

6. Métodos de Diferencias Finitas para la Solución de Ecuaciones Diferenciales Ordinarias

6.1. Introducción

En muchos problemas de ciencia e ingeniería se plantean ecuaciones diferenciales ordinarias (EDO). Por ejemplo, $y' = f(x, y)$ con la condición inicial $y(a) = c$. No siempre es factible hallar una solución analítica de tales ecuaciones. En este capítulo se revisan procedimientos numéricos de solución de EDO basados en la aproximación de los operadores de derivación por diferencias finitas. Tales procedimientos son también aplicables cuando se tienen sistemas de ecuaciones diferenciales de primer orden, como por ejemplo:

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}) \quad (6.1a)$$

con condiciones iniciales:

$$\mathbf{y}(t_0) = \mathbf{c} \quad (6.1b)$$

e incluso ecuaciones de orden superior: $y''' = f(x, y, y', y'')$, que con algunos cambios de variables pueden siempre convertirse en un sistema de ecuaciones diferenciales de primer orden. Así:

$$\begin{aligned} y'' &= x^2 - y' - z^2 \\ z'' &= x + z' + y^3 \end{aligned} \quad \text{con condiciones iniciales} \quad \begin{aligned} y(0) &= 0 & y'(0) &= 1 \\ z(0) &= 1 & z'(0) &= 0 \end{aligned}$$

es equivalente a:

$$\begin{aligned} u' &= x^2 - u - z^2 & y(0) &= 0 \\ v' &= x + v + y^3 & z(0) &= 1 \\ y' &= u & u(0) &= 1 \\ z' &= v & v(0) &= 0 \end{aligned} \quad \text{con condiciones iniciales}$$

Aunque la primera parte de este capítulo se refiere directamente al caso de las ecuaciones diferenciales ordinarias de primer orden, la sección 6.3 trata de procedimientos específicos para el importante caso de sistemas de ecuaciones diferenciales de segundo orden. La sección 6.5 se refiere a problemas de valor frontera, un tema que – con otros métodos – se trata también en capítulos siguientes.

6.2. Ecuaciones Diferenciales de Primer Orden

6.2.1 Método de Euler

Este es el método más simple para resolver EDO de primer orden $y' = f(x, y)$ con $y(a) = c$. El intervalo entre a y b se divide en subintervalos habitualmente iguales, de longitud h , de modo que $x_n = a + nh$. Haciendo $y_0 = c$ se determinan sucesivamente $y_1, y_2, y_3, y_4, \dots$ que son aproximaciones a los valores exactos $y(x_1), y(x_2), y(x_3), y(x_4), \dots$. Para ello $y'(x_i)$ se aproxima por $\Delta y_i / h = (y_{i+1} - y_i) / h$, de donde resulta la fórmula de recursión:

$$y_{i+1} = y_i + h f(x_i, y_i) \quad (6.2)$$

Este método es aplicable en situaciones en las que f podría ser una función bastante complicada, o podría ser el resultado de operaciones y decisiones no expresables por una simple fórmula, pero se presenta aquí un caso muy simple, con propósitos didácticos.

Supóngase que: $y' = y$ con condición inicial $y(0) = 1$. La solución exacta es en este caso conocida: $y = e^x$. Empleando el método de Euler: $y_{i+1} = y_i + h y_i = (1 + h) y_i$. Con dos distintos intervalos $h = 0.2$ y $h = 0.1$ se obtienen:

| Solución exacta | | Solución con $h = 0.2$ | | | Solución con $h = 0.1$ | | | |
|-----------------|-------|------------------------|-------|---------|------------------------|-------|---------|---------|
| i | x_i | $y(x_i) = e^{x_i}$ | y_i | $h f_i$ | $Error$ | y_i | $h f_i$ | $Error$ |
| 0 | 0 | 1.000 | 1.000 | 0.200 | 0 | 1.000 | 0.100 | 0 |
| 1 | 0.1 | 1.105 | | | | 1.100 | 0.110 | -0.005 |
| 2 | 0.2 | 1.221 | 1.200 | 0.240 | -0.021 | 1.210 | 0.121 | -0.011 |
| 3 | 0.3 | 1.350 | | | | 1.331 | 0.133 | -0.019 |
| 4 | 0.4 | 1.492 | 1.440 | 0.244 | -0.052 | 1.464 | 0.146 | -0.023 |
| 5 | 0.5 | 1.649 | | | | 1.610 | 0.161 | -0.039 |
| 6 | 0.6 | 1.822 | 1.728 | | -0.094 | 1.771 | | -0.051 |

Se observa que el error es aproximadamente proporcional a h y que en este caso crece con x .

El error de truncación local, es decir el error introducido en cada paso, es de $O(h^2)$. Sin embargo, como el número de pasos que se realizan para integrar la EDO en un intervalo dado es inversamente proporcional a h , el error global o total es de $O(h)$. También aquí podría emplearse la extrapolación de Richardson:

$$y(x) \approx [2 y(x, h) - y(x, 2h)] + O(h^2)$$

Esta expresión sería correcta para una *extrapolación pasiva*, es decir la que se hace para mejorar algunos resultados y típicamente no en cada paso. En cambio, una *extrapolación activa*, sería aquella que se realiza en cada paso, utilizándose los valores así mejorados en los sucesivos pasos del proceso. En ciertos casos la extrapolación pasiva puede ser más conveniente, por ser numéricamente más estable. Para el ejemplo anterior, con extrapolación pasiva se obtiene:

| x_i | $y(x_i) = e^{x_i}$ | $y(x, 0.2)$ | $y(x, 0.1)$ | $2 y(x, 0.1) - y(x, 0.2)$ | $Error$ |
|-------|--------------------|-------------|-------------|---------------------------|---------|
| 0.2 | 1.221 | 1.200 | 1.210 | 1.220 | -0.001 |
| 0.4 | 1.492 | 1.440 | 1.464 | 1.488 | -0.004 |
| 0.6 | 1.822 | 1.728 | 1.771 | 1.814 | -0.008 |

La solución de la ecuación $y' = f(x, y)$ depende de la condición inicial $y(a) = c$. Se tiene así como solución una familia de curvas o *trayectorias* $y(x, c)$, que en el intervalo de interés pueden ser convergentes o divergentes. Esta es una característica de la ecuación diferencial, no del procedimiento numérico empleado en su solución. Los

errores de redondeo o de truncación producen un cambio de trayectoria (y pueden entonces verse como equivalentes a resolver un problema con condiciones iniciales algo distintas). Las figuras bosquejan soluciones numéricas obtenidas para el caso en que las trayectorias divergen, en el que los errores tienden a acumularse; lo contrario ocurre cuando las trayectorias son convergentes.

Para $x = x_n$ el error acumulado en la solución numérica está dado por: $\varepsilon_n = y_n - y(x_n)$

$$\therefore \varepsilon_{n+1} - \varepsilon_n = (y_{n+1} - y_n) - (y(x_{n+1}) - y(x_n))$$

Para el método de Euler: $y_{n+1} - y_n = h f(x_n, y_n)$

Por otro lado, de la expansión en series de Taylor:

$$y(x_{n+1}) = y(x_n) + h y'(x_n) + \frac{1}{2} h^2 y''(x_n) + \dots$$

Se tiene que:

$$y(x_{n+1}) - y(x_n) = h f(x_n, y(x_n)) + O(h^2)$$

Remplazando en la primera expresión:

$$\varepsilon_{n+1} - \varepsilon_n = h [f(x_n, y_n) - f(x_n, y(x_n))] + O(h^2)$$

Pero: $f(x_n, y_n) = f(x_n, y(x_n)) + (y_n - y(x_n)) \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=y(x_n)}} + \dots$

Y por el teorema del valor medio:

$$f(x_n, y_n) - f(x_n, y(x_n)) = (y_n - y(x_n)) \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=\alpha}} = \varepsilon_n \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=\alpha}}$$

De donde:

$$\varepsilon_{n+1} \approx \varepsilon_n \left(1 + h \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=\alpha}} \right)$$

Si h es suficientemente pequeño y $\left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=\alpha}}$ es negativa el método de Euler es adecuado.

Si en cambio si $\partial f / \partial y > 0$ el error crece y el proceso sólo podría funcionar si el intervalo fuera suficientemente pequeño. Tal es el caso del ejemplo precedente, pero no el del ejemplo siguiente. Considérese la ecuación $y' = 1000(x^2 - y) + 2x$ con $y(0) = 0$. Es fácil verificar que la solución exacta es $y = x^2$, pero con el método de Euler se obtienen:

| k | x_k | $y(x_k)$ | y_k | ε_k |
|-----|-------|----------|---------|-----------------|
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0.01 | 0.0001 | 0 | -0.0001 |
| 2 | 0.02 | 0.0004 | 0.0012 | 0.0008 |
| 3 | 0.03 | 0.0009 | -0.0064 | -0.0073 |
| 4 | 0.04 | 0.0016 | 0.0672 | 0.0656 |
| 5 | 0.05 | 0.0025 | -0.5880 | -0.5905 |

Los resultados muestran oscilaciones típicas de una inestabilidad numérica. Reduciendo h no se elimina la inestabilidad; se requiere más bien cambiar de método.

