$ (asociado a un nodo negro) o $\frac{\partial^q\mathbf{g}}{\partial\mathbf{y}^q}$ de $F(\sigma)$ (asociado a un nodo blanco) con respecto a \mathbf{y} (y consecuentemente se añade el factor $\mathbf{y}' = \mathbf{f}$), a partir del árbol σ se obtiene un nuevo árbol, que será de orden una unidad mayor, añadiendo al vértice asociado al término $\frac{\partial^{m+n}\mathbf{f}}{\partial\mathbf{y}^m\partial\mathbf{z}^n}$ o $\frac{\partial^q\mathbf{g}}{\partial\mathbf{y}^q}$ que derivamos respecto de \mathbf{y} una rama ascendente que se conectará con un árbol τ (un nuevo nodo negro). En la Figura 2.2 pueden verse ejemplos de la aplicación de esta regla donde aparece el árbol con raíz inicial marcado en azul.

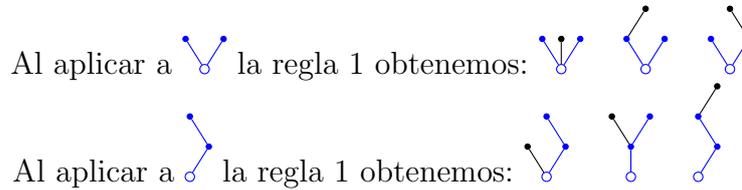


Figura 2.2: Ejemplo de aplicación de la regla 1

2. En el caso de que en la aplicación de la regla de Leibniz se derive un término $\frac{\partial^{m+n}\mathbf{f}}{\partial\mathbf{y}^m\partial\mathbf{z}^n}$ de $F(\sigma)$ con respecto a \mathbf{z} y se añadan consecuentemente los factores

$$(-\mathbf{g}_y\mathbf{f}_z)^{-1}\mathbf{g}_{yy}(\mathbf{f}, \mathbf{f}) \quad \text{y} \quad (-\mathbf{g}_y\mathbf{f}_z)^{-1}\mathbf{g}_y\mathbf{f}_y\mathbf{f},$$

que corresponden a \mathbf{z}' en su expresión (2.18), a partir del árbol σ se obtienen dos nuevos árboles, ambos de orden una unidad mayor, añadiendo al vértice negro asociado al término $\frac{\partial^{m+n}\mathbf{f}}{\partial\mathbf{y}^m\partial\mathbf{z}^n}$ que derivamos, una rama ascendente que se conectará para obtener uno de ellos con un árbol $[\tau, \tau]_{\mathbf{z}}$ y para obtener el otro con un árbol $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$. En la Figura 2.3 pueden verse ejemplos de la aplicación de esta regla donde aparece el árbol con raíz inicial marcado en azul.

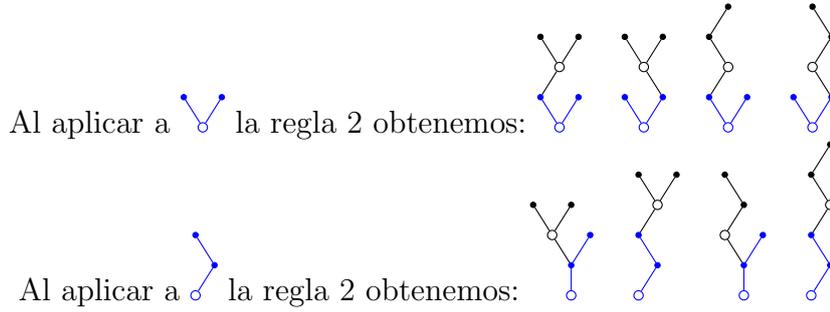


Figura 2.3: Ejemplo de aplicación de la regla 2

3. En el caso de que en la aplicación de la regla de Leibniz se derive un término $(-\mathbf{g}_y \mathbf{f}_z)^{-1}$ de $F(\sigma)$ con respecto a t , término que estará asociado a un nodo blanco de σ , aparecen los cuatro términos de (2.20)

- 1 : $(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy} (\mathbf{f}_z (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, \mathbf{f})$,
- 2 : $(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zy} ((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, \mathbf{f})$,
- 3 : $(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zz} ((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_{yy} (\mathbf{f}, \mathbf{f}))$,
- 4 : $(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_{zz} ((-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}, (-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y \mathbf{f}_y \mathbf{f})$,

donde $(-\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{u}$ será la diferencial elemental asociada al árbol w cuya raíz es el nodo blanco asociado al término $(-\mathbf{g}_y \mathbf{f}_z)^{-1}$ que se está derivando. En la parte izquierda de la Figura 2.4 puede verse un árbol σ donde el nodo blanco representa el nodo asociado al término $(-\mathbf{g}_y \mathbf{f}_z)^{-1}$ que se está derivando y v son todos los nodos y ramas que quedan al quitar de σ las ramas que parten del nodo blanco destacado (y que forman, por tanto, parte de w). En la parte derecha de la Figura 2.4 puede verse que a partir del árbol σ se obtienen cuatro nuevos árboles, todos de orden una unidad mayor, que podemos describir de la forma siguiente: en primer lugar, se duplica el nodo blanco asociado al término $(-\mathbf{g}_y \mathbf{f}_z)^{-1}$ y los dos nodos blancos resultantes se unen entre sí mediante un nuevo nodo negro intermedio.

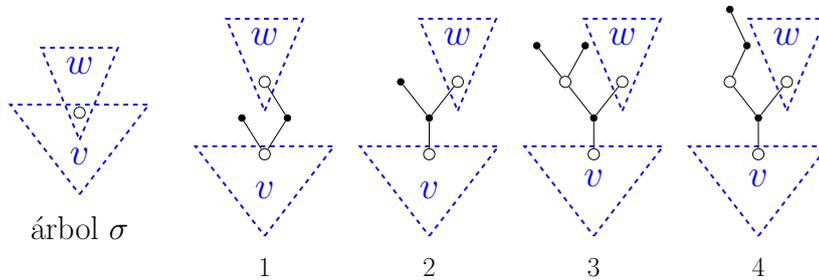


Figura 2.4: Esquema de la regla 3

Obtenemos uno de los cuatro árboles añadiendo al nodo blanco inferior una rama ascendente conectada a un árbol τ (término 1). Los otros tres árboles los obtenemos añadiendo al nuevo nodo negro una rama ascendente a la que conectamos un árbol τ (término 2), $[\tau, \tau]_{\mathbf{z}}$ (término 3) y $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ (término 4), respectivamente. En la Figura 2.5 pueden verse ejemplos de la aplicación de esta regla donde aparece el árbol con raíz inicial marcado en azul.

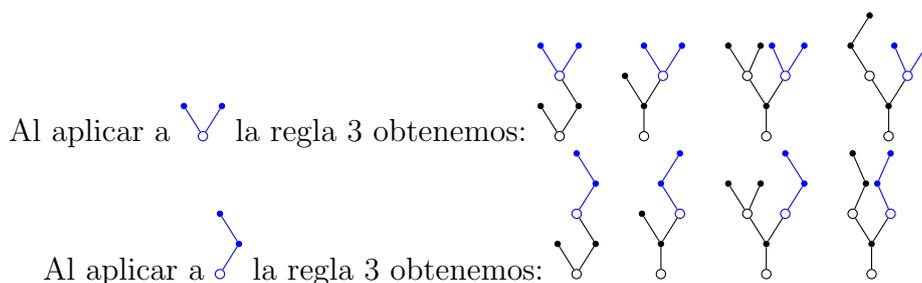


Figura 2.5: Ejemplo de aplicación de la regla 3

Como puede verse en la Figura 2.2 algunos de los árboles pueden repetirse al derivar una misma diferencial elemental o al diferenciar varias que corresponden a la misma derivada de \mathbf{z} o \mathbf{y} . Ya habíamos constatado este hecho en (2.21), donde algunos términos aparecen acompañados de coeficientes enteros distintos de la unidad.

Para distinguir los árboles que aparecen más de una vez y saber cuantas veces aparece cada diferencial elemental en las derivadas de \mathbf{z} e \mathbf{y} indicaremos el *orden de generación* de los vértices negros añadiendo a los árboles de orden 1, a saber τ , $[\tau, \tau]_{\mathbf{z}}$, $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$, una etiqueta 1 en uno de sus vértices negros de modo que de cada vértice blanco salga una rama cuyo vértice negro está sin etiquetar.

Cada vez que “derivamos” un árbol de acuerdo con las normas descritas anteriormente, pondremos a los árboles de orden 1 que se añaden una nueva etiqueta de modo que el criterio “de cada vértice blanco sale una rama cuyo vértice negro está sin etiquetar” se conserve. El etiquetado que se obtiene es evidentemente creciente a lo largo de cada rama.

Definición 2.4. Un árbol $t \in T_{\mathbf{y}}$ (o $u \in T_{\mathbf{z}}$), junto a un etiquetado monótono que repartirá $\rho(t)$ (o $\rho(u)$) etiquetas entre los vértices negros del árbol de forma que “de cada vértice blanco salga una rama cuyo vértice negro esté sin etiquetar” recibe el nombre de **árbol con etiquetado monótono**.

Las definiciones 2.2 (orden de un árbol) y 2.3 (diferenciales elementales) se extienden de forma natural a los árboles con etiquetado monótono.

Lema 2.1 (Generación de todos los árboles por derivación). *Todos los árboles de $T = T_{\mathbf{y}} \cup T_{\mathbf{z}}$ proceden de “derivar” repetidamente uno de los tres árboles siguientes: τ , $[\tau, \tau]_{\mathbf{z}}$ y $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$.*

Además si el árbol pertenece a $T_{\mathbf{y}}$ proviene necesariamente del árbol τ , y si el árbol pertenece a $T_{\mathbf{z}}$ proviene de los árboles $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$.

Demostración. Demostraremos el resultado aplicando un proceso de inducción completa al enunciado siguiente: los árboles con raíz negra de orden $\leq r$ provienen de “derivar” τ y los árboles con raíz blanca de orden $\leq r - 1$ proceden de “derivar” $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$, para $r \geq 2$. Para $r = 2$, los árboles de $T_{\mathbf{y}}$ de orden 2, que aparecen en la Figura 2.6, proceden de aplicar las reglas 1 y 2 al árbol τ mientras que los árboles de orden 1 de $T_{\mathbf{z}}$ son simplemente los propios árboles $[\tau, \tau]_{\mathbf{z}}$ y $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ de los que partimos.

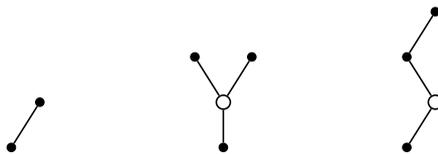


Figura 2.6: Árboles de orden 2 de $T_{\mathbf{y}}$

Supongamos (hipótesis de inducción) que los árboles con raíz negra de orden $\leq r$ provienen de “derivar” τ y los árboles con raíz blanca de orden $\leq r - 1$ proceden de “derivar” $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$, para un $r \geq 2$.

Sea $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}}$ con $\rho(t) = r + 1$. Si $(m, n) \neq (0, 1)$ tenemos que $\rho(t_i) \leq r$ y $\rho(u_i) \leq r - 1$ y, por hipótesis de inducción, sabemos como obtener por derivación los árboles t_i y u_i . Para obtener t por “derivación” de τ bastará “derivar” m veces la raíz aplicando la regla 1 y, a continuación, aplicar a cada uno de los m nodos negros generados los procesos de derivación necesarios (que conocemos por hipótesis de inducción) para obtener los árboles t_1, \dots, t_m , respectivamente, y “derivar” n veces la raíz aplicando la regla 2 en la versión que añade a la rama ascendente el árbol $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ según que u_i se obtenga a partir de uno u otro y, a continuación, aplicar a cada uno de los n nuevos árboles generados los procesos de derivación necesarios (que conocemos por hipótesis de inducción) para obtener los árboles u_1, \dots, u_n .

Sea $u = [t_1, \dots, t_m]_{\mathbf{z}} \in T_{\mathbf{z}}$ de orden $\rho(u) = r$. Si $m \geq 2$, tenemos que $\rho(t_i) \leq r$ y, por hipótesis de inducción, sabemos como obtener por derivación los árboles t_i . Para obtener u por “derivación” de $[\tau, \tau]_{\mathbf{z}}$ bastará “derivar” m veces la raíz aplicando la regla 1 y, a continuación, aplicar a cada uno de los m nodos

negros generados los procesos de derivación necesarios (que conocemos por hipótesis de inducción) para obtener los árboles t_1, \dots, t_m , respectivamente. En caso contrario, $m = 1$, $u = [t_1]_{\mathbf{z}}$ y $\rho(t_1) = r + 1$. Además, por la Definición 2.1, sabemos que $t_1 = [\tilde{t}_1, \dots, \tilde{t}_p, \tilde{u}_1, \dots, \tilde{u}_q]_{\mathbf{y}}$ con $(p, q) \neq (0, 1)$, luego, por lo visto en el párrafo anterior, sabemos como obtener por derivación t_1 . Si $p \geq 1$, obtenemos u a partir de $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ realizando en el nodo negro que sigue a la raíz el proceso explicado para t en el párrafo anterior aplicado a t_1 con la salvedad de que derivaremos una vez menos aplicando la regla 1 pues utilizaremos la preexistente rama ascendente finalizada en un nodo negro para la generación de \tilde{t}_1 . Si $p = 0$ y por tanto $q \geq 2$, obtenemos u a partir del árbol del que obtuviéramos \tilde{u}_1 aplicando la regla 3 en la versión 3 de la Figura 2.4 si \tilde{u}_2 se obtiene de $[\tau, \tau]_{\mathbf{z}}$ o en la versión 4 de la misma figura si \tilde{u}_2 se obtiene de $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ y realizando a continuación en el nodo negro que sigue a la raíz el proceso explicado para t en el párrafo anterior aplicado a t_1 con la salvedad de que derivaremos dos veces menos aplicando la regla 2 pues, para la generación de \tilde{u}_1 y \tilde{u}_2 utilizaremos los árboles $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ situados al final de las dos ramas que parten del nodo negro que aparecen tras la aplicación de la regla 3.

Veamos ahora el caso en que $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}}$ con $\rho(t) = r + 1$ y $(m, n) = (0, 1)$, es decir, $t = [u_1]_{\mathbf{y}}$. En este caso tenemos $\rho(u_1) = r$ y, por lo visto en el párrafo anterior, sabemos como obtener por derivación el árbol u_1 . Para obtener t por “derivación” de τ bastará “derivar” una vez la raíz aplicando la regla 2 en la versión que añade a la rama ascendente el árbol $[\tau, \tau]_{\mathbf{z}}$ o $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ según que u_1 se obtenga a partir de uno u otro y, a continuación, aplicar al nuevo árbol generado el proceso de derivación necesario que conocemos por el párrafo anterior para obtener u_1 .

Esto finaliza la prueba por inducción completa. \square

Nota. Los tres árboles de partida τ , $[\tau, \tau]_{\mathbf{z}}$ y $[[\tau]_{\mathbf{y}}]_{\mathbf{z}}$ son los únicos árboles de orden 1 y además corresponden con las diferenciales elementales de \mathbf{y}' y \mathbf{z}' como se puede ver en la Figura 2.1.

Nota. Obsérvese que los distintos procedimientos de construcción que figuran en la demostración anterior no son los únicos posibles y que, en la mayor parte de los casos, los árboles se pueden obtener de más de una manera.

