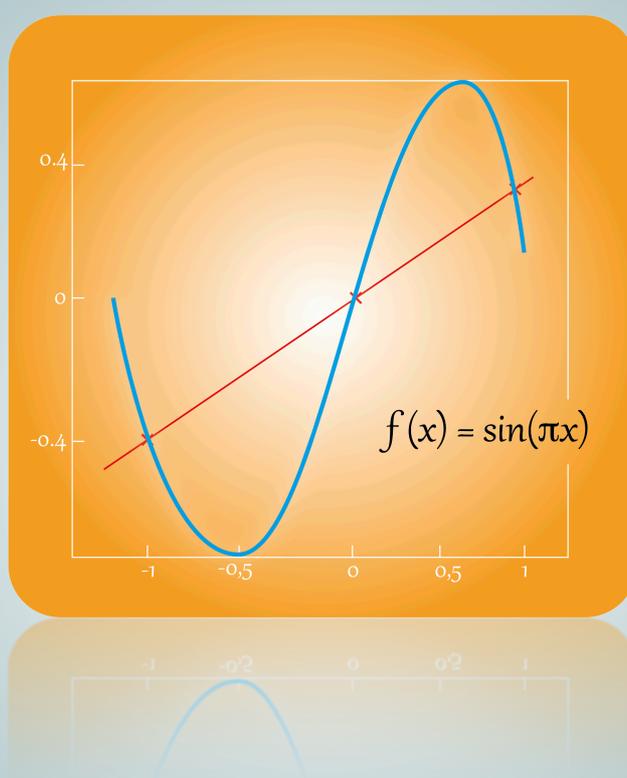


Métodos Numéricos

Capítulo 5. Resolución de problemas de valores iniciales



Carlos Beltrán Álvarez

Departamento de Matemáticas, Estadística y
Computación

Este tema se publica bajo Licencia:

[Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Capítulo 5

Resolución de problemas de valores iniciales

Recordemos el estudio del péndulo simple sin rozamiento, uno de los primeros que se acometen en los cursos de iniciación a la Física. Tenemos una bolita colgando de un hilo de masa despreciable de longitud l , que en un momento se desplaza de su posición vertical de equilibrio y se suelta. Las ecuaciones del movimiento son sencillas de obtener:

$$\frac{d^2}{dt^2} \theta + \frac{g}{l} \sin \theta = 0,$$

donde θ es el ángulo que forma la cuerda con la vertical. La forma habitual de resolver este problema es suponer que θ es muy pequeño, con lo que $\sin \theta \approx \theta$ y la ecuación del péndulo se convierte en una que podemos resolver analíticamente.

Si bien esta forma de resolver el problema tiene algunas virtudes, no deja de ser una aproximación bastante burda. Las suposiciones simplificadoras (cuerda sin masa, no hay rozamiento del aire, la bolita es puntual) son soportables, pero la suposición de que como $\sin \theta \approx \theta$, la solución al problema aproximado será parecida a la solución verdadera es cuanto menos sospechosa. Si además deseamos utilizar un péndulo simple con un ángulo ligeramente mayor, no tenemos otro remedio que tomar la ecuación con el término $\sin \theta$. Lamentablemente, ninguno de los métodos para resolver ecuaciones de segundo grado que el ser humano haya inventado ha conseguido encontrar una expresión (fórmula) directa, sencilla y satisfactoria de la solución al problema propuesto, porque de hecho no existe tal expresión “sencilla”, del mismo modo que muchas de las funciones elementales no tienen una primitiva con una fórmula sencilla como dice el Teorema 3.0.1.

En este capítulo responderemos a la pregunta: ¿cómo podemos resolver numéricamente las ecuaciones del movimiento del péndulo simple? Nótese que una cantidad ingente de problemas de la Física comparten las características del anterior. Entre los problemas que motivaron el estudio intensivo de métodos numéricos para la resolución de este tipo de cuestiones se encuentra el estudio de las trayectorias de cohetes y asteroides. Hoy en día la resolución de problemas de valores iniciales como el que abre este capítulo es una herramienta fundamental en el estudio del mundo que nos rodea, y existen métodos extremadamente precisos y confiables que nos permiten conocer en gran cantidad de situaciones cómo evolucionarán las características físicas de un sistema. Los métodos para predecir la climatología, para comprender la economía de los estados, para analizar la expansión de las enfermedades, para estudiar el control que ejerce nuestro cerebro sobre las distintas concentraciones de sustancias en la sangre, y un larguísimo etcétera, pasan por resolver problemas de valores iniciales como los que estudiamos en este capítulo, y por sus primos mayores (problemas con valores en la frontera y problemas de

ecuaciones en derivadas parciales) que no tenemos tiempo de tratar.

Será útil para el lector recordar los resultados teóricos elementales concernientes a los problemas de valores iniciales, con lo que incluimos un pequeño repaso en la Sección 5.1.

En este capítulo trataremos primero de resolver problemas de valores iniciales (o de Cauchy) de la forma

$$\begin{cases} \dot{y} = f(t, y) \\ y(a) \text{ conocido} \end{cases} \quad (5.1)$$

con $t \in [a, b]$.

5.1. Teoría elemental de los problemas de valor inicial

Comenzamos con algunas definiciones y resultados básicos que nos permitirán enunciar teoremas de existencia y unicidad de solución para un sistema de ecuaciones diferenciales ordinarias. A lo largo de esta sección consideraremos un conjunto Ω abierto de la forma $J \times U \subseteq \mathbb{R} \times \mathbb{R}^n$ con J un intervalo abierto y U un abierto de \mathbb{R}^n .

Definición 5.1.1 *Se dice que una función $f : \Omega \rightarrow \mathbb{R}^n$ satisface una condición de tipo Lipschitz en la variable y en un conjunto $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ si existe una constante $L > 0$ tal que*

$$(t, y_1), (t, y_2) \in \Omega \implies \|f(t, y_1) - f(t, y_2)\| \leq L\|y_1 - y_2\|.$$

Llamamos a la menor L que cumple esta propiedad “Constante de Lipschitz” para f .