Algunas alternativas, no siempre mejores, se presentan en las secciones siguientes.

6.2.2 Métodos de Runge Kutta

Estos métodos son, como el método de Euler, de *paso simple*. Es decir, sólo se requiere conocer y_n para determinar y_{n+1} . Las fórmulas de Runge – Kutta requieren evaluar $f(x, y)$ en diversos puntos apropiadamente ubicados en el intervalo $[x_n, x_{n+1} = x_n + h]$, ponderándose los resultados de modo de obtener un error de truncación del mayor orden posible.

Considérese el caso en que $f(x, y)$ se calcula en dos puntos en el intervalo $[x_n, x_{n+1}]$:

$$\begin{aligned}\hat{y} &= y_n + \alpha h f(x_n, y_n) \\ y_{n+1} &= y_n + \beta h f(x_n, y_n) + \gamma h f(x_n + \alpha h, \hat{y})\end{aligned}$$

Siendo

$$f(x_n + \alpha h, \hat{y}) = f(x_n, y_n) + \alpha h \left. \frac{\partial f}{\partial x} \right|_{\substack{x=x_n \\ y=y_n}} + (\hat{y} - y_n) \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=y_n}} + O(h^2)$$

Pero $\hat{y} - y_n = \alpha h f(x_n, y_n)$, por lo que se obtiene:

$$y_{n+1} = y_n + (\beta + \gamma) h f(x_n, y_n) + \alpha \gamma h^2 \left[\left. \frac{\partial f}{\partial x} \right|_{\substack{x=x_n \\ y=y_n}} + f(x_n, y_n) \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=y_n}} \right] + O(h^3)$$

Por otro lado, expandiendo $y(x_n + h)$ en series de Taylor:

$$y(x_n + h) = y(x_n) + h y'(x_n) + \frac{1}{2} h^2 y''(x_n) + O(h^3)$$

Pero: $y(x_n + h) \approx y_{n+1}$

$$y(x_n) \approx y_n$$

$$y'(x_n) \approx f(x_n, y(x_n)) \approx f(x_n, y_n)$$

$$y'' = \frac{\partial y'}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial y}{\partial x} \frac{\partial f}{\partial y} \Rightarrow y''(x_n) \approx \left. \frac{\partial f}{\partial x} \right|_{\substack{x=x_n \\ y=y_n}} + f(x_n, y_n) \left. \frac{\partial f}{\partial y} \right|_{\substack{x=x_n \\ y=y_n}}$$

Sustituyendo estas expresiones e identificando coeficientes se obtienen:

$$\beta + \gamma = 1$$

$$\alpha \gamma = \frac{1}{2}$$

Con un error de truncación local de $O(h^3)$ y global de $O(h^2)$.

De las infinitas alternativas de selección para $\alpha \beta \gamma$ tres son las más comunes:

La *fórmula del punto medio*:

$$\hat{y} = y_n + \frac{h}{2} f(x_n, y_n)$$

$$y_{n+1} = y_n + h f(x_n + \frac{1}{2} h, \hat{y})$$

El método de *Heun*, también conocido como *Euler modificado*:

$$\hat{y} = y_n + h f(x_n, y_n)$$

$$y_{n+1} = y_n + \frac{h}{2} [f(x_n, y_n) + f(x_n + h, \hat{y})]$$

Puede anotarse que si f no fuera función de y este método equivale a evaluar la integral:

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} y'(x) dx$$

por el método de los trapecios.

El método de *Ralston*:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + \frac{3}{4}h, y_n + \frac{3}{4}k_1h)$$

$$y_{n+1} = y_n + h(\frac{1}{3}k_1 + \frac{2}{3}k_2)$$

Análogamente, pueden obtenerse fórmulas con un error de truncación local de $O(h^4)$ y global de $O(h^3)$:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1h)$$

$$k_3 = f(x_n + h, y_n - k_1h + 2k_2h)$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 4k_2 + k_3)$$

o bien:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + \frac{1}{3}h, y_n + \frac{1}{3}k_1h)$$

$$k_3 = f(x_n + \frac{2}{3}h, y_n + \frac{2}{3}k_2h)$$

$$y_{n+1} = y_n + \frac{1}{4}h(k_1 + 3k_3)$$

Si f fuera independiente de y las expresiones precedentes equivalen al método de Simpson.

El método comúnmente denominado de Runge – Kutta es un proceso con error global de $O(h^4)$:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1h)$$

$$k_3 = f(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2h)$$

$$k_4 = f(x_n + h, y_n + k_3h)$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4)$$

Que también coincide con la regla de Simpson en el caso en que f no es función de y :

$$y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} y'(x) dx \approx \frac{h}{6} [f(x_n) + 4f(x_n + \frac{1}{2}h) + f(x_n + h)]$$

Como ejemplo de aplicación de este método, considérese la ecuación diferencial:

$$y' = (x + y)^{\frac{1}{2}}$$

con condición inicial $y(0.4) = 0.41$. Trabajando con $h = 0.4$, las operaciones en un paso serían:

$$\begin{aligned}
 k_1 &= f(x_n, y_n) = (x_n + y_n)^{\frac{1}{2}} = (0.4 + 0.41)^{\frac{1}{2}} = 0.9 \\
 k_2 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1h\right) = ((0.4 + 0.2) + (0.41 + 0.18))^{\frac{1}{2}} = 1.09087 \\
 k_3 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2h\right) = ((0.4 + 0.2) + (0.41 + 0.218174))^{\frac{1}{2}} = 1.10823 \\
 k_4 &= f(x_n + h, y_n + k_3h) = ((0.4 + 0.4) + (0.41 + 0.443292))^{\frac{1}{2}} = 1.285805 \\
 y_{n+1} &= y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4) = 0.41 + 0.438934 = 0.848934
 \end{aligned}$$

Otro método de Runge – Kutta de cuarto orden (global) es el método de *Gill*:

$$\begin{aligned}
 k_1 &= f(x_n, y_n) \\
 k_2 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1h\right) \\
 k_3 &= f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}(\sqrt{2} - 1)k_1h + \frac{1}{2}(2 - \sqrt{2})k_2h\right) \\
 k_4 &= f\left(x_n + h, y_n - \frac{1}{2}\sqrt{2}k_2h + \frac{1}{2}(2 + \sqrt{2})k_3h\right) \\
 y_{n+1} &= y_n + \frac{1}{6}h(k_1 + (2 - \sqrt{2})k_2 + (2 + \sqrt{2})k_3 + k_4)
 \end{aligned}$$

Y muchas otras alternativas son posibles.

6.2.3 Estimación del Error y Control del Paso

Al utilizar los procedimientos de paso simple descritos en las secciones precedentes, el intervalo h puede ajustarse en forma automática. Esto puede ser necesario al integrar ecuaciones cuya solución varía lentamente en algunas regiones y muy rápidamente en otras. Lo primero es poder estimar el error al integrar las ecuaciones con un cierto intervalo.

Supóngase que se resuelve una ecuación diferencial con un procedimiento de Runge – Kutta de cuarto orden (global) empleando un intervalo $2h$. La solución, que en lo que sigue se denomina y_{2h} , tendría un error de orden $(2h)^4$. Si paralelamente se obtuviera la solución con intervalo h , en lo que sigue y_h , ésta tendría un error de orden $(h)^4$. Podría entonces eliminarse el término dominante del error haciendo una extrapolación (pasiva). Para cada abscisa:

$$y_{\text{corregido}} = \frac{16y_h - y_{2h}}{15} = y_h + \frac{y_h - y_{2h}}{15}$$

Los resultados así obtenidos tendrían un error de orden $(h)^5$. Nótese que si se hiciera una extrapolación activa habría que considerar el orden del error local.

Una alternativa más conveniente sería aprovechar los resultados parciales al usar un proceso de orden m para determinar otra con un proceso de orden $m + 1$. Por ejemplo, en el procedimiento de Runge - Kutta Cash – Karp se emplean las fórmulas de cuarto orden (global):

$$y_{n+1} = y_n + h\left(\frac{37}{378}k_1 + \frac{250}{621}k_3 + \frac{125}{594}k_4 + \frac{512}{1771}k_6\right)$$

Y de quinto orden (global):

$$y_{n+1} = y_n + h\left(\frac{2825}{27648}k_1 + \frac{18575}{48384}k_3 + \frac{13525}{55296}k_4 + \frac{277}{14336}k_5 + \frac{1}{4}k_6\right)$$

Que solamente requieren seis evaluaciones de la función:

$$\begin{aligned}
 k_1 &= f(x_n, y_n) \\
 k_2 &= f\left(x_n + \frac{1}{5}h, y_n + \frac{1}{5}k_1h\right) \\
 k_3 &= f\left(x_n + \frac{3}{10}h, y_n + \frac{3}{40}k_1h + \frac{9}{40}k_2h\right) \\
 k_4 &= f\left(x_n + \frac{3}{5}h, y_n + \frac{3}{10}k_1h - \frac{9}{10}k_2h + \frac{6}{5}k_3h\right) \\
 k_5 &= f\left(x_n + h, y_n - \frac{11}{54}k_1h + \frac{5}{2}k_2h - \frac{70}{27}k_3h + \frac{35}{27}k_4h\right) \\
 k_6 &= f\left(x_n + \frac{7}{8}h, y_n + \frac{1631}{55296}k_1h + \frac{175}{512}k_2h + \frac{575}{13824}k_3h + \frac{44275}{110592}k_4h + \frac{253}{4096}k_5h\right)
 \end{aligned}$$

6.2.4 Métodos de Paso Múltiple

Los métodos tratados anteriormente permiten determinar y_{n+1} conociendo únicamente el valor de la función en el punto inmediatamente precedente. Como alternativa a estos métodos de paso simple pueden plantearse *métodos de paso múltiple* en los que el cálculo de y_{n+1} requiere utilizar $y_n, y_{n-1}, y_{n-2}, \dots$.

