

Resolución numérica de ecuaciones diferenciales algebraicas de índice 1

Jorge Rodríguez Domínguez

Resumen de mi Trabajo de Fin de Grado en el Grado de Matemáticas

1. Introducción

Muchos fenómenos de la vida real se pueden modelar a través de ecuaciones diferenciales, pues aparecen en una gran variedad de áreas como la Física, la Economía, la Biología, la Ingeniería... En general, no es posible resolver estas ecuaciones de manera exacta, por lo que es importante su resolución numérica mediante aproximaciones. Este trabajo se centra en la resolución numérica de sistemas de ecuaciones diferenciales algebraicos. La forma más general [3] de una ecuación diferencial algebraica (EDA) es un sistema de ecuaciones implícito de la forma

$$F(u', u, t) = 0, \quad (1)$$

siendo $F : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^N$ suficientemente regular. Una solución de (1) es una función $u : I \rightarrow \mathbb{R}^N$ derivable que satisface (1) en cada $t \in I$, siendo I un intervalo. La diferencia con los sistemas de ecuaciones diferenciales ordinarias es que, en general, el jacobiano de F con respecto a la variable u' es singular, lo que implica que las derivadas de alguna de las componentes puede no aparecer, dando lugar a que el sistema pueda contener ecuaciones algebraicas. El sistema (1) siempre se puede transformar en uno de la forma

$$F(u', u) = 0, \quad (2)$$

añadiendo t como incógnita artificial que satisface la ecuación $t' - 1 = 0$. Se denomina índice de diferenciación de un sistema de la forma (1) al menor número de veces que debemos diferenciar (1) para escribirlo en forma equivalente como un sistema de ecuaciones diferenciales ordinarias. En este trabajo vamos a considerar EDAs de la forma

$$y' = f(y, z), \quad (3a)$$

$$0 = g(y, z), \quad (3b)$$

con $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, que es un caso particular de (2), siendo $N = n + m$, $u = [y, z]^t$, $F(u', u) = [f(y, z) - y', g(y, z)]^t$. Si g es diferenciable y g_z es invertible en un entorno de la solución de (3), siendo

$$g_z(y, z) = \begin{bmatrix} \frac{\partial g_1}{\partial z_1}(y, z) & \dots & \frac{\partial g_1}{\partial z_m}(y, z) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial z_1}(y, z) & \dots & \frac{\partial g_m}{\partial z_m}(y, z) \end{bmatrix},$$

entonces (3) es un problema de índice 1, ya que, derivando (3b) con respecto a la variable independiente t obtenemos

$$0 = g_y(y, z)f(y, z) + g_z(y, z)z',$$

y como g_z es invertible en un entorno de la solución

$$z' = -g_z^{-1}(y, z)g_y(y, z)f(y, z).$$

De esta manera, podemos reescribir (3) como

$$\begin{bmatrix} y \\ z \end{bmatrix}' = \begin{bmatrix} f(y, z) \\ -g_z^{-1}(y, z)g_y(y, z)f(y, z) \end{bmatrix}$$

por tanto, el índice de diferenciación de (3) es $m = 1$. Un problema de valores iniciales es resolver (3) con condiciones iniciales $y(0) = y_0$, $z(0) = z_0$, siendo $y_0 \in \mathbb{R}^n$, $z_0 \in \mathbb{R}^m$ tales que $g(y_0, z_0) = 0$. Se dice que las condiciones iniciales que satisfacen esta relación son consistentes.

Este trabajo se centra en la resolución de problemas de condiciones iniciales para EDAs de índice 1 usando métodos Runge-Kutta implícitos. El primer método Runge-Kutta implícito [2] fue introducido por Cauchy (1824) con el fin de resolver numéricamente el siguiente problema de valor inicial

$$\begin{cases} y'(t) = f(t, y(t)), & t \in [t_0, t_0 + T], \\ y(t_0) = y_0, \end{cases}$$

La solución exacta de dicho problema verifica

$$y(t_{k+1}) = y(t_k) + \int_{t_k}^{t_{k+1}} f(t, y(t)) dt, \quad (4)$$

siendo t_0, t_1, \dots, t_N una partición del intervalo $[t_0, t_0 + T]$. Cauchy propuso la siguiente aproximación para esta expresión:

$$y_{k+1} = y_k + hf(t_k + \theta h, y_k + \Theta(y_{k+1} - y_k)),$$

siendo $h = t_{k+1} - t_k$, y_k una aproximación de $y(t_k)$, $0 \leq \theta$ y $\Theta \leq 1$. Para el caso $\theta = \Theta = 0$ se obtiene

$$y_{k+1} = y_k + hf(t_k, y_k),$$

que se conoce como el método de Euler. Para el caso $\theta = \Theta = 1$ obtenemos

$$y_{k+1} = y_k + hf(t_{k+1}, y_{k+1}),$$

conocido como el método de Euler implícito. C. Runge (1895) [1] decidió aproximar la integral de la expresión (4) con fórmulas de cuadraturas más sofisticadas como la del punto medio o el trapecio. Con el trabajo de Runge y de sus sucesores Heun y Kutta se desarrollaron métodos con más precisión y eficiencia, y se sentaron las bases de los conocidos métodos Runge-Kutta.

2. Método ε -encajado y Método State Space Form

Consideremos un método Runge-Kutta con matriz $A = (a_{ij})$ y vector de pesos b . Aplicando este método Runge-Kutta implícito a una ecuación diferencial algebraica, obtenemos una expresión como la siguiente:

$$\begin{aligned} y_k^{(i)} &= y_k + h \sum_{j=1}^s a_{ij} f(y_k^{(j)}, z_k^{(j)}), \quad i = 1, \dots, s, \\ 0 &= g(y_k^{(i)}, z_k^{(i)}), \quad i = 1, \dots, s, \\ y_{k+1} &= y_k + h \sum_{i=1}^s b_i f(y_k^{(i)}, z_k^{(i)}), \\ z_{k+1} &= \left(1 - \sum_{i,j=1}^s b_i \omega_{ij} \right) z_k + \sum_{i,j=1}^s b_i \omega_{ij} z_k^{(j)}. \end{aligned}$$