Definición 5.1.2 *Se dice que un conjunto $U \subseteq \mathbb{R}^n$ es convexo si $y_1, y_2 \in U$ implica que todo el segmento entre ambos puntos está contenido en U .*

Teorema 5.1.3 *Sea $f : \Omega \rightarrow \mathbb{R}^n$ con $\Omega = J \times U$ con U convexo tal que la norma de la Jacobiana con respecto a las variables y , $\|J_y f(t, y)\|$, está acotada por $L > 0$ para todos los i, j . Entonces, f satisface una condición de tipo Lipschitz en la variable y con constante de Lipschitz L .*

DEMOSTRACIÓN. Sean $p_1 = (t, y_1), p_2 = (t, y_2) \in \Omega$. Sea $\gamma(t) = (1-t)p_1 + tp_2$, $t \in [0, 1]$ el segmento que une ambos puntos. Entonces, por el Teorema Fundamental del Cálculo Vectorial,

$$f(p_2) = f(p_1) + \int_0^1 J_y f(\gamma(t)) \dot{\gamma}(t) dt,$$

de donde

$$\|f(p_2) - f(p_1)\| \leq \int_0^1 \|J_y f(\gamma(t)) \dot{\gamma}(t)\| dt \leq \int_0^1 \|J_y f(\gamma(t))\| \|\dot{\gamma}(t)\| dt \leq L \int_0^1 \|\dot{\gamma}(t)\| dt = L\|p_1 - p_2\|,$$

como queríamos. □

Vemos ahora un primer resultado de existencia y unicidad.

Teorema 5.1.4 *Sea $f : \Omega \rightarrow \mathbb{R}^n$ continua en Ω y de tipo Lipschitz en la variable y . Para cualquier condición inicial $y(t_0) = y_0$ (con $(t_0, y_0) \in \Omega$), existe una única solución $y(t)$ al problema $\dot{y}(t) = f(t, y)$, definida en algún intervalo temporal que contiene a t_0 . Adicionalmente, si f es de tipo C^k entonces la solución es de tipo C^{k+1} para $k = 1, 2, \dots, \infty$.*

DEMOSTRACIÓN.

La demostración de este teorema (una versión del Teorema de Piccard–Lindelöf) y del otros resultados teóricos concernientes a la resolución de problemas de valores iniciales se puede encontrar en cualquier libro de ecuaciones diferenciales, por ejemplo “Ordinary Differential Equations with Applications” de Chicone.

□

Nótese que el Teorema 5.1.4 solo garantiza la existencia y unicidad de solución en un intervalo temporal que puede ser muy pequeño. Es posible ver que la solución no está definida necesariamente para todo $t > t_0$, por ejemplo considerando el problema $y' = y^2$, $y(0) = 1$ que tiene como solución $y(t) = (1 - t)^{-1}$. Sin embargo, en ese caso necesariamente tenemos:

Teorema 5.1.5 *En las hipótesis del Teorema 5.1.4, si suponemos que f es de tipo C^1 , entonces hay un intervalo abierto maximal en el que se puede definir la solución $y(t)$, y si es de la forma (α, β) con $\beta \neq \infty$ entonces o bien la solución no está acotada cuando $t \rightarrow \beta$ o bien la solución se aproxima a la frontera de U cuando $t \rightarrow \beta$, y lo mismo se puede decir de α cambiando ∞ por $-\infty$.*

Damos finalmente otro teorema de existencia y unicidad global en el caso (bastante frecuente) de que $\Omega = J \times \mathbb{R}^n$ con J un intervalo.

Teorema 5.1.6 *En las condiciones del Teorema 5.1.4, si $\Omega = J \times \mathbb{R}^n$, y si f satisface una condición de tipo Lipschitz en la variable y en Ω , entonces la solución al problema de valores iniciales está definida en todo el intervalo J .*

La gran mayoría de las ecuaciones diferenciales ordinarias de la Física (exceptuando aquellos fenómenos en los que intervienen procesos violentos como choques) son de la forma $\dot{y} = f(t, y)$ con f de tipo C^∞ , con lo que se aplicará el teorema de existencia y unicidad 5.1.4. Los teoremas 5.1.5 y 5.1.6 nos ayudarán a veces a decidir qué está sucediendo cuando intentamos resolver un problema numéricamente y obtenemos soluciones divergentes o extrañas.

5.2. Reducción al caso de problemas de primer orden

En general, los problemas de valores iniciales de la Física son ecuaciones de segundo orden, en el sentido de que aparecen funciones, sus primeras y sus segundas derivadas. El ejemplo típico es el del oscilador armónico (clásico): $\ddot{x} = -kx$, pero el lector puede sin duda proponer otras cuantas ecuaciones del estilo.

Sin embargo, la gran mayoría de los métodos que se han desarrollado históricamente (y todos los que mencionaremos en este capítulo) están diseñados para problemas de primer orden, esto es, problemas como (5.1) en los que aparece solo la primera derivada aunque, eso sí, se permite que la función que se busca (la solución $y(t)$) sea vectorial, esto es $y(t) \in \mathbb{R}^n$ para algún $n \geq 1$. Por eso se suele hablar de sistemas de ecuaciones diferenciales ordinarias, viniendo el término “sistemas” del hecho de que podemos verlo como una colección de ecuaciones

$$\dot{y}_1 = f_1(t, y), \quad \dots, \quad \dot{y}_n = f_n(t, y).$$

Nótese que en general $f_1(t, y)$ no depende solo de t, y_1 sino de t, y_1, \dots, y_n , esto es las ecuaciones están ligadas entre sí.