Un método de este tipo resulta al aproximar y' en la ecuación $y' = f(x, y)$ por una diferencia central, con lo que se tiene:

$$y_{n+1} = y_{n-1} + 2h f(x_n, y_n)$$

Este es el *método explícito del punto medio* (en inglés conocido como *leapfrog*), un método de doble paso, puesto que la expresión para obtener y_{n+1} requiere y_{n-1} e y_n .

Algunas ventajas y desventajas de los métodos de pasos múltiples con relación a los métodos de un solo paso pueden citarse:

- Para el mismo número de evaluaciones de la función $f(x, y)$ se tiene un mayor orden en el error de truncación. Así por ejemplo, el método explícito del punto medio tiene un error local de $O(h^3)$, contra $O(h^2)$ en el método de Euler, aún cuando ambos requieren una sola nueva evaluación de la función $f(x, y)$ en cada paso.
- Se presentan dificultades para arrancar el proceso, no siendo suficiente conocer y_0 sino además uno o más valores adicionales y_1, y_2, \dots . Estos valores deben obtenerse con un procedimiento distinto, de orden igual o mayor que el método de paso múltiple a emplearse.
- Es en general difícil (aunque no imposible) cambiar el intervalo de integración, h (lo cual, en cambio, no es ningún problema en los métodos de paso simple). Esto podría reducir la relativa eficiencia de los métodos de pasos múltiples.
- La inestabilidad numérica es un problema más frecuente en los métodos de paso múltiple que en los métodos de paso simple. Éste y otros temas relacionados se revisan en la sección 6.2.5.

Entre los métodos de paso múltiple que se encuentran en la literatura están los métodos explícitos de Adams – Bashfort:

$$y_{n+1} = y_n + h (\beta_1 f_n + \beta_2 f_{n-1} + \beta_3 f_{n-2} + \dots + \beta_{k+1} f_{n-k})$$

Y los correspondientes métodos implícitos de Adams – Moulton:

$$y_{n+1} = y_n + h (\beta_0 f_{n+1} + \beta_1 f_n + \beta_2 f_{n-1} + \beta_3 f_{n-2} + \dots + \beta_k f_{n-k+1})$$

En estas expresiones f_n denota $f(x_n, y_n)$. Los coeficientes β pueden obtenerse considerando que:

$$y(x_{n+1}) - y(x_n) = \int_{x_n}^{x_{n+1}} \frac{\partial y}{\partial x} dx = \int_{x_n}^{x_{n+1}} f(x, y(x)) dx$$

Y aproximando $f(x, y)$ por un polinomio de interpolación escrito en términos de diferencias finitas hacia atrás:

$$y_{n+1} - y_n = h \left(f_n + \frac{1}{2} \nabla f_n + \frac{5}{12} \nabla^2 f_n + \frac{3}{8} \nabla^3 f_n + \dots \right) \quad (\text{A-B})$$

$$y_{n+1} - y_n = h \left(f_{n+1} - \frac{1}{2} \nabla f_{n+1} - \frac{1}{12} \nabla^2 f_{n+1} - \frac{1}{24} \nabla^3 f_{n+1} + \dots \right) \quad (\text{A-M})$$

Por otro lado, pueden escribirse expansiones en series de Taylor, obteniéndose las β por identificación de coeficientes. Así, para la expresión explícita con $k = 1$:

$$y_{n+1} = y_n + h (\beta_1 f_n + \beta_2 f_{n-1})$$

$$y(x_n + h) \approx y(x_n) + h (\beta_1 y'(x_n) + \beta_2 y'(x_n - h))$$

$$y(x_n + h) \approx y(x_n) + h \beta_1 y'(x_n) + h \beta_2 \left(y'(x_n) - h y''(x_n) + \frac{1}{2} h^2 y'''(x_n) - \dots \right)$$

$$y(x_n + h) \approx y(x_n) + h (\beta_1 + \beta_2) y'(x_n) + h^2 (-\beta_2) y''(x_n) + O(h^3)$$

Comparando con:

$$y(x_n + h) = y(x_n) + h y'(x_n) + \frac{1}{2} h^2 y''(x_n) + O(h^3)$$

Se tiene que: $\beta_1 + \beta_2 = 1$ y $-\beta_2 = \frac{1}{2}$, de donde: $\beta_1 = \frac{3}{2}$ y $\beta_2 = -\frac{1}{2}$, es decir:

$$y_{n+1} = y_n + \frac{h}{2} (3f_n - f_{n-1})$$

Algunos resultados similares se listan a continuación. El error de truncación local es de $O(h^{k+2})$ y el global de $O(h^{k+1})$:

Métodos de Adams – Bashfort (explícitos):

| k | $y_{n+1} = y_n + h (\beta_1 f_n + \beta_2 f_{n-1} + \beta_3 f_{n-2} + \dots + \beta_{k+1} f_{n-k})$ | |
|-----|---|--------------|
| 0 | $y_{n+1} = y_n + h f_n$ | <i>Euler</i> |
| 1 | $y_{n+1} = y_n + \frac{1}{2} h (3f_n - f_{n-1})$ | |
| 3 | $y_{n+1} = y_n + \frac{1}{24} h (55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$ | |

Métodos de Adams – Moulton (implícitos):

| k | $y_{n+1} = y_n + h (\beta_0 f_{n+1} + \beta_1 f_n + \beta_2 f_{n-1} + \beta_3 f_{n-2} + \dots + \beta_k f_{n-k+1})$ | |
|-----|---|--|
| 0 | $y_{n+1} = y_n + h f_{n+1}$ | <i>Euler inverso</i> |
| 1 | $y_{n+1} = y_n + \frac{1}{2} h (f_{n+1} + f_n)$ | <i>Crank Nicholson (regla trapezoidal)</i> |
| 3 | $y_{n+1} = y_n + \frac{1}{24} h (9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$ | |

6.2.5 Procedimientos Predictor – Corrector

El error de truncación de una fórmula implícita de este tipo es siempre menor que el error del correspondiente proceso explícito del mismo orden. Excepto en el caso trivial en que la función $f(x, y)$ es lineal, la solución de la(s) ecuación(es) que definen el proceso implícito requiere iterar. Con tal fin, se utiliza como primera aproximación el resultado obtenido con la fórmula explícita (*predictor*). Esta aproximación se refina utilizando repetidamente la correspondiente fórmula implícita (que en este contexto se denomina *corrector*). Por ejemplo:

$$y_{n+1}^{(0)} = y_n + h f(x_n, y_n)$$

$$y_{n+1}^{(i+1)} = y_n + h f(x_{n+1}, y_{n+1}^{(i)})$$

El superíndice se refiere en este caso a la iteración. Pueden, por supuesto, plantearse métodos Predictor – Corrector en los que las expresiones empleadas no sean las de Adams – Bashfort – Moulton. Tal es el caso del método de *Milne* (un procedimiento débilmente estable, que no se recomienda):

| | |
|---|---|
| | Error local |
| Predictor: $y_{n+1}^{(0)} = y_{n-3} + \frac{4}{3}h(2f_n - f_{n-1} + 2f_{n-2})$ | $O\left(\frac{14}{45}h^5 y''(\xi)\right)$ |
| Corrector: $y_{n+1}^{(i+1)} = y_{n-1} + \frac{1}{3}h(f_{n+1}^{(i)} + 4f_n + f_{n-1})$ | $O\left(-\frac{1}{90}h^5 y''(\xi)\right)$ |

Supóngase, por ejemplo, la ecuación: $y' = -xy^2$ con condición inicial $y(0) = y_0 = 2$. Se ha obtenido (con algún otro método) $y(0.1) \approx y_1 = 1.980198$.