Este método se conoce como ε -encajado. Sin embargo, la solución numérica (y_{k+1}, z_{k+1}) de este enfoque, generalmente no verificará $g(y_{k+1}, z_{k+1}) = 0$, pero esto puede solucionarse reemplazando la última condición por

$$0 = g(y_{k+1}, z_{k+1}). \quad (5)$$

Así, llegamos al *método state space form*

$$\begin{aligned} y_k^{(i)} &= y_k + h \sum_{j=1}^s a_{ij} f(y_k^{(j)}, z_k^{(j)}), \quad i = 1, \dots, s, \\ 0 &= g(y_k^{(i)}, z_k^{(i)}), \quad i = 1, \dots, s, \\ y_{k+1} &= y_k + h \sum_{i=1}^s b_i f(y_k^{(i)}, z_k^{(i)}), \\ 0 &= g(y_{k+1}, z_{k+1}). \end{aligned}$$

3. Implementación

A continuación explicamos lo necesario para implementar el método state space form en `python`. Recordemos que dicho método es el siguiente

$$\begin{aligned} y_k^{(i)} &= y_k + h \sum_{j=1}^s a_{ij} f(y_k^{(j)}, z_k^{(j)}), \quad i = 1, \dots, s, \\ 0 &= g(y_k^{(i)}, z_k^{(i)}), \quad i = 1, \dots, s, \\ y_{k+1} &= y_k + h \sum_{i=1}^s b_i f(y_k^{(i)}, z_k^{(i)}), \\ 0 &= g(y_{k+1}, z_{k+1}). \end{aligned}$$

Para poder calcular y_{k+1} , hemos de resolver el sistema de ecuaciones que forman las dos primeras condiciones. En general, tendremos que resolver un sistema no lineal, por lo que vamos a aplicar el método de Newton-Raphson. Sea $H : U \subset \mathbb{R}^q \rightarrow \mathbb{R}^q$ definida como $H(u) = [h_1(u), \dots, h_q(u)]$, $u \in U$ y supongamos que queremos resolver el sistema $h_i(u) = 0$, $i = 1, \dots, q$. El método de Newton-Raphson define de forma recursiva una sucesión $\{u_k\}$ de vectores dada por

$$u_k = u_{k-1} - [DH(u_{k-1})]^{-1} H(u_{k-1}), \quad k = 0, 1, 2, \dots$$

donde $DH(u_{k-1})$ es la matriz jacobiana de H evaluada en u_{k-1} . Este método puede no converger, pero en caso de hacerlo, converge a un cero de H . En la práctica no es necesario calcular la inversa de $DH(u)$ pues podemos hacer lo siguiente

$$\begin{aligned} u_k &= u_{k-1} - [DH(u_{k-1})]^{-1} H(u_{k-1}), \\ H(u_{k-1}) &= DH(u_{k-1})(u_{k-1} - u_k). \end{aligned}$$

Bastaría con resolver el sistema $DH(u_{k-1})x = H(u_{k-1})$. Si x_k es la solución de dicho sistema, tenemos que $x_k = u_{k-1} - u_k$, es decir, $u_k = u_{k-1} - x_k$. El sistema $(??)$ y $(??)$ es equivalente a encontrar un cero de la función $H : \mathbb{R}^{ns} \times \mathbb{R}^{ms} \rightarrow \mathbb{R}^{ns} \times \mathbb{R}^{ms}$ definida por $H(y^{(1)}, \dots, y^{(n)}, z^{(1)}, \dots, z^{(m)}) =$

$$\begin{bmatrix} \begin{bmatrix} y_1^1 - y_k^1 \\ \vdots \\ y_s^1 - y_k^1 \end{bmatrix} - \begin{bmatrix} a_{11} & \dots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \dots & a_{ss} \end{bmatrix} \cdot \begin{bmatrix} f^1(y_1^1, \dots, y_1^n, z_1^1, \dots, z_1^m) \\ \vdots \\ f^1(y_s^1, \dots, y_s^n, z_s^1, \dots, z_s^m) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} y_1^n - y_k^n \\ \vdots \\ y_s^n - y_k^n \end{bmatrix} - \begin{bmatrix} a_{11} & \dots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \dots & a_{ss} \end{bmatrix} \cdot \begin{bmatrix} f^n(y_1^1, \dots, y_1^n, z_1^1, \dots, z_1^m) \\ \vdots \\ f^n(y_s^1, \dots, y_s^n, z_s^1, \dots, z_s^m) \end{bmatrix} \\ g^1(y_1^1, \dots, y_1^n, z_1^1, \dots, z_1^m) \\ \vdots \\ g^1(y_s^1, \dots, y_s^n, z_s^1, \dots, z_s^m) \\ \vdots \\ g^m(y_1^1, \dots, y_1^n, z_1^1, \dots, z_1^m) \\ \vdots \\ g^m(y_s^1, \dots, y_s^n, z_s^1, \dots, z_s^m) \end{bmatrix},$$

donde $y^{(i)} = [y_1^i, \dots, y_s^i]^t$, $z^{(j)} = [z_1^j, \dots, z_s^j]^t$, $i = 1, \dots, n$, $j = 1, \dots, m$. Definimos $u = [y^{(1)}, \dots, y^{(n)}, z^{(1)}, \dots, z^{(m)}]^t$. La matriz jacobiana de H se puede ver como la siguiente matriz por bloques

$$DH(u) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

siendo

$$\begin{aligned} A_{11} &= \begin{bmatrix} I - hA \cdot F_{y_1}^1 & -hA \cdot F_{y_2}^1 & \dots & -hA \cdot F_{y_n}^1 \\ -hA \cdot F_{y_1}^2 & I - hA \cdot F_{y_2}^2 & \dots & -hA \cdot F_{y_n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -hA \cdot F_{y_1}^n & -hA \cdot F_{y_2}^n & \dots & I - hA \cdot F_{y_n}^n \end{bmatrix}, & A_{21} &= \begin{bmatrix} G_{y_1}^1 & G_{y_2}^1 & \dots & G_{y_n}^1 \\ G_{y_1}^2 & G_{y_2}^2 & \dots & G_{y_n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ G_{y_1}^m & G_{y_2}^m & \dots & G_{y_n}^m \end{bmatrix}, \\ A_{12} &= \begin{bmatrix} -hA \cdot F_{z_1}^1 & -hA \cdot F_{z_2}^1 & \dots & -hA \cdot F_{z_m}^1 \\ -hA \cdot F_{z_1}^2 & -hA \cdot F_{z_2}^2 & \dots & -hA \cdot F_{z_m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -hA \cdot F_{z_1}^n & -hA \cdot F_{z_2}^n & \dots & -hA \cdot F_{z_m}^n \end{bmatrix}, & A_{22} &= \begin{bmatrix} G_{z_1}^1 & G_{z_2}^1 & \dots & G_{z_m}^1 \\ G_{z_1}^2 & G_{z_2}^2 & \dots & G_{z_m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ G_{z_1}^m & G_{z_2}^m & \dots & G_{z_m}^m \end{bmatrix}, \end{aligned}$$