Los resultados teóricos de existencia y unicidad de la sección 5.1 son también para problemas de primer orden. Esto nos podría hacer pensar (o hacer temer) que fuese necesario desarrollar una teoría para ecuaciones de segundo orden, otra para ecuaciones de tercer orden y así sucesivamente. Afortunadamente, no tenemos que pasar por este proceso, pues un sencillo truco permite convertir ecuaciones de orden superior en ecuaciones de

primer orden. Ilustramos este hecho con un ejemplo concreto: la ecuación del péndulo simple con condiciones iniciales $\theta_0 = 1$, $\dot{\theta}_0 = 0$ es

$$\begin{cases} \ddot{\theta} + \frac{g}{l} \sin \theta = 0 \\ \theta(0) = 1, \dot{\theta}(0) = 0 \end{cases} \quad (5.2)$$

Para transformarlo en un sistema de primer orden, consideramos una función vectorial $y(t)$ con dos coordenadas, de forma que la primera sea $\theta(t)$ y la segunda sea $\dot{\theta}(t)$ donde $\theta(t)$ es la solución de (5.2). Esto es,

$$y(t) = \begin{pmatrix} \theta(t) \\ \dot{\theta}(t) \end{pmatrix}.$$

A continuación observamos que se cumple:

$$\dot{y}(t) = \begin{pmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} = \begin{pmatrix} \dot{\theta}(t) \\ \ddot{\theta}(t) \end{pmatrix} = \begin{pmatrix} y_2(t) \\ -\frac{g}{l} \sin y_1(t) \end{pmatrix}$$

Por lo tanto la función vectorial $y(t)$ es solución del problema:

$$\begin{cases} \dot{y} = \begin{pmatrix} y_2 \\ -\frac{g}{l} \sin y_1 \end{pmatrix} \\ y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{cases} \quad \text{esto es} \quad \begin{cases} \dot{y} = f(t, y) \\ y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{cases} \quad \text{con} \quad f(t, y) = \begin{pmatrix} y_2 \\ -\frac{g}{l} \sin y_1 \end{pmatrix} \quad (5.3)$$

que sí que entra en la teoría vista en la Sección 5.1 y de hecho por el Teorema 5.1.6 sabemos que tiene una solución única definida para todo $t > 0$.

Este procedimiento para convertir problemas de orden mayor que uno en problemas de primer orden es frecuentemente el primer paso que debemos tomar para resolver numéricamente un problema definido a partir de una o más ecuaciones diferenciales.

5.3. Métodos basados en la expansión de Taylor: Euler y Euler modificado.

Una colección de métodos para resolver el problema (5.1) se basa en el uso de la aproximación de Taylor (Teorema 1.3.1), o más bien de su versión para funciones de n variables que no hemos introducido aún.

5.3.1. Método de Euler

El método de Euler utiliza la versión más sencilla de la aproximación dada por el Teorema de Taylor:

$$y(t+h) \approx y(t) + h\dot{y}(t).$$

Con ello, y mirando (5.1), tenemos

$$y(t+h) \approx y(t) + hf(t, y(t)),$$

de donde, a partir del conocimiento de y en el momento t , podemos inferir el valor de y en el momento $t+h$. Como $y(a)$ es conocido, eligiendo h “pequeño” y repitiendo este proceso tenemos aproximaciones para $y(a+h)$, $y(a+2h)$, \dots , $y(b)$. Notamos ya que nuestra aproximación a la solución no será dada por una fórmula, sino por una sucesión de puntos que son los valores de y en una discretización del intervalo temporal.

En otras palabras, una vez elegimos h (normalmente de la forma $h = (b - a)/N$ donde $N \gg 1$ es algún número natural) generamos recursivamente una sucesión

$$w_0 = y(a), \quad w_{k+1} = w_k + hf(t + kh, w_k),$$

que satisface $w_k \approx y(a + kh)$.

Ejemplo 5.3.1 Consideremos el problema de valores iniciales en $t \in [0, 1]$:

$$\begin{cases} \dot{x} = 2t \\ x(0) = 0 \end{cases}$$

cuya solución exacta conocemos: $x(t) = t^2$. Tomando $h = 0,1$ veamos el resultado que nos ofrece el método de Euler: El resultado exacto es $x(1) = 1$ y nuestra aproximación es $x(1) \approx w_{10} = 0,9$, luego cometemos un error

$x(kh)$	w_k	Resultado
$x(0)$	w_0	0
$x(0,1)$	w_1	0
$x(0,2)$	w_2	0,02
$x(0,3)$	w_3	0,06
$x(0,4)$	w_4	0,12
$x(0,5)$	w_5	0,2
$x(0,6)$	w_6	0,3
$x(0,7)$	w_7	0,42
$x(0,8)$	w_8	0,56
$x(0,9)$	w_9	0,72
$x(1)$	w_{10}	0,9

de una milésima. Si disminuimos h , podemos hacer más pequeño este error.

Ejemplo 5.3.2 Consideremos ahora el problema de valores iniciales en $t \in [0, 0,5]$:

$$\begin{cases} \dot{x} = x \\ x(0) = 1 \end{cases}$$

cuya solución exacta conocemos: $x(t) = e^t$. Tomando $h = 0,1$ veamos el resultado que nos ofrece el método de Euler: El resultado exacto es $x(0,5) = \sqrt{e} = 1,6487212707 \dots$ y nuestra aproximación es $x(1) \approx w_5 = 1,61051$, luego cometemos un error de unas cuatro milésimas. Como en el caso anterior, si disminuimos h , podemos hacer más pequeño este error.

Ejemplo 5.3.3 Escribir en forma de problema de Cauchy de primer orden el problema:

$$\begin{cases} \ddot{x} = \frac{\cos(\sin(xt))}{10} \\ x(0) = 1, \quad x'(0) = -1 \end{cases}$$

$x(kh)$	w_k	Resultado
$x(0)$	w_0	1
$x(0,1)$	w_1	1,1
$x(0,2)$	w_2	1,21
$x(0,3)$	w_3	1,331
$x(0,4)$	w_4	1,4641
$x(0,5)$	w_5	1,61051

donde $t \in [0, 2]$, $x = x(t)$. Calcular aproximadamente el valor de $x(2)$ utilizando el método de Euler con paso $h = 1$. *Respuesta: consideramos la función $w = w(t)$ con coordenadas $w_1(t) = x(t)$, $w_2(t) = \dot{x}(t)$. Con ello, el problema se lee:*

$$\begin{cases} \dot{w} = F(t, w) = \begin{pmatrix} \frac{w_2}{10} \\ \cos(\sin(w_1 t)) \end{pmatrix} \\ w(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{cases}$$

El cálculo de $x(2)$ se sigue del de $w(2)$:

$$w(1) \approx w(0) + F(0, w(0)) = \begin{pmatrix} 0 \\ -0,9 \end{pmatrix},$$

$$w(2) \approx w(1) + F(1, w(1)) = \begin{pmatrix} -0,9 \\ -0,8 \end{pmatrix}.$$

Con ello, $x(2) \approx -0,9$.

