

Universidad de Oviedo

Facultad de Ciencias

Grado en Matemáticas

Trabajo de Fin de Grado

**Resolución Numérica de Ecuaciones Diferenciales  
Ordinarias mediante Métodos Multipaso**

Autora: Daniela Stromer Gutiérrez

---

Supervisado por:

Pablo Pérez Riera



## Agradecimientos

*A mi tutor, Pablo, por guiarme, explicarme y enseñarme de manera tan atenta.*

*A mis padres, por el apoyo incondicional que tantas veces me ayudó a seguir avanzando.*

*A mi tío Seve, a mi hermana y a mis amigos, por estar siempre.*

*A la Universidad de Oviedo y a la Universidad de Tübingen, por permitirme vivir una de las mejores etapas de mi vida tanto a nivel académico como personal.*



# Índice general

<b>Resumen</b>	<b>7</b>
<b>1. Introducción</b>	<b>9</b>
1.1. Los métodos de un paso	11
1.1.1. Los métodos Runge-Kutta	14
<b>2. Algunas familias notables de métodos multipaso</b>	<b>17</b>
2.1. Los métodos de Adams	20
2.1.1. Los métodos de Adams-Bashforth	20
2.1.2. Los métodos de Adams-Moulton	23
2.1.3. Fórmulacion de los métodos de Adams con diferencias regresivas	25
2.2. Los métodos de Nyström	31
2.3. Los métodos de Milne-Simpson	32
2.4. Los métodos BDF	33
2.5. Métodos con paso variable	35
<b>3. Consistencia, estabilidad y convergencia</b>	<b>37</b>
3.1. Convergencia	38
3.2. Estabilidad	39
3.2.1. Formulación como un método de un paso	40
3.2.2. Estabilidad de los métodos de Adams, de Nyström y de Milne-Simpson	48
3.2.3. Estabilidad de los métodos BDF	48
3.3. Consistencia y orden	49

3.3.1. Orden de los métodos de Adams, de Nyström y de Milne-Simpson . . . . .	52
3.3.2. Análisis del orden . . . . .	53
3.4. Análisis de la convergencia . . . . .	55
3.5. Métodos Predictor-Corrector . . . . .	59
3.5.1. Orden de un método predictor-corrector . . . . .	62
3.5.2. Estabilidad de un método predictor-corrector . . . . .	65
3.5.3. Implementación en Matlab . . . . .	69
<b>Bibliografía</b>	<b>71</b>

# Resumen

Este trabajo fin de grado está dedicado al estudio de los métodos multipaso de resolución numérica de ecuaciones diferenciales. El término multipaso hace referencia a que es necesario *recordar* las aproximaciones obtenidas en varios instantes previos para poder calcular la aproximación en el cada instante; se conocen también como métodos *con memoria*; en cambio, en los métodos de un solo paso, la aproximación en un instante se calcula a partir de la información disponible únicamente en el instante inmediatamente anterior. Cada una de estas familias tiene ventajas e inconvenientes; según como sea la ecuación diferencial (o el sistema diferencial) unos métodos pueden ser más eficaces que otros, aunque esta es una cuestión difícil de medir.

Hemos dividido el trabajo en tres capítulos. El primero es meramente introductorio: sirve para fijar la notación que vamos a utilizar y para recordar algunos conceptos básicos. En el Capítulo 2 construimos en detalle los métodos multipaso más conocidos, como los métodos de Adams (Bashforth y Moulton), de Nyström, de Milne-Simpson y los métodos BDF (del inglés *Backward Differentiation Formulae*).

El Capítulo 3 está dedicado al estudio de la convergencia de los métodos. Se puede decir que la noción de convergencia está basada en una idea simple: se trata de estudiar en cada instante la diferencia entre la solución exacta y la solución numérica. Sin embargo, el análisis de la convergencia es una tarea complicada; una forma de abordarlo consiste en introducir los conceptos de estabilidad y consistencia (u orden), que son habituales en Análisis Numérico, y establecer las relaciones entre estos conceptos y el de convergencia. El resultado fundamental es que la estabilidad y la consistencia son condiciones necesarias y suficientes para la convergencia. Eso justifica que se dedique también mucha atención a la caracterización de la estabilidad (en términos de la *condición de las raíces*) y de la consistencia. Al final, nos ocupamos de un tipo de métodos multipaso muy utilizados, conocidos como métodos *predictor-corrector*, porque reducen el coste computacional que conllevaría utilizar métodos implícitos pero sin disminuir su orden.





# Capítulo 1

## Introducción

Existen muchos fenómenos de interés científico que, cuando se modelan matemáticamente, dan lugar a problemas escritos en términos de sistemas de ecuaciones diferenciales. Estos problemas pueden ser de escalas muy distintas: desde escalas atómicas, en problemas de mecánica cuántica o reacciones químicas, hasta grandes escalas, como en la predicción del tiempo atmosférico, o incluso escalas cósmicas, como en el movimiento de objetos celestes. También la naturaleza de los fenómenos estudiados puede ser muy diversa, pues pueden estar relacionados con la química, la física, la biología, las finanzas...

Una ecuación diferencial ordinaria (ODE) describe cómo cambia una función; dicho de otro modo, una ecuación diferencial establece la relación que satisface la tasa instantánea de variación de una variable con respecto a otra de la cual depende. Así, la ecuación

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)), \quad t \geq t_0,$$

indica que, para cada  $t \geq t_0$ , la tasa de variación de  $\mathbf{y}$  en el instante  $t$  (o sea,  $\mathbf{y}'(t)$ ) coincide con el valor  $\mathbf{f}(t, \mathbf{y}(t))$ , donde  $\mathbf{f}$  se supone que es una función conocida. Utilizaremos letras en negrita para variables cuyo valor puede ser vectorial. Es habitual utilizar la letra  $t$  (inicial de *tiempo*) para referirse a la variable independiente porque históricamente muchos de los problemas estudiados modelan fenómenos que evolucionan con el tiempo. La variable  $\mathbf{y}$  es la variable dependiente y el valor  $\mathbf{y}(t)$  representa el *estado* de esa variable en el instante  $t$ . Es costumbre escribir la ecuación en la forma

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}), \quad t \geq t_0.$$

Esa es la forma *normal* de una ecuación diferencial ordinaria de primer orden. Por supuesto, existen ecuaciones de orden superior a 1, o escritas en forma implícita, pero no son objeto de estudio en este trabajo.

Generalmente se conoce el estado inicial del sistema, es decir, el valor de la variable dependiente

en el instante inicial  $t_0$ . Además, es natural restringir el estudio a intervalos acotados en el tiempo, sobre todo si se pretende hacer simulaciones numéricas. Por tanto, escribiremos el problema en la forma

$$\begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \boldsymbol{\eta}, \end{cases} \quad (1.1)$$

donde  $\boldsymbol{\eta}$  representa el estado en el instante  $t_0$ , que se supone conocido, y  $T$  es la longitud del intervalo  $I$  de tiempo. El problema (1.1) se dice que es un *problema de Cauchy* o *de valor inicial*.

Esta formulación engloba tanto ecuaciones escalares como sistemas de ecuaciones. Por ejemplo, el sistema de valor inicial

$$\begin{cases} y_1' = f_1(t, y_1, y_2), \\ y_2' = f_2(t, y_1, y_2), \\ y_1(t_0) = \eta_1, \\ y_2(t_0) = \eta_2. \end{cases} \quad t \in I = [t_0, t_0 + T],$$

se puede expresar en la forma (1.1) si se denota  $\mathbf{y} = (y_1, y_2)$  y

$$\begin{aligned} \mathbf{f} : I \times \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (t, \mathbf{y}) &\mapsto \mathbf{f}(t, \mathbf{y}) = \begin{pmatrix} f_1(t, y_1, y_2) \\ f_2(t, y_1, y_2) \end{pmatrix} \end{aligned}$$

Así, en general supondremos  $\mathbf{f} : I \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ , con  $m \geq 1$ . Cualquier función  $\mathbf{y} : I \rightarrow \mathbb{R}^m$  que verifique

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)), \quad \text{para cualquier } t \in I,$$

es una solución de la ecuación diferencial; si además satisface la condición inicial  $\mathbf{y}(t_0) = \boldsymbol{\eta}$ , entonces es una solución del problema.

La existencia y unicidad de solución se pueden garantizar si la función  $\mathbf{f}$  es continua en  $I \times \mathbb{R}^m$  y *lipschitziana* respecto de su segundo argumento, es decir, si existe alguna constante  $L > 0$  (llamada constante de Lipschitz) para la cual se cumpla lo siguiente:

$$t \in I, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m \quad \Rightarrow \quad \|\mathbf{f}(t, \mathbf{y}) - \mathbf{f}(t, \mathbf{z})\| \leq L \|\mathbf{y} - \mathbf{z}\|,$$

donde  $\|\cdot\|$  denota una norma en  $\mathbb{R}^m$ .

Para abreviar, representaremos por  $\mathcal{C}_L(I \times \mathbb{R}^m)$  al conjunto de las funciones continuas en  $I \times \mathbb{R}^m$  y lipschitzianas respecto del segundo argumento con constante  $L > 0$ . En lo que sigue, siempre supondremos siempre  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$  para alguna constante  $L > 0$ . En esas condiciones, se puede probar el resultado siguiente (véase [3]).

**Teorema 1.1** (de Cauchy-Lipschitz). Si  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$  para alguna constante  $L > 0$ , entonces para cada  $\boldsymbol{\eta} \in \mathbb{R}^m$  el problema de valor inicial (1.1)

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

posee una solución y sólo una.

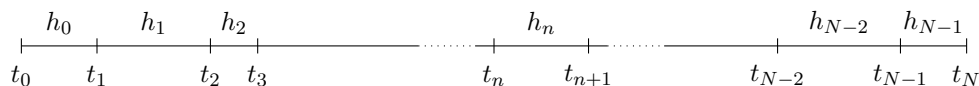
Ahora bien, aunque el problema tenga solución única, eso no significa que sea sencillo encontrarla. En general, esa es una tarea imposible salvo en pocos casos (cuando los sistemas son lineales, por ejemplo) y por ello es necesario recurrir a métodos numéricos que permitan aproximar la solución exacta del problema. Habitualmente, los métodos numéricos utilizan una *partición*  $\mathcal{P}$

$$\mathcal{P} = \{t_0, t_1, \dots, t_N : t_n < t_{n+1} \text{ para } n = 0, 1, \dots, N-1 \text{ y } t_N = t_0 + T\}$$

del intervalo  $I$  y en cada instante  $t_n$  de la partición proporcionan un valor aproximado  $\mathbf{y}_n$  del valor exacto  $\mathbf{y}(t_n)$ . Por tanto, la solución aproximada  $\{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_N\}$  es una *colección discreta* mientras que la solución exacta es una *función continua*  $\mathbf{y}$  definida en el intervalo  $I$ .