donde I es la matriz identidad de dimensión s y

$$\begin{aligned} F_{y_j}^i &= \text{diag} [\partial_{y_j} f^i(y_1^1, \dots, z_1^m), \dots, \partial_{y_j} f^i(y_s^1, \dots, z_s^m)]^t, & i &= 1, \dots, n, & j &= 1, \dots, n, \\ F_{z_j}^i &= \text{diag} [\partial_{z_j} f^i(y_1^1, \dots, z_s^m), \dots, \partial_{z_j} f^i(y_s^1, \dots, z_s^m)]^t, & i &= 1, \dots, n, & j &= 1, \dots, m, \\ G_{y_j}^i &= \text{diag} [\partial_{y_j} g^i(y_1^1, \dots, z_s^m), \dots, \partial_{y_j} g^i(y_s^1, \dots, z_s^m)]^t, & i &= 1, \dots, m, & j &= 1, \dots, n, \\ G_{z_j}^i &= \text{diag} [\partial_{z_j} g^i(y_1^1, \dots, z_s^m), \dots, \partial_{z_j} g^i(y_s^1, \dots, z_s^m)]^t, & i &= 1, \dots, m, & j &= 1, \dots, m. \end{aligned}$$

Dado un vector $u \in \mathbb{R}^k$, denotamos por $\text{diag}(u)$ a la matriz diagonal de dimensión k , cuya diagonal es precisamente u .

Todo esto es lo necesario para calcular $y_k^{(i)}$ y $z_k^{(i)}$, $i = 1, \dots, s$ que forman el cero de H . Para calcular y_{k+1} basta usar (??). Para calcular z_{k+1} basta aplicar el método de Newton-Raphson a g pero teniendo en cuenta que ahora y es conocido, por tanto la matriz jacobiana que vamos a usar es $g_z(y, z)$. El programa que hemos implementado para resolver (??) usando el método de Newton-Raphson necesita los siguientes parámetros de entrada

- los valores iniciales y_0 y z_0 sobre los que comenzaremos a aplicar el método.

- el intervalo temporal donde queremos resolver el sistema (??) y que denotaremos como $[t_0, t_f]$.
- el número de particiones N del intervalo $[t_0, t_f]$.
- las funciones $f(y, z)$ y $g(y, z)$.
- las matrices jacobianas $f_y(y, z)$, $f_z(y, z)$, $g_y(y, z)$ y $g_z(y, z)$.
- el número máximo de iteraciones que vamos a aplicar del método de Newton-Raphson y que denotamos como *itermax*.
- la tolerancia bajo la cual consideramos que tenemos convergencia al aplicar el método de Newton-Raphson y que denotamos como *tol*.
- la matriz A y el vector b del método Runge-Kutta que queremos aplicar.

Como valores de salida damos $[t_k, y_k, z_k]^t$, $k = 0, 1, \dots, N$ que han sido obtenidos en cada iteración del método, donde t_k son los instantes de tiempo donde calculamos la solución numérica.

Para implementar el método ε -encajado, podemos reutilizar todo lo que hemos aplicado para el método state space form, pues la única diferencia entre ambos métodos es como se calcula z_{k+1} . En el caso del método ε -encajado, debemos resolver las ecuaciones de la ecuación de antes pero cambiando la última condición por

$$z_{k+1} = \left(1 - \sum_{i,j=1}^s b_i \omega_{ij}\right) z_k + \sum_{i,j=1}^s b_i \omega_{ij} z_k^{(j)},$$

siendo ω_{ij} la componente (i, j) de la inversa de A .

4. Péndulo simple

Un péndulo simple [4] consta de una varilla de longitud l de masa despreciable, con uno de los extremos fijos y en el otro una partícula de masa m . En ausencia de rozamiento, las fuerzas que intervienen en el movimiento son la tensión de la varilla, en ambas direcciones del sistema de referencia, y el peso de la partícula. Sea x_1 y x_2 el sistema de coordenadas y $x_1(t)$ y $x_2(t)$ la posición del centro de gravedad del péndulo en el instante t . En la Figura 1 podemos ver un esquema del péndulo simple.

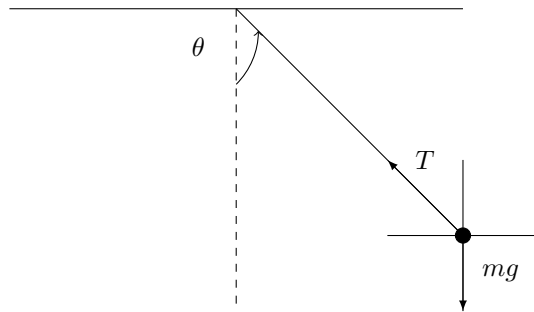


Figura 1: Esquema del péndulo simple.