Ejemplo 5.3.4 Realizamos un último ejemplo, en este caso en una función vectorial. Para ello tomamos el ejemplo del péndulo simple, como se describe en (5.3). Tomemos $l = g$. Entonces, el problema es:

$$\begin{cases} \dot{y} = \begin{pmatrix} y_2 \\ -\sin y_1 \end{pmatrix} \\ y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{cases} \quad (5.4)$$

Resolvemos este problema con el método de Euler con $h = 0,1$ en el intervalo temporal $[0, 0,5]$:

$y(kh)$	w_k	Resultado
$y(0)$	w_0	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$y(0,1)$	w_1	$\begin{pmatrix} 1 \\ -0,084147 \end{pmatrix}$
$y(0,2)$	w_2	$\begin{pmatrix} 0,9915852 \\ -0,168294 \end{pmatrix}$
$y(0,3)$	w_3	$\begin{pmatrix} 0,974755 \\ -0,25198 \end{pmatrix}$
$y(0,4)$	w_4	$\begin{pmatrix} 0,9495575 \\ -0,33474 \end{pmatrix}$
$y(0,5)$	w_5	$\begin{pmatrix} 0,91608 \\ -0,41605 \end{pmatrix}$

Por lo tanto la respuesta que obtenemos para la posición y velocidad en el momento $t = 0,5$ es:

$$\theta(0,5) \approx (w_5)_1 = 0,91608, \quad \dot{\theta}(0,5) = (w_5)_2 = -0,41605.$$

Al contrario que en los casos anteriores, y como nos sucederá casi siempre en las aplicaciones, en este caso no conocemos la respuesta exacta de modo que no podemos conocer de manera precisa el error que estamos cometiendo. Veremos que mediante el uso de algunos teoremas sí que podemos, en algunos casos, acotar este error.

Aunque el método de Euler es poco preciso, su sencillez lo hace más fácil de comprender que otros métodos más sofisticados. Más concretamente, tenemos el siguiente resultado:

Teorema 5.3.5 *Supongamos que la ecuación (5.1) es tal que f es de tipo C^1 en $[a, b] \times \mathbb{R}^n$, acotada, y de tipo Lipschitz en la variable y con constante de Lipschitz L . Sea $y(t)$ la solución (que existe y es única según el Teorema 5.1.6), y supongamos que existe una constante $M > 0$ con la propiedad de que $\|\dot{y}(t)\| \leq M$ para $t \in [a, b]$. Entonces la sucesión generada por el método de Euler satisface:*

$$\|y(t + kh) - w_k\| \leq \frac{hM}{2L}(e^{kLh} - 1).$$

DEMOSTRACIÓN.

La demostración se sale del ámbito de este curso, puede encontrarse por ejemplo en el libro de Burden y Faires. \square

El Teorema 5.3.5 nos dice que, bajo ciertas condiciones en la ecuación (5.1), si tomamos h suficientemente pequeño, el método de Euler aproximará correctamente la solución. Incluimos el siguiente corolario que nos permite calcular la constante M en algunos casos.

Corolario 5.3.6 *En el Teorema 5.3.5, podemos tomar*

$$L = \sup_{(t,y) \in [a,b] \times \mathbb{R}^n} \|J_y f(t, y)\|,$$

y

$$M = \sup_{(t,y) \in [a,b] \times \mathbb{R}^n} \left\| \frac{\partial f}{\partial t}(t, y) + J_y f(t, y)f(t, y) \right\| \leq \sup_{(t,y) \in [a,b] \times \mathbb{R}^n} \left(\left\| \frac{\partial f}{\partial t}(t, y) \right\| + L\|f(t, y)\| \right),$$

si este número es finito. Más aún, podemos cambiar en el supremo $[a, b] \times \mathbb{R}^n$ por $[a, b] \times U$ si sabemos que $y(t) \subseteq U$ para todo $t \in [a, b]$.

DEMOSTRACIÓN. Primero observamos por el Teorema 5.1.3 que podemos tomar L como se dice en el corolario. Segundo, notemos que si $y(t)$ es la solución de (5.1) entonces

$$\ddot{y}(t) = \frac{d}{dt}(\dot{y}(t)) = \frac{d}{dt}(f(t, y(t))) = \frac{\partial f}{\partial t}(t, y(t)) + J_y f(t, y(t))f(t, y(t)),$$

luego la afirmación del corolario es una consecuencia inmediata del Teorema 5.3.5. \square

Ejemplo 5.3.7 *En la ecuación del péndulo simple (5.3), $f(t, y)$ es de tipo C^1 (hecho es de tipo C^∞) y tenemos*

$$\frac{\partial f}{\partial t}(t, y) = 0, \quad J_y f = \begin{pmatrix} 0 & 1 \\ -\frac{g}{l} \cos y_1 & 0 \end{pmatrix},$$

con lo que por el Corolario 5.3.6 podemos tomar:

$$L = \sup_{(t,y) \in \mathbb{R} \times \mathbb{R}^2} \|J_f(t, y)\| = \max(1, g/l),$$

$$M = \sup_{(t,y) \in \mathbb{R} \times \mathbb{R}^2} \|J_y f(t,y) f(t,y)\| = \sup_{y \in \mathbb{R}^2} \left\| \begin{pmatrix} 0 & 1 \\ -\frac{g}{l} \cos y_1 & 0 \end{pmatrix} \begin{pmatrix} y_2 \\ -\frac{g}{l} \sin y_1 \end{pmatrix} \right\| = \sup_{y \in \mathbb{R}^2} \left\| \begin{pmatrix} -\frac{g}{l} \sin y_1 \\ -y_2 \frac{g}{l} \cos y_1 \end{pmatrix} \right\|.$$