Usando el lema precedente y teniendo en cuenta que cada derivación aumenta el orden de los árboles en una unidad, podemos escribir las derivadas de la solución exacta de la siguiente forma:

Teorema 2.3. *Para la solución exacta (\mathbf{y}, \mathbf{z}) de (2.1) tenemos que:*

$$\begin{aligned}\mathbf{y}^{(q)}(t_0) &= \sum_{t \in T_{\mathbf{y}}, \rho(t)=q} \alpha(t) F(t)(\mathbf{y}_0, \mathbf{z}_0), \\ \mathbf{z}^{(q)}(t_0) &= \sum_{u \in T_{\mathbf{y}}, \rho(u)=q} \alpha(u) F(u)(\mathbf{y}_0, \mathbf{z}_0),\end{aligned}$$

donde $\alpha(t)$ y $\alpha(u)$ son coeficientes enteros.

Nota. Los coeficientes $\alpha(t)$ y $\alpha(u)$ son en realidad el número de etiquetados monótonos distintos de cada árbol [5, p. 509].

2.3.2. Desarrollo de Taylor de la solución numérica generada con métodos RK semi-implícitos

En esta subsección trataremos de obtener una expresión similar a la del Teorema 2.3 para la solución numérica obtenida con métodos Runge-Kutta semi-implícitos de la forma (2.6). Adoptamos para ello la formulación (2.11) con la notación (2.10) y (2.9).

Considerando valores iniciales $\mathbf{y}_0, \mathbf{z}_0$ consistentes, si aplicamos un paso del método Runge-Kutta semi-implícito en su formulación (2.11) tenemos (con $\tilde{\mathbf{Y}}_0 = \mathbf{y}_0$)

$$\tilde{\mathbf{Y}}_i = \mathbf{y}_0 + h \sum_{j=1}^i \tilde{a}_{ij} \tilde{\mathbf{k}}_j, \quad i = 1, \dots, s, \quad (2.22a)$$

$$\mathbf{0} = \mathbf{g}(\tilde{\mathbf{Y}}_i), \quad i = 1, \dots, s. \quad (2.22b)$$

donde

$$\tilde{\mathbf{k}}_i = \mathbf{f}(\tilde{\mathbf{Y}}_{i-1}, \mathbf{Z}_i), \quad i = 1, \dots, s. \quad (2.23)$$

Por el Teorema 2.1 sabemos que dicha solución existe y es única al ser $\tilde{\mathbf{A}} = (\tilde{a}_{ij})$ definida en (2.10) invertible. Consideramos en esta subsección $\tilde{\mathbf{Y}}_i, \mathbf{Z}_i$ y $\tilde{\mathbf{k}}_i$ como funciones de h y calculamos a continuación sus derivadas en $h = 0$. Derivando (2.22a) respecto de h tenemos

$$\frac{d\tilde{\mathbf{Y}}_i}{dh} = \sum_{j=1}^i \tilde{a}_{ij} \tilde{\mathbf{k}}_j + h \sum_{j=1}^i \tilde{a}_{ij} \frac{d\tilde{\mathbf{k}}_j}{dh}, \quad (2.24)$$

donde se ha omitido la dependencia de h en $\tilde{\mathbf{Y}}_i, \mathbf{Z}_i, \tilde{\mathbf{k}}_i$ y sus derivadas. De ahora en adelante se omitirá, generalmente, esta dependencia. Como $\tilde{\mathbf{Y}}_i(0) = \mathbf{y}_0$

y $\mathbf{Z}_i(0) = \mathbf{z}_0$, evaluando (2.24) en $h = 0$ tenemos

$$\frac{d\tilde{\mathbf{Y}}_i}{dh}(0) = \sum_{j=1}^i \tilde{a}_{ij} \mathbf{f}(\mathbf{y}_0, \mathbf{z}_0). \quad (2.25)$$

De hecho, se prueba fácilmente por inducción que para $q \geq 0$

$$\frac{d^{q+1}\tilde{\mathbf{Y}}_i}{dh^{q+1}} = (q+1) \sum_{j=1}^i \tilde{a}_{ij} \frac{d^q \tilde{\mathbf{k}}_j}{dh^q} + h \sum_{j=1}^i \tilde{a}_{ij} \frac{d^{q+1} \tilde{\mathbf{k}}_j}{dh^{q+1}} \quad (2.26)$$

y, por tanto,

$$\frac{d^{q+1}\tilde{\mathbf{Y}}_i}{dh^{q+1}}(0) = (q+1) \sum_{j=1}^i \tilde{a}_{ij} \frac{d^q \tilde{\mathbf{k}}_j}{dh^q}(0). \quad (2.27)$$

Derivando (2.23) respecto de h tenemos

$$\frac{d\tilde{\mathbf{k}}_i}{dh} = \mathbf{f}_y(\tilde{\mathbf{Y}}_{i-1}, \mathbf{Z}_i) \frac{d\tilde{\mathbf{Y}}_{i-1}}{dh} + \mathbf{f}_z(\tilde{\mathbf{Y}}_{i-1}, \mathbf{Z}_i) \frac{d\mathbf{Z}_i}{dh}. \quad (2.28)$$

Para calcular $\frac{d\mathbf{Z}_i}{dh}(0)$ primero derivamos dos veces respecto de h la ecuación (2.22b)

$$\begin{aligned} \mathbf{0} &= \mathbf{g}_y(\tilde{\mathbf{Y}}_i) \frac{d\tilde{\mathbf{Y}}_i}{dh}, \\ \mathbf{0} &= \mathbf{g}_{yy}(\tilde{\mathbf{Y}}_i) \left(\frac{d\tilde{\mathbf{Y}}_i}{dh}, \frac{d\tilde{\mathbf{Y}}_i}{dh} \right) + \mathbf{g}_y(\tilde{\mathbf{Y}}_i) \frac{d^2\tilde{\mathbf{Y}}_i}{dh^2}, \end{aligned} \quad (2.29)$$

a continuación sustituimos en (2.29) $\frac{d^2\tilde{\mathbf{Y}}_i}{dh^2}$ por su expresión según (2.26) y tras esto $\frac{d\tilde{\mathbf{k}}_i}{dh}$ por su expresión según (2.28). Obtenemos entonces al evaluar en $h = 0$,

$$\begin{aligned} \mathbf{0} &= \mathbf{g}_{yy} \left(\frac{d\tilde{\mathbf{Y}}_i}{dh}(0), \frac{d\tilde{\mathbf{Y}}_i}{dh}(0) \right) + 2 \sum_{j=1}^i \tilde{a}_{ij} \mathbf{g}_y \mathbf{f}_y \frac{d\tilde{\mathbf{Y}}_{i-1}}{dh}(0) \\ &\quad + 2 \sum_{j=1}^i \tilde{a}_{ij} \mathbf{g}_y \mathbf{f}_z \frac{d\mathbf{Z}_i}{dh}(0), \end{aligned} \quad (2.30)$$

donde de nuevo se han omitido los argumentos de \mathbf{f} , \mathbf{g} y sus derivadas. De ahora en adelante se omitirán los argumentos siempre que estos sean \mathbf{y}_0 o \mathbf{z}_0 .

Utilizando ahora (2.27) llegamos a

$$\begin{aligned} \mathbf{0} = \mathbf{g}_{\mathbf{y}\mathbf{y}} \left(\sum_{j=1}^i \tilde{a}_{ij} \mathbf{f}, \sum_{j=1}^i \tilde{a}_{ij} \mathbf{f} \right) + 2 \sum_{j=1}^i \tilde{a}_{ij} \mathbf{g}_{\mathbf{y}\mathbf{f}_y} \sum_{k=1}^i \tilde{a}_{i-1,k} \mathbf{f} \\ + 2 \sum_{j=1}^i \tilde{a}_{ij} \mathbf{g}_{\mathbf{y}\mathbf{f}_z} \frac{d\mathbf{Z}_i}{dh}(0). \end{aligned} \quad (2.31)$$

Si la matriz $\tilde{\mathcal{A}}$ es invertible y $\mathbf{g}_{\mathbf{y}\mathbf{f}_z}$ es invertible en un entorno de la solución, despejando $\frac{d\mathbf{Z}_i}{dh}(0)$ de (2.31) tenemos

$$\begin{aligned} \frac{d\mathbf{Z}_i}{dh}(0) = \frac{1}{2} \sum_{j,k,l} \tilde{w}_{ij} \tilde{a}_{jk} \tilde{a}_{jl} (-\mathbf{g}_{\mathbf{y}\mathbf{f}_z})^{-1} \mathbf{g}_{\mathbf{y}\mathbf{y}}(\mathbf{f}, \mathbf{f}) \\ + \sum_j \tilde{a}_{i-1,j} (-\mathbf{g}_{\mathbf{y}\mathbf{f}_z})^{-1} \mathbf{g}_{\mathbf{y}\mathbf{f}_y} \mathbf{f}, \end{aligned} \quad (2.32)$$

donde \tilde{w}_{ij} denota el elemento (i, j) de la inversa de la matriz de coeficientes del método Runge-Kutta $\tilde{\mathcal{A}}$.

Si evaluamos (2.28) en $h = 0$ y sustituimos $\frac{d\tilde{\mathbf{Y}}_i}{dh}(0)$ y $\frac{d\mathbf{Z}_i}{dh}(0)$ por sus expresiones según (2.25) y (2.32) tenemos

$$\begin{aligned} \frac{d\tilde{\mathbf{k}}_i}{dh}(0) = \sum_{j=1}^i \tilde{a}_{i-1,j} \mathbf{f}_y \mathbf{f} + \frac{1}{2} \sum_{j,k,l} \tilde{w}_{ij} \tilde{a}_{jk} \tilde{a}_{jl} \mathbf{f}_z (-\mathbf{g}_{\mathbf{y}\mathbf{f}_z})^{-1} \mathbf{g}_{\mathbf{y}\mathbf{y}}(\mathbf{f}, \mathbf{f}) \\ + \sum_j \tilde{a}_{i-1,j} \mathbf{f}_z (-\mathbf{g}_{\mathbf{y}\mathbf{f}_z})^{-1} \mathbf{g}_{\mathbf{y}\mathbf{f}_y} \mathbf{f}. \end{aligned} \quad (2.33)$$

Derivando repetidamente (2.28) y (2.29) respecto de h y utilizando (2.27) obtenemos las expresiones de las derivadas sucesivas respecto de h de \mathbf{Z}_i y $\tilde{\mathbf{k}}_i$ evaluadas en $h = 0$ como combinación lineal de diferenciales elementales con coeficientes dependientes del método Runge-Kutta.

Los coeficientes que acompañan a cada diferencial elemental podrán asociarse al árbol asociado a la diferencial elemental. El coeficiente que acompaña a la diferencial elemental $F(\sigma)$ será el producto de $\gamma(\sigma)$, que será un número racional independiente del método Runge-Kutta, y $\phi_i(\sigma)$ que será una expresión dependiente de los coeficientes del método Runge-Kutta. Definimos a continuación $\gamma(\sigma)$ y $\phi_i(\sigma)$, que suelen recibir el nombre de *función de densidad* y *peso elemental* de σ , respectivamente.

Definición 2.5. El **peso elemental** $\phi_i(t)$ o $\phi_i(u)$ asociado a los árboles de T queda definido recursivamente de la siguiente forma:

- $\phi_i(\tau) = 1,$
- $\phi_i(t) = \sum_{j_1, \dots, j_m} a_{ij_1} \cdots a_{ij_m} \phi_{j_1}(t_1) \cdots \phi_{j_m}(t_m) \phi_i(u_1) \cdots \phi_i(u_n),$ si
 $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}},$
- $\phi_i(u) = \sum_{j, k_1, \dots, k_m} \tilde{w}_{ij} \tilde{a}_{jk_1} \cdots \tilde{a}_{jk_m} \phi_{k_1}(t_1) \cdots \phi_{k_m}(t_m),$ si
 $u = [t_1, \dots, t_m]_{\mathbf{z}} \in T_{\mathbf{z}}.$

Definición 2.6. La **función de densidad** $\gamma(t)$ o $\gamma(u)$ asociada a los árboles de T queda definida recursivamente de la siguiente forma:

- $\gamma(\tau) = 1,$
- $\gamma(t) = \rho(t) \gamma(t_1) \cdots \gamma(t_m) \gamma(u_1) \cdots \gamma(u_n),$ si
 $t = [t_1, \dots, t_m, u_1, \dots, u_n]_{\mathbf{y}} \in T_{\mathbf{y}},$
- $\gamma(u) = \frac{1}{\rho(u) + 1} \gamma(t_1) \cdots \gamma(t_m)$ si $u = [t_1, \dots, t_m]_{\mathbf{z}} \in T_{\mathbf{z}}.$

Teorema 2.4. Las derivadas de $\tilde{\mathbf{k}}_i$ y \mathbf{Z}_i satisfacen

$$\frac{d^q \tilde{\mathbf{k}}_i}{dh^q}(0) = \frac{1}{q+1} \sum_{t \in T_{\mathbf{y}}, \rho(t)=q+1} \alpha(t) \gamma(t) \phi_i(t) F(t)(\mathbf{y}_0, \mathbf{z}_0),$$

$$\frac{d^q \mathbf{Z}_i}{dh^q}(0) = \sum_{u \in T_{\mathbf{z}}, \rho(u)=q} \alpha(u) \gamma(u) \phi_i(u) F(u)(\mathbf{y}_0, \mathbf{z}_0),$$

donde los coeficientes enteros $\alpha(t)$ y $\alpha(u)$ son los mismos que los del Teorema 2.3.

Nota. Para obtener una demostración de este teorema basta adaptar la que se puede encontrar en [4, p. 61].

Utilizando los coeficientes que acabamos de definir, el siguiente teorema nos ofrece una expresión para la solución numérica similar a la del Teorema 2.3 para la solución exacta.

Teorema 2.5. Para la solución numérica $(\mathbf{y}_1, \mathbf{z}_1)$ de (2.22) calculada tras un paso de longitud h con el método Runge-Kutta semi-explicito (2.6) tenemos:

$$\frac{d^{q+1}\mathbf{y}_1}{dh^{q+1}}(0) = \sum_{t \in T_{\mathbf{y}}, \rho(t)=q+1} \alpha(t)\gamma(t) \left(\sum_{i=1}^s b_i \phi_i(t) \right) F(t)(\mathbf{y}_0, \mathbf{z}_0), \quad (2.34)$$

$$\frac{d^{q+1}\mathbf{z}_1}{dh^{q+1}}(0) = \sum_{u \in T_{\mathbf{z}}, \rho(u)=q} \alpha(u)\gamma(u)\phi_s(u)F(u)(\mathbf{y}_0, \mathbf{z}_0) \quad (2.35)$$

donde los coeficientes enteros $\alpha(t)$ y $\alpha(u)$ son los mismos que los del Teorema 2.3.

Demostración. Recordemos que la solución numérica de los métodos Runge-Kutta semi-explicitos viene dada por $\mathbf{y}_1 = \tilde{\mathbf{Y}}_s$ y $\mathbf{z}_1 = \mathbf{Z}_s$. Utilizando (2.27), tenemos que las derivadas en $h = 0$ de $\mathbf{y}_1 = \tilde{\mathbf{Y}}_s$ vienen dadas por

$$\frac{d^{q+1}\mathbf{y}_1}{dh^{q+1}}(0) = (q+1) \sum_{i=1}^s \tilde{a}_{sj} \frac{d^q \tilde{\mathbf{k}}_i}{dh^q}(0) = (q+1) \sum_{i=1}^s b_i \frac{d^q \tilde{\mathbf{k}}_i}{dh^q}(0). \quad (2.36)$$

Si sustituimos en (2.36) $\frac{d^q \tilde{\mathbf{k}}_i}{dh^q}(0)$ por su expresión según el Teorema 2.4 tenemos

$$\frac{d^{q+1}\mathbf{y}_1}{dh^{q+1}}(0) = \sum_{t \in T_{\mathbf{y}}, \rho(t)=q+1} \alpha(t)\gamma(t) \left(\sum_{i=1}^s b_i \phi_i(t) \right) F(t)(\mathbf{y}_0, \mathbf{z}_0).$$

Las derivadas en $h = 0$ de $\mathbf{z}_1 = \mathbf{Z}_s$ se obtienen directamente de Teorema 2.4. \square

2.3.3. Condiciones de orden

En esta subsección obtendremos finalmente las condiciones de orden de los métodos Runge-Kutta semi-explicitos utilizando los resultados obtenidos anteriormente en la sección. La solución numérica de los métodos Runge-Kutta semi-explicitos en su formulación (2.11) viene dada por $\mathbf{y}_1 = \tilde{\mathbf{Y}}_s$.