Sea  $\mathcal{P} = \{t_0, t_1, \dots, t_N\}$  una *partición* del intervalo  $I$ . La diferencia  $h_n = t_{n+1} - t_n$  es el *paso* que se necesita para avanzar desde  $t_n$  hasta  $t_{n+1}$ ; así,  $N$  representa el número de pasos que damos, y ese número depende de cada partición. Denotaremos por  $h$  al *diámetro* de la partición  $\mathcal{P}$ :

$$h = \|\mathcal{P}\| = \max\{t_{n+1} - t_n : 0 \leq n \leq N-1\} = \max\{h_n : 0 \leq n \leq N-1\}.$$



En los métodos numéricos, la aproximación  $\mathbf{y}_n$  de la solución en el instante  $t_n$  se calcula utilizando una o varias aproximaciones anteriores consecutivas. Cuando solo interviene la aproximación en el instante inmediatamente anterior, se dice que el método es de *un paso*; si se utiliza la aproximación en los dos instantes previos, entonces el método es de *dos pasos*; en general, un método es de  $k$  *pasos* si la aproximación en el instante  $t_n$  se obtiene mediante una fórmula en la que interviene la aproximación en el instante  $t_{n-k}$  y, posiblemente, algunas de las calculadas en los instantes intermedios  $t_{n-1}, t_{n-2}, \dots, t_{n-k+1}$ .

## 1.1. Los métodos de un paso

Como hemos dicho, un método de un paso calcula una aproximación  $\mathbf{y}_n$  mediante una fórmula en la que interviene la aproximación  $\mathbf{y}_{n-1}$ , pero no las anteriores. El método más sencillo es el

método de Euler explícito. Existen varias formas de motivar cómo se deduce su ley de recurrencia; por ejemplo, si  $\mathbf{y}$  es una solución de la ecuación diferencial del problema de Cauchy (1.1), entonces

$$\mathbf{y}(t_{n+1}) = \mathbf{y}(t_n) + \int_{t_n}^{t_{n+1}} \mathbf{y}'(t) dt = \mathbf{y}(t_n) + \int_{t_n}^{t_{n+1}} \mathbf{f}(t, \mathbf{y}(t)) dt. \quad (1.2)$$

Así, una forma de aproximar  $\mathbf{y}(t_{n+1})$  consiste en aproximar cada sumando del lado derecho de la aproximación anterior. Si  $\mathbf{y}_n$  es conocido, en realidad la tarea se reduce a aproximar la integral definida y eso puede hacerse mediante una fórmula de cuadratura de tipo interpolatorio; en el método de Euler explícito esa fórmula es la del *rectángulo a la izquierda*, representada en la Figura 1.1.

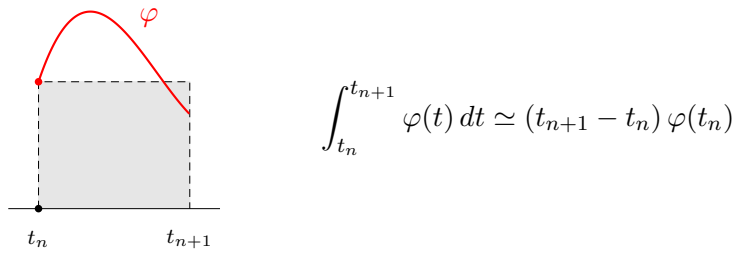


Figura 1.1: Interpretación geométrica de la fórmula del *rectángulo a la izquierda*.

Si en la expresión (1.2) sustituimos  $\mathbf{y}(t_n)$  por  $\mathbf{y}_n$  y utilizamos la fórmula de cuadratura con

$$\begin{aligned} \varphi : I &\longrightarrow \mathbb{R}^m \\ t &\longmapsto \varphi(t) = \mathbf{f}(t, \mathbf{y}(t)) \end{aligned}$$

entonces obtenemos

$$\mathbf{y}(t_{n+1}) \simeq \mathbf{y}_n + (t_{n+1} - t_n) \varphi(t_n) = \mathbf{y}_n + (t_{n+1} - t_n) \mathbf{f}(t_n, \mathbf{y}(t_n)) \simeq \mathbf{y}_n + (t_{n+1} - t_n) \mathbf{f}(t_n, \mathbf{y}_n).$$

Por tanto, el valor  $\mathbf{y}_{n+1}$  definido por

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

puede tomarse como una aproximación del valor exacto  $\mathbf{y}(t_{n+1})$ .

A partir de una aproximación  $\mathbf{y}_0 = \boldsymbol{\eta}_h$  de  $\mathbf{y}(t_0) = \boldsymbol{\eta}$  y una partición  $\mathcal{P}$  del intervalo, el método de Euler explícito queda definido por la siguiente ley de recurrencia:

$$\begin{cases} \mathbf{y}_0 = \boldsymbol{\eta}_h \\ \mathbf{y}_{n+1} = \mathbf{y}_n + h_n \mathbf{f}(t_n, \mathbf{y}_n), \quad n = 0, 1, 2, \dots, N-1. \end{cases} \quad (1.3)$$

Como vemos en la ley de recurrencia, para calcular  $\mathbf{y}_{n+1}$ , que es la aproximación de la solución

en el instante  $t_{n+1}$ , solo se utiliza la información disponible en el instante  $t_n$ , pero no en los anteriores. Por eso se dice que es un método de un paso (o sin memoria).

En general, un método de un paso es un algoritmo que proporciona una aproximación de la solución en un cierto instante a partir de la información del instante anterior. Su ley de recurrencia suele escribirse en la forma

$$\begin{cases} \mathbf{y}_0 = \boldsymbol{\eta}_h \\ \mathbf{y}_{n+1} = \mathbf{y}_n + h_n \boldsymbol{\phi}(t_n, \mathbf{y}_n, h_n), \end{cases} \quad n = 0, 1, \dots, N-1. \quad (1.4)$$

La función  $\boldsymbol{\phi}$  se denomina *función incremento*; cada método queda determinado por la elección de la función  $\boldsymbol{\phi}$ , que generalmente se construye a partir de  $\mathbf{f}$  y sus derivadas. Así, en el método de Euler explícito, la función incremento está definida por

$$\boldsymbol{\phi}(t, \mathbf{y}, h) = \mathbf{f}(t, \mathbf{y})$$

y en este caso particular no depende del paso  $h$ . Otro ejemplo de método de un paso es el *método de Euler modificado*, cuya ley es

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n \mathbf{f}(t_n + \frac{1}{2}h_n, \mathbf{y}_n + \frac{1}{2}h_n \mathbf{f}(t_n, \mathbf{y}_n)), \quad (1.5)$$

que puede deducirse a partir de la fórmula de cuadratura *del punto medio*; en este caso la función incremento es

$$\boldsymbol{\phi}(t, \mathbf{y}, h) = \mathbf{f}(t + \frac{1}{2}h, \mathbf{y} + \frac{1}{2}h \mathbf{f}(t, \mathbf{y})).$$

Como ya hemos mencionado, para cada partición  $\mathcal{P}$ , un método numérico proporciona una colección de aproximaciones  $\{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_N\}$ . La discrepancia entre cada valor exacto  $\mathbf{y}(t_n)$  y su aproximación  $\mathbf{y}_n$  es el error en  $t_n$ , que denotaremos por  $\mathbf{e}_n$ :

$$\mathbf{e}_n = \mathbf{y}(t_n) - \mathbf{y}_n, \quad n = 0, 1, 2, \dots, N.$$

Observemos que el error inicial es  $\mathbf{e}_0 = \mathbf{y}(t_0) - \mathbf{y}_0 = \boldsymbol{\eta} - \boldsymbol{\eta}_h$ , es decir, es el error cometido por utilizar  $\boldsymbol{\eta}_h$  como estimación del valor de la condición inicial.

Usualmente se necesitará que, en todos los puntos de una partición  $\mathcal{P}$ , el error (en norma) no supere una cota prefijada; por esa razón será necesario analizar el comportamiento asintótico de  $\max\{\|\mathbf{e}_n\| : 1 \leq n \leq N\}$  cuando utilizamos particiones cada vez más finas ( $h = \|\mathcal{P}\| \rightarrow 0$ ) y disminuimos el error inicial ( $\|\boldsymbol{\eta} - \boldsymbol{\eta}_h\| \rightarrow 0$ ). De forma un tanto imprecisa, diremos que un método de un paso tiene *orden de convergencia* al menos  $p$  si se verifica que

$$\max\{\|\mathbf{e}_n\| : 0 \leq n \leq N\} = \mathcal{O}(h^p), \quad \text{siempre que } \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\| = \mathcal{O}(h^p).$$

Se puede probar que el método de Euler explícito tiene orden de convergencia 1 si la función

$\mathbf{f}$  es de clase  $\mathcal{C}^1$  en  $I \times \mathbb{R}^m$ . Y el método de Euler modificado es de orden 2 bajo condiciones adecuadas de regularidad.