Ahora utilizemos un poquito de nuestros conocimientos físicos para acotar este último término: la velocidad y_2 es máxima cuando el péndulo se encuentra en posición vertical, y en este caso por la conservación de la energía tenemos (llamando h_0 a la diferencia entre las alturas al comienzo y en ese momento):

$$\frac{1}{2} m y_2^2 = m g h_0 = m g l (1 - \cos 1),$$

por lo que tenemos

$$y_2^2 \leq 2gl(1 - \cos 1),$$

y podemos tomar

$$M = \frac{g}{l} \sup_{y \in \mathbb{R}^2} \sqrt{\sin^2 y_1 + 4g^2 l^2 (1 - \cos 1)^2 \cos^2 y_2} \leq \frac{g}{l} \sqrt{1 + 0,85g^2 l^2}.$$

Por el Corolario 5.3.6, y tomando $l = 1$ m para simplificar los cálculos, podemos entonces decir que con longitud de paso h tras k pasos el error cometido por el método es a lo sumo

$$\|y(t + kh) - w_k\| \leq \frac{hM}{2L} (e^{kLh} - 1) \leq \frac{h\sqrt{1 + 0,85g^2}}{2} (e^{kgh} - 1).$$

En particular, si queremos averiguar el error cometido al calcular la posición del péndulo en el momento $t = 1$ con paso $h = 1/N$, esto es, tras $k = N$ pasos, lo podemos acotar por

$$\frac{\sqrt{1 + 0,85g^2}}{2N} (e^g - 1).$$

Esto es, tomando por ejemplo $N = 10^7$ cometeremos un error de a lo sumo (aprox.) tres milésimas. Nótese que este error es en la norma de y , esto es, hablamos de la raíz cuadrada de la suma de los cuadrados de los errores en la posición y la velocidad angulares.

Lamentablemente en muchos otros problemas no se cumplen las condiciones del Teorema 5.3.5 (ni de su corolario). Adicionalmente, el teorema no cuenta con la aparición de errores de redondeo (solo es válido en las condiciones, nunca satisfechas, de la posibilidad de realizar cálculos exactos). En general, en la resolución de problemas como (5.1) es imposible garantizar la precisión de los cálculos, y solo para algunos casos muy concretos se ha conseguido obtener teoremas que garanticen que la verdadera solución es parecida a la numérica. Sin embargo, con los métodos que veremos más adelante, se ha comprobado repetidas veces en la práctica que es posible llevar a cabo tareas complejÍsimas (como hacer aterrizar una nave en Venus, pongamos) utilizando estos cálculos, incluso en la presencia de errores de redondeo.

Una implementación en Matlab del método de Euler sería:

```
function [t,w]=odeeuler(f,a,b,ya,N)
% Resuelve \dot y=f(t,y), y(a)=ya, en el intervalo [a,b] con el
% metodo de Euler con paso h=(b-a)/N.
% En la variable t se devuelven los tiempos y en la y los valores.
% De esta forma, w(k) sera (aprox.) y(t=a+kh).
```

```
h=(b-a)/N;
t=linspace(a,b,N+1);
```

```
w=zeros(length(ya),length(t)); % Para que tenga la dimension correcta desde el principio
w(:,1)=ya;
for k=2:N+1
    ya=ya+h*f(a,ya); % Ahora ya= y(a+kh) .
    w(:,k)=ya;
    a=a+h;
end
```

5.3.2. Método de Euler modificado

Dado el problema (5.1), el método de Euler modificado (llamado también método del punto medio, o método de Runge–Kutta de orden 2) consiste en aproximar $y(kh)$ por w_k , donde h es un número pequeño, $w_0 = y_0$ y

$$w_{k+1} = w_k + hf(kh + h/2, w_k + (h/2)f(kh, w_k)).$$

De este modo, al igual que en el caso del método de Euler, generamos una secuencia w_0, w_1, w_2, \dots que proporciona valores aproximados para los valores exactos $y(0), y(h), y(2h), \dots$. El origen de este método es simplemente partir del método de Euler

$$w_{k+1} = w_k + hf(kh, w_k),$$

y sustituir el valor temporal kh y el valor espacial w_k donde se está evaluando f por el valor temporal a mitad del intervalo $[kh, (k + 1)h]$ y la aproximación dada por el método de Euler para el valor espacial en la mitad de ese intervalo.

Ejemplo 5.3.8 Consideremos el siguiente problema de Cauchy, que describe el movimiento de un péndulo simple de longitud ℓ sin rozamiento cuando desplazamos la masa un ángulo de $\pi/2$ radianes de la vertical y lo dejamos caer.

$$\theta'' = \frac{\ell}{m} \sin(\theta), \quad \theta(0) = \pi/2, \quad \theta'(0) = 0.$$

Se pide calcular aproximadamente la posición de la masa tras 1 segundo asumiendo que $\ell = 3,27m$, $g = 9,81m/s^2$. *Solución: reescribimos el problema en forma de primer orden:*

$$W(t) = \begin{pmatrix} \theta(t) \\ \theta'(t) \end{pmatrix} \Rightarrow W'(t) = \begin{pmatrix} W_2(t) \\ -\frac{1}{3} \sin W_1(t) \end{pmatrix}, \quad W(0) = \begin{pmatrix} \frac{\pi}{2} \\ 0 \end{pmatrix},$$

con lo que tenemos un problema de Cauchy con

$$f(t, W) = \begin{pmatrix} W_2 \\ -\frac{1}{3} \sin W_1 \end{pmatrix}.$$

A continuación elegimos $h = 0,1$ y definimos

$$w_0 = W(0) = \begin{pmatrix} \frac{\pi}{2} \\ 0 \end{pmatrix}, \quad w_{k+1} = w_k + hf(kh + h/2, w_k + (h/2)f(kh, w_k)),$$

de forma que según lo estudiado arriba tenemos $w_k \approx W(kh)$. Dado que queremos $W(1)$, tendremos que calcular $w_0, w_1, \dots, w_9, w_{10} = W(1)$. Podemos hacer esto con la ayuda del ordenador, obteniendo:

t	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$\theta(t) = W_1(t)$	1.5708	1.5691	1.5641	1.5558	1.5441	1.5291	1.5108	1.4891	1.4642	1.4359	1.4043
$\theta'(t) = W_2(t)$	0	-0.0333	-0.0667	-0.1000	-0.1333	-0.1666	-0.1999	-0.2332	-0.2664	-0.2995	-0.3324

Con ello, después de 1 segundo, $\theta(t = 1) \approx 1,4642$, que era lo que nos pedía el enunciado.