Teorema 2.6. *El error local $\delta\mathbf{y}_h(t_0) = \mathbf{y}_1 - \mathbf{y}(t_0 + h)$ verifica*

$$\delta\mathbf{y}_h(t_0) = O(h^r), \quad P(t_0)\delta\mathbf{y}_h(t_0) = O(h^{r+1}), \quad (2.37)$$

con $P(t_0) = I - \mathbf{f}_z(\mathbf{g}_y \mathbf{f}_z)^{-1} \mathbf{g}_y(\mathbf{y}_0, \mathbf{z}_0)$, si se verifican las condiciones

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)} \quad (2.38)$$

para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) \leq r-1$ y para aquellos de orden $\rho(t) = r$ que no son de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

Nota. Por el resultado de convergencia del Teorema (2.2) las condiciones de orden anteriores implican que el error global de la componente \mathbf{y} es $O(h^r)$.

Demostración. Supongamos que se verifican las condiciones

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)}$$

para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) \leq r-1$ y para aquellos de orden $\rho(t) = r$ que no son de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$. Entonces, para $q+1 \leq r-1$, la expresión (2.34) pasa a ser

$$\frac{d^{q+1} \mathbf{y}_1}{dh^{q+1}}(0) = \sum_{t \in T_{\mathbf{y}}, \rho(t)=q+1} \alpha(t) F(t)(\mathbf{y}_0, \mathbf{z}_0),$$

lo que coincide con la expresión de $\mathbf{y}^{(q+1)}(t_0)$ obtenida en el Teorema 2.3, luego queda probado que $\delta\mathbf{y}_h(t_0) = O(h^r)$.

Como $P(t_0)\mathbf{f}_z(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{0}$ tenemos que

$$P(t_0)F(t)(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{0} \quad \text{para } t = [u]_{\mathbf{y}} \text{ con } u \in T_{\mathbf{z}},$$

de manera que los árboles de orden r de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$ no impondrán condiciones sobre los coeficientes del método y, dado que (2.38) se verifica para todos los árboles de orden r que no son de esa forma, se cumple también la segunda relación: $P(t_0)\delta\mathbf{y}_h(t_0) = O(h^{r+1})$. \square

Por tanto, en virtud del Teorema 2.2 y del Teorema 2.6, si queremos orden de convergencia r bastará con exigir la condición (2.38) para todos los árboles de $T_{\mathbf{y}}$ de orden menor o igual que $r-1$ y para aquellos de orden r que no sean de la forma $[u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

De orden 1 sólo hay un árbol en $T_{\mathbf{y}}$: τ (que no es de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$). De las definiciones 2.5 (peso elemental) y 2.6 (función de densidad) tenemos

$$\phi_i(\tau) = 1 \quad \text{y} \quad \gamma(\tau) = 1.$$

Por tanto, la condición (2.38) es para el árbol τ

$$\sum_{i=1}^s b_i = 1, \quad (2.39)$$

que será la condición que habrá que exigir para que el método Runge-Kutta semi-implícito tenga orden de convergencia 1.

De orden 2 hay tres árboles en $T_{\mathbf{y}}$: $[\tau]_{\mathbf{y}}$, $[[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}$ y $[[[\tau]_{\mathbf{y}}]_{\mathbf{z}}]_{\mathbf{y}}$. De ellos, los dos últimos son de la forma $[u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$. Para el otro árbol, $[\tau]_{\mathbf{y}}$, tenemos

$$\begin{aligned} \phi_i([\tau]_{\mathbf{y}}) &= \sum_j a_{ij} \phi_j(\tau) = \sum_j a_{ij}, \\ \gamma([\tau]_{\mathbf{y}}) &= \rho([\tau]_{\mathbf{y}}) \gamma(\tau) = 2 \cdot 1 = 2, \end{aligned}$$

y, por tanto, la condición (2.38) es

$$\sum_{i,j} b_i a_{ij} = \frac{1}{2},$$

que, con la notación $c_i = \sum_{j=1}^s a_{ij}$ se reescribe

$$\sum_i b_i c_i = \frac{1}{2}. \quad (2.40)$$

El método Runge-Kutta semi-implícito tendrá orden 2 si se verifican las condiciones (2.39) y (2.40). Si queremos orden 3, además de estas dos condiciones será necesario exigir que se verifique la condición (2.38) para los dos árboles de orden 2 de la forma $[u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$: $[[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}$ y $[[[\tau]_{\mathbf{y}}]_{\mathbf{z}}]_{\mathbf{y}}$ y para aquellos de orden 3 que no tengan esa forma.

Para el árbol $[[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}$ tenemos

$$\begin{aligned} \phi_i([[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}) &= \phi_i([\tau, \tau]_{\mathbf{z}}) = \sum_{j,k,l} \tilde{w}_{ij} \tilde{a}_{jl} \tilde{a}_{kl} \phi_k(\tau) \phi_l(\tau) = \sum_{j,k,l} \tilde{w}_{ij} \tilde{a}_{jl} \tilde{a}_{kl}, \\ \gamma([[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}) &= \rho([[\tau, \tau]_{\mathbf{z}}]_{\mathbf{y}}) \frac{\gamma(\tau) \gamma(\tau)}{\rho([\tau, \tau]_{\mathbf{z}}) + 1} = 2 \cdot \frac{1}{2} = 1. \end{aligned}$$

Por tanto, la condición (2.38) es

$$\sum_{i,j,k,l} b_i \tilde{w}_{ij} \tilde{a}_{jl} \tilde{a}_{il} = 1,$$

que, teniendo en cuenta que $c_{i+1} = \sum_{j=1}^s \tilde{a}_{ij}$ con $i = 1, \dots, s-1$ y denotando $c_{s+1} = \sum_{j=1}^s \tilde{a}_{sj}$ es

$$\sum_{i,j,k,l} b_i \tilde{w}_{ij} c_{j+1}^2 = 1. \quad (2.41)$$

De forma similar, para el árbol $[[[\tau]_{\mathbf{y}}]_{\mathbf{z}}]_{\mathbf{y}}$ tenemos

$$\begin{aligned} \phi_i\left([\tau]_{\mathbf{y}}\right) &= \phi_i\left([\tau]_{\mathbf{y}}\right) = \sum_{j,k} \tilde{w}_{ij} \tilde{a}_{jk} \phi_k\left([\tau]_{\mathbf{y}}\right) = \sum_{j,k} \tilde{w}_{ij} \tilde{a}_{jk} c_k, \\ \gamma\left([\tau]_{\mathbf{y}}\right) &= \rho\left([\tau]_{\mathbf{y}}\right) \frac{\gamma([\tau]_{\mathbf{y}})}{\rho([\tau]_{\mathbf{y}}) + 1} = 2 \cdot \frac{2}{2} = 2, \end{aligned}$$

y, por tanto, la condición (2.38) es

$$\sum_{i,j,k} b_i \tilde{w}_{ij} \tilde{a}_{jk} c_k = \frac{1}{2}. \quad (2.42)$$

Como $(b_1, \dots, b_s) = (0, \dots, 1) \cdot \tilde{\mathcal{A}}$, la condición (2.41) se verifica automáticamente

$$\sum_{i,j,k,l} b_i \tilde{w}_{ij} c_{j+1}^2 = (0, \dots, 1) \cdot \tilde{\mathcal{A}} \cdot \tilde{\mathcal{A}}^{-1} \cdot (c_2^2, \dots, c_s^2, c_{s+1}^2)^T = c_{s+1}^2 = 1. \quad (2.43)$$

En la última igualdad se ha utilizado que $c_{s+1} = \sum_{j=1}^s \tilde{a}_{sj} = \sum_{i=1}^s b_i = 1$, pues ya hemos pedido que se verifique (2.39).

La condición (2.42) se verifica también automáticamente pues, utilizando de nuevo que $(b_1, \dots, b_s) = (0, \dots, 1) \cdot \tilde{\mathcal{A}}$, tenemos

$$\begin{aligned} \sum_{i,j,k} b_i \tilde{w}_{ij} \tilde{a}_{jk} c_k &= (0, \dots, 1) \cdot \tilde{\mathcal{A}} \cdot \tilde{\mathcal{A}}^{-1} \cdot \left(\sum_k \tilde{a}_{1k} c_k, \dots, \sum_k \tilde{a}_{sk} c_k \right)^T \\ &= \sum_k \tilde{a}_{sk} c_k = \sum_{k=1}^s b_k c_k = \frac{1}{2}, \end{aligned} \quad (2.44)$$

donde la última igualdad es consecuencia de que ya hemos exigido la condición (2.40).

n^0	árbol	orden	condición de orden
1	\cdot	1	$\sum b_i = 1$
2		2	$\sum b_i c_i = \frac{1}{2}$
3		3	$\sum b_i c_i^2 = \frac{1}{3}$
4		3	$\sum b_i a_{ij} c_j = \frac{1}{6}$
5		3	$\sum b_i c_i \tilde{w}_{ij} c_{j+1}^2 = \frac{2}{3}$
6		3	$\sum b_i c_i \tilde{w}_{ij} \tilde{a}_{jk} c_k = \frac{1}{3}$
7		3	$\sum b_i \tilde{w}_{ij} c_{j+1}^2 \tilde{w}_{ik} c_{k+1}^2 = \frac{4}{3}$
8		3	$\sum b_i \tilde{w}_{ij} c_{j+1}^2 \tilde{w}_{ik} \tilde{a}_{kl} c_l = \frac{2}{3}$
9		3	$\sum b_i \tilde{w}_{ij} \tilde{a}_{jk} c_k \tilde{w}_{il} \tilde{a}_{lm} c_m = \frac{1}{3}$
10		3	$\sum b_i a_{ij} \tilde{w}_{jk} c_{k+1}^2 = \frac{1}{3}$
11		3	$\sum b_i a_{ij} \tilde{w}_{jk} \tilde{a}_{kl} c_l = \frac{1}{6}$

Tabla 2.1: Condiciones de orden 1, 2 y 3

Cálculos similares a los anteriores aplicados a los árboles de orden 3 que no son de la forma $[u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$ dan lugar a las condiciones 3–11 de la Tabla 2.1.

Por la forma en que se ha construido $\tilde{\mathcal{A}}$ a partir de \mathcal{A} se verifica

$$\mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} = \left[\begin{array}{ccc|c} 0 & \cdots & 0 & 0 \\ 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \end{array} \right]$$

y, en consecuencia,

$$(b_1, \dots, b_s) \cdot \mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} = (b_2, \dots, b_s, 0).$$

Veamos que esto implica que las dos últimas condiciones de la Tabla 2.1 se verifican automáticamente. Para la condición 10 de la Tabla 2.1, observamos que

$$\begin{aligned} \sum_{i,j,k} b_i a_{ij} \tilde{w}_{jk} c_{k+1}^2 &= (b_1, \dots, b_s) \cdot \mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} \cdot (c_2^2, \dots, c_s^2, c_{s+1}^2)^T \\ &= (b_2, \dots, b_s, 0) \cdot (c_2^2, \dots, c_s^2, c_{s+1}^2)^T = \sum_{i=2}^s b_i c_i^2 = \frac{1}{3}. \end{aligned}$$

En la última igualdad se ha utilizado que

$$\sum_{i=2}^s b_i c_i^2 = \sum_{i=1}^s b_i c_i^2 = \frac{1}{3}$$

puesto que $c_1 = 0$ y ya hemos pedido que se verifique la condición 3 de la Tabla 2.1.

La condición 11 de la Tabla 2.1 se verifica también automáticamente pues

$$\begin{aligned} \sum_{i,j,k,l} b_i a_{ij} \tilde{w}_{jk} \tilde{a}_{kl} c_l &= (b_1, \dots, b_s) \cdot \mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} \cdot \left(\sum_l \tilde{a}_{1l} c_l, \dots, \sum_l \tilde{a}_{sl} c_l \right)^T \\ &= (b_2, \dots, b_s, 0) \cdot \left(\sum_l a_{2l} c_l, \dots, \sum_l a_{s+1,l} c_l \right)^T \\ &= \sum_{i=2, j=1}^s b_i a_{ij} c_j = \frac{1}{6}. \end{aligned}$$

En la última igualdad se ha utilizado que

$$\sum_{i=2, j=1}^s b_i a_{ij} c_j = \sum_{i,j=1}^s b_i a_{ij} c_j = \frac{1}{6}$$

pues $a_{1j} = 0$ para $j = 1, \dots, s$ y ya hemos pedido que se verifique la condición 4 de la Tabla 2.1.

A continuación damos un resultado que generaliza lo sucedido en (2.43) y (2.44) a todo árbol de la forma $[u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$ y otro resultado que generaliza lo sucedido con las condiciones 10–11 de la Tabla 2.1 a todo árbol de la forma $[[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$. En la Figura 2.7 puede verse la estructura de los árboles de la forma $[u]_{\mathbf{y}}$ y $[[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

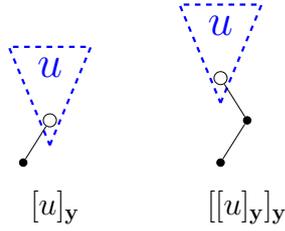


Figura 2.7: Estructura de los árboles

Lema 2.2. *Si se verifican las condiciones*

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)}$$

para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) \leq r-1$ y para aquellos de orden $\rho(t) = r$ que no son de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, entonces se verifica para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) = r$.

Demostración. Sea $t \in T_{\mathbf{y}}$ de orden $\rho(t) = r$. Obviamente si $t \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, la condición $\sum_{i=1}^s b_i \phi_i(t) = 1/\gamma(t)$ se verifica por hipótesis.