### 1.1.1. Los métodos Runge-Kutta

Los métodos Runge-Kutta son una familia de métodos que permiten alcanzar un orden de convergencia alto a base de reevaluaciones únicamente de la función  $\mathbf{f}$  (no intervienen sus derivadas). Un ejemplo de método de Runge-Kutta es el método de Euler modificado; su ley de recurrencia (1.5) puede reescribirse en la forma siguiente

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_n, \mathbf{y}_n), \\ \mathbf{k}_2 &= \mathbf{f}\left(t_n + \frac{1}{2}h_n, \mathbf{y}_n + \frac{1}{2}h_n\mathbf{k}_1\right), \\ \mathbf{y}_{n+1} &= \mathbf{y}_n + h_n\mathbf{k}_2. \end{aligned} \tag{1.6}$$

En general, un método de Runge-Kutta es un método de un paso cuya ley de recurrencia es de la forma

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h_n(b_1\mathbf{k}_1 + b_2\mathbf{k}_2 + \dots + b_q\mathbf{k}_q),$$

con

$$\begin{cases} \mathbf{k}_1 = \mathbf{f}(t_n + c_1h_n, \mathbf{y}_n + h_n(a_{11}\mathbf{k}_1 + a_{12}\mathbf{k}_2 + \dots + a_{1q}\mathbf{k}_q)), \\ \mathbf{k}_2 = \mathbf{f}(t_n + c_2h_n, \mathbf{y}_n + h_n(a_{21}\mathbf{k}_1 + a_{22}\mathbf{k}_2 + \dots + a_{2q}\mathbf{k}_q)), \\ \vdots \\ \mathbf{k}_q = \mathbf{f}(t_n + c_qh_n, \mathbf{y}_n + h_n(a_{q1}\mathbf{k}_1 + a_{q2}\mathbf{k}_2 + \dots + a_{qq}\mathbf{k}_q)). \end{cases}$$

Cada método de Runge-Kutta queda definido por el número  $q$  de etapas y el conjunto de coeficientes  $\{a_{ij}\}$ ,  $\{b_i\}$  y  $\{c_i\}$ . Estos coeficientes suelen disponerse en la forma siguiente, denominada *tabla de Butcher*:

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

La tabla de Butcher del método de Euler modificado, escrito en la forma (1.6) es

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

También el método de Euler explícito puede considerarse como un método de Runge-Kutta de 1 etapa, con tabla de Butcher

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

Si la matriz  $A = \{a_{ij}\}$  es estrictamente triangular inferior, los métodos de Runge-Kutta son explícitos; en caso contrario son implícitos. Los métodos explícitos son fáciles de programar y proporcionan buenas aproximaciones cuando el problema no presenta rigideces numéricas; cuando sí las hay (el problema se dice *rígido* o *stiff*) es necesario recurrir a métodos de Runge-Kutta implícitos (muy costosos computacionalmente) o a métodos multipaso. Estos son los métodos que vamos a estudiar en este trabajo.





## Capítulo 2

# Algunas familias notables de métodos multipaso

Como comentamos en el capítulo anterior, un método numérico de  $k$  pasos para resolver numéricamente un problema de valor inicial es un algoritmo que permite calcular una aproximación en el instante  $t_n$  mediante una fórmula en la que interviene la aproximación en el instante  $t_{n-k}$  y, posiblemente, algunas de las calculadas en los instantes intermedios  $t_{n-1}, t_{n-2}, \dots, t_{n-k+1}$ .

$$\begin{array}{ccccccc} & y_{n-k} & y_{n-k+1} & & y_{n-2} & y_{n-1} & y_n \\ \cdots & | & | & \cdots & | & | & | \\ & t_{n-k} & t_{n-k+1} & & t_{n-2} & t_{n-1} & t_n \\ & & & & & & \cdots \end{array}$$

Los métodos multipaso se pueden diseñar combinando diversas técnicas de aproximación. A modo de ejemplo, vamos a mostrar cómo se puede construir un método de 2 pasos utilizando técnicas de interpolación polinómica. Consideremos un problema de valor inicial escalar (por simplicidad para la exposición) escrito en la forma habitual

$$\begin{cases} y' = f(t, y), & t \in I = [t_0, t_0 + T], \\ y(t_0) = \eta, \end{cases} \quad (2.1)$$

con  $f : I \times \mathbb{R} \rightarrow \mathbb{R}$ . Denotemos por  $\mathcal{P} = \{t_0, t_1, \dots, t_N\}$  una partición del intervalo  $I$ . Para cualquier solución  $y$  del problema se puede escribir

$$y(t_n) = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} y'(t) dt = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(t, y(t)) dt. \quad (2.2)$$

Esta expresión es equivalente a (1.2), pues no hay más que un cambio de índices entre ambas (aparte de que ahora estamos suponiendo que las funciones son escalares). En un método de 2 pasos, para calcular la aproximación  $y_n$  solo podremos utilizar las aproximaciones  $y_{n-1}$  e  $y_{n-2}$ , que supondremos conocidas porque ya habrán sido calculadas.

Según la expresión (2.2), una forma de aproximar  $y(t_n)$  consiste en aproximar el lado derecho de esa identidad, es decir, cada uno de los dos sumandos. Para aproximar  $y(t_{n-1})$  disponemos de  $y_{n-1}$ , así que la tarea se reduce a estimar la integral definida. Observemos que la función que integramos es  $t \mapsto f(t, y(t))$ : es una función de una sola variable en la que interviene la solución exacta, que será desconocida. Podemos reemplazar esta función de integración por un polinomio de interpolación calculado a partir de las aproximaciones  $y_{n-1}$  e  $y_{n-2}$ , es decir,

$$\int_{t_{n-1}}^{t_n} f(t, y(t)) dt \simeq \int_{t_{n-1}}^{t_n} p(t) dt, \quad (2.3)$$

donde  $p$  es el polinomio de menor grado que satisface

$$p(t_{n-2}) = f(t_{n-2}, y_{n-2}), \quad p(t_{n-1}) = f(t_{n-1}, y_{n-1}).$$

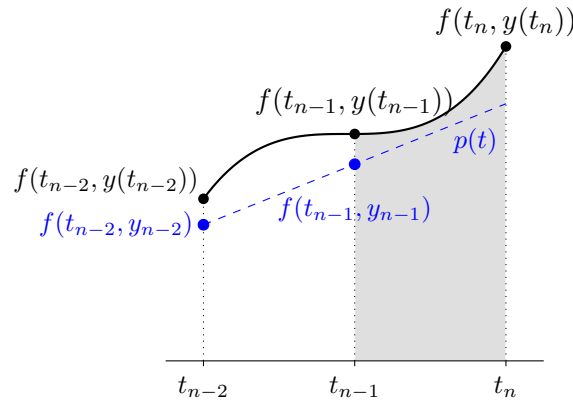


Figura 2.1: Construcción de un método de 2 pasos

El polinomio  $p$  queda determinado por dos puntos (es una recta, véase la Figura 2.1) y puede calcularse fácilmente:

$$\begin{aligned} p(t) &= f(t_{n-2}, y_{n-2}) + (t - t_{n-2}) \frac{f(t_{n-1}, y_{n-1}) - f(t_{n-2}, y_{n-2})}{t_{n-1} - t_{n-2}} \\ &= f(t_{n-1}, y_{n-1}) \frac{t - t_{n-2}}{t_{n-1} - t_{n-2}} + f(t_{n-2}, y_{n-2}) \frac{t - t_{n-1}}{t_{n-2} - t_{n-1}}. \end{aligned} \quad (2.4)$$

Si denotamos por  $L_{n-1}$  y  $L_{n-2}$  a los polinomios

$$L_{n-1}(t) = \frac{t - t_{n-2}}{t_{n-1} - t_{n-2}}, \quad L_{n-2}(t) = \frac{t - t_{n-1}}{t_{n-2} - t_{n-1}}, \quad (2.5)$$

entonces la expresión (2.4) se puede escribir en esta forma equivalente:

$$p(t) = f(t_{n-1}, y_{n-1}) L_{n-1}(t) + f(t_{n-2}, y_{n-2}) L_{n-2}(t).$$

Podemos utilizar esta fórmula para aproximar el valor de la integral en (2.2) y por tanto proponer un valor aproximado de  $y(t_n)$ :

$$\begin{aligned} y(t_n) &= y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(t, y(t)) dt \\ &\simeq y_{n-1} + \int_{t_{n-1}}^{t_n} p(t) dt \\ &= y_{n-1} + f(t_{n-1}, y_{n-1}) \int_{t_{n-1}}^{t_n} L_{n-1}(t) dt + f(t_{n-2}, y_{n-2}) \int_{t_{n-1}}^{t_n} L_{n-2}(t) dt. \end{aligned}$$

Es importante resaltar que las dos integrales que aparecen en esa expresión no dependen de los valores aproximados  $y_{n-1}$  e  $y_{n-2}$ ; el valor de esas integrales depende únicamente de la construcción de la partición, en virtud de (2.5). El método quedaría así: definimos  $y_n$  mediante

$$y_n = y_{n-1} + f(t_{n-1}, y_{n-1}) \int_{t_{n-1}}^{t_n} L_{n-1}(t) dt + f(t_{n-2}, y_{n-2}) \int_{t_{n-1}}^{t_n} L_{n-2}(t) dt. \quad (2.6)$$

En particular, si la partición  $\mathcal{P}$  es equiespaciada, es decir, si

$$t_n = t_0 + n h, \quad n = 0, 1, 2, \dots, N, \quad \text{con } h = \frac{T}{N},$$

entonces

$$\begin{aligned} \int_{t_{n-1}}^{t_n} L_{n-1}(t) dt &= \int_{t_{n-1}}^{t_n} \frac{t - t_{n-2}}{t_{n-1} - t_{n-2}} dt = \frac{3}{2} h, \\ \int_{t_{n-1}}^{t_n} L_{n-2}(t) dt &= \int_{t_{n-1}}^{t_n} \frac{t - t_{n-1}}{t_{n-2} - t_{n-1}} dt = -\frac{1}{2} h, \end{aligned}$$

y en ese caso la expresión de  $y_n$  dada por (2.6) es más sencilla:

$$y_n = y_{n-1} + h \left( \frac{3}{2} f(t_{n-1}, y_{n-1}) - \frac{1}{2} f(t_{n-2}, y_{n-2}) \right). \quad (2.7)$$

Este es el método de Adams-Bashforth de 2 pasos. De la misma forma, podríamos construir un método de 3 pasos; en este caso, la solución aproximada  $y_n$  en un instante  $t_n$  se calcula mediante una fórmula en la que intervienen las tres aproximaciones anteriores:  $y_{n-1}$ ,  $y_{n-2}$  e  $y_{n-3}$ . El proceso se puede generalizar para un número de pasos  $k$  cualquiera. Esto lo haremos para las diferentes familias de métodos multipaso que estudiaremos a lo largo este capítulo.