Ejemplo 5.3.9 Escribe en forma de problema de Cauchy el siguiente problema:

$$\begin{cases} \ddot{x} = 2^{-x} + t \\ x(0) = 2, \quad x'(0) = 1 \end{cases}$$

definido para $t > 0$. Calcula aproximadamente el valor de $x(1)$ utilizando el método de Euler modificado con paso $h = 1$. *Respuesta: consideramos la función $w = w(t)$ con coordenadas $w_1(t) = x(t)$, $w_2(t) = \dot{x}(t)$. Con ello, el problema se lee:*

$$\begin{cases} \dot{w} = F(t, w) = (2^{-w_1} + t) \\ w(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \end{cases}$$

El cálculo de $x(1)$ se sigue del de $w(1)$:

$$\begin{aligned} w(1) &\approx w(0) + F(0,5, w(0) + 0,5F(0, w(0))) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + F\left(0,5, \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0,5\begin{pmatrix} 1 \\ 2^{-2} \end{pmatrix}\right), \\ &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} + F\left(0,5, \begin{pmatrix} 2,5 \\ 1,125 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1,125 \\ 2^{-2,5} + 0,5 \end{pmatrix} = \begin{pmatrix} 3,125 \\ 1,6768 \end{pmatrix}. \end{aligned}$$

Con ello, $x(2) \approx 3,125$.

5.4. Métodos de Runge–Kutta

La estrella de los métodos para resolver el problema (5.1) es el (los) llamado(s) método(s) de Runge–Kutta. Estos son una familia de métodos basados en una interpretación diferente del Teorema de Taylor en varias variables, y hay numerosas versiones, cada una de las cuales tiene sus virtudes y defectos. El método de Euler modificado que hemos estudiado es de hecho el método de la familia Runge–Kutta más sencillo.

En lugar de explayarnos sobre la derivación de estos métodos y sobre las distintas variantes (cuestión que nos llevaría mucho tiempo y esfuerzo), escribimos únicamente el llamado Runge–Kutta de orden 4 (RK4 para los amigos), que es el método más utilizado en la práctica para la resolución de problemas de valores iniciales. El método se define como sigue:

$$\begin{aligned} t_0 &= a \\ w_0 &= y(a) \\ \forall k \geq 0 \text{ (mientras se tenga } t_{k+1} \leq b) : \\ k_1 &= hf(t_k, w_k) \\ k_2 &= hf\left(t_k + \frac{h}{2}, w_k + \frac{k_1}{2}\right) \\ k_3 &= hf\left(t_k + \frac{h}{2}, w_k + \frac{k_2}{2}\right) \\ k_4 &= hf(t_{k+1}, w_k + k_3) \\ w_{k+1} &= w_k + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ t_{k+1} &= t_k + h. \end{aligned}$$

Una implementación en Matlab del método de Runge–Kutta 4 sería:

```

function [t,w]=odeRK4(f,a,b,ya,N)
% Resuelve \dot y=f(y), y(a)=ya, en el intervalo [a,b] con el
% metodo RK4 con paso h=(b-a)/N.
% En la variable t se devuelven los tiempos y en la y los valores.
% De esta forma, w(k) sera (aprox.) y(t=a+kh).

h=(b-a)/N;
t=linspace(a,b,N+1);
w=zeros(length(ya),length(t)); % Para que tenga la dimension correcta desde el principio
w(:,1)=ya;
for k=2:N+1
    k1=h*f(a,ya);
    k2=h*f(a+h/2,ya+k1/2);
    k3=h*f(a+h/2,ya+k2/2);
    a=a+h;
    k4=h*f(a,ya+k3);
    ya=ya+(k1+2*k2+2*k3+k4)/6; % Ahora ya= y(a+kh).
    w(:,k)=ya;
end

```

Debemos señalar que en las aplicaciones prácticas es frecuente utilizar métodos llamados multipaso, esto es métodos tales que el tamaño h del paso cambia con el tiempo. La idea es actuar de un modo similar a como se indicó en la sección 3.3.1: primero ejecutamos un método de Runge-Kutta con un paso dado, luego con un paso la mitad de pequeño, y si el resultado es muy similar concluimos que la aproximación es buena y seguimos adelante. De hecho, es más habitual usar el mismo tamaño de paso pero usar dos métodos distintos (un método RK4 y otro RK5 que no hemos comentado, por ejemplo) y comparar los resultados. El tamaño del paso, a su vez, se incrementará si los resultados son muy parecidos y se disminuirá en otro caso. Los detalles precisos de este proceso, que se pueden organizar de muchas maneras diferentes, se escapan al alcance de este curso, pero incluimos el código Matlab de una versión popular de este proceso, llamado algoritmo de Runge-Kutta-Fehlberg. Nótese que en él aparecen muchas constantes (por ejemplo los coeficientes multiplicando a distintas magnitudes, o el 0,84 en la fórmula de delta) cuyo valor no hemos justificado en absoluto, pero cuyo cálculo ha sido realizado por expertos del área y comprobado en multitud de ocasiones y de problemas prácticos.

```

function [t,w]=odeRKFehlberg(f,a,b,ya,hmin,hmax,epsilon)
% Resuelve \dot y=f(y), y(a)=ya, en el intervalo [a,b] con el
% metodo RK Fehlberg.
% En la variable t se devuelven los tiempos y en la y los valores.
% De esta forma, w(k) sera (aprox.) y(t=a+kh).
% Epsilon marca la precision deseada.

h=hmax;
w=zeros(length(ya),1); % Para que tenga las filas correctas desde el principio
w(:,1)=ya;
contador=2;
t=a;
while (contador>0) && (b-a)>0
    k1=h*f(a,ya);
    k2=h*f(a+h/4,ya+k1/4);
    k3=h*f(a+3/8*h,ya+3/32*k1+9/32*k2);
    k4=h*f(a+12/13*h,ya+1932/2197*k1-7200/2197*k2+7296/2197*k3);
    k5=h*f(a+h,ya+439/216*k1-8*k2+3680/513*k3-845/4104*k4);
    k6=h*f(a+h/2,ya-8/27*k1+2*k2-3544/2565*k3+1859/4104*k4-11/40*k5);