Separando las componentes de la tensión obtenemos:

$$\begin{aligned} T_{x_1} &= -T \sin \theta, \\ T_{x_2} &= -T \cos \theta, \end{aligned}$$

siendo T el módulo de la tensión. Nótese que $x_1^2 + x_2^2 = l^2$, de donde deducimos que $x_1 = l \sin \theta$ y $x_2 = l \cos \theta$, y por tanto:

$$T_{x_1} = -T \frac{x_1}{l},$$

$$T_{x_2} = -T \frac{x_2}{l},$$

Aplicando las leyes de Newton, llegamos a

$$mx_1'' = -T \frac{x_1}{l}, \quad (6a)$$

$$mx_2'' = -T \frac{x_2}{l} + mg, \quad (6b)$$

siendo g la aceleración de la gravedad. Derivando x_1 y x_2 respecto de t obtenemos

$$x_1' = l\theta' \cos \theta, \quad x_2' = -l\theta' \sin \theta, \quad (7a)$$

$$x_1'' = l\theta'' \cos \theta - l(\theta')^2 \sin \theta, \quad x_2'' = -l\theta'' \sin \theta - l(\theta')^2 \cos \theta. \quad (7b)$$

Sustituyendo x_1'' y x_2'' en (6) obtenemos

$$l\theta'' \cos \theta - l(\theta')^2 \sin \theta = -\frac{T}{lm} \sin \theta, \quad (8a)$$

$$-l\theta'' \sin \theta - l(\theta')^2 \cos \theta = -\frac{T}{lm} l \cos \theta + g. \quad (8b)$$

Simplificando y reagrupando términos llegamos a

$$\left(\frac{T}{l} - m(\theta')^2\right) l \sin \theta + ml\theta'' \cos \theta = 0, \quad (9a)$$

$$\left(\frac{T}{l} - m(\theta')^2\right) l \cos \theta - ml\theta'' - mg = 0. \quad (9b)$$

Multiplicando (9a) por $\cos \theta$, (9b) por $-\sin \theta$ y sumando ambas igualdades, obtenemos

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

Despejando θ'' obtenemos

$$\theta'' = -\frac{g}{l} \sin \theta, \quad (10)$$

que es la conocida EDO de segundo orden para el péndulo simple. Nuestro objetivo ahora es reformular este problema para convertirlo en una EDA de índice 1. En primer lugar, podemos transformar el sistema (6) en uno de primer orden de la siguiente manera

$$x_1' = v_1, \quad x_2' = v_2, \quad (11a)$$

$$v_1' = -\frac{T x_1}{lm}, \quad v_2' = -\frac{T x_2}{lm} + g. \quad (11b)$$

Además, el sistema debe cumplir $x_1^2 + x_2^2 = l^2$. Así, el sistema final es

$$x_1' = v_1, \quad x_2' = v_2, \quad (12a)$$

$$v_1' = -\frac{T x_1}{lm}, \quad v_2' = -\frac{T x_2}{lm} + g, \quad (12b)$$

$$0 = x_1^2 + x_2^2 - l^2. \quad (12c)$$

El problema (12) no es una EDA de índice 1, pues la variable T no aparece en la ecuación algebraica, pero es posible reformular dicho sistema para que sí lo sea. Derivando (12c) respecto de la variable independiente obtenemos

$$x_1 v_1 + x_2 v_2 = 0. \quad (13)$$

Volviendo a derivar en (13) tenemos

$$\begin{aligned} 0 &= x'_1 v_1 + x_1 v'_1 + x'_2 v_2 + x_2 v'_2, \\ &= v_1^2 + x_1 v'_1 + v_2^2 + x_2 v'_2. \end{aligned}$$

Usando (12b) llegamos a

$$\begin{aligned} 0 &= v_1^2 + v_2^2 + x_1 \left(-\frac{T x_1}{lm} \right) + x_2 \left(-\frac{T x_2}{lm} + g \right), \\ &= v_1^2 + v_2^2 - \frac{T}{ml} (x_1^2 + x_2^2) + g x_2, \\ &= v_1^2 + v_2^2 - \frac{Tl}{m} + g x_2, \end{aligned}$$

de donde concluimos

$$m(v_1^2 + v_2^2) - Tl + g m x_2 = 0.$$

De esta manera, conseguimos transformar (12) en el siguiente sistema de ecuaciones

$$x'_1 = v_1, \quad x'_2 = v_2, \quad (14a)$$

$$v'_1 = -\frac{T x_1}{lm}, \quad v'_2 = -\frac{T x_2}{lm} + g, \quad (14b)$$

$$0 = m(v_1^2 + v_2^2) - Tl + g m x_2, \quad (14c)$$

que tras definir las variables $y = [x_1, x_2, v_1, v_2]^t$ y $z = T$, constituye una EDA de índice 1. Las funciones que definen dicha EDA son

$$\begin{aligned} f(y, z) &= \left[v_1, v_2, -\frac{z x_1}{lm}, -\frac{z x_2}{lm} + g \right]^t, \\ g(y, z) &= m(v_1^2 + v_2^2) - z l + g m x_2. \end{aligned}$$

Nótese que $g_z(y, z) = -l \neq 0$. Vamos a resolver la EDA definida en (14), para ello damos como parámetros

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

Fijamos los valores iniciales

$$y(0) = [0, 1, 0, 0]^t, \quad z(0) = 0.$$

El valor $z(0)$ ha sido elegido para que las condiciones iniciales sean consistentes, es decir, para que se verifique la condición $g(y(0), z(0)) = 0$. El tiempo de integración considerado es $[0, 5]$ y la longitud del paso de malla es $h = 0,0125$. También hemos fijado $itermax = 100$ y $tol = 10^{-12}$. En la Figura 2 podemos ver como evoluciona con el tiempo la variable y . La gráfica superior corresponde a la posición del péndulo (primera y segunda componente de y). La gráfica inferior corresponde a la velocidad del péndulo (tercera y cuarta componente de y).

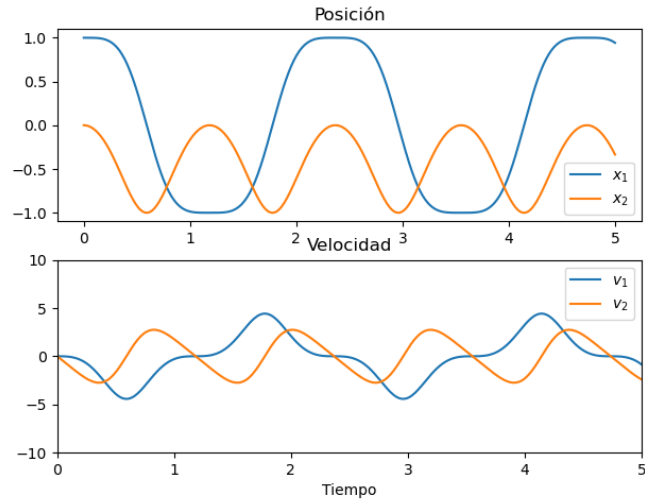


Figura 2: Evolución de la posición y la velocidad con el tiempo.