## 2.1. Los métodos de Adams

Como hemos comentado, la solución del problema de valor inicial (2.1) satisface la relación (2.2)

$$y(t_n) = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(t, y(t)) dt.$$

Una forma de construir un método numérico para aproximar la solución del problema consiste en reemplazar esa integral por una fórmula de cuadratura de tipo interpolatorio, de forma que se obtiene una expresión del tipo

$$y(t_n) = y(t_{n-1}) + (t_n - t_{n-1}) \sum_{j=1}^m \omega_j f(\tau_j, y(\tau_j)) + E(f),$$

donde  $\{\tau_1, \tau_2, \dots, \tau_m\}$ ,  $\{\omega_1, \omega_2, \dots, \omega_m\}$  y  $E(f)$  son, respectivamente, los nodos, los pesos y el error de la fórmula de cuadratura. La dificultad reside en cómo aproximar los valores de la solución en los nodos elegidos, es decir, cómo reemplazar  $\{y(\tau_1), y(\tau_2), \dots, y(\tau_m)\}$  por aproximaciones sin incurrir en un gran coste computacional. En los métodos de Runge-Kutta (véase [1]), los nodos de la fórmula se encuentran habitualmente en el intervalo  $[t_{n-1}, t_n]$  y cada  $y(\tau_j)$  se aproxima a su vez por una nueva fórmula de cuadratura con los mismos nodos  $\{\tau_1, \tau_2, \dots, \tau_m\}$ . Un punto de vista diferente consiste en utilizar como nodos los puntos en los que se ha calculado previamente la solución (que habrá que guardar en la memoria). Hay varias formas de llevar esto a cabo. En los métodos de Adams-Bashforth de  $k$  pasos, los nodos son los  $k$  instantes anteriores, es decir, son  $\{t_{n-1}, t_{n-2}, \dots, t_{n-k}\}$ , de forma que el nodo  $t_n$  no interviene en los cálculos: el método es explícito. En cambio, en los métodos de Adams-Moulton, se usa también el nodo  $t_n$  y los métodos resultan ser implícitos. Veremos esto con detalle a continuación.

### 2.1.1. Los métodos de Adams-Bashforth

Los métodos de Adams-Bashforth se construyen partiendo de la identidad (2.2):

$$y(t_n) = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(t, y(t)) dt.$$

Al igual que para el método de dos pasos, se puede estimar el valor de  $y(t_n)$  mediante la aproximación de cada uno de los sumandos del lado derecho de esa identidad. Consideramos entonces  $y_{n-1}$  como aproximación de  $y(t_{n-1})$ , y un polinomio de interpolación  $p_{k-1}(t)$  para reemplazar la función de integración  $f(t, y(t))$ , es decir,

$$y(t_n) \simeq y_{n-1} + \int_{t_{n-1}}^{t_n} p_{k-1}(t) dt,$$

donde, en este caso,  $p_{k-1}$  es el polinomio de interpolación de Lagrange asociado a los  $k$  nodos anteriores  $\{t_{n-1}, t_{n-2}, \dots, t_{n-k}\}$  y a los valores  $\{f(t_{n-1}, y_{n-1}), f(t_{n-2}, y_{n-2}), \dots, f(t_{n-k}, y_{n-k})\}$ . Para simplificar la escritura, denotaremos  $f_{n-i} = f(t_{n-i}, y_{n-i})$  para  $i = 1, 2, \dots, k$ .

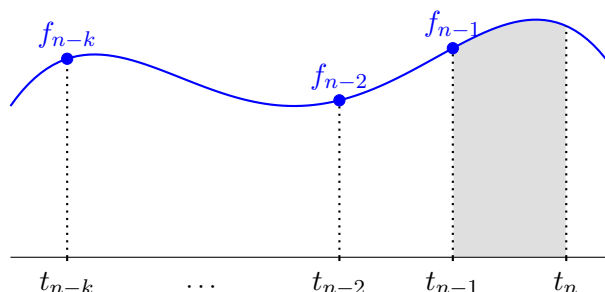


Figura 2.2: Nodos y valores de interpolación en el método de Adams-Bashforth.

Este polinomio puede escribirse en la forma (véase [3])

$$p_{k-1}(t) = f_{n-1} L_{n-1}(t) + f_{n-2} L_{n-2}(t) + \dots + f_{n-k} L_{n-k}(t) = \sum_{i=1}^k f_{n-i} L_{n-i}(t),$$

donde, para cada  $i = 1, 2, \dots, k$ , hemos denotado por

$$L_{n-i}(t) = \prod_{\substack{j=1 \\ j \neq i}}^k \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}}$$

al denominado polinomio de Lagrange del nodo  $t_{n-i}$ . El conjunto  $\{L_{n-1}(t), L_{n-2}(t), \dots, L_{n-k}(t)\}$  se conoce como la base de Lagrange del espacio vectorial de los polinomios de grado menor o igual que  $k-1$  en la variable  $t$ , que denotaremos por  $\mathbb{P}_{k-1}[t]$ .

Razonando como en el caso del método de 2 pasos, podemos escribir

$$y(t_n) \simeq y_{n-1} + \sum_{i=1}^k f_{n-i} \int_{t_{n-1}}^{t_n} L_{n-i}(t) dt.$$

En el método de Adams-Bashforth de dos pasos que construimos en la sección anterior suponíamos conocidas las dos aproximaciones inmediatamente anteriores  $y_{n-1}$  e  $y_{n-2}$ . Ahora necesitamos disponer de las  $k$  aproximaciones  $\{y_{n-1}, y_{n-2}, \dots, y_{n-k}\}$  para poder calcular las evaluaciones  $f_{n-i} = f(t_{n-i}, y_{n-i})$ . Cabe resaltar que las  $k$  integrales que aparecen en esa expresión no dependen de los valores aproximados  $\{y_{n-1}, y_{n-2}, \dots, y_{n-k}\}$ , sino únicamente de la elección de los nodos, en virtud de la expresión de cada polinomio  $L_{n-i}(t)$ . En lo que sigue supondremos que los nodos forman una partición equiespaciada del intervalo  $I = [t_0, t_0 + T]$ . Así, definiremos  $y_n$

mediante

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p_{k-1}(t) dt = y_{n-1} + \sum_{i=1}^k f_{n-i} \int_{t_{n-1}}^{t_n} L_{n-i}(t) dt. \quad (2.8)$$

Podemos reordenar los sumandos del lado derecho para que aparezcan en orden creciente:

$$y_n = y_{n-1} + \sum_{i=0}^{k-1} f_{n-k+i} \int_{t_{n-1}}^{t_n} L_{n-k+i}(t) dt, \quad (2.9)$$

En esta formulación, parece que el coeficiente que multiplica a cada evaluación  $f_{n-k+i}$  depende de  $n$ ; de ser así, el coste computacional del método sería enorme porque en cada instante  $t_n$  deberíamos calcular  $k$  coeficientes definidos por medio de una expresión integral. Sin embargo, se puede probar que tales coeficientes son independientes de  $n$ ; para ello, vamos a hacer el cambio de variable  $s \rightarrow t = t_{n-1} + s h$  en la integral definida:

$$\begin{aligned} \int_{t_{n-1}}^{t_n} L_{n-k+i} dt &= \int_{t_{n-1}}^{t_n} \prod_{\substack{j=0 \\ j \neq i}}^{k-1} \frac{t - t_{n-k+j}}{t_{n-k+i} - t_{n-k+j}} dt \\ &= \int_0^1 \prod_{\substack{j=0 \\ j \neq i}}^{k-1} \frac{t_{n-1} + s h - t_{n-k+j}}{t_{n-k+i} - t_{n-k+j}} h dt \\ &= h \int_0^1 \prod_{\substack{j=0 \\ j \neq i}}^{k-1} \frac{t_{n-k} + (k-1)h + s h - t_{n-k} - j h}{t_{n-k} + i h - t_{n-k} - j h} dt \\ &= h \int_0^1 \prod_{\substack{j=0 \\ j \neq i}}^{k-1} \frac{k-1-j+s}{i-j} dt. \end{aligned}$$

Notemos que no solo el intervalo no depende de  $n$ , sino que tampoco lo hace el valor de la integral. Si denotamos

$$b_{k-i}^{(k-1)} = \int_0^1 \prod_{\substack{j=0 \\ j \neq i}}^{k-1} \frac{k-1-j+s}{i-j} dt, \quad i = 0, 1, 2, \dots, k-1, \quad (2.10)$$

entonces la expresión (2.9) equivale a

$$y_n = y_{n-1} + h \sum_{i=0}^{k-1} b_{k-i}^{(k-1)} f_{n-k+i}. \quad (2.11)$$

Para cada elección del número natural  $k$ , el método que hemos construido es explícito puesto que la fórmula que define a  $y_n$  es explícita; además, se trata de un método de  $k$  pasos porque para calcular  $y_n$  intervienen a lo sumo los nodos  $y_{n-1}, y_{n-2}, \dots, y_{n-k}$  y este último es necesario

ya que el coeficiente de  $f_{n-k} = f(t_{n-k}, y_{n-k})$  no se anula:

$$b_k^{(k-1)} = \frac{(-1)^{k-1}}{(j-1)!} \int_0^1 \prod_{j=1}^{k-1} (k-1-j+s) dt.$$

Por ejemplo, para  $k=2$  resulta  $b_1^{(1)} = \frac{3}{2}$  y  $b_2^{(1)} = -\frac{1}{2}$ , que es lo que habíamos obtenido en (2.7).

Es preciso notar que calcular cada uno de los coeficientes  $b_{k-i}^{(k-1)}$  usando (2.10) es laborioso. Más adelante mostraremos que es posible encontrar una formulación equivalente a (2.11), que permite obtener un algoritmo sencillo para calcular dichos coeficientes.

### 2.1.2. Los métodos de Adams-Moulton

Los métodos de Adams-Moulton también se construyen partiendo de la identidad (2.2):

$$y(t_n) = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(t, y(t)) dt$$

Aproximaremos entonces  $y(t_n)$  mediante la aproximación de cada uno de los sumandos del lado derecho de esa identidad: consideraremos  $y_{n-1}$  como aproximación de  $y(t_{n-1})$  y en la integral reemplazaremos  $f(t, y(t))$  por un polinomio de interpolación  $p_k(t)$ :

$$y(t_n) \simeq y_{n-1} + \int_{t_{n-1}}^{t_n} p_k(t) dt.$$

En este caso,  $p_k$  será el polinomio de interpolación de Lagrange asociado al conjunto de  $k+1$  nodos  $\{t_n, t_{n-1}, \dots, t_{n-k}\}$  y  $k+1$  valores  $\{f_n, f_{n-1}, \dots, f_{n-k}\}$ . Depende de  $f_n = f(t_n, y_n)$ , que es

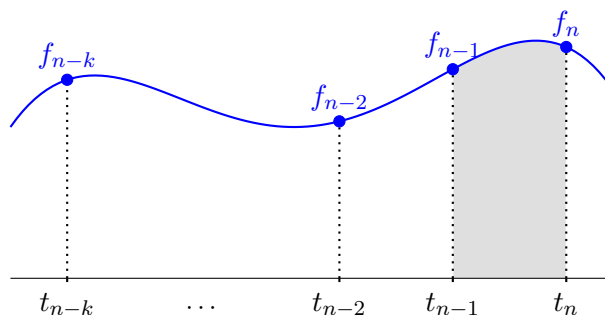


Figura 2.3: Nodos y valores de interpolación en el método de Adams-Moulton. El valor en  $t_n$  es  $f_n = f(t_n, y_n)$ , que es desconocido.

desconocido, pero podemos usando la formulación de Lagrange:

$$p_k(t) = f_n L_n(t) + f_{n-1} L_{n-1}(t) + \dots + f_{n-k} L_{n-k}(t) = \sum_{i=0}^k f_{n-i} L_{n-i}(t),$$

donde, para cada  $i = 0, 1, 2, \dots, k$ , hemos denotado por

$$L_{n-i}(t) = \prod_{\substack{j=0 \\ j \neq i}}^k \frac{t - t_{n-j}}{t_{n-i} - t_{n-j}}.$$

El conjunto  $\{L_n(t), L_{n-1}(t), \dots, L_{n-k}(t)\}$  es la base de Lagrange del espacio  $\mathbb{P}_k[t]$ . Aunque es un caso singular, se puede elegir  $k = 0$ ; el único nodo sería  $t_n$  y el polinomio  $p_0(t)$  tomaría el valor constante  $f_n = f(t_n, y_n)$ . Supondremos en lo que sigue  $k \geq 1$  y analizaremos el caso  $k = 0$  más adelante.