Con el predictor $y_{n+1}^{(0)} = y_n + \frac{1}{2}h(3f_n - f_{n-1})$ se tiene:

$$y_2^{(0)} = y_1 + 0.05 \left[-3x_1(y_1)^2 + x_0(y_0)^2 \right] =$$

$$= 1.980198 + 0.05 \left[-0.3 \cdot (1.980198)^2 - 0 \right] = 1.921380$$

y luego con el corrector (en este caso la regla trapezoidal, también conocida como método de Crank Nicholson):

$$y_{n+1} = y_n + \frac{1}{2}h(f_{n+1} + f_n)$$

se tiene:

$$y_2^{(1)} = y_1 + 0.05 \left[-x_2(y_2^{(0)})^2 + x_1(y_1)^2 \right] =$$

$$= 1.980198 + 0.05 \left[-0.2 \cdot (1.921380)^2 - 0.1 \cdot (1.980198)^2 \right] = 1.923675$$

Y en forma similar:

$$y_2^{(2)} = 1.923587$$

$$y_2^{(3)} = 1.923590$$

$$y_2^{(4)} = 1.923590$$

El proceso es en este caso convergente. La solución exacta es en este caso $y = 2/(1+x^2)$, de donde $y(0.2) = 1.923077$

Si el intervalo de integración, h , es demasiado grande, se observan importantes diferencias entre la aproximación inicial obtenida con el predictor, $y_{n+1}^{(0)}$, y el valor corregido, $y_{n+1}^{(k)}$. En tal caso la convergencia es lenta, e incluso podría no producirse. Por otro lado, diferencias insignificantes implican que h es innecesariamente pequeño.

6.2.6 Consistencia, Estabilidad y Convergencia

Al emplear una ecuación de diferencias en lugar de la ecuación diferencial, se introducen errores de truncación en cada paso.

Por ejemplo, en el método de Euler, la expresión:

$$y(x+h) = y(x) + h y'(x) + \frac{h^2}{2} y''(x) + \dots$$

se reemplaza por:

$$y_{n+1} = y_n + h f(x_n, y_n)$$

teniéndose un error de orden h^2 . A esto deben agregarse los errores debidos a la aritmética imperfecta del ordenador.

Evidentemente no es posible determinar estos errores, pero si pueden hacerse algunas estimaciones.

Al emplear un método numérico para resolver EDO, se espera que éste sea *convergente*, lo que significa que al trabajar con intervalos, h , cada vez más pequeños, las soluciones deben aproximar cada vez más a la solución exacta. Para que el procedimiento numérico sea convergente debe ser *consistente* y *estable*.

Consistencia significa que en el límite $h \rightarrow 0$ la ecuación de diferencias que define el método numérico resulta formalmente la ecuación diferencial. Refiriéndose nuevamente al método de Euler, que puede escribirse:

$$f(x_n, y_n) = \frac{y_{n+1} - y_n}{h}$$

se observa que:

$$\lim_{h \rightarrow 0} \frac{y(x_n + h) - y(x_n)}{h} = y'(x_n)$$

Si en cambio se escribiera, por ejemplo, $y_{n+1} = 2y_n + h f(x_n, y_n)$ no habría consistencia.

Para que el procedimiento numérico sea convergente no es suficiente que sea consistente. Se requiere además que sea estable. Un método es *estable* cuando los errores de truncación y de redondeo, al propagarse durante el proceso, son siempre pequeños en comparación con la solución exacta.

CONSISTENCIA + ESTABILIDAD \Leftrightarrow CONVERGENCIA

Alternativamente, podría decirse que un método es estable si, para condiciones iniciales típicas y siempre que los errores de truncación y de redondeo sean pequeños, se tiene convergencia.

Un ejemplo de inestabilidad numérica

Para una observación inicial sobre el tema de estabilidad, considérese la ecuación diferencial $y' = -y$, con condición inicial $y(0) = 1$, cuya solución exacta es $y = e^{-x}$. El método de Euler podría ser apropiado en este caso. Sin embargo, supóngase que se emplea la "regla explícita del punto medio":

$$y_{n+1} = y_{n-1} + 2h f(x_n, y_n)$$

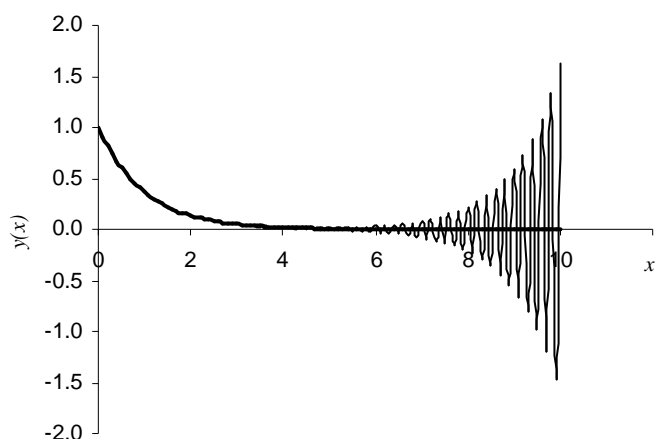
con intervalo $h = 0.1$, es decir:

$$y_{n+1} = y_{n-1} - 0.2 y_n$$

Para iniciar el proceso no es suficiente conocer $y_0 = y(0) = 1$. Se requiere además y_1 . Supóngase que se calcula $y_1 = e^{-h} = e^{-0.1} = 0.90483741803596$, con 13 cifras significativas correctas. Trabajando con aritmética de doble precisión se obtienen:

| i | x_i | $y(x_i)$ | y_i | $e_i = y_i - y(x_i)$ |
|-----|-------|----------|-----------|----------------------|
| 0 | 0.0 | 1.000000 | 1.000000 | 0.000000 |
| 1 | 0.1 | 0.904837 | 0.904837 | 0.000000 |
| 2 | 0.2 | 0.818731 | 0.819033 | 0.000302 |
| 3 | 0.3 | 0.740818 | 0.741031 | 0.000213 |
| ... | | | | |
| 97 | 9.7 | 0.000061 | -1.199394 | -1.199455 |
| 98 | 9.8 | 0.000055 | 1.325440 | 1.325385 |
| 99 | 9.9 | 0.000050 | -1.464482 | -1.464532 |
| 100 | 10 | 0.000045 | 1.618337 | 1.618291 |

La solución exacta es exponencialmente decreciente, como se indica en línea gruesa en la figura siguiente. Sin embargo, el procedimiento produce los resultados que se presentan en línea más delgada. Después de aproximadamente $x = 5$ se obtienen valores con signos alternados, con amplitud cada vez mayor (que podría llegar a rebasar el número máximo que puede almacenarse). Esto es una inestabilidad numérica.



Si se reduce h o se aumenta el número de cifras en los cálculos el problema puede posponerse, pero no evitarse. ¿A qué se debe esta inestabilidad?

Considérese la ecuación diferencial:

$$y' = \lambda y$$

para la cual el método explícito del punto medio resulta:

$$y_{n+1} = y_{n-1} + 2\lambda h y_n$$

Esta es una ecuación de diferencias lineal de orden 2. Definiendo $\alpha = \lambda h$, la ecuación característica es:

$$r^2 = 1 + 2\alpha r$$

con las raíces:

$$r_1 = \alpha + (1 + \alpha^2)^{\frac{1}{2}} = 1 + \alpha + \frac{1}{2}\alpha^2 + O(\alpha^4)$$

$$r_2 = -\frac{1}{r_1}$$

Al comparar la primera de estas expresiones con la expansión en series:

$$e^\alpha = 1 + \alpha + \frac{1}{2}\alpha^2 + \frac{1}{6}\alpha^3 + O(\alpha^4)$$

se concluye que:

$$r_1 = e^\alpha - \frac{1}{6}\alpha^3 + O(\alpha^4)$$

$$r_1^n \approx e^{\alpha n} = e^{\lambda n h} = e^{\lambda x_n}$$

$$r_2^n = (-1)^n r_1^{-n} \approx (-1)^n e^{-\lambda x_n}$$

y en consecuencia:

$$y_n = C_1 r_1^n + C_2 r_2^n \approx C_1 e^{\lambda x_n} + (-1)^n C_2 e^{-\lambda x_n}$$

La solución de la ecuación de diferencias tiene dos “modos componentes”, el primero de los cuales corresponde a la solución correcta, mientras que el segundo no debería existir. En teoría se debería tener $C_1 = 1$ y $C_2 = 0$, de modo que $y_n \approx e^{\lambda x_n}$, pero no es así, ni siquiera al iniciarse el proceso. Efectivamente, para el caso $\alpha = \lambda h = -0.1$, los valores iniciales hacen que:

$$y_0 = C_1 + C_2 = 1$$

$$y_1 = C_1 r_1 + C_2 r_2 = e^{-0.1}$$

de donde, trabajando en doble precisión:

$$C_2 = (e^{-0.1} - r_1) / (r_2 - r_1) \approx 7.469947 \cdot 10^{-5}$$

y suponiendo que las operaciones siguientes se efectuaran con una aritmética infinitamente precisa:

$$C_2 r_2^{100} \approx 7.469947 \cdot 10^{-5} e^{100(0.1)} = 1.645365$$

Nótese que este resultado es similar al error observado para $x = 10$. Por otro lado, aún cuando inicialmente se tuviera $C_2 = 0$, los errores numéricos introducirían la componente extraña. El factor $(-1)^n$ explica la alternancia de signos del error.