Para las condiciones iniciales impuestas, el centro de masa del péndulo comienza en la posición $(1,0)$ y se deja caer, pues las velocidades iniciales son cero. Debido a que no hay rozamiento, el centro de masa llega a la posición $(-1,0)$. A partir de este momento, la trayectoria del péndulo comienza a repetirse. En la Figura 3 podemos ver la posición del péndulo en diferentes instantes de tiempo.

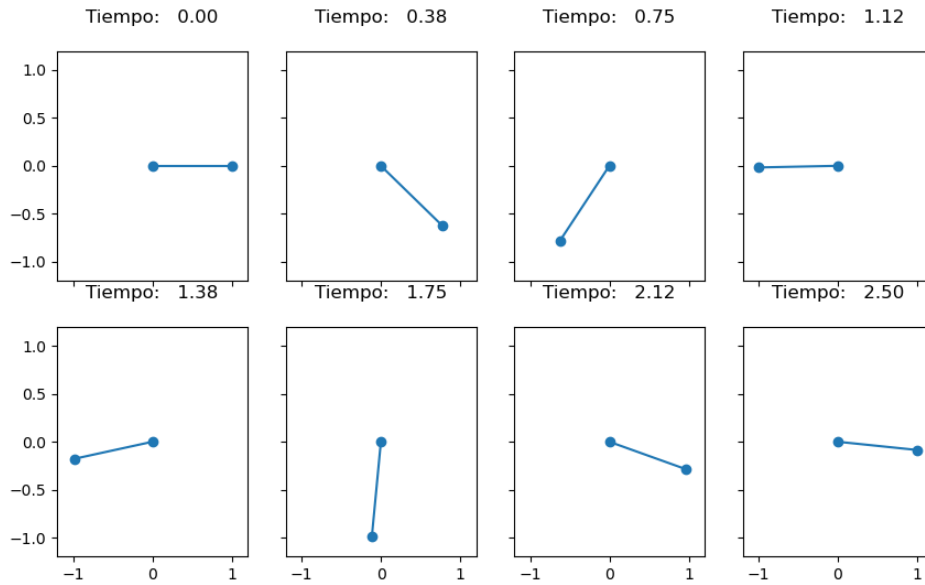


Figura 3: Animación del péndulo simple.

En la Figura 4 podemos ver como evoluciona con el tiempo la variable z , que corresponde a la tensión del péndulo.

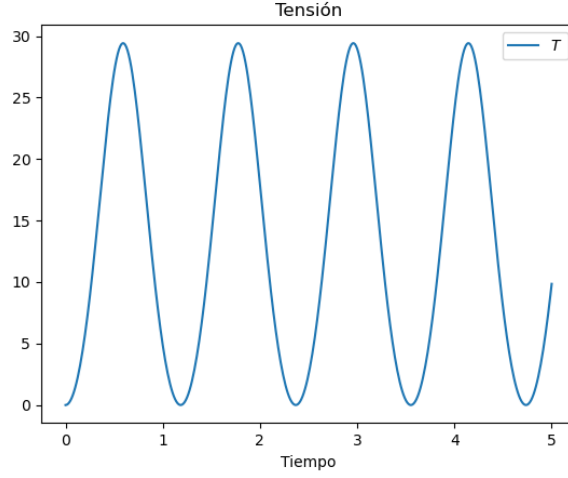


Figura 4: Evolución de la tensión con el tiempo.

Ambas Figuras han sido obtenidas al aplicar el método ε -encajado junto con el método Radau IIA de 3 etapas, que como sabemos es de orden 5.

En la Figura 5 vemos con que exactitud se cumplen las igualdades (12c) y (14c). En la gráfica superior vemos como se comporta la condición $g(y, z)$. En la gráfica inferior, vemos como se comporta $x_1^2 + x_2^2 - l^2$. La igualdad (14c) es prácticamente del orden de precisión de la máquina.

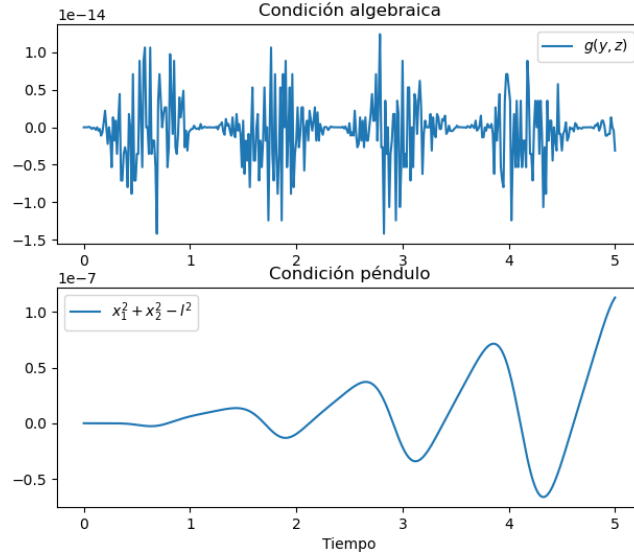


Figura 5: Evolución de la condición algebraica con el tiempo.

Usando la función `solve_ivp` de `python`, podemos resolver (10) y tomar la solución numérica como referencia para calcular el error que comete el programa que hemos implementado. Hemos fijado en la función `solve_ivp` una tolerancia de 10^{-12} para el error relativo y absoluto y hemos aplicado el método Radau de orden 5. En la Figura 6 representamos el error cometido en la variables y y z usando la norma infinito frente al paso de malla en una escala logarítmica. Los pasos de malla utilizados han sido $h = \frac{1}{2^i 100}$, para $i = 1, \dots, 5$.