Razonando como en el caso de los métodos de Adams-Bashforth, la aproximación  $y_n$  de  $y(t_n)$  se define mediante la fórmula siguiente:

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p_k(t) dt = \sum_{i=0}^k f_{n-i} \int_{t_{n-1}}^{t_n} L_{n-i}(t) dt, \quad (2.12)$$

o equivalentemente, mediante una reordenación de sumandos, como

$$y_n = y_{n-1} + \sum_{i=0}^k f_{n-k+i} \int_{t_{n-1}}^{t_n} L_{n-k+i}(t) dt.$$

Vamos a probar que el coeficiente que acompaña a cada  $f_{n-k+i}$  no depende de  $n$ ; si hacemos el cambio de variable  $s \rightarrow t = t_n + sh$ , entonces resulta lo siguiente:

$$\begin{aligned} \int_{t_{n-1}}^{t_n} L_{n-k+i,k} dt &= \int_{t_{n-1}}^{t_n} \prod_{\substack{j=0 \\ j \neq i}}^k \frac{t - t_{n-k+j}}{t_{n-k+i} - t_{n-k+j}} dt \\ &= \int_{-1}^0 \prod_{\substack{j=0 \\ j \neq i}}^k \frac{t_n + sh - t_{n-k+j}}{t_{n-k+i} - t_{n-k+j}} h ds \\ &= h \int_{-1}^0 \prod_{\substack{j=0 \\ j \neq i}}^k \frac{t_n + sh - t_n + (k-j)h}{t_n - (k-i)h - t_n + (k-j)h} ds \\ &= h \int_{-1}^0 \prod_{\substack{j=0 \\ j \neq i}}^k \frac{k-j+s}{i-j} ds. \end{aligned}$$

Si denotamos

$$\bar{b}_{k-i}^{(k)} = \int_{-1}^0 \prod_{\substack{j=0 \\ j \neq i}}^k \frac{k-j+s}{i-j} ds, \quad i = 0, 1, 2, \dots, k, \quad (2.13)$$



entonces el método que estamos construyendo equivale a

$$y_n = y_{n-1} + h \sum_{i=0}^k \bar{b}_{k-i}^{(k)} f(t_{n-k+i}, y_{n-k+i}). \quad (2.14)$$

Claramente es un método implícito, porque el coeficiente de  $f(t_n, y_n)$  es no nulo:

$$\bar{b}_0^{(k)} = \int_{-1}^0 \prod_{j=0}^{k-1} \frac{k-j+s}{k-j} ds > 0$$

Además, el nodo  $t_{n-k}$  interviene en la fórmula porque

$$\bar{b}_k^{(k)} = \int_{-1}^0 \prod_{j=1}^k \frac{k-j+s}{-j} ds \neq 0.$$

Por tanto, para cada  $k \geq 1$  el método es de  $k$  pasos. ¿Y qué sucede si  $k = 0$ ? En ese caso ya hemos dicho que  $p_0(t) = f_n$  (es constante) y se obtiene de forma inmediata

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p_0(t) dt = y_{n-1} + h f_n,$$

que es un método de un paso: es el método de Euler implícito.

En los métodos implícitos es necesario resolver en cada instante la ecuación implícita (2.14) que define a  $y_n$ . De hecho, es necesario analizar previamente si tal ecuación posee solución y si es única. En general, eso es cierto siempre que el paso sea suficientemente pequeño, bajo la hipótesis de que la función  $f$  sea lipschitziana respecto de su segundo argumento. La resolución de (2.14) suele llevarse a cabo mediante un método iterativo (punto fijo, Newton...), lo que incrementa el coste computacional. Sin embargo, existen técnicas numéricas que evitan este inconveniente y hace que estos métodos sean muy utilizados. De ello hablaremos en la Sección 3.5.

### 2.1.3. Fórmulacion de los métodos de Adams con diferencias regresivas

Las formulaciones anteriores de los métodos de Adams-Bashforth y de Adams-Moulton están basadas en la expresión de Lagrange de los polinomios de interpolación  $p_{k-1}$  y  $p_k$ . En el caso de que los nodos estén igualmente espaciados, mostraremos que la expresión de estos polinomios en términos de las diferencias regresivas de Newton permite obtener los coeficientes  $b_{k-i}^{(k-1)}$  y  $\bar{b}_{k-i}^{(k)}$  de (2.10) y (2.13) de una forma bastante sencilla.

Denotemos

$$\begin{aligned} f_i &= f(t_{n-i}, y_{n-i}), \\ \nabla f_i &= f_i - f_{i-1}, \\ \nabla^j f_i &= \nabla^{j-1} f_i - \nabla^{j-1} f_{i-1}. \end{aligned}$$

Con esta notación, el polinomio de interpolación de Lagrange  $p_{k-1}$  asociado a los  $k$  nodos  $\{t_{n-1}, t_{n-2}, \dots, t_{n-k}\}$  y a los valores  $\{f_{n-1}, f_{n-2}, \dots, f_{n-k}\}$  se puede escribir en la forma siguiente (véase [3]):

$$\begin{aligned} p_{k-1}(t) &= f_{n-1} + \frac{s}{1!} \nabla f_{n-1} + \frac{s(s+1)}{2!} \nabla^2 f_{n-1} + \dots + \frac{s(s+1) \cdots (s+k-2)}{k!} \nabla^{k-1} f_{n-1} \\ &= \sum_{j=0}^{k-1} (-1)^j \binom{-s}{j} \nabla^j f_{n-1} \end{aligned}$$

con  $s = \frac{t-t_{n-1}}{h}$ . Si usamos esta representación de  $p_{k-1}(t)$  en la formulación (2.8) del método de Adams-Bashforth y hacemos un cambio de variable en la integral, resulta

$$\begin{aligned} y_n &= y_{n-1} + \int_{t_{n-1}}^{t_n} p_{k-1}(t) dt \\ &= y_{n-1} + \int_0^1 \sum_{j=0}^{k-1} (-1)^j \binom{-s}{j} \nabla^j f_{n-1} h ds \\ &= y_{n-1} + h \sum_{j=0}^{k-1} \nabla^j f_{n-1} (-1)^j \int_0^1 \binom{-s}{j} ds \\ &= y_{n-1} + h \sum_{j=0}^{k-1} \gamma_j \nabla^j f_{n-1}, \end{aligned} \tag{2.15}$$

donde hemos denotado por  $\gamma_j$  a la integral siguiente

$$\gamma_j = (-1)^j \int_0^1 \binom{-s}{j} ds. \tag{2.16}$$

Obviamente, las formulaciones (2.11) y (2.15) son equivalentes, en el sentido de que definen el mismo método. Por tanto, debe haber alguna relación entre los coeficientes  $b_{k-i}^{(k-1)}$  de la primera y los  $\gamma_j$  de la segunda. Las diferencias regresivas satisfacen las relaciones

$$\nabla^j f_n = \sum_{i=0}^j (-1)^i \binom{j}{i} f_{n-i}, \tag{2.17}$$

con  $j \geq 0$ ; si utilizamos estas relaciones en (2.15) obtenemos

$$\begin{aligned}
 y_n &= y_{n-1} + h \sum_{j=0}^{k-1} \gamma_j \nabla^j f_{n-1} \\
 &= y_{n-1} + h \sum_{j=0}^{k-1} \gamma_j \sum_{i=0}^j (-1)^i \binom{j}{i} f_{n-1-i} \\
 &= y_{n-1} + h \sum_{i=0}^{k-1} (-1)^i \sum_{j=i}^{k-1} \binom{j}{i} \gamma_j f_{n-1-i}.
 \end{aligned} \tag{2.18}$$

Por otro lado, podemos reordenar los sumandos en (2.11):

$$y_n = y_{n-1} + h \sum_{i=0}^{k-1} b_{k-i}^{(k-1)} f_{n-k+i} = y_{n-1} + h \sum_{i=0}^{k-1} b_{i+1}^{(k-1)} f_{n-1-i}.$$

Si comparamos ambas expresiones, se deduce

$$b_{i+1}^{(k-1)} = (-1)^i \sum_{j=i}^{k-1} \binom{j}{i} \gamma_j, \quad i = 0, 1, 2, \dots, k-1. \tag{2.19}$$

Esto permite calcular los coeficientes  $\{b_1^{(k-1)}, b_2^{(k-1)}, \dots, b_k^{(k-1)}\}$  a partir de  $\{\gamma_0, \gamma_1, \dots, \gamma_{k-1}\}$ , o viceversa (resolviendo un sistema lineal). Estos últimos se pueden obtener mediante la expresión (2.16), pero existe un procedimiento alternativo que es mucho más sencillo; definimos una función *generatriz*  $G$  cuya expansión en serie de potencias tenga por coeficientes a los  $\gamma_i$ , es decir,

$$G(t) = \sum_{i=0}^{+\infty} \gamma_i t^i \tag{2.20}$$

Si reemplazamos cada coeficiente  $\gamma_i$  por su expresión en (2.16) y operamos *formalmente*, obtenemos

$$G(t) = \sum_{i=0}^{+\infty} t^i (-1)^i \int_0^1 \binom{-s}{i} ds = \int_0^1 \sum_{i=0}^{+\infty} \binom{-s}{i} (-t)^i ds.$$

La función integrando del término derecho es la serie de potencias de  $(1-t)^{-s}$ ; por tanto

$$G(t) = \int_0^1 (1-t)^{-s} ds = \left[ -\frac{(1-t)^{-s}}{\ln(1-t)} \right]_{s=0}^{s=1} = -\frac{t}{(1-t)\ln(1-t)}.$$