En caso contrario, si existe $\mu \in T_{\mathbf{z}}$ tal que $t = [\mu]_{\mathbf{y}}$, por la Definición 2.1 sabemos que existen $m \geq 1$ y t_1, \dots, t_m tales que $\mu = [t_1, \dots, t_m]_{\mathbf{z}}$; además, en caso de que $m = 1$, $t_1 \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

Por la Definición 2.2 (orden de un árbol) tenemos que

$$\rho(t_1) + \dots + \rho(t_m) = \rho(\mu) + 1 = \rho(t) = r$$

y por tanto o bien $\rho(t_i) < r$ para todo $i = 1, \dots, m$, o bien $m = 1$ y $\rho(t_1) = r$, lo que implica que $t_1 \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$. En ambos casos tenemos por hipótesis que para todo $i = 1, \dots, m$

$$\sum_{i=1}^s b_i \phi_i(t_i) = \frac{1}{\gamma(t_i)}. \quad (2.45)$$

Por las definiciones 2.5 (peso elemental) y 2.6 (función de densidad) tenemos

$$\begin{aligned} \phi_i(t) &= \phi_i([u]_{\mathbf{y}}) = \phi_i(\mu) = \sum_{j, k_1, \dots, k_m} \tilde{w}_{ij} \tilde{a}_{jk_1} \cdots \tilde{a}_{jk_m} \phi_{k_1}(t_1) \cdots \phi_{k_m}(t_m), \\ \gamma(t) &= \rho(t) \gamma(\mu) = \rho(t) \frac{1}{\rho(\mu) + 1} \gamma(t_1) \cdots \gamma(t_m) = \gamma(t_1) \cdots \gamma(t_m), \end{aligned} \quad (2.46)$$

pues $\rho(t) = \rho(\mu) + 1$. Como además $(b_1, \dots, b_s) = (0, \dots, 1) \cdot \tilde{\mathcal{A}}$, tenemos

$$\begin{aligned} \sum_i b_i \phi_i(t) &= \sum_{i,j,k_1,\dots,k_m} b_i \tilde{w}_{ij} \tilde{a}_{jk_1} \cdots \tilde{a}_{jk_m} \phi_{k_1}(t_1) \cdots \phi_{k_m}(t_m) \\ &= (0, \dots, 1) \cdot \tilde{\mathcal{A}} \cdot \tilde{\mathcal{A}}^{-1}(\xi_1, \dots, \xi_s)^T = \xi_s, \end{aligned}$$

donde

$$\xi_j = \sum_{k_1,\dots,k_m} \tilde{a}_{jk_1} \cdots \tilde{a}_{jk_m} \phi_{k_1}(t_1) \cdots \phi_{k_m}(t_m) = \prod_{q=1}^m \left(\sum_{k=1}^s \tilde{a}_{jk} \phi_k(t_q) \right).$$

Evaluando la expresión anterior en $j = s$ y utilizando primero (2.45) y luego (2.46) tenemos

$$\xi_s = \prod_{q=1}^m \left(\sum_{k=1}^s b_k \phi_k(t_q) \right) = \prod_{q=1}^m \left(\frac{1}{\gamma(t_q)} \right) = \frac{1}{\gamma(t)}$$

y por tanto

$$\sum_i b_i \phi_i(t) = \frac{1}{\gamma(t)}.$$

□

Corolario 2.1. *Si se verifican las condiciones*

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)}$$

para aquellos árboles de orden $\rho(t) \leq r$ que no son de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, entonces la misma condición se verifica para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) \leq r$.

Demostración. Basta aplicar inducción sobre r y utilizar el Lema 2.2 □

Lema 2.3. *Si se verifican las condiciones*

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)}$$

para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) = r$ que no sea de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, ni de la forma $t = [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, entonces se verifica para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) = r$ de la forma $t = [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

Demostración. Sea $t \in T_{\mathbf{y}}$ de orden $\rho(t) = r$ de la forma $t = [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, es decir, existe $\mu \in T_{\mathbf{z}}$ tal que $t = [[\mu]_{\mathbf{y}}]_{\mathbf{y}}$. Por la Definición 2.1 sabemos que existen $m \geq 1$ y t_1, \dots, t_m tales que $\mu = [t_1, \dots, t_m]_{\mathbf{z}}$ y además, en caso de que $m = 1$, $t_1 \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

Por la Definición 2.2 (orden de un árbol) tenemos que

$$\rho(t_1) + \dots + \rho(t_m) = \rho(\mu) + 1 = \rho([\mu]_{\mathbf{y}}) = \rho(t) - 1 = r - 1$$

y denotando $\theta = [t_1, \dots, t_m]_{\mathbf{y}}$ tenemos

$$\rho(\theta) = \rho(t_1) + \dots + \rho(t_m) + 1 = \rho(t) = r.$$

Como además $\theta \neq [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$ (ya que si $m = 1$ tenemos que $t_1 \neq [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$) y $\theta \neq [u]_{\mathbf{y}}$ tenemos por hipótesis que

$$\sum_{i=1}^s b_i \phi_i(\theta) = \frac{1}{\gamma(\theta)}. \quad (2.47)$$

Por la Definición 2.5 (peso elemental) tenemos

$$\begin{aligned} \phi_i(t) &= \phi_i([[\mu]_{\mathbf{y}}]_{\mathbf{y}}) = \sum_j a_{ij} \phi_j([\mu]_{\mathbf{y}}) = \sum_j a_{ij} \phi_j(\mu) \\ &= \sum_{j,k,l_1,\dots,l_m} a_{ij} \tilde{w}_{jk} \tilde{a}_{kl_1} \cdots \tilde{a}_{kl_m} \phi_{l_1}(t_1) \cdots \phi_{l_m}(t_m) \\ &= \sum_{j,k} a_{ij} \tilde{w}_{jk} \xi_k, \end{aligned}$$

donde

$$\xi_j = \sum_{l_1,\dots,l_m} \tilde{a}_{jl_1} \cdots \tilde{a}_{jl_m} \phi_{l_1}(t_1) \cdots \phi_{l_m}(t_m),$$

y, por la Definición 2.6 (función de densidad), tenemos

$$\begin{aligned} \gamma(t) &= \rho(t) \rho([\mu]_{\mathbf{y}}) \gamma(\mu) = \rho(t) \rho([\mu]_{\mathbf{y}}) \frac{1}{\rho(\mu) + 1} \gamma(t_1) \cdots \gamma(t_m) \\ &= \rho(t) \gamma(t_1) \cdots \gamma(t_m) = \rho(\theta) \gamma(t_1) \cdots \gamma(t_m) = \gamma(\theta), \end{aligned} \quad (2.48)$$

pues $\rho([\mu]_{\mathbf{y}}) = \rho(\mu) + 1$. Como además $(b_1, \dots, b_s) \cdot \mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} = (b_2, \dots, b_s, 0)$

tenemos

$$\begin{aligned} \sum_i b_i \phi_i(t) &= \sum_{i,j,k} b_i a_{ij} \tilde{w}_{jk} \xi_k \\ &= (b_1, \dots, b_s) \cdot \mathcal{A} \cdot \tilde{\mathcal{A}}^{-1} \cdot (\xi_1, \dots, \xi_s)^T \\ &= (b_2, \dots, b_s, 0) \cdot (\xi_1, \dots, \xi_s)^T = \sum_{k=2}^s b_k \xi_{k-1}, \end{aligned}$$

y dado que para $j = 1, \dots, s-1$, $\xi_j = \phi_{j+1}(\theta)$ y $\phi_1(\theta) = 0$, pues $a_{1l} = 0$ para $l = 1, \dots, s$, tenemos

$$\sum_i b_i \phi_i(t) = \sum_{k=2}^s b_k \xi_{k-1} = \sum_{k=2}^s b_k \phi_k(\theta) = \sum_{k=1}^s b_k \phi_k(\theta)$$

Utilizando ahora (2.47) y (2.48) concluimos

$$\sum_i b_i \phi_i(t) = \sum_{k=1}^s b_k \phi_k(\theta) = \frac{1}{\gamma(\theta)} = \frac{1}{\gamma(t)}.$$

□

El resultado que exponemos a continuación no es más que una síntesis del Corolario 2.1 y del Lema 2.3.

Corolario 2.2. *Si se verifican las condiciones*

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)}$$

para aquellos árboles de orden $\rho(t) \leq r$ que no sean de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, ni de la forma $t = [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, entonces la misma condición se verifica para todo árbol $t \in T_{\mathbf{y}}$ de orden $\rho(t) \leq r$.

Tras lo que acabamos de ver, el Teorema 2.6 se escribe ahora como

Teorema 2.7. *El error local $\delta \mathbf{y}_h(t_0) = \mathbf{y}_1 - \mathbf{y}(t_0 + h)$ viene dado por*

$$\delta \mathbf{y}_h(t_0) = O(h^{r+1}) \quad (2.49)$$

si se verifican las condiciones

$$\sum_{i=1}^s b_i \phi_i(t) = \frac{1}{\gamma(t)} \quad (2.50)$$

para aquellos árboles de orden $\rho(t) \leq r$ que no sean de la forma $t = [u]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$, ni de la forma $t = [[u]_{\mathbf{y}}]_{\mathbf{y}}$ con $u \in T_{\mathbf{z}}$.

Nota. De nuevo, por el resultado de convergencia del Teorema 2.2, las condiciones de orden anteriores implican que el método Runge-Kutta es convergente de orden r .

Las condiciones de orden (2.50) se pueden obtener directamente a partir de la estructura de los árboles utilizando el siguiente algoritmo, que supone una alternativa a la Definición 2.5.

Algoritmo (Formación de las condiciones de orden para un árbol dado t). Añadimos a cada nodo del árbol un índice de sumación i, j, k, \dots . El primer miembro de la condición de orden es el sumatorio sobre todos los índices de un producto con factores

$$\begin{array}{ll} b_i & \text{si } i \text{ es el índice de la raíz,} \\ a_{ij} & \text{si el nodo negro } j \text{ está justo encima del nodo negro } i, \\ \tilde{a}_{ij} = a_{i+1,j} & \text{si el nodo negro } j \text{ está justo encima del nodo blanco } i, \\ \tilde{w}_{ij} & \text{si el nodo blanco } j \text{ está justo encima del nodo negro } i. \end{array}$$

donde i y j representan dos índices cualesquiera. El segundo miembro de la condición de orden es $\frac{1}{\gamma(t)}$, y se calcula haciendo uso de la Definición 2.6.

Si estamos interesados en obtener una aproximación de la variable algebraica \mathbf{z} , se pueden establecer también condiciones de orden similares. Para ello hay que comparar los desarrollos de Taylor de la solución exacta $\mathbf{z}(t_0 + h)$ y de la solución numérica \mathbf{z}_1 que se han obtenido en el Teorema 2.3 y el Teorema 2.4, respectivamente. Se puede ver en [3] que, para que el error local de la variable algebraica verifique

$$\delta \mathbf{z}_h(t_0) = \mathbf{z}_1 - \mathbf{z}(t_0 + h) = O(h^{r+1}),$$

hay que pedir que $\delta \mathbf{y}_h(t_0) = O(h^{r+1})$ y que los desarrollos de Taylor de $\mathbf{z}(t_0 + h)$ y \mathbf{z}_1 coincidan hasta los términos $O(h^r)$. En [3, p. 544] se puede encontrar una tabla con las condiciones de orden adicionales que hay que exigir para que los desarrollos de Taylor coincidan hasta los términos de orden $O(h^3)$. Todas ellas corresponden a árboles en los que la raíz es un nodo blanco.

Capítulo 3

Una aplicación: sistemas mecánicos con restricciones

En este capítulo describimos una familia importante de problemas que pueden ser descritos mediante ecuaciones diferenciales algebraicas de índice 2 de la forma (2.1): los sistemas mecánicos con restricciones. Además, vamos a ver como, para estos problemas, los métodos Runge-Kutta semi-implícitos pueden implementarse de manera que sólo es necesario resolver sistemas lineales.

3.1. Descripción del problema

Las ecuaciones de Lagrange que describen el movimiento de un sistema multicuerpo con coordenadas generalizadas $\mathbf{q}(t) = (q_1(t), \dots, q_n(t))^T$ y velocidades generalizadas $\mathbf{v}(t) = \mathbf{q}'(t) = (q'_1(t), \dots, q'_n(t))^T$ son

$$\frac{d}{dt} \left(\frac{\partial L}{\partial q'_i} \right) - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots, n, \quad (3.1)$$

donde $L = T - U$ es la correspondiente función Lagrangiana, $U = U(\mathbf{q})$ es la energía potencial y $T = T(\mathbf{q}, \mathbf{v})$ es la energía cinética, que en numerosas aplicaciones viene dada por

$$T = \frac{1}{2} \mathbf{v}^T M(\mathbf{q}) \mathbf{v}. \quad (3.2)$$

Cuando el sistema mecánico está sujeto a m restricciones $\mathbf{g}_j(\mathbf{q}(t)) = 0$, $j = 1, \dots, m$, se considera la nueva función Lagrangiana

$$L = T - U - \lambda_1 \mathbf{g}_1(\mathbf{q}) - \dots - \lambda_m \mathbf{g}_m(\mathbf{q}), \quad (3.3)$$

en la que los *multiplicadores de Lagrange* $\lambda_1, \dots, \lambda_m$ se consideran como coordenadas generalizadas adicionales. El hecho de que el Lagrangiano (3.3) no dependa de las derivadas de los multiplicadores hace que las ecuaciones (3.1) para las nuevas coordenadas $\lambda_1, \dots, \lambda_m$ no sean otra cosa que las restricciones buscadas $\mathbf{g}_j(\mathbf{q}(t)) = 0$, $1 \leq j \leq m$.

Las ecuaciones (3.1) que describen el movimiento de un sistema mecánico con restricciones que no dependen de las velocidades (restricciones holonómicas) pueden escribirse como

$$\mathbf{q}' = \mathbf{v}, \quad (3.4a)$$

$$M(\mathbf{q})\mathbf{v}' = \mathbf{f}(\mathbf{q}, \mathbf{v}) - G^T(\mathbf{q})\boldsymbol{\lambda}, \quad (3.4b)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{q}), \quad (3.4c)$$

donde $\mathbf{g} = (g_1, \dots, g_m)^T$, $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)^T$, $G(\mathbf{q}) = \mathbf{g}_q(\mathbf{q})$, $M(\mathbf{q})$ es la matriz de masa generalizada, que se supone definida positiva, y $\mathbf{f}(\mathbf{q}, \mathbf{v})$ son las fuerzas aplicadas, i.e. las que no son fuerzas de restricción. Tanto en el desarrollo anterior como en lo que sigue se ha omitido la dependencia explícita de t en \mathbf{q} , \mathbf{v} y $\boldsymbol{\lambda}$. Un análisis más detallado de la obtención de estas ecuaciones del movimiento puede encontrarse en cualquier libro de mecánica clásica, por ejemplo en [1, Part II] o [9, Capítulo 7].

Si multiplicamos formalmente (3.4b) por $M^{-1}(\mathbf{q})$, el sistema (3.4) es de la forma (1.7) con $(\mathbf{q}, \mathbf{v}, \boldsymbol{\lambda})$ en los papeles de $(\mathbf{y}, \mathbf{z}, \mathbf{u})$. Para que el sistema sea de índice 3 debemos pedir que $G(\mathbf{q})M^{-1}(\mathbf{q})G^T(\mathbf{q})$ sea invertible en un entorno de la solución. Esta condición se verifica si las restricciones (3.4c) son independientes, es decir, si las filas de $G(\mathbf{q})$ son linealmente independientes.

Para obtener un sistema de la forma (2.1) derivamos (3.4c) respecto de t y obtenemos

$$\mathbf{0} = G(\mathbf{q})\mathbf{v}. \quad (3.5)$$

El sistema (3.4a), (3.4b), (3.5) es de la forma (2.1) con (\mathbf{q}, \mathbf{v}) en el papel de \mathbf{y} y $\boldsymbol{\lambda}$ en el papel de \mathbf{z} . Para que el sistema sea de índice 2 debemos pedir, al igual que en el caso de índice 3, que $G(\mathbf{q})M^{-1}(\mathbf{q})G^T(\mathbf{q})$ sea invertible en un entorno de la solución.