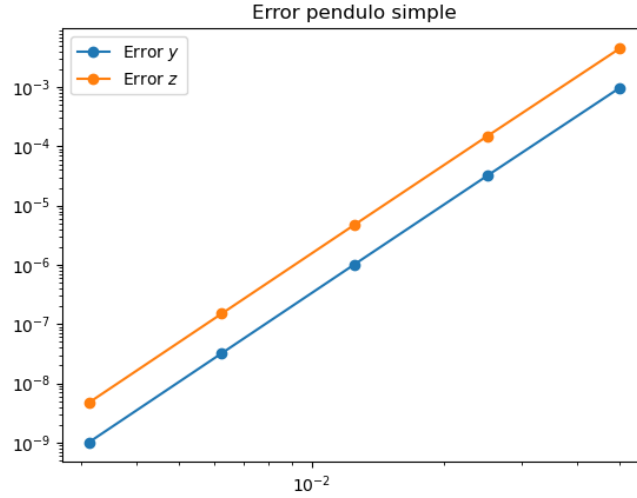


Figura 6: Error en las variables y y z en función del paso de malla.

Como el método Radau IIA de 3 etapas es rígidamente preciso y de orden 5, la recta errores frente al paso de malla (en escala logarítmica) de y y de z tiene pendiente aproximadamente 5. En el Cuadro 1 recogemos el orden aproximado del error en cada variable para diferentes pasos de malla.

h	Orden y	Orden z
0,05	—	—
0,025	4,968	4,966
0,0125	4,993	4,988
0,0625	4,998	4,998
0,003125	4,999	5,000

Cuadro 1: Órdenes aproximados en la variables y y z del péndulo simple.

5. Amplificador de transistores

A menudo, las ecuaciones diferenciales algebraicas pueden aparecer como problemas de la forma $Mu' = \varphi(u)$, donde M es una matriz constante y singular. Como un ejemplo de ello, consideramos el circuito eléctrico [3] representado en la Figura 7.

Un circuito eléctrico es una conexión entre componentes electrónicos que transportan corriente eléctrica. El circuito considerado posee cinco nodos, representados por vértices y conectados por aristas, llamadas ramas. Posee además seis resistencias, tres condensadores y un transistor.

En el circuito de la Figura 7 aparecen además, $U_e(t)$ que es la entrada de voltaje, $U_b = 6$ que es el voltaje de salida. Se representa por $U_i(t)$ al voltaje del nodo i -ésimo, $i = 1 \dots, 5$.

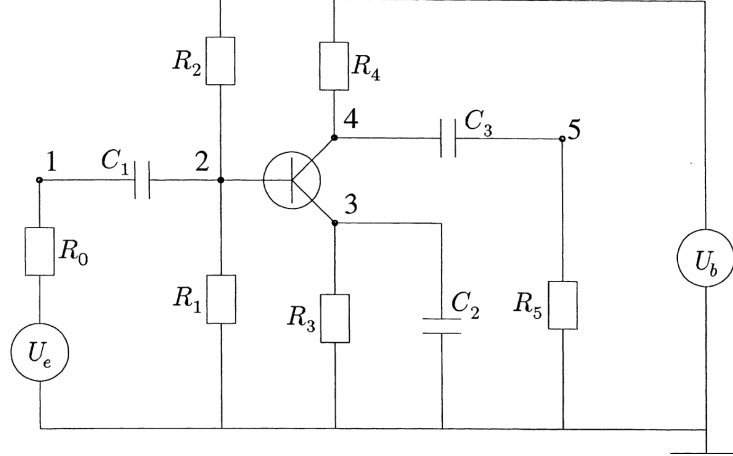


Figura 7: Amplificador de transistores.

La ley de Ohm establece que la corriente a través de una resistencia satisface $I = U/R$. Sin embargo, a través de un condensador satisface $I = C \cdot dU/dt$, donde R y C son constantes y U es el voltaje, que depende del tiempo t . El transistor actúa como amplificador en el siguiente sentido: la diferencia de voltajes entre los nodos 2 y 3 induce una intensidad $0,01h(U_2 - U_3)$ en la rama que conecta los nodos 2 y 3, mientras que en la rama que conecta los nodos 3 y 4 es de $0,99h(U_2 - U_3)$, siendo h una función no lineal.

La ley de Kirchhoff establece que la suma de las corrientes que entran en un nodo es nula, por lo que aplicando esta ley a cada uno de los nodos obtenemos el siguiente sistema de ecuaciones diferenciales:

$$\text{Nodo 1: } \frac{U_e}{R_0} - \frac{U_1}{R_0} + C_1(U_2' - U_1') = 0, \quad (15a)$$

$$\text{Nodo 2: } \frac{U_b}{R_2} - U_2 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) + C_1(U_1' - U_2') - 0,01h(U_2 - U_3) = 0, \quad (15b)$$

$$\text{Nodo 3: } h(U_2 - U_3) - \frac{U_3}{R_3} - C_2U_3' = 0, \quad (15c)$$

$$\text{Nodo 4: } \frac{U_b}{R_4} - \frac{U_4}{R_4} + C_3(U_5' - U_4') - 0,99h(U_2 - U_3) = 0, \quad (15d)$$

$$\text{Nodo 5: } -\frac{U_5}{R_5} + C_3(U_4' - U_5') = 0. \quad (15e)$$

Damos como datos

$$\begin{aligned} h(U) &= 10^{-6} \left(\exp \left(\frac{U}{0,026} \right) - 1 \right), \\ R_0 &= 1000, \quad R_i = 9000, \quad i = 1, \dots, 5, \\ C_k &= k \cdot 10^{-6}, \quad k = 1, 2, 3. \end{aligned}$$

Vamos a considerar que el voltaje de entrada viene dado por la función $U_e(t) = 0,4 \sin(200\pi t)$. Podemos escribir las ecuaciones (15) como un sistema de ecuaciones diferenciales de la forma $Mu' = \varphi(u)$, siendo $u = [U_1, U_2, U_3, U_4, U_5]^t$ y

$$M = \begin{bmatrix} -C_1 & C_1 & & & \\ C_1 & C_1 & & & \\ & & -C_2 & & \\ & & & -C_3 & C_3 \\ & & & C_3 & -C_3 \end{bmatrix},$$

$$\varphi(u) = \begin{bmatrix} \frac{U_1}{R_0} - \frac{U_e}{R_0} \\ 0,01h(U_2 - U_3) - \frac{U_b}{R_2} + U_2 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \\ \frac{U_3}{R_3} - h(U_2 - U_3) \\ 0,99h(U_2 - U_3) - \frac{U_b}{R_4} + \frac{U_4}{R_4} \\ \frac{U_5}{R_5} \end{bmatrix}.$$