Así, la función *generatriz*  $G$  satisface la relación

$$-\frac{\ln(1-t)}{t} G(t) = \frac{1}{1-t}.$$

Los cálculos anteriores son válidos si  $t \in [-r, r]$  con  $r \in (0, 1)$ . Por otro lado sabemos que, en ese

caso, son válidas las identidades siguientes:

$$-\frac{\ln(1-t)}{t} = 1 + \frac{t}{2} + \frac{t^2}{3} + \frac{t^3}{4} + \dots$$

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \dots$$

luego utilizando además la definición (2.20) se deduce

$$\left(1 + \frac{t}{2} + \frac{t^2}{3} + \frac{t^3}{4} + \dots\right) (\gamma_0 + \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3 + \dots) = 1 + t + t^2 + t^3 + \dots$$

y por tanto, identificando los coeficientes de cada  $t^i$  en ambos lados, resulta

$$\gamma_i + \frac{\gamma_{i-1}}{2} + \frac{\gamma_{i-2}}{3} + \dots + \frac{\gamma_1}{i} + \frac{\gamma_0}{i+1} = 1.$$

Luego hemos obtenido una ley de recurrencia para calcular los coeficientes  $\{\gamma_i\}$  de forma sencilla. Por ejemplo, para  $i = 0$  obtenemos  $\gamma_0 = 1$ , y para  $i = 1$  obtenemos  $\frac{\gamma_0}{2} + \gamma_1 = 1$ , luego  $\gamma_1 = \frac{1}{2}$ . Reiterando este proceso construimos la Tabla 2.1. Ahora, volviendo a la ecuación (2.19), podemos

$j$	0	1	2	3	4	5	6	7	8
$\gamma_j$	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$	$\frac{19087}{60480}$	$\frac{5257}{17280}$	$\frac{1070017}{3628800}$

Tabla 2.1: Coeficientes  $\{\gamma_i\}$  para los métodos de Adams-Bashforth

obtener los coeficientes  $\{b_{i+1}^{(k-1)}\}$ , y ello nos permite escribir algunos métodos de Adams-Bashforth para cada elección de  $k$ :

$$k = 1: \quad y_n = y_{n-1} + h f_{n-1},$$

$$k = 2: \quad y_n = y_{n-1} + \frac{h}{2}(3 f_{n-1} - f_{n-2}),$$

$$k = 3: \quad y_n = y_{n-1} + \frac{h}{12}(23 f_{n-1} - 16 f_{n-2} + 5 f_{n-3}),$$

$$k = 4: \quad y_n = y_{n-1} + \frac{h}{24}(55 f_{n-1} - 59 f_{n-2} + 37 f_{n-3} - 9 f_{n-4}).$$

Cabe notar que para  $k = 1$  se obtiene el método de Euler explícito.

La misma idea se puede usar para obtener los coeficientes de los métodos de Adams-Moulton. En este caso, el polinomio  $p_k$  que interpola en los  $k + 1$  nodos  $\{t_n, t_{n-1}, \dots, t_{n-k}\}$  con valores  $\{f_n, f_{n-1}, \dots, f_{n-k}\}$  se puede escribir utilizando la formulación con diferencias divididas:

$$p_k(t) = f_n + \frac{s}{1!} \nabla f_n + \frac{s(s+1)}{2!} \nabla^2 f_n + \dots + \frac{s(s+1) \dots (s+k-1)}{k!} \nabla^k f_n,$$

con  $s = \frac{t-t_n}{h}$ . Así, reemplazando esta expresión en (2.12), se obtiene

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p_k(t) dt = y_{n-1} + h \sum_{j=0}^k \gamma_j^* \nabla^j f_n, \quad (2.21)$$

donde hemos definido  $\gamma_j^* = (-1)^j \int_{-1}^0 \binom{-s}{j} ds$ . Si usamos de nuevo las fórmulas (2.17), obtenemos

$$\begin{aligned} y_n &= y_{n-1} + \int_{t_{n-1}}^{t_n} p_k(t) dt \\ &= y_{n-1} + h \sum_{j=0}^k \gamma_j^* \nabla^j f_n \\ &= y_{n-1} + h \sum_{j=0}^k \gamma_j^* \sum_{i=0}^j (-1)^i \binom{j}{i} f_{n-i} \\ &= y_{n-1} + h \sum_{i=0}^k (-1)^i \sum_{j=i}^k \gamma_j^* \binom{j}{i} f_{n-i}, \end{aligned}$$

de manera que, si comparamos con (2.14), se deduce

$$\bar{b}_i^{(k)} = (-1)^i \sum_{j=i}^k \gamma_j^* \binom{j}{i}. \quad (2.22)$$

Para calcular los coeficientes  $\{\gamma_j^*\}$ , podemos definir la función *generatriz*  $G^*$

$$G^*(t) = \sum_{i=0}^{+\infty} \gamma_i^* t^i. \quad (2.23)$$

Si sustituimos cada  $\gamma_j^*$  por su expresión, obtenemos

$$\begin{aligned} G^*(t) &= \sum_{i=0}^{+\infty} t^i (-1)^i \int_{-1}^0 \binom{-s}{i} ds \\ &= \int_{-1}^0 \sum_{i=0}^{+\infty} \binom{-s}{i} (-t)^i ds \\ &= \int_{-1}^0 (1-t)^{-s} ds \\ &= -\frac{t}{\ln(1-t)}, \end{aligned}$$

con  $t \in (-1, 1)$ . Deducimos que la función *generatriz*  $G^*$  verifica la identidad siguiente:

$$-\frac{\ln(1-t)}{t} G^*(t) = 1.$$

Nuevamente estamos suponiendo  $t \in [-r, r]$  con  $r \in (0, 1)$ .

Por otro lado sabemos que

$$-\frac{\ln(1-t)}{t} = 1 + \frac{t}{2} + \frac{t^2}{3} + \frac{t^3}{4} + \dots$$

luego utilizando además la definición (2.23), se deduce

$$(1 + \frac{t}{2} + \frac{t^2}{3} + \frac{t^3}{4} + \dots)(\gamma_0^* + \gamma_1^*t + \gamma_2^*t^2 + \gamma_3^*t^3 + \dots) = 1,$$

y por tanto

$$\begin{aligned} \gamma_0^* &= 1, \\ \gamma_i^* + \frac{\gamma_{i-1}^*}{2} + \frac{\gamma_{i-2}^*}{3} + \dots + \frac{\gamma_1^*}{i} + \frac{\gamma_0^*}{i+1} &= 0, \quad \text{para } i \geq 1. \end{aligned}$$

Así obtenemos una ley de recurrencia para calcular los coeficientes  $\{\gamma_i^*\}$ : para  $i = 0$  tenemos  $\gamma_0^* = 1$ , para  $i = 1$  obtenemos entonces  $\gamma_1^* - \frac{1}{2} = 0$ , luego  $\gamma_1^* = \frac{1}{2}$ , y reiterando de esta forma determinaríamos los demás. Sin embargo, para simplificar dichos cálculos, cabe notar que esta no es la única forma de calcular los coeficientes  $\gamma_j^*$ , pues fácilmente podemos relacionar las funciones generatrices  $G^*$  y  $G$ :

$$\left. \begin{aligned} -\frac{\ln(1-t)}{t} G(t) &= \frac{1}{1-t} \\ -\frac{\ln(1-t)}{t} G^*(t) &= 1 \end{aligned} \right\} \Rightarrow G^*(t) = (1-t)G(t),$$

luego podemos deducir

$$\begin{aligned} \gamma_0^* &= \gamma_0, \\ \gamma_i^* &= \gamma_i - \gamma_{i-1}, \quad \text{para } i \geq 1. \end{aligned}$$

Así, tomando los datos de la Tabla 2.1, obtenemos los primeros términos de la sucesión  $\{\gamma_i^*\}$ , como se muestra en la Tabla 2.2. Finalmente, los coeficientes  $\{\bar{b}_i^{(k)}\}$  se obtienen mediante la

$j$	0	1	2	3	4	5	6	7	8
$\gamma_j^*$	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$	$-\frac{863}{60480}$	$-\frac{275}{24192}$	$-\frac{33953}{3628800}$

Tabla 2.2: Coeficientes  $\{\gamma_j^*\}$  para los métodos de Adams-Moulton

relación (2.22). Los métodos de Adams-Moulton de hasta 4 pasos son los siguientes:

$$\begin{aligned} k = 0 : \quad & y_n = y_{n-1} + h f_n, \\ k = 1 : \quad & y_n = y_{n-1} + \frac{h}{2} (f_n + f_{n-1}), \\ k = 2 : \quad & y_n = y_{n-1} + \frac{h}{12} (5 f_n + 8 f_{n-1} - f_{n-2}), \\ k = 3 : \quad & y_n = y_{n-1} + \frac{h}{24} (9 f_n + 19 f_{n-1} - 5 f_{n-2} + f_{n-3}). \end{aligned}$$

Es preciso notar que, en el caso de los métodos implícitos,  $k$  solo representa el número de pasos si  $k \geq 1$ ; realmente,  $k$  es el número de nodos distintos de  $t_n$  en los que interpola el polinomio  $p_k$ . Es decir, en el caso  $k = 0$ , el polinomio  $p_0$  interpola únicamente en el instante  $t_n$ ; es un método de un paso: se trata del método de Euler implícito. Para  $k = 1$  el polinomio  $p_k$  interpola en  $t_n$  y en  $t_{n-1}$ ; también es un método de un paso, es el método del trapecio.

## 2.2. Los métodos de Nyström

Los métodos de Nyström (1925), al igual que los métodos de Adams-Bashforth, son métodos explícitos. Se construyen de manera similar: se parte de la identidad

$$y(t_n) = y(t_{n-2}) + \int_{t_{n-2}}^{t_n} f(t, y(t)) dt.$$

y se define la aproximación  $y_n$  mediante

$$y_n = y_{n-2} + \int_{t_{n-2}}^{t_n} p_{k-1}(t) dt$$

La única diferencia es que, en la identidad inicial, ahora hemos considerado el intervalo  $[t_{n-2}, t_n]$  en lugar de  $[t_{n-1}, t_n]$ .