Si consideramos valores iniciales consistentes \mathbf{q}_0 , \mathbf{v}_0 , $\boldsymbol{\lambda}_0$ y aplicamos ahora un paso de longitud h con un método Runge-Kutta semi-implícito (2.6) al

sistema de índice 2

$$\mathbf{q}' = \mathbf{v}, \quad (3.6a)$$

$$M(\mathbf{q})\mathbf{v}' = \mathbf{f}(\mathbf{q}, \mathbf{v}) - G^T(\mathbf{q})\boldsymbol{\lambda}, \quad (3.6b)$$

$$\mathbf{0} = G(\mathbf{q})\mathbf{v}, \quad (3.6c)$$

con $G(\mathbf{q}) M^{-1}(\mathbf{q}) G^T(\mathbf{q})$ invertible en un entorno de la solución, obtenemos las etapas intermedias

$$\mathbf{Q}_1 = \mathbf{q}_0, \quad \mathbf{V}_1 = \mathbf{v}_0, \quad (3.7a)$$

$$\mathbf{Q}_i = \mathbf{q}_0 + h \sum_{j=1}^{i-1} a_{ij} \mathbf{V}_j, \quad i = 2, \dots, s, \quad (3.7b)$$

$$\mathbf{V}_i = \mathbf{v}_0 + h \sum_{j=1}^{i-1} a_{ij} \mathbf{V}'_j, \quad i = 2, \dots, s,$$

$$\mathbf{0} = G_i \mathbf{V}_i, \quad i = 1, \dots, s, \quad (3.7c)$$

con

$$M_i \mathbf{V}'_i = \mathbf{f}_i - G_i^T \boldsymbol{\Lambda}_i, \quad 1 \leq i \leq s-1, \quad (3.7d)$$

donde se ha utilizado la notación simplificada $M_i = M(\mathbf{Q}_i)$, $G_i = G(\mathbf{Q}_i)$ y $\mathbf{f}_i = \mathbf{f}(\mathbf{Q}_i, \mathbf{V}_i)$. La solución numérica tras avanzar un paso con el método $\mathbf{q}_1, \mathbf{v}_1, \boldsymbol{\lambda}_1$ viene dada entonces por las ecuaciones

$$\mathbf{q}_1 = \mathbf{q}_0 + h \sum_{i=1}^s b_i \mathbf{V}_i, \quad (3.8a)$$

$$\mathbf{v}_1 = \mathbf{v}_0 + h \sum_{i=1}^s b_i \mathbf{V}'_i,$$

$$\mathbf{0} = G(\mathbf{q}_1)\mathbf{v}_1, \quad (3.8b)$$

con

$$M_s \mathbf{V}'_s = \mathbf{f}_s - G_s^T \boldsymbol{\Lambda}_s. \quad (3.8c)$$

Como los valores iniciales son consistentes, para $i = 1$ (3.7c) se verifica por hipótesis. Para los sucesivos pasos, (3.7c) también se satisfará automáticamente para $i = 1$ pues a $\mathbf{q}_n, \mathbf{v}_n$ se les habrá exigido en el paso anterior que $\mathbf{0} = G(\mathbf{q}_n)\mathbf{v}_n$.

Si sustituimos a continuación en (3.7c) las expresiones de \mathbf{Q}_2 y \mathbf{V}_2 dadas por (3.7b), tenemos junto a (3.7d) para $i = 1$ un sistema lineal en \mathbf{V}'_1 y $\mathbf{\Lambda}_1$ que puede ser resuelto. Una vez hemos calculado \mathbf{V}'_1 y $\mathbf{\Lambda}_1$, podemos obtener \mathbf{V}_2 explícitamente (\mathbf{Q}_2 ya se conocía).

Podemos repetir este proceso en las sucesivas etapas; en la etapa $(i+1)$ -ésima deberemos resolver el sistema lineal en \mathbf{V}'_i y $\mathbf{\Lambda}_i$

$$\begin{bmatrix} M_i & G_i^T \\ G_{i+1} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}'_i \\ \mathbf{\Lambda}_i \end{bmatrix} = \begin{bmatrix} \mathbf{f}_i \\ \mathbf{r}_i \end{bmatrix}, \quad (3.9)$$

donde

$$\mathbf{r}_i = \frac{-1}{ha_{i+1,i}} \mathbf{G}_{i+1} \left(\mathbf{v}_0 + h \sum_{j=1}^{i-1} a_{i+1,j} \mathbf{V}'_j \right).$$

El hecho de que $M(\mathbf{q})$ sea regular, $G(\mathbf{q})$ tenga rango máximo y la matriz $G(\mathbf{q})M^{-1}(\mathbf{q})G^T(\mathbf{q})$ sea invertible en un entorno de la solución, garantiza que la matriz de (3.9) es invertible y que el sistema lineal tiene solución única.

A continuación, obtenemos \mathbf{V}'_s y $\mathbf{\Lambda}_s$ resolviendo el sistema formado por (3.8c) y por la ecuación que se obtiene al sustituir en (3.8b) las expresiones de \mathbf{q}_1 y \mathbf{v}_1 dadas por (3.8a). El sistema es

$$\begin{bmatrix} M_s & G_s^T \\ G(\mathbf{q}_1) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}'_s \\ \mathbf{\Lambda}_s \end{bmatrix} = \begin{bmatrix} \mathbf{f}_s \\ \mathbf{r}_s \end{bmatrix}, \quad (3.10)$$

donde

$$\mathbf{r}_s = \frac{-1}{hb_s} G(\mathbf{q}_1) \left(\mathbf{v}_0 + h \sum_{i=1}^{s-1} b_i \mathbf{V}'_i \right).$$

Como en las etapas intermedias, antes de resolver el sistema ya conocemos el valor de \mathbf{q}_1 . Por tanto, para finalizar basta calcular de forma explícita \mathbf{v}_1 de (3.8a).

Como ya se explicó en la descripción de los métodos semi-implícitos, la solución numérica de la variable algebraica, que en este caso es $\mathbf{\lambda}_1$, no se puede determinar y por tanto para obtener una aproximación a $\mathbf{\lambda}(t+h)$ se considerarán métodos con $c_s = 1$ y tomaremos $\lambda_1 = \mathbf{\Lambda}_s$.

3.2. Algunos Ejemplos

En esta sección describiremos tres sistemas mecánicos con restricciones: el péndulo matemático, un brazo plano articulado de dos grados de libertad y un sistema de siete cuerpos conocido como “Andrews’ squeezing mechanism”.

3.2.1. Péndulo Matemático

El péndulo descrito en la Figura 1.1 en coordenadas cartesianas (x, y) junto a la restricción considerada en el Capítulo 1, $x^2 + y^2 - l^2 = 0$, tiene como función Lagrangiana

$$L = \frac{m}{2}((x')^2 + (y')^2) + mgy - \lambda(x^2 + y^2 - l^2),$$

donde, al igual que en el Capítulo 1, m es la masa puntual suspendida al final de una varilla sin masa de longitud l bajo la influencia de una gravedad g . Por tanto, las ecuaciones del movimiento (3.1) son en este caso

$$x' = u, \quad (3.11a)$$

$$y' = v, \quad (3.11b)$$

$$mu' = -2x\lambda, \quad (3.11c)$$

$$mv' = mg - 2y\lambda, \quad (3.11d)$$

$$0 = x^2 + y^2 - l^2. \quad (3.11e)$$

Para este problema, el significado físico de λ , que era la tensión de la varilla que mantiene la masa puntual a una distancia constante del origen, sigue siendo la misma, excepto que queda dividida por $2l$. Esto produce una pequeña diferencia en la presentación de las nuevas ecuaciones (3.11c)–(3.11d) frente a las ecuaciones (1.10) que obtuvimos mediante las leyes de Newton.

El sistema anterior, viene dado por la forma general (3.4) con $\mathbf{q} = [x, y]^T$, $\mathbf{v} = [u, v]^T$, $\boldsymbol{\lambda} = \lambda$,

$$M(\mathbf{q}) = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}, \quad \mathbf{f}(\mathbf{q}, \mathbf{v}) = \begin{bmatrix} 0 \\ mg \end{bmatrix},$$

$$G(\mathbf{q}) = [2x, 2y] \quad \text{y} \quad \mathbf{g}(\mathbf{q}) = x^2 + y^2 - l^2.$$

Nótese que

$$G(\mathbf{q}) M^{-1}(\mathbf{q}) G^T(\mathbf{q}) = \frac{2}{m}(x^2 + y^2)$$

es invertible en un entorno de la solución y, por tanto, el sistema (3.11) es un sistema de EDAs de índice 3.

Si, al igual que en el Capítulo 1, derivamos la restricción (3.11e) respecto de t obtenemos

$$0 = xu + yv$$

que, junto a las restantes ecuaciones de (3.11), configuran un sistema de EDAs de índice 2, que coincide con la formulación de (3.6) y a esta formulación es a

la que podemos aplicar los métodos Runge-Kutta semi-implícitos estudiados en el Capítulo 2.

En la simulación que realizamos utilizaremos los parámetros

$$m = 1kg, \quad l = 1m \quad y \quad g = 9.81m/s^2,$$

y los valores iniciales

$$\mathbf{q}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}(0) = \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \quad y \quad \lambda(0) = 22,905.$$

La elección de $\lambda(0)$ se ha realizado para partir de condiciones iniciales consistentes.

3.2.2. Brazo plano articulado de dos grados de libertad

Consideramos en esta subsección un brazo plano articulado de dos grados de libertad con un camino predefinido para su extremo libre. El brazo está formado por dos varillas rígidas de masas m_1 y m_2 y longitudes l_1 y l_2 . Un extremo de una de las varillas (que denominamos primera) está fijado al origen de coordenadas y el otro está conectado a la segunda varilla permitiendo rotaciones en el plano x - y . Las coordenadas generalizadas del sistema mecánico que utilizaremos serán $\mathbf{q} = [\theta_1, \theta_2]^T$, siendo θ_1 el ángulo que forma la primera varilla con el eje horizontal y θ_2 el ángulo que forma la segunda varilla con respecto a la primera. Todo lo descrito anteriormente puede observarse en el esquema que aparece en la Figura 3.1.

Las coordenadas cartesianas del enlace (la “articulación”) entre ambas varillas vienen dadas por

$$x_1 = l_1 \cos(\theta_1), \quad y_1 = l_1 \sin(\theta_1),$$

y las del extremo libre son

$$x_2 = x_1 + l_2 \cos(\theta_1 + \theta_2), \quad y_2 = y_1 + l_2 \sin(\theta_1 + \theta_2).$$

La restricción viene dada por el camino predefinido para el extremo libre, cuya segunda coordenada cartesiana fijamos en función del tiempo como

$$y_2(t) = \sin^2(t/2). \tag{3.12}$$

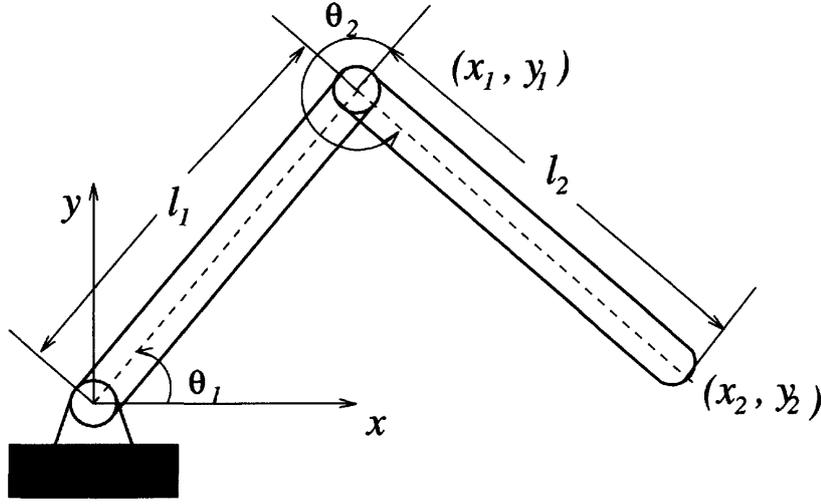


Figura 3.1: Brazo plano articulado de dos grados de libertad (de [2, p. 287]).

Como la restricción depende del tiempo, hemos de cambiar (3.4c) por $\mathbf{0} = \mathbf{g}(\mathbf{q}, t)$ y al derivar respecto de t obtenemos

$$\mathbf{0} = G(\mathbf{q}, t)\mathbf{v} + \frac{\partial \mathbf{g}}{\partial t}(\mathbf{q}, t)$$

en lugar de (3.6c). Además, de forma general, se deberá añadir la dependencia de $G = \mathbf{g}_{\mathbf{q}}$ respecto de t . En nuestro caso,

$$g(\mathbf{q}, t) = y_2(t) - \sin^2(t/2) = l_1 \sin \theta_1 + l_2 \sin(\theta_1 + \theta_2) - \sin^2(t/2), \quad (3.13a)$$

$$G(\mathbf{q}, t) = [l_1 \cos \theta_1 + l_2 \cos(\theta_1 + \theta_2), \quad l_2 \cos(\theta_1 + \theta_2)], \quad (3.13b)$$

luego G no depende del tiempo.

A la hora de implementar el método Runge-Kutta semi-implícito de la forma descrita por (3.7)–(3.8) se han de modificar los vectores \mathbf{r}_i , $1 \leq i \leq s$, que pasan a estar definidos por

$$\mathbf{r}_i = \frac{-1}{ha_{i+1,i}} \left(\mathbf{G}_{i+1} \left(\mathbf{v}_0 + h \sum_{j=1}^{i-1} a_{i+1,j} \mathbf{V}'_j \right) + \mathbf{g}_{ti} \right), \quad i = 1, \dots, s-1,$$

$$\mathbf{r}_s = \frac{-1}{hb_s} \left(G(\mathbf{q}_1) \left(\mathbf{v}_0 + h \sum_{i=1}^{s-1} b_i \mathbf{V}'_j \right) + \mathbf{g}_{ts} \right),$$

donde $\mathbf{g}_{ti} = \mathbf{g}_t(\mathbf{Q}_i, t_0 + c_i h)$, $1 \leq i \leq s$.

De manera análoga, con modificaciones mínimas, se podría considerar el caso en que tanto \mathbf{f} como M dependan de t .

Las ecuaciones de Lagrange que describen el movimiento del brazo con coordenadas generalizadas $\mathbf{q} = [\theta_1, \theta_2]^T$ vienen dadas por (3.4) y se obtiene un sistema de EDAs de índice 3 y por (3.6), si se quiere trabajar con EDAs de índice 2. Las funciones \mathbf{g} y G para este ejemplo han sido descritas en (3.13) y la matriz de masa y la fuerza vienen dadas por [2, p. 287]

$$M = \begin{bmatrix} \frac{m_1 l_1^2}{3} + m_2 \left(l_1^2 + \frac{l_2^2}{3} + l_1 l_2 c_2 \right) & m_2 \left(\frac{l_2^2}{3} + \frac{l_1 l_2 c_2}{2} \right) \\ m_2 \left(\frac{l_2^2}{3} + \frac{l_1 l_2 c_2}{2} \right) & \frac{m_2 l_2^2}{3} \end{bmatrix},$$

$$\mathbf{f} = \begin{bmatrix} -\frac{m_1 g l_1 c_1}{2} - m_2 g \left(l_1 c_1 + \frac{l_2 c_{12}}{2} \right) \\ -\frac{m_2 g l_2 c_{12}}{2} \end{bmatrix} + \begin{bmatrix} \frac{m_2 l_1 l_2 s_2 (2\theta_1' \theta_2' + (\theta_2')^2)}{2} \\ -\frac{m_2 l_1 l_2 s_2 (\theta_1')^2}{2} \end{bmatrix}.$$

donde $c_i = \cos \theta_i$, $s_i = \sin \theta_i$, $c_{12} = \cos(\theta_1 + \theta_2)$, $s_{12} = \sin(\theta_1 + \theta_2)$.

En la simulación que realizamos utilizaremos los parámetros

$$m_1 = m_2 = 36kg, \quad l_1 = l_2 = 1m, \quad g = 9.81m/s^2,$$

y los valores iniciales

$$\theta_1(0) = 70^\circ \approx 1.2217 \quad \theta_2(0) = -140^\circ \approx -2.4435 \quad \theta_1'(0) = \theta_2'(0) = 0.$$

Tanto los valores de los parámetros como las condiciones iniciales han sido tomados de [2, p. 287].