El procedimiento no funciona para λ negativo, porque en ese caso el modo extraño $(-1)^n C_2 e^{-\lambda x}$ tiende a crecer, mientras que la solución $C_1 e^{\lambda x}$ es decreciente. Sin embargo, si sirve para el caso (poco práctico) en que λ es positivo y la solución es exponencialmente creciente.

Región de estabilidad absoluta

Puede obtenerse valiosa información con relación al comportamiento de un método numérico de solución de EDO al observarse cómo funciona con una ecuación de la forma $y' = \lambda y$, donde λ es una variable compleja, con condición inicial $y(0) = 1$. Con referencia a esta ecuación, se define la *región de estabilidad absoluta* de un método numérico como el lugar geométrico de los λh para los que la solución es finita cuando $n \rightarrow \infty$. Esto es equivalente a decir que las raíces de la ecuación característica satisfacen $|r_i| \leq 1$.

Para $h \rightarrow 0$ se requiere que una de las raíces sea igual a 1 (esto es un requisito para que el método sea consistente). Sin embargo, si dos o más raíces tienen módulo igual a 1 el método es *débilmente estable* y con frecuencia tiene problemas numéricos. Este es el caso del método explícito del punto medio, del que se trató en el acápite anterior.

Considérese, por ejemplo, el método de Euler:

$$y_{n+1} = y_n + h f(x_n, y_n)$$

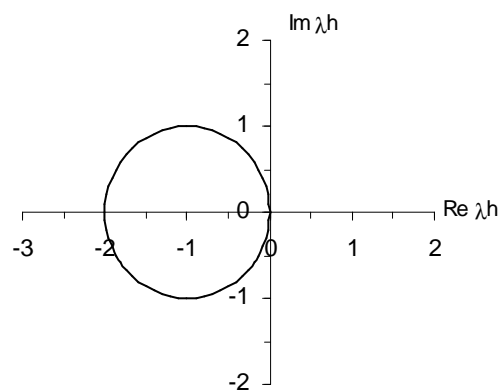
que aplicado a la ecuación $y' = \lambda y$ resulta:

$$y_{n+1} = y_n + \lambda h y_n = (1 + \lambda h) y_n$$

La solución de esta ecuación de diferencias es de la forma $y_n = r^n$. Al sustituir esta solución en la ecuación de diferencias se obtiene la ecuación característica:

$$r = 1 + \lambda h$$

La región de estabilidad absoluta queda definida por: $|1 + \lambda h| \leq 1$. Esta es el área dentro de una circunferencia de radio 1 y con centro en $(-1, 0)$:



Región de estabilidad absoluta - Euler

Puede concluirse que el método de Euler es apropiado para integrar ecuaciones cuya solución es exponencialmente decreciente ($\lambda < 0$), pero no para el caso en que la solución es oscilatoria y no amortiguada (λ imaginaria pura).

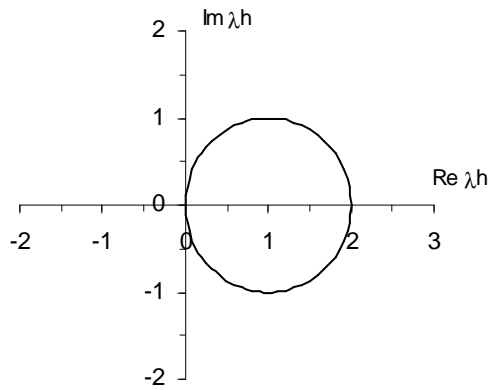
Análogamente para el método de Euler inverso:

$$y_{n+1} = y_n + h f(x_{n+1}, y_{n+1}) = y_n + \lambda h y_{n+1}$$

se obtiene:

$$r = \frac{1}{1 - \lambda h}$$

La región de estabilidad absoluta, en la que $|r| \leq 1$, es en este caso la región externa al círculo de radio 1 y con centro en (1,0).



Región de estabilidad absoluta
Euler Inverso

El método de Euler inverso sería apropiado para el caso en que la solución es oscilatoria, amortiguada o no. No debería emplearse para integrar ecuaciones cuya solución es exponencialmente creciente, ya que por consideraciones de estabilidad requeriría un paso grande, lo que estaría en conflicto con los objetivos de precisión.

Todos los procesos de Runge Kutta del mismo orden tienen la misma región de estabilidad absoluta. Por ejemplo, considérese un proceso de segundo orden, método de Heun, también llamado Euler modificado:

$$\hat{y} = y_n + h f(x_n, y_n)$$

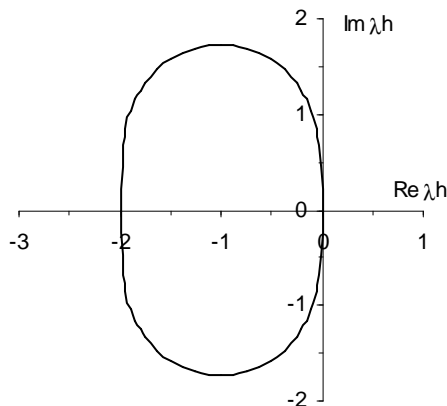
$$y_{n+1} = y_n + \frac{h}{2} [f(x_n, y_n) + f(x_{n+1}, \hat{y})]$$

que aplicado a la ecuación $y' = \lambda y$ resulta:

$$\hat{y} = y_n + h \lambda y_n = (1 + \lambda h) y_n$$

$$y_{n+1} = y_n + \frac{h}{2} [\lambda y_n + \lambda (y_n + h \lambda y_n)] = \left[1 + \lambda h + \frac{1}{2}(\lambda h)^2\right] y_n$$

y por lo tanto $r = 1 + \lambda h + \frac{1}{2}(\lambda h)^2$. La condición $|r| \leq 1$ es satisfecha por todos los puntos dentro del elipsoide que se muestra en la figura siguiente:



Región de estabilidad absoluta
Runge Kutta - Segundo Orden

Otros métodos de Runge Kutta de segundo orden, como el método del punto medio o el método de Ralston, tienen exactamente la misma ecuación característica y por lo tanto la misma región de estabilidad absoluta.

Para el "clásico" método de Runge Kutta:

$$\begin{aligned} k_1 &= h f(x_n, y_n) \\ k_2 &= h f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \\ k_3 &= h f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2\right) \\ k_4 &= h f(x_n + h, y_n + k_3) \\ y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

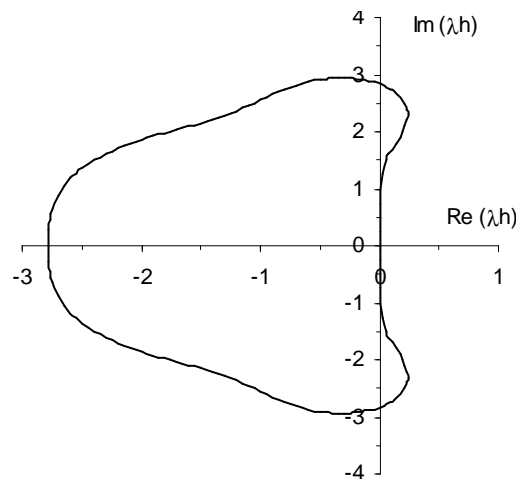
se obtienen en este caso:

$$\begin{aligned} k_1 &= \lambda h y_n \\ k_2 &= \lambda h \left(1 + \frac{1}{2}\lambda h\right) y_n \\ k_3 &= \lambda h \left(1 + \frac{1}{2}\lambda h + \frac{1}{4}(\lambda h)^2\right) y_n \\ k_4 &= \lambda h \left(1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \frac{1}{4}(\lambda h)^3\right) y_n \\ y_{n+1} &= \left(1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \frac{1}{6}(\lambda h)^3 + \frac{1}{24}(\lambda h)^4\right) y_n \end{aligned}$$

Y resulta:

$$r = 1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \frac{1}{6}(\lambda h)^3 + \frac{1}{24}(\lambda h)^4$$

La expresión es la misma para cualquier otro método de Runge Kutta de cuarto orden (global): La región de estabilidad absoluta es aquella dentro del límite indicado en la figura siguiente:



Región de estabilidad absoluta
Runge - Kutta 4º Orden

Las mismas ideas pueden también aplicarse a métodos de pasos múltiples. Por ejemplo, para la regla explícita del punto medio, a la que se hizo referencia al inicio de esta sección:

$$y_{n+1} = y_{n-1} + 2h f(x_n, y_n) = y_{n-1} + 2\lambda h y_n$$

se obtienen las raíces de la ecuación característica:

$$r = \lambda h \pm \left(1 + (\lambda h)^2\right)^{\frac{1}{2}}$$

y para tener $|r_i| \leq 1$ se requiere que λh sea imaginaria pura, en el rango entre $-i$ y $+i$. Debe observarse que este método es débilmente estable, porque para $h \rightarrow 0$ se tiene $|r_1| = |r_2| = 1$.