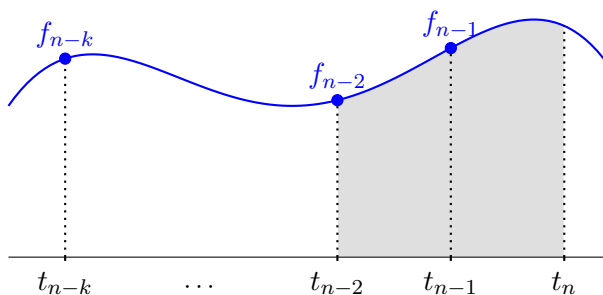


Figura 2.4: Nodos y valores de interpolación en el método de Nyström.

Siguiendo un procedimiento análogo al de los métodos de Adams-Bashforth en (2.11), podemos

escribir el método en la forma (véase [6])

$$y_n = y_{n-2} + h \sum_{j=0}^{k-1} \kappa_j \nabla^j f_{n-1},$$

con

$$\kappa_j = (-1)^j \int_{-1}^1 \binom{-s}{j} ds. \quad (2.24)$$

Estos coeficientes se muestran en la Tabla 2.3.

$j$	0	1	2	3	4	5	6	7	8
$\kappa_j$	2	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{29}{90}$	$\frac{14}{45}$	$\frac{1139}{3780}$	$\frac{41}{140}$	$\frac{32377}{113400}$

Tabla 2.3: Coeficientes  $\{\kappa_j\}$  para los métodos de Nyström.

Así, para  $k = 1$ , se obtiene el método

$$y_n = y_{n-2} + 2h f_{n-1},$$

que es conocido como el método del punto medio. Si  $k = 2$  también se obtiene lo mismo porque  $\kappa_1 = 0$ . Para  $k = 3$  resulta

$$y_n = y_{n-2} + h \left( \frac{7}{3} f_{n-1} - \frac{2}{3} f_{n-2} + \frac{1}{3} f_{n-3} \right).$$

### 2.3. Los métodos de Milne-Simpson

Los métodos de Milne-Simpson se asemejan a los métodos de Nyström porque parten de la misma identidad

$$y(t_n) = y(t_{n-2}) + \int_{t_{n-2}}^{t_n} f(t, y(t)) dt$$

pero se diferencian en que la función  $f$  se aproxima por el polinomio de interpolación  $p_k$  que interpola en los  $k + 1$  nodos  $\{t_n, t_{n-1}, \dots, t_{n-k}\}$  con valores  $\{f_n, f_{n-1}, \dots, f_{n-k}\}$ ; así, la aproximación  $y_n$  se define mediante

$$y_n = y_{n-2} + \int_{t_{n-2}}^{t_n} p_k(t) dt.$$



Por esta razón, en general se trata de métodos implícitos, como los de Adams-Moulton. Algunos ejemplos sencillos son los siguientes:

$$\begin{aligned} k = 0: \quad y_n &= y_{n-2} + 2h f_n, \\ k = 1: \quad y_n &= y_{n-2} + 2h f_{n-1}, \\ k = 2: \quad y_n &= y_{n-2} + h \left( \frac{1}{3} f_n + \frac{4}{3} f_{n-1} + \frac{1}{3} f_{n-2} \right). \end{aligned} \tag{2.25}$$

En el caso  $k = 0$ , se trata del método de Euler implícito con paso  $2h$ . Para  $k = 1$ , de nuevo nos encontramos con el método *del punto medio*. El método obtenido para  $k = 2$  se denomina método de Milne, y es una generalización de la fórmula de cuadratura de Simpson. Para  $k = 3$  se obtiene el mismo método que si  $k = 2$  (véase [6]).

## 2.4. Los métodos BDF

Las familias de métodos que hemos presentado hasta ahora tienen en común que en todos los casos se parte de una identidad de la forma

$$y(t_n) = y(t_{n-r}) + \int_{t_{n-r}}^{t_n} f(t, y(t)) dt,$$

y a continuación reemplazan la función  $f$  por el polinomio de interpolación  $p_{k-1}$  (métodos explícitos de Adams-Bashforth y Nyström) o por  $p_k$  (métodos implícitos de Adams-Moulton y Milne-Simpson). Son métodos que se basan en la integración numérica de la función  $f$  mediante una fórmula de cuadratura de tipo interpolatorio.

Existen otras técnicas para construir métodos de varios pasos. En los métodos que vamos a considerar ahora, denominados métodos BDF (del inglés *Backward Differentiation Formulae*), se utilizan fórmulas de derivación numérica. Como referencia seguiremos las ideas expuestas en [8]. Así, partimos directamente de la ecuación diferencial

$$y' = f(t, y).$$

En el instante  $t_n$ , supongamos que disponemos de las aproximaciones  $\{y_{n-1}, y_{n-2}, \dots, y_{n-k}\}$  de la solución en los  $k$  instantes anteriores  $\{t_{n-1}, t_{n-2}, \dots, t_{n-k}\}$ . Sea  $q_k$  el polinomio de interpolación asociado a los  $k+1$  pares  $\{(t_{n-i}, y_{n-i}) : i = 0, 1, 2, \dots, k\}$ ; obsérvese que hemos incluido el nodo  $t_n$  y el valor  $y_n$ , que es justamente el valor que deseamos calcular. Este valor se obtiene imponiendo que el polinomio  $q_k$  satisfaga la ecuación diferencial al menos en el instante  $t_n$ :

$$q_k'(t_n) = f(t_n, q_k(t_n)) = f(t_n, y_n)$$

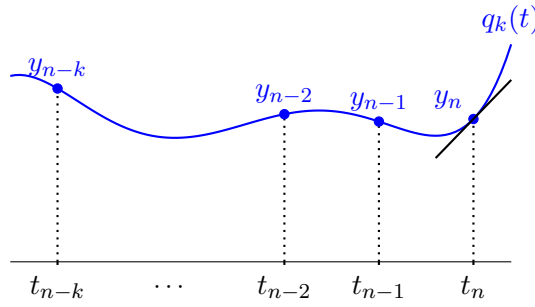


Figura 2.5: Nodos y valores para la construcción del polinomio de interpolación  $q_k$  en el método BDF de  $k$  pasos. La pendiente de la recta tangente en el instante  $t_n$  debe ser igual a  $f(t_n, y_n)$ .

A modo de ejemplo, vamos a ilustrar cómo se construye el método BDF de un paso. Primero, denotamos por  $q_1$  al polinomio que pasa por los puntos  $(t_{n-1}, y_{n-1})$  y  $(t_n, y_n)$ , es decir,

$$q_1(t) = y_{n-1} + (t - t_{n-1}) \frac{y_n - y_{n-1}}{t_n - t_{n-1}}.$$

Este polinomio depende de  $y_n$ , que es la aproximación que queremos obtener. Para ello, imponemos la condición

$$q_1'(t_n) = f(t_n, y_n). \quad (2.26)$$

Luego, si calculamos  $q_1'(t_n)$  y sustituimos su expresión en (2.26), obtenemos

$$\frac{y_n - y_{n-1}}{t_n - t_{n-1}} = f(t_n, y_n) \iff y_n = y_{n-1} + (t_n - t_{n-1}) f(t_n, y_n) = y_{n-1} + h_{n-1} f(t_n, y_n),$$

que es el método de Euler implícito.

En general, si la partición  $\mathcal{P}$  es equiespaciada, entonces la ley de recurrencia del método BDF de  $k$  pasos se puede escribir en la forma

$$\sum_{j=1}^k \frac{1}{j} \nabla^j y_n = h f_n. \quad (2.27)$$

Para  $k = 2$  resulta

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

y para  $k = 3$  se obtiene

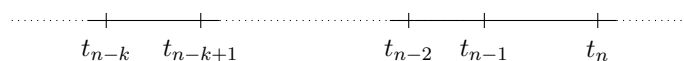
$$\frac{11}{6} y_n - 3 y_{n-1} + \frac{3}{2} y_{n-2} - \frac{1}{3} y_{n-3} = h f_n.$$

## 2.5. Métodos con paso variable

Para construir los métodos de Adams, de Nyström, de Milne-Simpson y BDF hemos supuesto, salvo en casos muy particulares, que la partición  $\mathcal{P}$  está formada por nodos equidistantes. Esta limitación puede ser muy exigente porque puede obligar a elegir un paso excesivamente pequeño si se desea que sea válido para obtener una buena aproximación en todo el intervalo; en los algoritmos modernos, el tamaño del paso se va acomodando en cada instante a la variación de la solución a lo largo del proceso de integración para cumplir con los requisitos de la precisión exigida.

Los métodos que hemos visto hasta ahora también se pueden utilizar en el caso de que la partición  $\mathcal{P} = \{t_0, t_1, \dots, t_N\}$  no esté formada por nodos equidistantes. En esta situación, es evidente que los pasos  $h_n = t_{n+1} - t_n$ , con  $n = 0, 1, 2, \dots, N-1$ , no son todos iguales.

Para calcular una aproximación en el instante  $t_n$  mediante un método numérico de  $k$  pasos, precisamos de la aproximación en el instante  $t_{n-k}$  y, si cabe, de algunas de las calculadas en los instantes intermedios  $t_{n-1}, t_{n-2}, \dots, t_{n-k+1}$ . De esta forma, estamos considerando un subconjunto de  $k+1$  nodos de la partición  $\mathcal{P}$ :



A modo de ejemplo, vamos a mostrar cómo se construye el método de Adams-Bashforth de 2 pasos, con paso variable. Para calcular la aproximación  $y_n$  en  $t_n$  intervienen las aproximaciones en los instantes  $t_{n-2}$  y  $t_{n-1}$ , como se muestra en la Figura 2.6. Denotamos por  $p_1$  al polinomio de

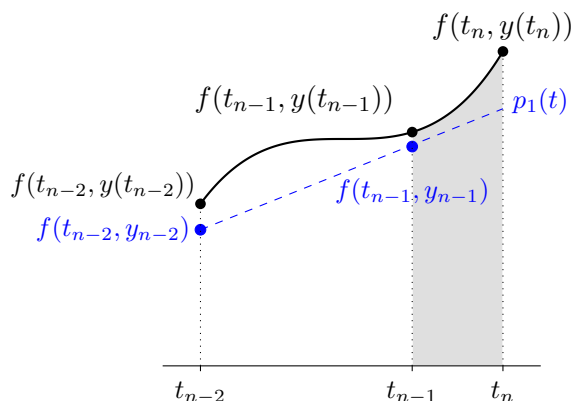


Figura 2.6: Construcción de un método de 2 pasos con paso variable. Obsérvese que los intervalos  $[t_{n-2}, t_{n-1}]$  y  $[t_{n-1}, t_n]$  tienen distinta longitud.

interpolación que pasa por los puntos  $(t_{n-2}, f_{n-2})$  y  $(t_{n-1}, f_{n-1})$ . La aproximación  $y_n$  se define del modo siguiente:

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p_1(t) dt.$$

Como  $p_1$  es una recta, en este caso su expresión se calcula fácilmente:

$$p_1(t) = f_{n-2} + \frac{f_{n-1} - f_{n-2}}{t_{n-1} - t_{n-2}} (t - t_{n-2}).$$

Se obtiene

$$\begin{aligned} y_n &= y_{n-1} + \int_{t_{n-1}}^{t_n} p_1(t) dt \\ &= y_{n-1} + h_{n-1} \left( \left(1 + \frac{1}{2} \frac{h_{n-1}}{h_{n-2}}\right) f_{n-1} - \frac{1}{2} \frac{h_{n-1}}{h_{n-2}} f_{n-2} \right). \end{aligned}$$

Esta fórmula es la generalización de la que habíamos escrito en (2.7):

$$y_n = y_{n-1} + h \left( \frac{3}{2} f(t_{n-1}, y_{n-1}) - \frac{1}{2} f(t_{n-2}, y_{n-2}) \right).$$

Este es el método de Adams-Bashforth de 2 pasos con paso variable. En los demás casos, el procedimiento es análogo al del método correspondiente con paso constante; sin embargo, las fórmulas que obtenemos no son tan sencillas porque los coeficientes dependen de las longitudes de los subintervalos.