```

```

R=norm(k1/360-128/4275*k3-2197/75240*k4+k5/50+2/55*k6)/h;
if R<epsilon % Estamos en condiciones de continuar
    a=a+h;
    ya=ya+25/216*k1+1408/2565*k3+2197/4104*k4-k5/5;
    w(:,contador)=ya;
    contador=contador+1;
    delta=0.84*(epsilon/R)^(1/4);
    if delta<.1 % Elegimos el paso nuevo con varios casos
        h=h*.1;
    elseif delta>4
        h=min([hmax,b-a,h*4]);
    else
        h=min([hmax,b-a,delta*h]);
    end
    t=[t,a];
else
    h=h/2; % Si no podemos continuar con ese h, lo dividimos por dos
end
if (h<hmin) && (a~=b)
    error('Se ha rebasado el valor minimo de h');
end
end
end

```

Finalizamos esta sección indicando que en Matlab existen diversas funciones para resolver sistemas de ecuaciones diferenciales, como `ode23` ó `ode45`. En Octave, se puede utilizar la función `lsode`. En general, estos programas están basados en variantes de Runge–Kutta.

5.5. Exercises. ODEs

Exercise 5.1

Discuss, for each of the following functions, if they are Lipschitz. In the case they are, compute or estimate their Lipschitz constant.

- Defined in \mathbb{R} :

$$f_1(x) = |x| \quad f_2(x) = \frac{1}{1+x^2} \quad f_3(x) = e^x \quad f_4(x) = \frac{\sin(x)}{x}$$

- Defined in $(0, 1]$:

$$g_1(x) = x^2 \quad g_2(x) = e^{x^2} \quad g_3(x) = \frac{1}{\sqrt{x}} \quad g_4(x) = \sqrt{x} \quad g_5(x) = \frac{1}{\log(\sin x)}$$

- Functions of several variables, from \mathbb{R}^2 to \mathbb{R}^2 :

$$h_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{xy}{1+x^2} \\ \exp(-x^2 - y^2) \end{pmatrix} \quad h_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ \frac{1}{1+x^2+y^4} \end{pmatrix} \quad h_3 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sin(x+y) \\ \cos(x+y^2) \end{pmatrix}$$

Exercise 5.2

Write down as first order problems the following initial value problems:

$$\begin{cases} m\ddot{x} = -kx \\ x(0) = \ell \\ \dot{x}(0) = 0 \end{cases} \quad \begin{cases} \ddot{r} = -\frac{GM_{\text{Sun}}}{\|r\|^3} r \\ v(0) = \begin{pmatrix} 150 \cdot 10^9 \\ 0 \end{pmatrix} \\ \dot{v}(0) = 30000 \end{cases} \quad \begin{cases} m\ddot{x} = -mg + \frac{ke^{-\tau t}}{x^2} \\ x(0) = 1 \\ \dot{x}(0) = 0 \end{cases}$$

The first one describes a classical harmonic oscillator; the second one describes approximately the orbit of Earth in the solar orbit (r is a two dimensional vector, G is the universal gravitational constant, and the Sun is still at the origin); the third problem is a simplification of the problem of having a light-weight object floating in the air by the force of a strong current of air that decreases with time.

Exercise 5.3

Compute (with pencil and paper) the value of $y(1)$ with Euler's method, with $h = 1/2$ and then with $h = 1/4$ in the following problems:

$$\begin{cases} x + y' = y \\ y(0) = 1 \end{cases}, \quad \begin{cases} f' = f^2 \\ f(0) = 2 \end{cases}, \quad \begin{cases} f' = tf \\ f(0) = -1 \end{cases}, \quad \begin{cases} tf' = f - t \\ f(1) = 0 \end{cases}, \quad \begin{cases} y' = te^{3t} - 2y \\ y(0) = 0 \end{cases}$$

When possible, compute the exact solution and the errors made by each method, as well as the theoretical bound for this error. Repeat the exercise, this time using modified Euler's method without theoretical error bound.

Exercise 5.4

Same question as before, but now with the following problems

$$\begin{cases} \theta'' + \sin \theta = 0 \\ \theta(0) = \frac{\pi}{6}, \theta'(0) = 0 \end{cases}, \quad \begin{cases} y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y \\ y(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{cases}, \quad \begin{cases} y' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} y + \|y\|^2 y \\ y(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{cases}$$

Exercise 5.5

Learn how to use the programs `odeeuler`, `odetaylor`, `odeRK4` and `odeRK Fehlberg` that have been provided to you by the professor. Solve each of the initial value problems of problems 5.3 and 5.4 with each of the solvers given to you by the professor. Which of the solvers is more efficient?

Exercise 5.6

Learn how to use Matlab's inner functions `ode45` and `ode23`. Compare the quality of the solutions given by these solvers with those given by `odeRK Fehlberg`. Which of the methods is faster?

Exercise 5.7

Use your programs to compare the results of the two following problems:

$$\begin{cases} \theta'' + \sin \theta = 0 \\ \theta(0) = \frac{\pi}{6}, \theta'(0) = 0 \end{cases}, \quad \begin{cases} \theta'' + \theta = 0 \\ \theta(0) = \frac{\pi}{6}, \theta'(0) = 0 \end{cases}$$

when $t \in [0, 10]$. This is the comparison of the (undamped) simple pendulum and its usual analytical simplification. Note that there is a notorious difference between the two solutions. What happens if we reduce or increase the initial angle?