Un mejor método de pasos múltiples es la regla trapezoidal o método de Crank – Nicholson:

$$y_{n+1} = y_n + \frac{1}{2}h [f(x_n, y_n) + f(x_{n+1}, y_{n+1})]$$

que para el caso $y' = \lambda y$ resulta:

$$y_{n+1} = y_n + \frac{1}{2}\lambda h (y_n + y_{n+1})$$

de donde:

$$\left(1 - \frac{1}{2}\lambda h\right)r = \left(1 + \frac{1}{2}\lambda h\right)$$

y por lo tanto se requiere que: $|r| = \left| \frac{1 + \frac{1}{2}\lambda h}{1 - \frac{1}{2}\lambda h} \right| \leq 1$, es decir $\text{Re}(\lambda h) \leq 0$.

6.3. Métodos para EDO de Segundo Orden

Las ecuaciones diferenciales de segundo orden siempre pueden describirse como un sistema de ecuaciones de primer orden. Sin embargo, es más eficiente emplear métodos más específicos, como los que se describen a continuación.

6.3.1 Runge Kutta

Para resolver EDO de la forma $y'' = f(x, y, y')$, con condiciones iniciales $y(a) = y_0$ e $y'(a) = y'_0$ se encuentran en la literatura modificaciones de los métodos de Runge Kutta como, por ejemplo, el proceso de cuarto orden (global):

$$\begin{aligned} k_1 &= h f(x_n, y_n, y'_n) \\ k_2 &= h f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}h y'_n + \frac{1}{8}h k_1, y' + \frac{1}{2}k_1\right) \\ k_3 &= h f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}h y'_n + \frac{1}{8}h k_1, y' + \frac{1}{2}k_2\right) \\ k_4 &= h f\left(x_n + h, y_n + h y'_n + \frac{1}{2}h k_3, y' + k_3\right) \\ y'_{n+1} &= y'_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ y_{n+1} &= y_n + h \left[y'_n + \frac{1}{6}(k_1 + k_2 + k_3) \right] \end{aligned}$$

6.3.2 Método de Diferencia Central

Este procedimiento se basa en sustituir las derivadas por sus aproximaciones con diferencias centrales. Así al resolver:

$$m\ddot{u} + ku = f(t)$$

se puede aproximar la segunda derivada con: $\ddot{u}_n = \frac{u_{n-1} - 2u_n + u_{n+1}}{(\Delta t)^2}$

de donde resulta: $u_{n+1} = 2u_n - u_{n-1} + \frac{(\Delta t)^2}{m}(f(t) - k u_n)$.

Este método puede también escribirse en la forma sumada:

$$m\ddot{u}_n = f(t_n) - ku_n$$

$$\dot{u}_{n+\frac{1}{2}} = \dot{u}_{n-\frac{1}{2}} + \ddot{u}_n \Delta t$$

$$u_{n+1} = u_n + \dot{u}_{n+\frac{1}{2}} \Delta t$$

cuya interpretación física es simple.

Cuando la ecuación es de la forma: $m\ddot{u} + c\dot{u} + ku = f(t)$ conviene aproximar la primera derivada mediante:

$$\dot{u}_n = \frac{u_{n+1} - u_{n-1}}{2\Delta t}$$

Y no con una diferencia hacia atrás, que a primera vista podría parecer una mejor elección. Esta consideración se relaciona con el tema de estabilidad, que se trata más adelante. Con la aproximación antes mencionada se obtiene:

$$\left(\frac{m}{(\Delta t)^2} + \frac{c}{2\Delta t} \right) u_{n+1} = f(t_n) - \left(k - \frac{2m}{(\Delta t)^2} \right) u_n - \left(\frac{m}{(\Delta t)^2} - \frac{c}{2\Delta t} \right) u_{n-1}$$

6.3.3 Método de Newmark

La familia de métodos de Newmark se basa en aproximaciones de la forma:

$$\dot{u}_{n+1} = \dot{u}_n + \Delta t [\delta \ddot{u}_n + (1 - \delta) \ddot{u}_{n+1}]$$

$$u_{n+1} = u_n + \Delta t \dot{u}_n + (\Delta t)^2 \left[\alpha \ddot{u}_n + \left(\frac{1}{2} - \alpha \right) \ddot{u}_{n+1} \right]$$

El caso con $\delta = \frac{1}{2}$ y $\alpha = \frac{1}{6}$ corresponde a suponer que \ddot{u} (la aceleración) varía linealmente en el intervalo. Aparentemente esa elección sería mejor que $\delta = \frac{1}{2}$ y $\alpha = \frac{1}{4}$ (lo que físicamente se interpretaría como suponer una aceleración constante promedio). Sin embargo, esta última elección es más apropiada, también por consideraciones de estabilidad, a las que se hace referencia más adelante.

Al remplazar las aproximaciones en:

$$m\ddot{u}_{n+1} + c\dot{u}_{n+1} + ku_{n+1} = f(t_{n+1})$$

se obtiene: $\hat{k}u_{n+1} = \hat{f}_{n+1}$

donde:

$$\hat{k} = k + a_0 m + a_1 c$$

$$\hat{f}_{n+1} = f_{n+1} + m(a_0 u_n + a_2 \dot{u}_n + a_3 \ddot{u}_n) + c(a_1 u_n + a_4 \dot{u}_n + a_5 \ddot{u}_n)$$

y los coeficientes $a_0 \dots a_7$ son:

$$a_0 = \frac{1}{\alpha(\Delta t)^2} \quad a_1 = \frac{\delta}{\alpha\Delta t} \quad a_2 = \frac{1}{\alpha\Delta t}$$

$$a_3 = \frac{1}{2\alpha} - 1 \quad a_4 = \frac{\delta}{\alpha} - 1 \quad a_5 = \frac{\Delta t}{2} \left(\frac{\delta}{\alpha} - 2 \right)$$

$$a_6 = \Delta t(1 - \delta) \quad a_7 = \delta\Delta t$$

Finalmente se pueden determinar los nuevos valores de \dot{u} y \ddot{u} mediante:

$$\ddot{u}_{n+1} = a_0(u_{n+1} - u_n) - a_2\dot{u}_n - a_3\ddot{u}_n$$

$$\dot{u}_{n+1} = \dot{u}_n + a_6\ddot{u}_n + a_7\ddot{u}_{n+1}$$

6.3.4 Estabilidad y Precisión

Todos los métodos para resolver EDO de segundo orden pueden ser escritos en la forma:

$$\mathbf{u}_{n+1} = \mathbf{A} \mathbf{u}_n + \mathbf{b} f_n$$

donde \mathbf{u}_n representa ahora el conjunto de resultados que describen la respuesta en el paso n . Por ejemplo, para el método de diferencia central:

$$\begin{Bmatrix} u_{n+1} \\ u_n \end{Bmatrix} = \begin{pmatrix} \frac{2 - (\omega \Delta t)^2}{1 + \beta \omega \Delta t} & -\frac{1 - \beta \omega \Delta t}{1 + \beta \omega \Delta t} \\ 1 & 0 \end{pmatrix} \begin{Bmatrix} u_n \\ u_{n-1} \end{Bmatrix} + \begin{Bmatrix} \frac{(\Delta t)^2}{1 + \beta \omega \Delta t} \\ 0 \end{Bmatrix} f_n$$

Los errores se comportan en forma análoga, pudiendo demostrarse que:

$$\boldsymbol{\varepsilon}_{n+1} = \mathbf{A} \boldsymbol{\varepsilon}_n$$

y por lo tanto, para que los errores se reduzcan (es decir, para que el método sea estable) se requiere que:

$$\rho(\mathbf{A}) = \max |\lambda_i| \leq 1$$

En esta expresión $\rho(\mathbf{A})$ es el *radio espectral* y los λ_i son los valores característicos de \mathbf{A} . Para el caso en que $\beta = 0$ y se utiliza el método de diferencia central, la condición $\rho(\mathbf{A}) \leq 1$ se cumple cuando:

$$\Delta t \leq \frac{2}{\omega} = \frac{T}{\pi}$$

Se dice entonces que el método de diferencia central es condicionalmente estable. Si Δt excede el límite antes indicado se produce una inestabilidad numérica. El límite corresponde a algo más que tres puntos por período, lo que sin embargo resulta insuficiente para tener buena precisión. Con $\Delta t = T/10$ se introducen errores del orden de 3% en cada paso, mientras que con $\Delta t = T/20$ los errores son del orden de 1%.

Para la familia de métodos de Newmark pueden también obtenerse las condiciones de estabilidad:

$$\delta \geq 0.50$$

$$\alpha \geq 0.25(0.5 + \delta)^2$$

Con una apropiada selección de los parámetros (lo habitual es $\delta = \frac{1}{2}$ y $\alpha = \frac{1}{6}$) se tiene un procedimiento *incondicionalmente estable*, es decir estable para cualquier selección de Δt . Nuevamente, el intervalo de integración está limitado por la precisión.