3.2.3. Andrews' Squeezing mechanism

En esta subsección vamos a describir el mecanismo plano de siete cuerpos conocido como "Andrews' Squeezing mechanism". Este mecanismo, que puede verse en la Figura 3.2, consiste en un conjunto de siete cuerpos rígidos, que denotamos por K_1, \dots, K_7 , enlazados entre ellos por articulaciones y conectados a un muelle, moviéndose todo ello en el plano $x-y$ en ausencia

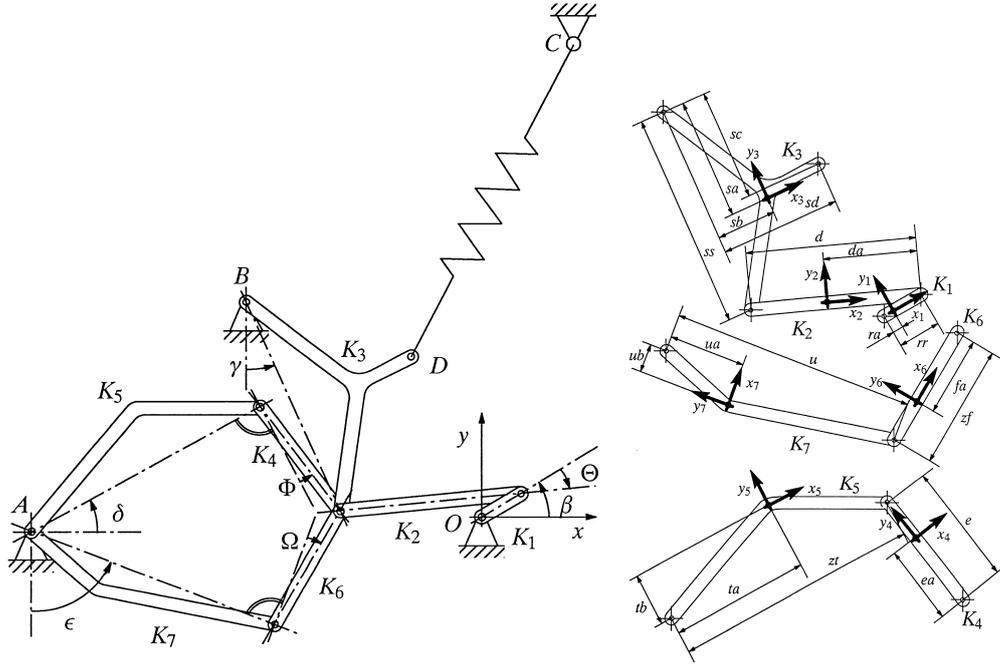


Figura 3.2: Andrews' squeezing mechanism (de [12, pp. 11 y 13])

de fricción. Este ejemplo mecánico cobró relevancia a través del trabajo de Giles [8] y Manning [10], que lo propusieron como problema test.

El mecanismo está fijado al origen de coordenadas, el punto O de la Figura 3.2, y a otros tres puntos, denotados por A , B y C en la misma figura. Denotamos por (x_a, y_a) las coordenadas del punto A y por (x_b, y_b) y (x_c, y_c) las de los puntos B y C , respectivamente. Por otra parte, en el origen de coordenadas O , hay un motor que transmite al mecanismo un par motor constante de magnitud mom . El muelle, que une el punto D del cuerpo K_3 con el punto fijo C , tiene constante de elasticidad c_0 y longitud en reposo ℓ_0 . Denotamos por C_i , $i = 1, \dots, 7$, el centro de gravedad del cuerpo K_i y además con origen en este centro se describe un par de ejes coordenados, que tienen como finalidad medir el ángulo total de rotación w_i del cuerpo K_i . La masa de cada cuerpo K_i , se denotará por m_i y su inercia por I_i .

Las coordenadas generalizadas del sistema mecánico que utilizaremos serán $\mathbf{q} = [\beta, \Theta, \gamma, \Phi, \delta, \Omega, \varepsilon]^T$, es decir, los ángulos

$$q_1 = \beta, \quad q_2 = \Theta, \quad q_3 = \gamma, \quad q_4 = \Phi, \quad q_5 = \delta, \quad q_6 = \Omega, \quad q_7 = \varepsilon. \quad (3.14)$$

La energía cinética del sistema multicuerpo viene dada por

$$T = \sum_{j=1}^7 m_j \frac{(x'_j)^2 + (y'_j)^2}{2} + \sum_{j=1}^7 I_j \frac{(\omega'_j)^2}{2}, \quad (3.15)$$

donde (x_j, y_j) son las coordenadas cartesianas de cada uno de los siete centros de gravedad C_j . Los valores de $x_j, y_j, (x'_j)^2 + (y'_j)^2$ y ω'_j vienen dados (utilizando la notación de la Figura 3.2 para las longitudes) en función de las coordenadas generalizadas por

$$\begin{aligned} C_1 : \quad x_1 &= ra \cdot \cos \beta, \\ y_1 &= ra \cdot \sin \beta, \\ (x'_1)^2 + (y'_1)^2 &= ra^2 \cdot (\beta')^2, \\ \omega'_1 &= \beta', \end{aligned}$$

$$\begin{aligned} C_2 : \quad x_2 &= rr \cdot \cos \beta - da \cdot \cos(\beta + \Theta), \\ y_2 &= rr \cdot \sin \beta - da \cdot \sin(\beta + \Theta), \\ (x'_2)^2 + (y'_2)^2 &= (rr^2 - 2 \cdot rr \cdot da \cdot \cos \Theta + da^2) \cdot (\beta')^2 \\ &\quad + 2 \cdot (-rr \cdot da \cdot \cos \Theta + da^2) \cdot \beta' \cdot \Theta' + da^2 \cdot (\Theta')^2, \\ w'_2 &= \beta' + \Theta', \end{aligned}$$

$$\begin{aligned} C_3 : \quad x_3 &= xb + sa \cdot \sin \gamma + sb \cdot \cos \gamma, \\ y_3 &= yb - sa \cdot \cos \gamma + sb \cdot \sin \gamma, \\ (x'_3)^2 + (y'_3)^2 &= (sa^2 + sb^2) \cdot (\gamma')^2, \\ w'_3 &= \gamma', \end{aligned}$$

$$\begin{aligned} C_4 : \quad x_4 &= xa + zt \cdot \cos \delta + (e - ea) \cdot \sin(\Phi + \delta), \\ y_4 &= ya + zt \cdot \sin \delta - (e - ea) \cdot \cos(\Phi + \delta), \\ (x'_4)^2 + (y'_4)^2 &= (e - ea)^2 \cdot (\Phi')^2 \\ &\quad + 2 \cdot ((e - ea)^2 + zt \cdot (e - ea) \cdot \sin \Phi) \cdot \Phi' \cdot \delta' \\ &\quad + (zt^2 + (e - ea)^2 + 2 \cdot zt \cdot (e - ea) \cdot \sin \Phi) \cdot (\delta')^2, \\ w'_4 &= \Phi' + \delta', \end{aligned}$$

$$\begin{aligned} C_5 : \quad x_5 &= xa + ta \cdot \cos \delta - tb \sin \delta, \\ y_5 &= ya + ta \cdot \sin \delta + tb \cos \delta, \\ (x'_5)^2 + (y'_5)^2 &= (ta^2 + tb^2) \cdot (\delta')^2, \\ w'_5 &= \delta', \end{aligned}$$

$$\begin{aligned}
C_6 : \quad x_6 &= xa + u \cdot \sin \varepsilon + (zf - fa) \cdot \cos(\Omega + \varepsilon), \\
y_6 &= ya - u \cdot \cos \varepsilon + (zf - fa) \cdot \sin(\Omega + \varepsilon), \\
(x'_6)^2 + (y'_6)^2 &= (zf - fa)^2 \cdot (\Omega')^2 \\
&\quad + 2 \cdot ((zf - fa)^2 - u \cdot (zf - fa) \cdot \sin \Omega) \cdot \Omega' \cdot \varepsilon' \\
&\quad + ((zf - fa)^2 - 2 \cdot u \cdot (zf - fa) \cdot \sin \Omega + u^2) \cdot (\varepsilon')^2, \\
w'_6 &= \Omega' + \varepsilon',
\end{aligned}$$

$$\begin{aligned}
C_7 : \quad x_7 &= xa + ua \cdot \sin \varepsilon - ub \cos \varepsilon, \\
y_7 &= ya - ua \cdot \cos \varepsilon - ub \sin \varepsilon, \\
(x'_7)^2 + (y'_7)^2 &= (ua^2 + ub^2) \cdot (\varepsilon')^2, \\
w'_7 &= \varepsilon',
\end{aligned}$$

La energía potencial del sistema mecánico se debe al motor situado en el origen de coordenadas y al muelle que une los puntos C y D . Está dada, aplicando la ley de Hooke, por

$$U = -mom \cdot \beta + c_0 \frac{(\ell - \ell_0)^2}{2}, \quad (3.16)$$

donde

$$\ell = \sqrt{(xd - xc)^2 + (yd - yc)^2}$$

es la distancia entre C y D , y

$$\begin{aligned}
xd &= xb + sc \cdot \sin \gamma + sd \cdot \cos \gamma \\
yd &= yb - sc \cdot \cos \gamma + sd \cdot \sin \gamma
\end{aligned}$$

son las coordenadas del punto D .

Las restricciones algebraicas del sistema mecánico se deben a las tres cadenas cinemáticas. La primera cadena conecta O con B a través de K_1 , K_2 y K_3 , lo que provoca las dos restricciones

$$\begin{aligned}
rr \cdot \cos \beta - d \cdot \cos(\beta + \Theta) - ss \cdot \sin \gamma &= xb, \\
rr \cdot \sin \beta - d \cdot \sin(\beta + \Theta) + ss \cdot \cos \gamma &= yb.
\end{aligned}$$

Las otras dos conectan O con A , una a través de K_1 , K_2 , K_4 y K_5 , que provoca las restricciones

$$\begin{aligned}
rr \cdot \cos \beta - d \cdot \cos(\beta + \Theta) - e \cdot \sin(\Phi + \delta) - zt \cdot \cos \delta &= xa, \\
rr \cdot \sin \beta - d \cdot \sin(\beta + \Theta) + e \cdot \cos(\Phi + \delta) - zt \cdot \sin \delta &= ya,
\end{aligned}$$

y la otra a través de K_1 , K_2 , K_6 y K_7 , generando las condiciones

$$\begin{aligned} rr \cdot \cos \beta - d \cdot \cos(\beta + \Theta) - zf \cdot \cos(\Omega + \varepsilon) - u \cdot \sin \varepsilon &= xa, \\ rr \cdot \sin \beta - d \cdot \sin(\beta + \Theta) - zf \cdot \sin(\Omega + \varepsilon) + u \cdot \cos \varepsilon &= ya. \end{aligned}$$

Estas seis ecuaciones representan, para las coordenadas generalizadas \mathbf{q} , las restricciones del sistema mecánico $\mathbf{g}(\mathbf{q}) = 0$ y derivando respecto de t obtenemos $\mathbf{0} = G(\mathbf{q})\mathbf{v}$. Esto junto a la expresión de la energía cinética T de (3.15) y la energía potencial U de (3.16) nos permite obtener las ecuaciones del movimiento de Lagrange de (3.1) utilizando que

$$L = T - U - \lambda_1 \mathbf{g}_1(\mathbf{q}) - \dots - \lambda_6 \mathbf{g}_6(\mathbf{q}).$$

Además, podemos obtener un sistema de EDAs de índice 3 dado por (3.4) o un sistema de EDAs índice 2 dado por (3.6). En ambos los elementos de $M(\mathbf{q})$ vienen dados por

$$m_{ij} = \frac{\partial^2 T}{\partial v_i \partial v_j},$$

y los de $\mathbf{f}(\mathbf{q}, \mathbf{v})$ por

$$\mathbf{f}_i(\mathbf{q}, \mathbf{v}) = \frac{\partial(T - U)}{\partial q_i} - \sum_{j=1}^7 \frac{\partial^2(T - U)}{\partial v_i \partial q_j} \cdot v_j,$$

puesto que

$$\frac{d}{dt} \left(\frac{\partial(T - U)}{\partial v_i} \right) = \sum_{j=1}^7 \frac{\partial^2(T - U)}{\partial v_i \partial q_j} \cdot v_j + \sum_{j=1}^7 \frac{\partial^2(T - U)}{\partial v_i \partial v_j} \cdot v_j'$$

y U no depende de las velocidades.

En la simulación que realizamos utilizaremos los parámetros dados por

$$\begin{bmatrix} xa \\ ya \end{bmatrix} = \begin{bmatrix} -0.06934 \\ -0.00227 \end{bmatrix}, \quad \begin{bmatrix} xb \\ yb \end{bmatrix} = \begin{bmatrix} -0.03635 \\ 0.03273 \end{bmatrix}, \quad \begin{bmatrix} xc \\ yc \end{bmatrix} = \begin{bmatrix} 0.014 \\ 0.072 \end{bmatrix},$$

la magnitud del par motor $mom = 0.033$ y por las tablas 3.1 y 3.2.

$d = 0.028$	$da = 0.0115$	$e = 0.02$
$ea = 0.01421$	$zf = 0.02$	$fa = 0.01421$
$rr = 0.007$	$ra = 0.00092$	$ss = 0.035$
$sa = 0.01874$	$sb = 0.01043$	$sc = 0.018$
$sd = 0.02$	$zt = 0.04$	$ta = 0.02308$
$tb = 0.00916$	$u = 0.04$	$ua = 0.01228$
$ub = 0.00449$	$c_0 = 4530$	$\ell_0 = 0.07785$

Tabla 3.1: Parámetros Geométricos.

masas	inercias
$m_1 = 0.04325$	$I_1 = 2.194 \cdot 10^{-6}$
$m_2 = 0.00365$	$I_2 = 4.410 \cdot 10^{-7}$
$m_3 = 0.02373$	$I_3 = 5.255 \cdot 10^{-6}$
$m_4 = 0.00706$	$I_4 = 5.667 \cdot 10^{-7}$
$m_5 = 0.07050$	$I_5 = 1.169 \cdot 10^{-5}$
$m_6 = 0.00706$	$I_6 = 5.667 \cdot 10^{-7}$
$m_7 = 0.05498$	$I_7 = 1.912 \cdot 10^{-5}$

Tabla 3.2: Parámetros de los siete cuerpos.

Los valores iniciales para las coordenadas generalizadas serán

$$\begin{aligned}
\beta(0) &= -0.0617138900142764496358948458001, \\
\Theta(0) &= 0, \\
\gamma(0) &= 0.455279819163070380255912382449, \\
\Phi(0) &= 0.222668390165885884674473185609, \\
\delta(0) &= 0.487364979543842550225598953530, \\
\Omega(0) &= -0.222668390165885884674473185609, \\
\varepsilon(0) &= 1.23054744454982119249735015568,
\end{aligned}$$

y para las velocidades generalizadas serán

$$\beta'(0) = \Theta'(0) = \gamma'(0) = \Phi'(0) = \delta'(0) = \Omega'(0) = \varepsilon'(0) = 0.$$

Para que las condiciones iniciales sean consistentes tomaremos, además,

$$\begin{aligned}\lambda_1 &= 98.5668703962410896057654982170, \\ \lambda_2 &= -6.12268834425566265503114393122,\end{aligned}$$

y nulos el resto de multiplicadores de Lagrange.

Tanto la formulación del problema, como los valores de los parámetros y de las condiciones iniciales han sido tomados de [5, pp. 530–536].

3.3. Resultados numéricos

Para obtener los resultados numéricos que se presentan en esta sección hemos implementado en MATLAB un método Runge-Kutta semi-implícito para integrar numéricamente los sistemas de EDAs de índice 2 de cada uno de los problemas introducidos en la sección anterior. En el Apéndice A puede verse el código de algunos de los programas que se han realizado.

El método Runge-Kutta semi-implícito que hemos implementado para integrar numéricamente los problemas de la forma (3.6) que hemos descrito en la sección anterior es un método de orden 4 conocido como HEM4 y que fue propuesto en [3]. En dicho trabajo se analiza la construcción de métodos RK semi-implícitos de orden 4 utilizando las condiciones de orden expuestas en el Capítulo 2 de esta memoria, que, junto con algunas condiciones simplificadoras adicionales, permiten encontrar conjuntos de coeficientes que las satisfacen. En la Tabla 3.3 puede verse el tablero de Butcher que contiene los coeficientes del método HEM4.