Exercise 5.8

Using the theory explained in the classroom, write down as a numerical ODE problem the movement of a cannon ball which suffers from air resistance at a magnitude which has two terms: one of them is linear on the modulus of the speed and the other one is quadratic on the same quantity. Each of these terms must be multiplied by a constant that depends on the air, the characteristics of the cannon ball, etc. Produce a matlab function which upon input the initial speed, angle and constants for resistance outputs the trajectory of the cannon ball. Try different scenarios of speeds, constants, angles. . .

Exercise 5.9

Now, repeat the same problem as (5.8) but assuming that the friction is also proportional to the air pressure which decreases with altitude according to the barometric formula:

$$P(h) = P_0 e^{-\frac{\rho M h}{RT_0}},$$

where h is the altitude in meters, $M = 0,0289644 \text{ kg/mol}$ is the molar mass of Earth's air, $R = 8,31432 \text{ N} \cdot \text{m}/(\text{K} \cdot \text{mol})$ is the universal gas constant and $T_0 = 288,15 \text{ K}$ is the averaged temperature at sea level. Does this change a lot the trajectories of your cannon balls?

Exercise 5.10

Consider a baloon whose radius R and total density ρ are known, and write down the equations of the movement when it is put into a large tank of water. Write down a program that is able to guess the trajectory of the baloon according to different initial conditions (the baloon falls to the water, or it starts in the equilibrium position, or it is initially submerged, or it is slightly deviated from the equilibrium) and considering just vertical movement. You can consider that the baloon has constant shape, and that it is perfectly spherical, although other forms are also interesting. If we neglect the effects of friction, the baloon when slightly deviated from the equilibrium will present some kind of oscillating behavior. Is it the same as the harmonic oscillator? Is it the same as the simple pendulum? Try to produce a minimum square approximation to the speed of the baloon using sine functions with the appropriate frequency. Can you then give a formula for the height of the baloon at every moment that approximates its movement?

Exercise 5.11

A number (25) of negative charges are disposed in a square grid, the integer coordinate points in $[2, 2] \times [-2, 2]$. If a positive charge is dropped at $(0, 25, 0, 6)$ with zero speed, will it scape the set $\{x > 0\}$? What if the original speed is $(0, 1)$ (same position)?

Exercise 5.12

A little ball falls to the floor from 1 m height. Each time it hits the ground, it loses 10% of its energy, and additionally while it is in the air it suffers from air resistance proportional to its speed with proportionality constant 0,0001. What percentage of its original energy does the ball keep at its highest position after 1, 2 and 3 times hitting the floor? Is it the same percentage as if the ball is dropped from 10 meters?

Exercise 5.13

A typical predator–prey problem is the following: two species x and y share the same territory. The prey, x , is eaten by the predator, y . At time t we have $x(t)$ preys and $y(t)$ predators. Because the probability that a predator and a prey meet is higher when both are high quantities, and the result of an encounter is that x diminishes while y increases (since animals are more likely to reproduce when they are correctly fed), and using some natural hypotheses on the evolution of these populations in the absence of encounters, we are led to the so-called Lotka–Volterra equations:

$$\begin{cases} \dot{x} = \alpha x - \beta xy \\ \dot{y} = \gamma xy - \delta y \end{cases}$$

where $\alpha, \beta, \gamma, \delta$ are positive constants. The constants α and γ represent respectively the evolution of each species in the absence of the other (so $y(t)$ tends to 0 and $x(t)$ tends to ∞), and the nonlinear terms correspond, as said above, to the mutual encounters. Let us take some reasonable values for the constants, for example $\alpha = 0,5$, $\beta = 1,2$, $\gamma = 0,9$, $\delta = 0,8$. Assuming that $x(0) = 1000$ and $y(0) = 10$, what happens in the long term? How many predators and preys are there at $t = 20$? Many applied problems in economics and human behavior can be modeled using this kind of predator–prey system.

Exercise 5.14

The professor has given to you a lot of programs that allow us to predict the movement of planets in different circumstances. Learn how to use them till you get familiar enough with them. Then, reproduce the results described in the notes. Finally, using modifications of these programs discuss what happens for the following circumstances:

- A planet such as Jupiter but as close to the Sun as Mercury is, has a moon orbiting at distance 1 million kilometers, initially at the opposite side of the Sun. Can our planet keep its moon or will it be devoured by the star? You can neglect the effects of general relativity in this and the other questions.
- Two equally sized planets orbit their common center of gravity. One Lagrange point of the system is that center. Is it stable? What kind of trajectories would objects in that point describe?
- Describe the trajectory of a point in the stable L4 Lagrange point of the Earth–Moon system.
- In the books by Arthur C. Clarke “An Space Odyssey” the spacecraft Discovery One is put to rest at a Lagrange point of the system Jupiter–Io (the first book uses Saturn’s moon Japetus instead of Jupiter’s Io). Which could be this Lagrange point? Will the behavior of objects at that point be similar to the one of the L4 point in Earth–Moon as in the previous exercise?
- A frequent dream in certain kind of people includes the catastrophic event that the moon and the Earth suddenly stop spinning around their common gravity center. As a result, they start to approach rapidly. Which trajectories do they follow? How long does it take them to collapse? If at time 0 an object is situated in the middle of the segment joining both bodies, will it first hit the Earth or the Moon?

Exercise 5.15

The arguably most famous system of ordinary differential equations is the so-called Lorentz system, a simplified model of atmospheric convection in a tube. The model is famous since it is one fundamental example of what is nowadays called Chaos Theory. The equations are:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y \\ \dot{z} = xy - \beta z \end{cases}$$

where one can take the constants with different values. Try to use $\sigma = 10$, $\rho = 40$, $\beta = 8/3$ and start at the point $(0, 1, 0)$. You can see the result using the command `plot3`. What you are now seeing is the real reason for the name “butterfly effect”.

Only recently it has been proved by Warwick Tucker that the behavior of the solutions is as wild as it seems (namely, that the graphic that you are seeing is not the effect of arithmetical errors). This chaotic behavior is also present in the study of the many-body problem in Newtonian mechanics (needless to say, the problem is even more difficult using Einstein’s general relativity), where “many” means “three or more”.