Dos procesos a los que también se hace referencia en la literatura son los métodos de Houbolt y de Wilson. Ambos métodos son también (para una selección apropiada de sus parámetros) incondicionalmente estables, pero acumulan mucho más errores que el de Newmark de aceleración constante y por tanto no son tan convenientes.

6.4. Sistemas de Ecuaciones Diferenciales Ordinarias

6.4.1 Integración Directa

Cuando las EDO son no lineales, y en consecuencia no es factible la descomposición modal a la que se hace referencia en la sección siguiente, cabe la posibilidad de integrar directamente las ecuaciones en su forma original. Las expresiones son prácticamente

las mismas usadas para el caso de una sola ecuación diferencial, excepto que todas las operaciones son realizadas con matrices. Así, por ejemplo, para resolver el sistema de EDO de primer orden:

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$$

Podría emplearse el método de Runge Kutta de cuarto orden:

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(x_n, \mathbf{y}_n) \\ \mathbf{k}_2 &= \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}\mathbf{k}_1h\right) \\ \mathbf{k}_3 &= \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}\mathbf{k}_2h\right) \\ \mathbf{k}_4 &= \mathbf{f}(x_n + h, \mathbf{y}_n + \mathbf{k}_3h) \\ \mathbf{y}_{n+1} &= \mathbf{y}_n + \frac{1}{6}h(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) \end{aligned}$$

(como es habitual, las minúsculas en letras negritas denotan matrices columna).

Análogamente, para resolver el sistema de ecuaciones diferenciales de segundo orden:

$$\mathbf{M} \ddot{\mathbf{u}} + \mathbf{K} \mathbf{u} = \mathbf{f}(t)$$

podría emplearse el método de diferencia central:

$$\begin{aligned} \mathbf{M} \ddot{\mathbf{u}}_n &= \mathbf{f}(t_n) - \mathbf{K} \mathbf{u}_n \\ \dot{\mathbf{u}}_{n+\frac{1}{2}} &= \dot{\mathbf{u}}_{n-\frac{1}{2}} + \ddot{\mathbf{u}}_n \Delta t \\ \mathbf{u}_{n+1} &= \mathbf{u}_n + \dot{\mathbf{u}}_{n+\frac{1}{2}} \Delta t \end{aligned}$$

El método de diferencia central es particularmente eficiente para el caso de EDO no lineales si la matriz \mathbf{M} es diagonal (una aproximación frecuente en ingeniería). Puede anotarse que en el caso de ecuaciones no lineales no requiere obtenerse \mathbf{K} explícitamente, sino más bien el producto $\mathbf{K} \mathbf{u}$ (lo que puede ser notoriamente más fácil).

Otra posibilidad para resolver ecuaciones de la forma $\mathbf{M} \ddot{\mathbf{u}}_{n+1} + \mathbf{C} \dot{\mathbf{u}}_{n+1} + \mathbf{K} \mathbf{u}_{n+1} = \mathbf{f}_{n+1}$ sería el método de Newmark, en el que \mathbf{u}_{n+1} se obtiene de:

$$\begin{aligned} \hat{\mathbf{K}} \mathbf{u}_{n+1} &= \hat{\mathbf{f}}_{n+1} \\ \hat{\mathbf{K}} &= \mathbf{K} + a_0 \mathbf{M} + a_1 \mathbf{C} = \mathbf{L} \mathbf{U} \\ \hat{\mathbf{f}}_{n+1} &= \mathbf{f}_{n+1} + \mathbf{M}(a_0 \mathbf{u}_n + a_2 \dot{\mathbf{u}}_n + a_3 \ddot{\mathbf{u}}_n) + \mathbf{C}(a_1 \mathbf{u}_n + a_4 \dot{\mathbf{u}}_n + a_5 \ddot{\mathbf{u}}_n) \end{aligned}$$

y luego:

$$\begin{aligned} \ddot{\mathbf{u}}_{t+\Delta t} &= a_0 (\mathbf{u}_{t+\Delta t} - \mathbf{u}_t) - a_2 \dot{\mathbf{u}}_t - a_3 \ddot{\mathbf{u}}_t \\ \dot{\mathbf{u}}_{t+\Delta t} &= \dot{\mathbf{u}}_t + a_6 \ddot{\mathbf{u}}_t + a_7 \ddot{\mathbf{u}}_{t+\Delta t} \end{aligned}$$

Los coeficientes $a_0 \dots a_7$ son los mismos de la sección 6.3.3.

Al resolver EDO no lineales el método de Newmark, en la forma antes descrita, requeriría una nueva factorización $\hat{\mathbf{K}} = \mathbf{L} \mathbf{U}$ en cada paso. El proceso podría mejorarse descomponiendo $\hat{\mathbf{K}}$ en dos partes y pasando los términos no lineales al segundo miembro, como parte de $\hat{\mathbf{f}}_{n+1}$.

6.4.2 Descomposición Modal

Cuando el sistema de ecuaciones diferenciales es lineal, y particularmente si las matrices que definen el sistema de ecuaciones diferenciales son simétricas, el procedimiento más eficiente se basa en la descomposición modal. Para mostrar en que

consiste la descomposición modal, supóngase que se tiene el sistema de ecuaciones diferenciales:

$$\mathbf{B}\dot{\mathbf{x}} + \mathbf{A}\mathbf{x} = \mathbf{f}(t)$$

Las matrices \mathbf{A} y \mathbf{B} son simétricas (y muy frecuentemente definidas positivas). La descomposición modal emplea los vectores característicos ϕ_i obtenidos de:

$$\mathbf{A}\phi_i = \lambda_i \mathbf{B}\phi_i$$

Estos vectores constituyen una base completa, y por lo tanto la solución $\mathbf{x}(t)$ puede siempre expresarse como una combinación lineal de los referidos vectores:

$$\mathbf{x}(t) = \sum a_j \phi_j$$

Nótese que, siendo las matrices \mathbf{A} y \mathbf{B} constantes, los vectores característicos ϕ_i no son función de tiempo. Sin embargo, las componentes no pueden en general ser constantes, puesto que la solución no los es. Por lo tanto:

$$\dot{\mathbf{x}}(t) = \sum \dot{a}_j(t) \phi_j$$

Al sustituir las expresiones anteriores en el sistema de ecuaciones diferenciales se tiene:

$$\sum \dot{a}_j \mathbf{B}\phi_j + \sum a_j \mathbf{A}\phi_j = \mathbf{f}(t)$$

Conviene ahora recordar las condiciones de ortogonalidad:

$$\begin{aligned} \phi_s^* \mathbf{B}\phi_r &= \delta_{rs} \\ \phi_s^* \mathbf{A}\phi_r &= \lambda_r \delta_{rs} \end{aligned}$$

Para simplificar la presentación, se ha supuesto que los vectores característicos están normalizados respecto a la matriz \mathbf{B} .

Al pre multiplicar las ecuaciones por ϕ_i^T :

$$\sum \dot{a}_j \phi_i^T \mathbf{B}\phi_j + \sum a_j \phi_i^T \mathbf{A}\phi_j = \phi_i^T \mathbf{f}(t)$$

Se observa que la mayor parte de los productos que se tienen en cada suma son cero. Sólo son distintos de cero aquellos en los que los dos índices i,j son iguales. En consecuencia se obtienen ecuaciones "desacopladas", independientes para cada componente $a_j(t)$:

$$\dot{a}_i + \lambda_i a_i = \phi_i^T \mathbf{f}(t)$$

Por otro lado, si las condiciones iniciales son $\mathbf{x}(0) = \mathbf{x}_0$, entonces:

$$\mathbf{x}_0 = \mathbf{x}(0) = \sum a_j(0) \phi_j$$

y por lo tanto, al premultiplicar por $\phi_i^T \mathbf{B}$:

$$a_i(0) = \phi_i^T \mathbf{B}\mathbf{x}_0$$

Resolver n de estas ecuaciones desacopladas es mucho más fácil que resolver un solo sistema de ecuaciones diferenciales de orden n (puede aquí hacerse un paralelismo con el caso de ecuaciones algebraicas). Además, en muchas situaciones prácticas se observa que las componentes $a(t)$ asociadas a los mayores valores característicos son

poco importantes, y pueden despreciarse. Por ello, se justifica plenamente el trabajo de determinar previamente los valores característicos.

Las ecuaciones desacopladas puede resolverse con algunos de los procedimientos numéricos tratados en la sección 6.2. Por ejemplo, si los valores característicos fueran negativos podría emplearse el método de Euler. En el caso en que las funciones $\mathbf{f}(t)$ son simples podría también pensarse en una solución analítica. Obtenidas cada una de las componentes como función de tiempo, puede regresarse a la expresión:

$$\mathbf{x}(t) = \sum a_j \phi_j$$

para hallar la solución $\mathbf{x}(t)$.