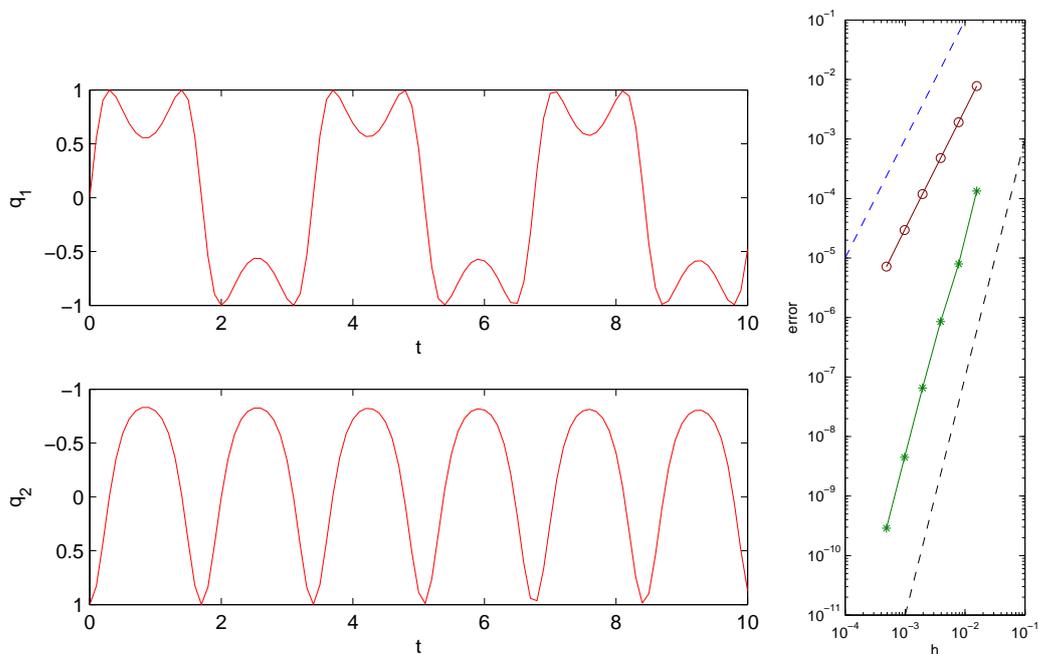
0					
$\frac{3}{10}$	$\frac{3}{10}$				
$\frac{4-\sqrt{6}}{10}$	$\frac{1+\sqrt{6}}{30}$	$\frac{11-4\sqrt{6}}{30}$			
$\frac{4+\sqrt{6}}{10}$	$\frac{-79-31\sqrt{6}}{150}$	$\frac{-1-4\sqrt{6}}{30}$	$\frac{24+11\sqrt{6}}{25}$		
1	$\frac{14+5\sqrt{6}}{6}$	$\frac{-8+7\sqrt{6}}{6}$	$\frac{-9-7\sqrt{6}}{4}$	$\frac{9-\sqrt{6}}{4}$	
	0	0	$\frac{16-\sqrt{6}}{36}$	$\frac{16+\sqrt{6}}{36}$	$\frac{1}{9}$

Tabla 3.3: Tablero de Butcher del método RK semi-implícito HEM4

3.3.1. Péndulo matemático

Tanto las ecuaciones del movimiento del péndulo como los valores de los parámetros y de las condiciones iniciales utilizadas se han fijado en la Subsección 3.2.1. El tiempo de integración que se ha considerado es $[0, 10]$ y con longitud de paso $h = 0.1$ los resultados son ya suficientemente precisos.

En la Figura 3.3a puede observarse para $h = 0.01$ la evolución con respecto al tiempo de las dos componentes de la posición, en la parte superior la de $q_1 = x$ y en la inferior la de $q_2 = y$. Nótese que en la gráfica de la parte inferior el eje vertical se encuentra con sentido contrario a lo habitual debido a que en la modelización del péndulo también se realizó esa inversión en el eje vertical.



(a) Evolución de $q_1 = x$ y $q_2 = y$ con el tiempo.

(b) Orden.

Figura 3.3: Resultados numéricos del péndulo

Pasamos a analizar ahora, valiéndonos de la Figura 3.3b, el orden del método HEM4 al integrar numéricamente el péndulo para las variables diferenciales $\mathbf{y} = (\mathbf{q}, \mathbf{v})$ (en línea verde continua uniendo asteriscos) y para la variable algebraica $\mathbf{z} = \lambda$ (en línea roja continua uniendo círculos). Las líneas discontinuas azul y negra son rectas de referencia que tienen, respectivamente, pendientes 2 y 4.

La gráfica muestra en escala doblemente logarítmica el máximo error en la solución numérica a lo largo de toda la integración, medido en la norma euclídea de \mathbb{R}^4 para la variable diferenciable y de \mathbb{R} para la variable algebraica, frente a la longitud del paso de integración. Las seis longitudes de paso utilizadas van desde $h = 1/64$ (la más grande) a $h = 1/2048$, reduciéndose a la mitad para pasar de una a la siguiente. Sólo se ha medido el error en los tiempos comunes a todos los pasos. Como solución de referencia se ha tomado la obtenida con el mismo método y paso $h = 1/10240$, que es suficientemente pequeño frente a las longitudes de paso utilizadas en la gráfica.

El orden buscado lo podemos apreciar gráficamente a través de la pendiente de cada una de las “rectas” (roja y verde) obtenidas en la gráfica. Si el orden del método es p el error global se comportará aproximadamente como $E(h) \approx Kh^p$. Por tanto la pendiente aproximada de la recta entre dos puntos correspondientes a pasos $h_1 < h_2$ es

$$\frac{\log(E(h_2)) - \log(E(h_1))}{\log(h_2) - \log(h_1)} \approx \frac{\log(K) + p \log(h_2) - \log(K) - p \log(h_1)}{\log(h_2) - \log(h_1)} = p,$$

el orden del método. En otras palabras, al dividir h por 2,

$$E(h/2) \approx K(h/2)^p = K * \frac{h^p}{2^p} = E(h) \frac{1}{2^p}$$

y, en consecuencia, el error global se divide por 2^p .

En nuestro caso la línea verde, asociada a la variable diferencial $\mathbf{y} = (\mathbf{q}, \mathbf{v})$, tiene pendiente aproximadamente 4 lo que coincide con el orden teórico del método, puesto que en la construcción de HEM4 en [3] se ha pedido que se satisfagan las condiciones de orden procedentes de igualar hasta los términos $O(h^4)$ los desarrollos de Taylor de \mathbf{y}_1 e $\mathbf{y}(t_0 + h)$. Para la línea roja, asociada a la variable algebraica $\mathbf{z} = \lambda$, la pendiente es únicamente 2. Esto es debido a que los desarrollos de Taylor de \mathbf{z}_1 y $\mathbf{z}(t_0 + h)$ coinciden, por construcción del método, sólo hasta términos $O(h^2)$.

3.3.2. Brazo plano articulado de dos grados de libertad

Tanto las ecuaciones del movimiento del brazo plano articulado como los valores de los parámetros y de las condiciones iniciales utilizadas en la simulación se han fijado en la Subsección 3.2.2. El tiempo de integración que se ha considerado es $[0, 10]$ y se han utilizado longitudes de paso $h = 0.02, 0.01, 0.005$. Con $h = 0.02$ los resultados todavía no son lo suficientemente precisos pero

para paso $h = 0.01$ y 0.005 se obtiene la gráfica que se muestra en la Figura 3.4 en la que puede observarse el camino restringido de (x_2, y_2) que coincide con el que aparece en [2, p. 288].

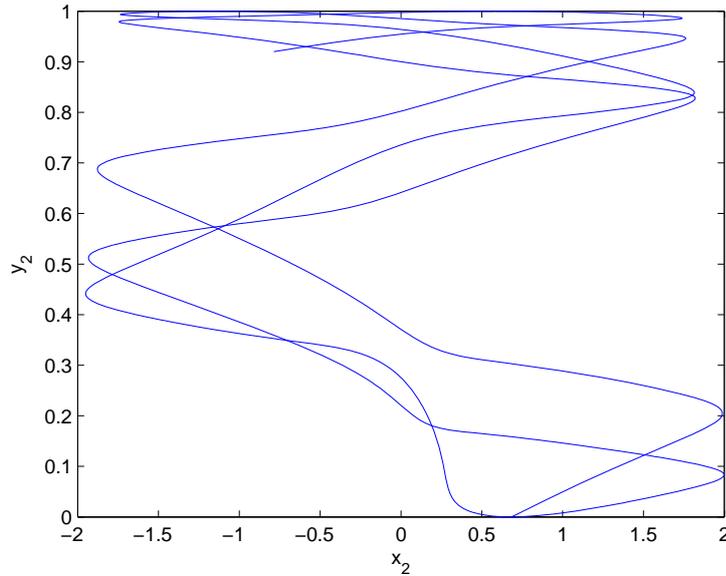


Figura 3.4: Posición del extremo “libre” en el plano $x-y$

En la Figura 3.5 mostramos tres gráficas. En la situada en la parte inferior aparece en azul la solución de la restricción teórica (3.13a) $y_2(t) = \sin^2(t/2)$ (con algunos de sus puntos marcados con asteriscos) y en rojo la aproximación numérica a y_2 obtenida en la simulación. En la gráfica del medio aparece en verde el valor absoluto de la diferencia entre ambos valores, es decir, el error cometido al evaluar la restricción original (3.13a) a lo largo de la solución numérica. En la gráfica superior se muestra en azul el error cometido al evaluar la restricción $G \cdot \mathbf{v}$ con G dado por (3.13b), que se impone realmente en la formulación de índice 2 del problema. Nótese que mientras que en el primer caso, al evaluar la restricción original, los mayores errores son del orden de 10^{-3} , en el segundo caso, al evaluar la restricción que efectivamente hemos exigido a la solución numérica, los mayores errores son del orden de 10^{-15} , lo que corresponde al tamaño de los errores de redondeo.

Si quisiéramos que la solución numérica verificara de forma exacta (salvo errores de redondeo) la restricción original, existe una formulación diferente propuesta en [6] que sigue siendo de índice 2. La formulación difiere de (3.6) en que se exige, además, la restricción $0 = \mathbf{g}(\mathbf{q})$ y se añade un multiplicador

de Lagrange adicional μ que aparece en la ecuación (3.6a), que pasa a ser

$$\mathbf{q}' = \mathbf{v} + G^T \mu.$$

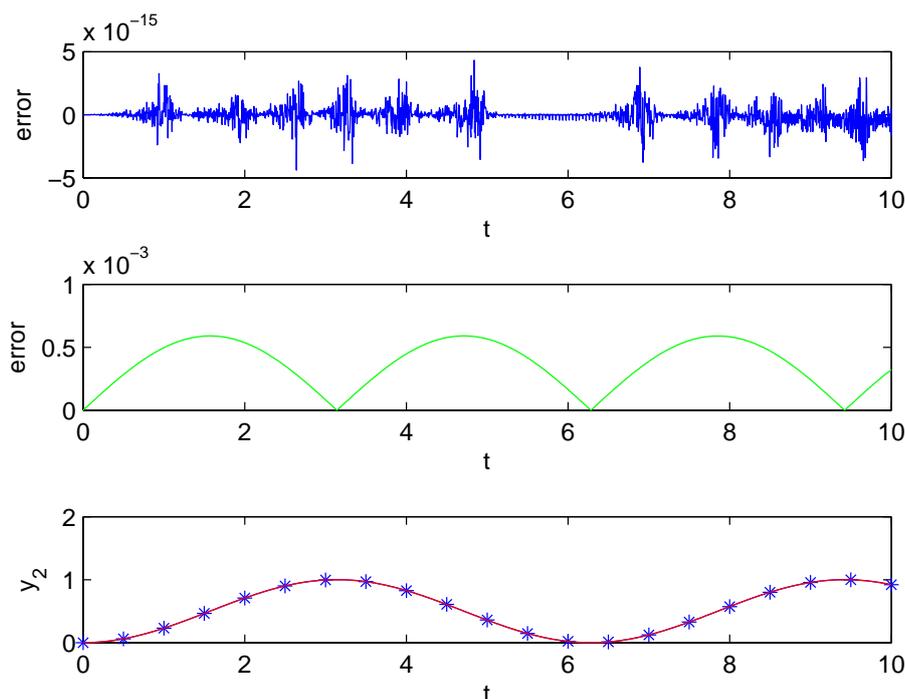


Figura 3.5: Comparación de y_2 y su aproximación numérica.

3.3.3. Andrews' Squeezing mechanism

Tanto las ecuaciones del movimiento del Andrews' Squeezing mechanism como los valores de los parámetros y de las condiciones iniciales que se han utilizado en los experimentos numéricos se han fijado en la Subsección 3.2.3. El tiempo de integración que se ha considerado es $[0, 0.03]$, como en [5], y con longitud de paso $h = 6 \cdot 10^{-5}$ se obtiene la gráfica que se muestra en la Figura 3.6a, en la que puede observarse la evolución de los ángulos (excepto ε) con el tiempo. La única razón para no representar la evolución con el tiempo del ángulo ε es conseguir que la Figura 3.6a reproduzca fielmente la gráfica que aparece en [5, p. 536].

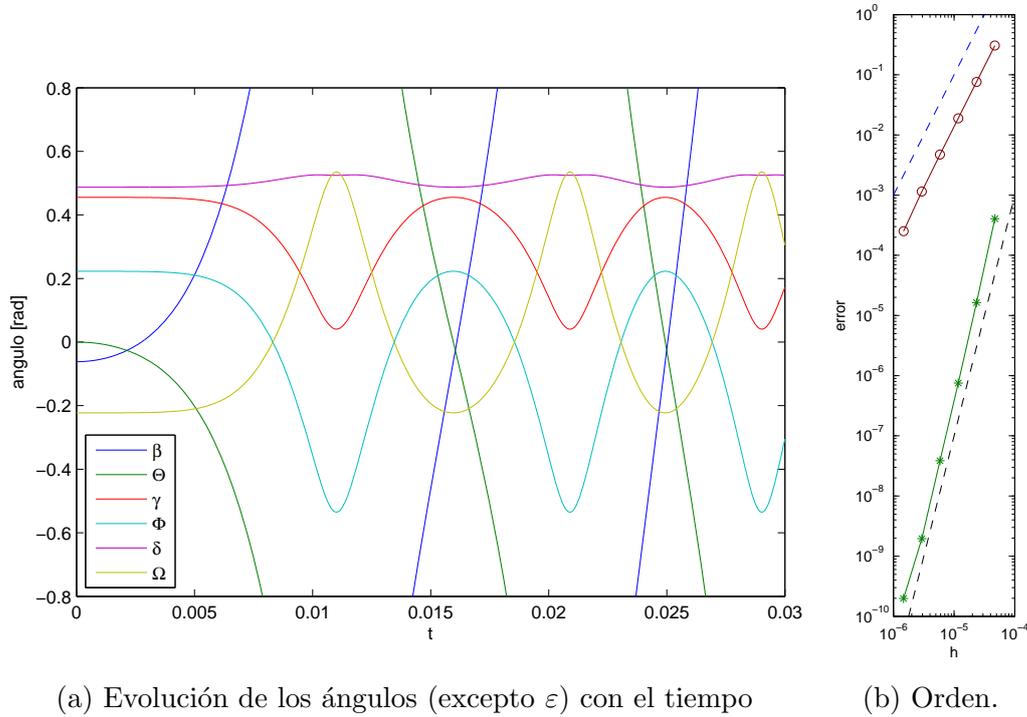


Figura 3.6: Resultados numéricos del Andrews' Squeezing mechanism

Como en el caso del péndulo, pasamos a analizar ahora, valiéndonos de la Figura 3.6b, el orden del método HEM4 al integrar numéricamente el Andrews' Squeezing mechanism para las variables diferenciales $\mathbf{y} = (\mathbf{q}, \mathbf{v})$ (en línea verde continua uniendo asteriscos) y para la variable algebraica $\mathbf{z} = \boldsymbol{\lambda}$ (en línea roja continua uniendo círculos). Las líneas discontinuas azul y negra son rectas de referencia que tienen, también ahora, pendientes 2 y 4, respectivamente.