La matriz M es una matriz singular de rango 3. Podemos transformar las ecuaciones (15) en un sistema de la forma

$$y' = f(y, z), \quad (16a)$$

$$0 = g(y, z), \quad (16b)$$

mediante el siguiente cambio de variable

$$y_1 = U_1 - U_2, \quad y_2 = U_3, \quad y_3 = U_4 - U_5, \quad z_1 = U_1, \quad z_2 = U_4.$$

Sustituyendo estas nuevas variables en (15) obtenemos [4]

$$0 = \frac{U_e - z_1}{R_0} - C_1 y'_1, \quad (17a)$$

$$0 = \frac{U_b}{R_2} + (y_1 - z_1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right) + C_1 y'_1 - 0,01h(z_1 - y_1 - y_2), \quad (17b)$$

$$0 = h(z_1 - y_1 - y_2) - \frac{y_2}{R_3} - C_2 y'_2, \quad (17c)$$

$$0 = \frac{U_b - z_2}{R_4} - C_3 y'_3 - 0,99h(z_1 - y_1 - y_2), \quad (17d)$$

$$0 = \frac{y_3 - z_2}{R_5} + C_3 y'_3. \quad (17e)$$

Despejando y'_1, y'_2, y'_3 de (17a), (17c) y (17e) respectivamente y definiendo $y = [y_1, y_2, y_3]^t$, $z = [z_1, z_2]^t$, obtenemos un sistema de la forma (16) donde

$$f(t, y, z) = \left[\frac{U_e - z_1}{R_0 C_1}, \frac{h(z_1 - y_1 - y_2)}{C_2} - \frac{y_2}{C_2 R_3}, \frac{z_2 - y_3}{C_3 R_5} \right]^t. \quad (18)$$

Para obtener g basta sumar (17a) con (17b) y (17d) con (17e), lo que nos da

$$0 = \frac{U_e - z_1}{R_0} + \frac{U_b}{R_2} + (y_1 - z_1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right) - \frac{h(z_1 - y_1 - y_2)}{100} := g_1(t, y, z), \quad (19a)$$

$$0 = \frac{U_b - z_2}{R_4} - 0,99h(z_1 - y_1 - y_2) + \frac{y_3 - z_2}{R_5} := g_2(t, y, z). \quad (19b)$$

Así, $g(t, y, z) = [g_1(t, y, z), g_2(t, y, z)]^t$. El sistema (18), (19) es de la forma (16) pero tal y como está escrito no es una EDA de índice 1 como la de la Definición ?? , pues U_e depende de la variable independiente t .

Para transformarlo en una EDA de índice 1, basta introducir la variable $y_4 = t$ y definir $\tilde{y} = [y_1, y_2, y_3, y_4]^t$, llegando a

$$f(\tilde{y}, z) = \left[\frac{U_e(y_4) - z_1}{R_0 C_1}, \frac{h(z_1 - y_1 - y_2)}{C_2} - \frac{y_2}{C_2 R_3}, \frac{z_2 - y_3}{C_3 R_5}, 1 \right]^t, \quad (20a)$$

$$g(\tilde{y}, z) = \begin{bmatrix} \frac{U_e(y_4) - z_1}{R_0} + \frac{U_b}{R_2} + (y_1 - z_1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right) - \frac{h(z_1 - y_1 - y_2)}{100} \\ \frac{U_b - z_2}{R_4} - 0,99h(z_1 - y_1 - y_2) + \frac{y_3 - z_2}{R_5} \end{bmatrix}. \quad (20b)$$

Para que los valores iniciales sean consistentes deben verificar $\varphi_1(u(0)) + \varphi_2(u(0)) = 0$ y $\varphi_4(u(0)) + \varphi_5(u(0)) = 0$, siendo φ_i la componente i -ésima de φ . Esto es

$$\begin{aligned} 0 &= \frac{U_1(0)}{R_0} - \frac{U_e(0)}{R_0} + 0,01h(U_2(0) - U_3(0)) - \frac{U_b(0)}{R_2} + U_2(0) \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \\ 0 &= 0,99h(U_2(0) - U_3(0)) - \frac{U_b(0)}{R_4} + \frac{U_4(0)}{R_4} + \frac{U_5(0)}{R_5} \end{aligned}$$

Si tomamos como condiciones iniciales $U_2(0) = U_3(0)$, tenemos que $h(U_2(0) - U_3(0)) = 0$ y en consecuencia

$$\begin{aligned} 0 &= \frac{U_1(0)}{R_0} - \frac{U_e(0)}{R_0} - \frac{U_b}{R_2} + U_2(0) \left(\frac{1}{R_1} + \frac{1}{R_2} \right), \\ 0 &= -\frac{U_b}{R_4} + \frac{U_4(0)}{R_4} + \frac{U_5(0)}{R_5}. \end{aligned}$$

Nótese que $U_e(0) = 0$. Si además tomamos $U_1(0) = U_5(0) = 0$ como condiciones iniciales, obtenemos

$$\begin{aligned} 0 &= -\frac{U_b}{R_2} + U_2(0) \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \\ 0 &= -\frac{U_b}{R_4} + \frac{U_4(0)}{R_4}. \end{aligned}$$

Despejando $U_2(0)$ y $U_3(0)$ se tiene

$$U_2(0) = \frac{U_b R_1}{R_1 + R_2}, \quad U_4(0) = U_b.$$

Por tanto, las condiciones iniciales elegidas han sido

$$U_1(0) = 0, \quad U_2(0) = U_3(0) = \frac{U_b R_1}{R_1 + R_2}, \quad U_4(0) = U_b, \quad U_5(0) = 0.$$

Pasamos ahora a aplicar el método ε -encajado al amplificador de transistores. Las condiciones iniciales para las variables \tilde{y} y z elegidas son

$$\begin{aligned} \tilde{y}(0) &= \left[-\frac{U_b R_1}{R_1 + R_2}, \frac{U_b R_1}{R_1 + R_2}, U_b, 0 \right]^t, \\ z(0) &= [0, U_b]^t. \end{aligned}$$

El tiempo de integración considerado es $[0, 1/5]$ y las gráficas que se muestran corresponden al paso de malla $h = 0,0002$. También hemos fijado $itermax = 100$ y $tol = 10^{-12}$. En la Figura 8 vemos la evolución del voltaje de entrada y de U_5 con el tiempo. Dicha gráfica coincide con la que aparece en [3, p. 379]. Los resultados son similares a los que se presentan en el TFG de la UVA [4] usando métodos de Rosenbrock.