Un caso frecuente es aquel en el que $\mathbf{f}(t) = \mathbf{f}_0 g(t)$ y la función $g(t)$ está definida por una colección de valores numéricos correspondientes a valores t uniformemente espaciados. En un intervalo cualquiera puede hacerse la aproximación: $g(t) = at + b$, de donde, para cada una de las ecuaciones desacopladas se tiene:

$$\dot{a}_i + \lambda_i a_i = \phi_i^T \mathbf{f}_0 (at + b)$$

Cuya solución puede obtenerse fácilmente (sumando la solución homogénea $C e^{-\lambda t}$ y la particular, de la forma $At + B$):

Conociendo a_i al inicio del intervalo, puede obtenerse la constante C y calcularse entonces el valor de a_i al finalizar el intervalo. El proceso se repite análogamente para los sucesivos intervalos.

Cabe anotar que este procedimiento es incondicionalmente estable. En realidad, sería exacto si la función fuera efectivamente lineal por intervalos. La aproximación está en haber descrito la función $g(t)$ con un número finito de valores numéricos, a partir de los cuales se está interpolando linealmente.

Las mismas ideas pueden aplicarse a sistemas lineales de ecuaciones diferenciales de segundo orden: $\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f}$. En este caso podrían primero determinarse los valores y vectores característicos del problema: $\mathbf{K}\phi = \omega^2 \mathbf{M}\phi$. Aquí se ha supuesto que \mathbf{K} y \mathbf{M} son no sólo simétricas, sino también definidas positivas, por lo que los $\lambda = \omega^2$ son positivos.

Nuevamente, la solución puede escribirse como una combinación lineal de los vectores característicos (que en lo que sigue se han supuesto normalizados respecto a \mathbf{M}):

$$\mathbf{u}(t) = \sum a_j(t) \phi_j$$

Los vectores característicos satisfacen las condiciones de ortogonalidad:

$$\begin{aligned} \phi_s^* \mathbf{M} \phi_r &= \delta_{rs} \\ \phi_s^* \mathbf{K} \phi_r &= \omega_r^2 \delta_{rs} \end{aligned}$$

Pero, salvo en algunos casos particulares, no podría afirmarse algo similar con la matriz \mathbf{C} . Sin embargo, hay aplicaciones en que la inclusión del término $\mathbf{C}\dot{\mathbf{u}}$ es sólo un artificio para introducir disipación en las ecuaciones lineales, que realmente podría haberse considerado con una \mathbf{K} variable, dependiente de \mathbf{u} . En este contexto, podría introducirse directamente la disipación en las ecuaciones desacopladas, no siendo necesario determinar la matriz \mathbf{C} :

$$\ddot{a}_i + 2\beta\omega_i\dot{a}_i + \omega_i^2 a_i = \phi_i^T \mathbf{f}(t)$$

Estas ecuaciones pueden también ser resueltas numéricamente, por procedimientos tales como el método de diferencia central o el método de Newmark de aceleración constante.

En el caso frecuente en el que $\mathbf{f}(t) = \mathbf{f}_0 g(t)$, estando $g(t)$ definida por una colección de valores numéricos a intervalos constantes, puede seguirse un procedimiento similar al descrito antes para ecuaciones diferenciales de primer orden.

6.4.3 Consideraciones Adicionales

En el acápite 6.3.4 se mencionó que al resolver una ecuación diferencial de segundo orden el intervalo está controlado por precisión y no por estabilidad. La situación es diferente cuando se resuelven grandes sistemas de EDO. Los comentarios siguientes parecieran referirse a sistemas de EDO lineales, pero realmente son también aplicables a ecuaciones no lineales, que pueden ser linearizadas localmente.

Al integrar directamente un sistema de EDO se están haciendo operaciones equivalentes a integrar las ecuaciones desacopladas; simplemente el sistema de referencia es distinto. El procedimiento será estable cuando el intervalo de integración cumpla las condiciones de estabilidad con todos y cada uno de los modos componentes. Por lo tanto, para un método como el de la diferencia central:

$$\Delta t \leq \frac{2}{\omega_{m\acute{a}x}} = \frac{T_{m\acute{i}n}}{\pi}$$

Cuando se tiene un comportamiento no lineal los períodos T tienden a crecer (y los ω tienden a reducirse), por lo que la estimación del Δt sobre la base de las condiciones iniciales es en general suficiente.

Por otro lado:

$$\omega_1 \leq \omega_2 \leq \dots \ll \omega_{m\acute{a}x}$$

$$T_1 \geq T_2 \geq \dots \gg T_{m\acute{i}n}$$

y habitualmente al cumplir la condición de estabilidad se tendrán cientos y tal vez miles de puntos por período para los modos asociados a las menores frecuencias, que son los importantes en la respuesta. En resumen, al resolver grandes sistemas de EDO con un procedimiento condicionalmente estable, satisfacer la condición de estabilidad implica que se tiene también precisión.

En cambio, al emplear un método incondicionalmente estable es la precisión la que siempre controla el intervalo de integración. Éste debe escogerse de modo que se integren con suficiente precisión todos aquellos modos cuya participación en la respuesta es significativa.

Considere el sistema de ecuaciones diferenciales $\ddot{\mathbf{u}} + \mathbf{A}\mathbf{u} = \mathbf{0}$ en la que la matriz \mathbf{A} es:

$$\mathbf{A} = \begin{pmatrix} 200.5 & 199.5 \\ 199.5 & 200.5 \end{pmatrix}$$

Los valores característicos de la matriz \mathbf{A} son $\lambda_1 = \omega_1^2 = 1$ y $\lambda_2 = \omega_2^2 = 400$. Supóngase que las condiciones iniciales son:

$$\mathbf{u}(0) = \begin{Bmatrix} 1.01 \\ -0.99 \end{Bmatrix} \quad \dot{\mathbf{u}}(0) = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

En este caso es fácil obtener la solución exacta:

$$u = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} \cos t + 0.01 \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \cos 20t$$

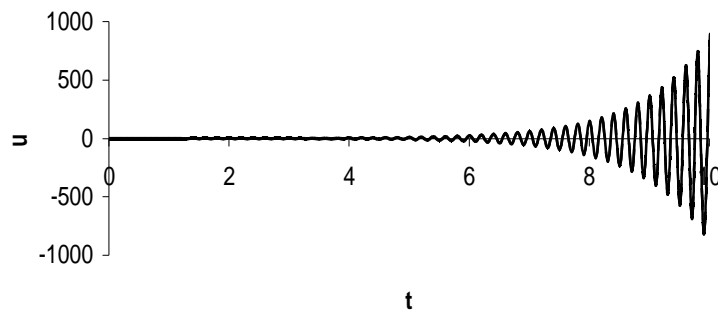
El primer modo, con $\omega_1 = 1$ y $T_1 = 2\pi$ es el importante en la solución. La contribución del segundo modo, con $\omega_2 = 20$ y $T_2 = \pi/10$ es comparativamente pequeña.

Supóngase ahora que se usa el proceso de diferencia central:

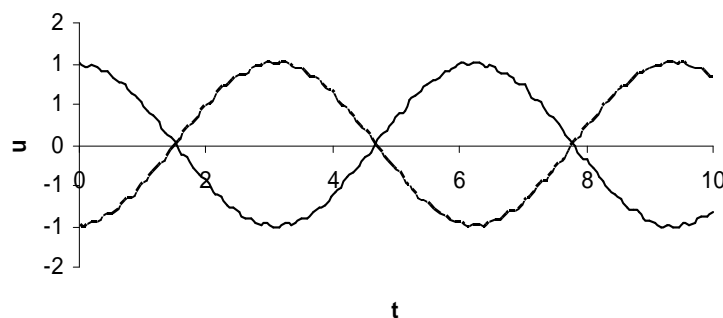
$$\begin{aligned} \ddot{\mathbf{u}}_n &= -\mathbf{A}\mathbf{u}_n \\ \dot{\mathbf{u}}_{n+1/2} &= \dot{\mathbf{u}}_{n-1/2} + \Delta t \ddot{\mathbf{u}}_n \\ \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \dot{\mathbf{u}}_{n+1/2} \end{aligned}$$

con condiciones iniciales $\mathbf{u}_0 = \begin{Bmatrix} 1.01 \\ -0.99 \end{Bmatrix}$ y $\dot{\mathbf{u}}_{1/2} = \dot{\mathbf{u}}(0) + \frac{1}{2} \Delta t \ddot{\mathbf{u}}_0 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$

Para integrar apropiadamente el primer modo sería suficiente considerar un Δt del orden de $T_1/20 \approx 0.3$. Sin embargo, es el segundo modo, poco importante en la respuesta, el que en este caso controla la estabilidad. Se requiere reducir el intervalo, de modo que: $\Delta t < 2/\omega_2 = T_2/\pi = 0.1$ lo que indirectamente hace que se obtenga precisión en la integración de la componente significativa. En las figuras siguientes se muestran resultados obtenidos con $\Delta t = 0.1001$ (el procedimiento es inestable) y con $\Delta t = 0.09$ (procedimiento estable). Los resultados serían aún más precisos si $T_1 \gg T_n$, como ocurre típicamente al resolver grandes sistemas de ecuaciones diferenciales.



Resultados obtenidos con $\Delta t = 0.1001$



Resultados obtenidos con $\Delta t = 0.09$