La gráfica muestra en escala doblemente logarítmica el máximo error en la solución numérica a lo largo de toda la integración, medido en la normas euclideas de \mathbb{R}^{14} y \mathbb{R}^6 respectivamente, frente a la longitud del paso de integración. Las seis longitudes de paso utilizadas van desde $h = 0.03/640$ (la más grande) a $h = 0.03/20480$, reduciéndose a la mitad para pasar de una a la siguiente. De nuevo, sólo se ha medido el error en los tiempos comunes a todas las integraciones. Como solución de referencia se ha tomado la obtenida con el mismo método y paso $h = 0.03/51200$, que es suficientemente pequeño frente a las longitudes de paso utilizadas en la gráfica.

Al igual que en el caso del péndulo, la línea verde, asociada a la variable diferencial $\mathbf{y} = (\mathbf{q}, \mathbf{v})$, tiene pendiente aproximadamente 4, orden teórico

del método. Para la línea roja, asociada a la variable algebraica $\mathbf{z} = \boldsymbol{\lambda}$, la pendiente vuelve a ser únicamente 2. Como ya comentamos en la Subsección 3.3.1, esto se debe a que los coeficientes del método HEM4 satisfacen las condiciones de orden relativas a los desarrollos de Taylor de \mathbf{y}_1 y $\mathbf{y}(t_0 + h)$ hasta los términos $O(h^4)$, pero sólo satisfacen las condiciones de orden relativas a los desarrollos de Taylor de \mathbf{z}_1 y $\mathbf{z}(t_0 + h)$ hasta los términos $O(h^2)$.

Bibliografía

- [1] V.I. Arnol'd, *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York (1978).
- [2] U.M. Ascher & L.R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. SIAM, Philadelphia (1998).
- [3] V. Brasey & E. Hairer, *Half-explicit Runge-Kutta methods for differential-algebraic systems of index 2*, SIAM J. Numer. Anal. 30 (1993) 538–552.
- [4] E. Hairer, Ch. Lubich & M. Roche, *The numerical solution of differential-algebraic systems by Runge-Kutta methods*. Springer Lecture Notes in Math. 1409. Springer-Verlag, Berlin (1989).
- [5] E. Hairer & G. Wanner, *Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems*. Second edition. Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin (1996).
- [6] C.W. Gear, G.K. Gupta & B. Leimkuhler, *Automatic integration of Euler-Lagrange equations with constraints*, J. Comp. Appl. Math., 12-13 (1985) 77–90.
- [7] C.W. Gear, *Differential-algebraic equations, indices, and integral algebraic equations*, SIAM J. Numer. Anal. 27 (1990) 1527–1534.
- [8] D.R.A Giles, *A comparison of three problem-oriented simulation programs for dynamic mechanical systems*. Thesis, Univ. Waterloo, Ontario (1978).
- [9] J.B. Marion & S.T. Thornton *Classical Dynamics of Particles and Systems*. Fifth edition, Harcourt Brace Jovanovich, Orlando, Florida (1988).

- [10] D.W. Manning, *A computer technique for simulating dynamic multibody systems based on dynamic formalism*. Thesis, Univ. Waterloo, Ontario (1981).
- [11] R. Riaza, *Differential-algebraic systems: analytical aspects and circuit applications*. World Scientific (2008).
- [12] W. Schiehlen, *Multibody systems handbook*. Springer Verlag, Berlin (1990).

Apéndice A

Código MATLAB

En este apéndice mostramos el código de los programas de MATLAB que hemos realizado para la implementación del método HEM4 y su aplicación a los tres problemas que hemos tomado como ejemplo en la Sección 3.2 y cuyos resultados numéricos mostramos en la Sección 3.3: el péndulo matemático, un brazo plano articulado de dos grados de libertad y el sistema de siete cuerpos “Andrews’ squeezing mechanism”.

A.1. Programa principal: `hem4.m`

Este es el programa principal que implementa, con paso fijo, el método Runge-Kutta semi-implícito HEM4 para la integración numérica de sistemas de EDAs de índice 2 asociados a sistemas mecánicos con restricciones. Necesita de una función auxiliar que, dependiendo del problema, contenga los parámetros y las funciones necesarias.

```
function [tt,qq,vv,ll]=hem4(modelo,t0,tf,N,q0,v0,l0)

% Programa principal para integrar numéricamente sistemas de
% EDAs de índice 2 asociados a sistemas mecánicos con
% restricciones, utilizando el metodo Runge-Kutta
% semi-implícito HEM4
%
% Uso: [tt,qq,vv,ll]=hem4(modelo,t0,tf,N,q0,v0,l0)
%
% modelo es el nombre de la función auxiliar para cada problema
% en la que se encuentran los parámetros y las funciones M, G,
```

```

% f, g_t
%
% t0 es el tiempo inicial
%
% tf es el tiempo final
%
% N es el número de pasos que se dan con el método
%
% q0, v0 y l0 son las condiciones iniciales
%
% tt son los tiempos de integración
%
% qq, vv y ll contienen por columnas las componentes
% correspondientes a las coordenadas generalizadas,
% velocidades generalizadas y multiplicadores de Lagrange,
% respectivamente, de la solución numérica para cada
% paso
%
sqr6=sqrt(6);
s=5;% número de etapas del método
% Coeficientes del método
A=[0 ,0 ,0 ,0 ,0 ;
  3/10 ,0 ,0 ,0 ,0 ;
  (1+sqr6)/30 , (11-4*sqr6)/30,0 ,0 ,0 ;
  (-79-31*sqr6)/150, (-1-4*sqr6)/30, (24+11*sqr6)/25,0 ,0 ;
  (14+5*sqr6)/6 , (-8+7*sqr6)/6 , (-9-7*sqr6)/4 , (9-sqr6)/4 ,0];
b=[0 ,0 , (16-sqr6)/36 , (16+sqr6)/36, 1/9];
c=[0, 3/10, (4-sqr6)/10, (4+sqr6)/10, 1];
A=[A;b];
%
AsD=diag(diag(A,-1),-1);
AsD=AsD(:,1:s);
AsD=A-AsD; % A con subdiagonal de 0s.
%
h=(tf-t0)/N;% longitud de paso
tt=t0+(0:N)*h;% tiempos de integración
dimq=length(q0);
diml=length(l0);
qq=zeros(dimq,N+1);
vv=qq;
ll=zeros(diml,N+1);
qq(:,1)=q0;
vv(:,1)=v0;
ll(:,1)=l0;
for pp=2:N+1
  T=[tt(pp-1)+c*h, tt(pp)];
  Q=zeros(dimq,s+1);
  V=Q;
  dV=V;

```

```

L=zeros(diml,s);
% Etapa 1
Q(:,1)=q0;
V(:,1)=v0;
% Resto de etapas y calculo de q1,v1,l1.
for ii=1:s
    Q(:,ii+1)=Q(:,1) + h*V(:,1:s)*A(ii+1,:)' ;
    M=feval(modelo,1,T(ii),Q(:,ii),V(:,ii));
    G1=feval(modelo,2,T(ii),Q(:,ii),V(:,ii));
    G2=feval(modelo,2,T(ii+1),Q(:,ii+1),V(:,ii));
    f=feval(modelo,3,T(ii),Q(:,ii),V(:,ii));
    gt=feval(modelo,4,T(ii),Q(:,ii),V(:,ii));

    r=- (G2*(V(:,1)+h*dV(:,1:s)*AsD(ii+1,:)' )+ gt)/...
        (h*A(ii+1,ii));
    AS=[M,G1';G2,zeros(diml)];
    bS=[f;r];
    xS=AS\bS;
    dV(:,ii)=xS(1:dimq);
    L(:,ii)=xS(dimq+1:end);
    V(:,ii+1)=V(:,1)+h*dV(:,1:s)*A(ii+1,:)' ;
end
q0=Q(:,s+1);
v0=V(:,s+1);
l0=L(:,s);
qq(:,pp)=q0;
vv(:,pp)=v0;
ll(:,pp)=l0;
end
end

```

A.2. Pendulo.m

Función que contiene los parámetros y funciones del péndulo matemático.

```
function Y=Pendolo(etiqueta,t,q,v)
% Pendulo

% Parámetros
m=1;
l=1;
g=9.81;
% Funciones
switch etiqueta
    case 1 % M(q)
        Y=[m, 0; 0 , m];
    case 2 % G(q)
        Y=[2*q(1), 2*q(2)];
    case 3 % f(q,v)
        Y=[0;m*g];
    case 4 % gt(q,t)
        Y=0;
end
end
```

A.3. Brazo.m

Función que contiene los parámetros y funciones del brazo plano articulado de dos grados de libertad.

```
function Y=Brazo(etiqueta,t,q,v)
% Brazo

% Parámetros
m1=36;
m2=36;
l1=1;
l2=1;
g=9.81;
%
c1=cos(q(1));
c2=cos(q(2));
s1=sin(q(1));
s2=sin(q(2));
c12=cos(q(1)+q(2));
s12=sin(q(1)+q(2));
% Funciones
switch etiqueta
    case 1 % M(q)
        Y=[m1*l1^2/3+m2*(l1^2+l2^2/3+l1*l2*c2), ...
            m2*(l2^2/3+l1*l2*c2/2);
            m2*(l2^2/3+l1*l2*c2/2), m2*l2^2/3];
    case 2 % G(q)
        Y=[l1*c1+l2*c12, l2*c12];
    case 3 % f(q,v)
        Y=[-m1*g*l1*c1/2-m2*g*(l1*c1+l2*c12/2);
            -m2*g*l2*c12/2] ...
            + [m2*l1*l2*s2*(2*v(1)*v(2)+v(2)^2)/2;
            -m2*l1*l2*s2*v(1)^2/2];
    case 4 % gt(q,t)
        Y=-sin(t/2)*cos(t/2);
end
end
```

A.4. Andrews.m

Función que contiene los parámetros y funciones del sistema de siete cuerpos “Andrews’ squeezing mechanism”.

```
function Y=Andrews(etiqueta,t,q,v)
% Andrew's squeezing mechanism

% Parámetros
mom=0.033;
% Puntos fijos
xa=-0.06934;
ya=-0.00227;
xb=-0.03635;
yb=0.03273;
xc=0.014;
yc=0.072;
% Parámetros geométricos
d =0.028;   da=0.0115;   e=0.02;
ea=0.01421; zf=0.02;   fa=0.01421;
rr=0.007;   ra=0.00092; ss=0.035;
sa=0.01874; sb=0.01043; sc=0.018;
sd=0.02;   zt=0.04;   ta=0.02308;
tb=0.00916; u=0.04;   ua=0.01228;
ub=0.00449; c0=4530;   l0=0.07785;
%
sibe=sin(q(1)); cobe=cos(q(1));
sith=sin(q(2)); coth=cos(q(2));
siga=sin(q(3)); coga=cos(q(3));
siph=sin(q(4)); coph=cos(q(4));
side=sin(q(5)); code=cos(q(5));
siom=sin(q(6)); coom=cos(q(6));
siep=sin(q(7)); coep=cos(q(7));
%
sibeth=sin(q(1)+q(2)); cobeth=cos(q(1)+q(2));
siphde=sin(q(4)+q(5)); cophde=cos(q(4)+q(5));
siomep=sin(q(6)+q(7)); coomep=cos(q(6)+q(7));
%
bep=v(1);   thp=q(2);
php=v(4);   dep=q(5);
omp=v(6);   epp=q(7);
%
M1=0.04325;
M2=0.00365;
M3=0.02373;
M4=0.00706;
M5=0.07050;
```

```

M6=0.00706;
M7=0.05498;
%
I1=2.194E-6;
I2=4.410E-7;
I3=5.255E-6;
I4=5.667E-7;
I5=1.169E-5;
I6=5.667E-7;
I7=1.912E-5;
%
xd=sd*coga+sc*sigaxb;
yd=sd*sigax-sc*coga+yb;
lang=sqrt((xd-xc)^2+(yd-yc)^2);
force=-c0*(lang-l0)/lang;
fx=force*(xd-xc);
fy=force*(yd-yc);
% Funciones
switch etiqueta
case 1 % M(q)
    Y=zeros(7,7);
    % Matriz simétrica y tridiagonal
    % Elementos de la diagonal y subdiagonal no nulos.
    Y(1,1)=M1*ra^2+M2*(rr^2-2*da*rr*coth+da^2)+I1+I2;
    Y(2,1)=M2*(da^2-da*rr*coth)+I2;
    Y(2,2)=M2*da^2+I2;
    Y(3,3)=M3*(sa^2+sb^2)+I3;
    Y(4,4)=M4*(e-ea)^2+I4;
    Y(5,4)=M4*((e-ea)^2+zt*(e-ea)*siph)+I4;
    Y(5,5)=M4*(zt^2+2*zt*(e-ea)*siph+(e-ea)^2)...
        +M5*(ta^2+tb^2)+I4+I5;
    Y(6,6)=M6*(zf-fa)^2+I6;
    Y(7,6)=M6*((zf-fa)^2-u*(zf-fa)*siom)+I6;
    Y(7,7)=M6*((zf-fa)^2-2*u*(zf-fa)*siom+u^2)...
        +M7*(ua^2+ub^2)+I6+I7;
    % Elementos de la supradiagonal por simetría
    Y(1,2)=Y(2,1);
    Y(4,5)=Y(5,4);
    Y(6,7)=Y(7,6);
case 2 % G(q)
    Y=zeros(6,7);
    Y(1,1)=-rr*sibe+d*sibeth;
    Y(1,2)=d*sibeth;
    Y(1,3)=-ss*coga;
    Y(2,1)=rr*cobe-d*cobeth;
    Y(2,2)=-d*cobeth;
    Y(2,3)=-ss*sigax;
    Y(3,1)=-rr*sibe+d*sibeth;
    Y(3,2)=d*sibeth;

```

```

Y(3,4)=-e*cophde;
Y(3,5)=-e*cophde+zt*side;
Y(4,1)=rr*cobe-d*cobeth;
Y(4,2)=-d*cobeth;
Y(4,4)=-e*siphde;
Y(4,5)=-e*siphde-zt*code;
Y(5,1)=-rr*sibe+d*sibeth;
Y(5,2)=d*sibeth;
Y(5,6)=zf*siomep;
Y(5,7)=zf*siomep-u*coep;
Y(6,1)=rr*cobe-d*cobeth;
Y(6,2)=-d*cobeth;
Y(6,6)=-zf*coomep;
Y(6,7)=-zf*coomep-u*siep;
case 3 % f(q,v)
    Y=zeros(7,1);
    Y(1)=mom-M2*da*rr*thp*(thp+2*bep)*sith;
    Y(2)=M2*da*rr*bep^2*sith;
    Y(3)=fx*(sc*coga-sd*sig)+fy*(sd*coga+sc*sig);
    Y(4)=M4*zt*(e-ea)*dep^2*coph;
    Y(5)=-M4*zt*(e-ea)*php*(php+2*dep)*coph;
    Y(6)=-M6*u*(zf-fa)*epp^2*coom;
    Y(7)=M6*u*(zf-fa)*omp*(omp+2*epp)*coom;
case 4 % gt(q,t)
    Y=zeros(6,1);
end
end

```