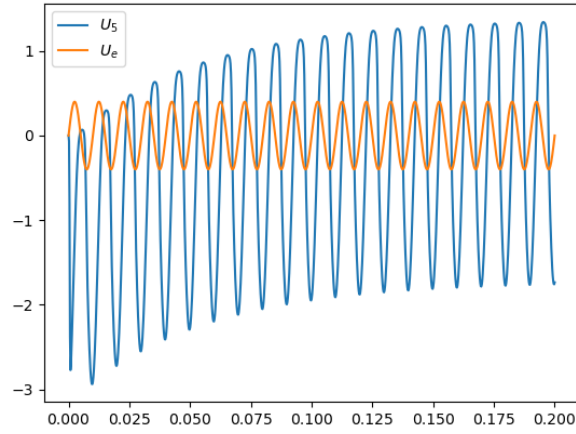


Figura 8: Evolución de U_e y U_5 con el tiempo.

Para aplicar el método ε -encajado necesitamos que la matriz A del método RK que queremos aplicar sea invertible. En el caso del amplificador de transistores hemos aplicado el método Lobatto IIIC de 3 etapas, que es un método de orden 4. En la Figura 9 podemos ver como evolucionan el resto de voltajes con el tiempo.

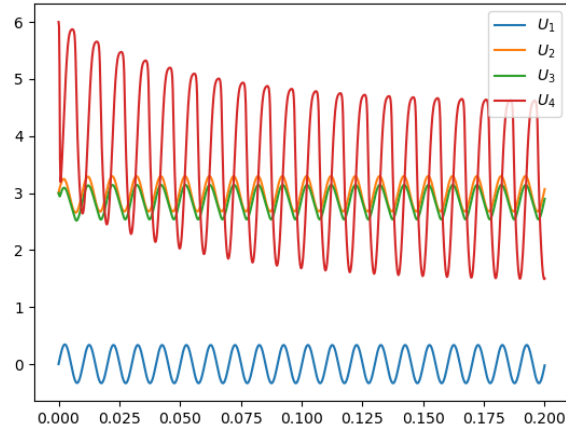


Figura 9: Evolución de los voltajes U_i , $i = 1, \dots, 4$, con el tiempo.

También es importante ver como evoluciona la condición algebraica a medida que avanza el tiempo. En la Figura 10 podemos ver este comportamiento al aplicar el método ε -encajado. En la gráfica superior vemos como evoluciona la condición $g_1(y, z)$. En la gráfica inferior vemos como evoluciona la condición $g_2(y, z)$. Ambas condiciones son del orden de precisión de la máquina.

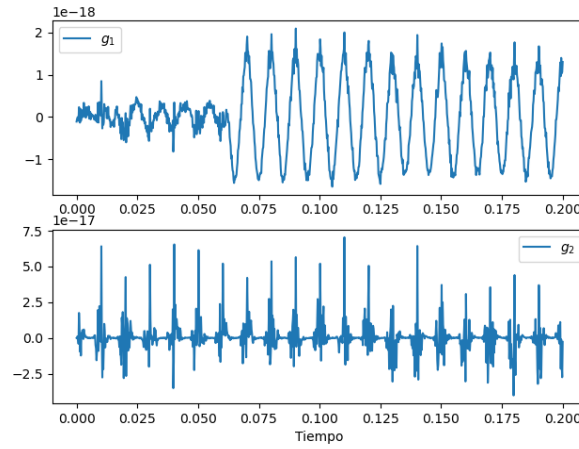


Figura 10: Evolución de la condición algebraica con el tiempo al aplicar el método ε -encajado.

En la Figura 11 hemos representado lo mismo pero aplicando el método state space form. De nuevo, ambas condiciones son del orden de precisión de la máquina.

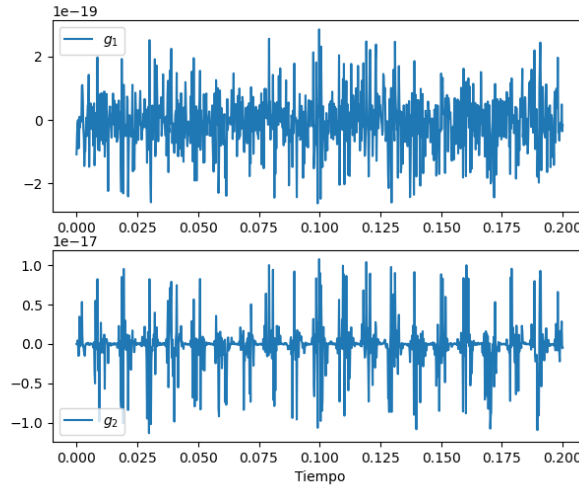


Figura 11: Evolución de la condición algebraica con el tiempo al aplicar el método state space form.

En el Cuadro 2 hemos recogido los tiempos de ejecución de CPU (medido en segundos) al aplicar el método ε -encajado y el método state space form para diferentes pasos de malla. El método ε -encajado es ligeramente más rápido.

h	ε -encajado	State Space Form
0,02	0,4049	0,4388
0,001	0,6597	0,7144
0,005	1,2179	1,3287
0,0025	2,2024	2,3886

Cuadro 2: Comparación de tiempos de ejecución para diferentes pasos de malla.

Referencias

- [1] J.C. Butcher. *A history of Runge-Kutta methods*, volume 20. 1996.
- [2] Ernst Hairer, Syvert Norsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8. 01 1993.
- [3] Ernst Hairer and Gerhard Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, volume 14. 01 1996.
- [4] José Ignacio Segovia Martín. *Introducción a la integración numérica de ecuaciones diferenciales algebraicas de índice 1*. <http://uvadoc.uva.es/handle/10324/40283>, 2019.