En la técnica que hemos ilustrado con ese ejemplo, los coeficientes del método considerado son variables: los coeficientes se ajustan a la partición. De hecho, se dice que el método es de *coeficientes variables*. Una técnica alternativa es la siguiente: para avanzar desde  $t_{n-1}$  hasta  $t_n$  con un paso  $h$ , se obtienen las aproximaciones de  $f(t_n - ih, y(t_n - ih))$ , para  $i = 2, 3, \dots, k$ , mediante un polinomio de interpolación adecuado y luego se usan esas aproximaciones con el método de  $k$  pasos de coeficientes constantes. En cualquiera de los dos casos, un aspecto importante es la elección del paso en cada instante, es decir, qué paso debe elegirse para que el error se mantenga por debajo de una cierta cota a lo largo de todo el intervalo  $I = [t_0, t_0 + T]$ . Eso queda fuera de los objetivos de este trabajo, pero el lector interesado puede consultar el Capítulo 11 de [2].

## Capítulo 3

# Consistencia, estabilidad y convergencia

Todos los métodos que se han construido en el capítulo anterior tienen en común que su ley de recurrencia se puede escribir en la forma siguiente:

$$\sum_{i=0}^k \alpha_{k-i} \mathbf{y}_{n-i} = h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}), \quad \text{con } n \geq k. \quad (3.1)$$

Un método de este tipo se dice *lineal* porque la expresión (3.1) solo contiene combinaciones lineales de las aproximaciones  $\mathbf{y}_{n-i}$  y de las evaluaciones  $\mathbf{f}(t_{n-i}, \mathbf{y}_{n-i})$ , para  $i = 0, 1, 2, \dots, k$ , en contraposición con lo que ocurre en los métodos Runge-Kutta. No debe confundirse esta linealidad con la de la ecuación diferencial: es el método el que es lineal, pero la ecuación diferencial no tiene por qué serlo. Se supone que el coeficiente de  $\mathbf{y}_n$  es distinto de 0, es decir,  $\alpha_k \neq 0$ ; de hecho, es habitual imponer la condición de *normalización*  $\alpha_k = 1$  (basta dividir entre  $\alpha_k$ ). El método se dice de  $k$  pasos porque para determinar  $\mathbf{y}_n$  con la ley de recurrencia intervienen a lo sumo las aproximaciones en los  $k$  instantes anteriores; es necesario que  $\mathbf{y}_{n-k}$  o  $\mathbf{f}(t_{n-k}, \mathbf{y}_{n-k})$  intervengan en la fórmula (3.1), es decir,  $\alpha_0$  y  $\beta_0$  no pueden ser ambos iguales a 0; así supondremos también  $|\alpha_0| + |\beta_0| \neq 0$ . Si  $\beta_k = 0$ , el método es explícito; en caso contrario es implícito.

Cada método queda unívocamente determinado por las colecciones de coeficientes  $\{\alpha_0, \alpha_1, \dots, \alpha_k\}$  y  $\{\beta_0, \beta_1, \dots, \beta_k\}$ . Existe una forma más *compacta* ([2,8]) de escribir la ley de recurrencia; dada una colección  $\{\mathbf{x}_n : n = 0, 1, 2, \dots\}$ , denotaremos por  $E$  al operador de *incremento progresivo* o *desplazamiento* que actúa en la forma

$$E\mathbf{x}_n = \mathbf{x}_{n+1}, \quad E^2\mathbf{x}_n = E(E\mathbf{x}_n) = \mathbf{x}_{n+2}, \quad \dots$$

**Definición 3.1.** *Los polinomios*

$$\rho(\zeta) = \sum_{i=0}^k \alpha_i \zeta^i, \quad \sigma(\zeta) = \sum_{i=0}^k \beta_i \zeta^i, \quad \text{con } \zeta \in \mathbb{C},$$

se denominan primer y segundo polinomios característicos asociados al método (3.1).

Con esta notación, la expresión (3.1) equivale a

$$\rho(E)\mathbf{y}_{n-k} = h\sigma(E)\mathbf{f}(t_{n-k}, \mathbf{y}_{n-k}), \quad \text{con } n \geq k.$$

Por esa razón, nos referiremos al método (3.1) como el método  $(\rho, \sigma)$ .

Finalmente, es evidente que, en un método de  $k$  pasos, es necesario proporcionar las  $k$  aproximaciones  $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}$  en los primeros  $k$  instantes; en general, se utiliza un método de un paso de la misma familia ([8]).

### 3.1. Convergencia

En lo que sigue, consideraremos un problema de valor inicial escrito en la forma

$$\begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), & t \in I = [t_0, t_0 + T], \\ \mathbf{y}(t_0) = \boldsymbol{\eta}, \end{cases} \quad (3.2)$$

con  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$  para alguna constante  $L > 0$ . Bajo esas hipótesis, el problema tiene una única solución (Teorema 1.1, de Cauchy-Lipschitz).

Lo que nos planteamos ahora es comparar, en cada uno de los instantes  $t_n$  de una partición  $\mathcal{P}$ , el valor de la solución exacta  $\mathbf{y}(t_n)$  con el valor  $\mathbf{y}_n$  que proporciona el método multipaso.

**Definición 3.2.** Se dice que el método lineal  $(\rho, \sigma)$  de  $k$  pasos (3.1) es convergente si para cada problema de valor inicial (3.2), con  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$ , y si los valores de arranque  $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}$  satisfacen

$$\lim_{h \rightarrow 0^+} \max\{\|\mathbf{y}(t_n) - \mathbf{y}_n\| : t_n \in \mathcal{P} \text{ con } 0 \leq n \leq k-1\} = 0,$$

se verifica

$$\lim_{h \rightarrow 0^+} \max\{\|\mathbf{y}(t_n) - \mathbf{y}_n\| : t_n \in \mathcal{P} \text{ con } n \geq k\} = 0.$$

La diferencia  $\mathbf{e}_n = \mathbf{y}(t_n) - \mathbf{y}_n$  se denomina error de discretización global y representa la discrepancia entre el valor exacto y el valor aproximado. En esencia, la definición de convergencia significa que los errores globales se pueden hacer tan pequeños como se desee siempre que los errores iniciales sean suficientemente pequeños.

Desafortunadamente, es complicado probar la convergencia de un método de forma directa. Para ello, utilizaremos las nociones de *estabilidad* y *consistencia*. En Análisis Numérico, la estabilidad estudia el efecto que los errores de redondeo (o las perturbaciones en los datos) producen en la aproximación que proporciona un algoritmo. La consistencia se refiere a la *precisión* del método

cuando calcula un valor aproximado a partir de valores exactos. Estudiaremos y caracterizaremos estos conceptos en las secciones siguientes.

## 3.2. Estabilidad

**Definición 3.3.** Dado un método  $(\rho, \sigma)$ , se denomina operador lineal en diferencias, que denotaremos por  $\mathcal{L}_h$ , al operador que actúa sobre el elemento  $\mathbf{z}_n$  de una colección en la forma siguiente

$$\mathcal{L}_h(\mathbf{z}_n) = \sum_{i=0}^k \alpha_{k-i} \mathbf{z}_{n-i} - h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i}), \quad n \geq k \quad (3.3)$$

Este operador lineal se puede escribir de manera equivalente si utilizamos la expresión del operador *desplazamiento*:

$$\mathcal{L}_h(\mathbf{z}_n) = \sum_{i=0}^k \alpha_{k-i} \mathbf{z}_{n-i} - h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i}) = \rho(E) \mathbf{z}_{n-k} - h \sigma(E) \mathbf{f}(t_{n-k}, \mathbf{z}_{n-k}).$$

Además, una condición necesaria y suficiente para que una colección  $\{\mathbf{z}_n\}$  satisfaga la ley de recurrencia (3.1) del método  $(\rho, \sigma)$  es que se cumpla la *ecuación en diferencias*

$$\mathcal{L}_h(\mathbf{z}_n) = \mathbf{0}, \quad n \geq k. \quad (3.4)$$

**Definición 3.4.** Decimos que el operador en diferencias  $\mathcal{L}$  es estable si existen  $h^* > 0$  y  $M > 0$  tales que para cualquier partición uniforme  $\mathcal{P} = \{t_0 < t_1 < \dots < t_N = t_0 + T\}$  con diámetro  $h \leq h^*$  y cualesquiera tres colecciones  $\{\mathbf{u}_n\}_{n=0}^N$ ,  $\{\mathbf{v}_n\}_{n=0}^N$  y  $\{\boldsymbol{\omega}_n\}_{n=0}^N$  que para  $k \leq n \leq N$  verifiquen

$$\begin{aligned} \mathcal{L}_h(\mathbf{u}_n) = \mathbf{0}, \quad \text{es decir,} \quad \sum_{i=0}^k \alpha_{k-i} \mathbf{u}_{n-i} &= h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{u}_{n-i}), \\ \mathcal{L}_h(\mathbf{v}_n) = \boldsymbol{\omega}_n, \quad \text{es decir,} \quad \sum_{i=0}^k \alpha_{k-i} \mathbf{v}_{n-i} &= h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{v}_{n-i}) + \boldsymbol{\omega}_n, \end{aligned} \quad (3.5)$$

se cumple

$$\|\mathbf{v}_n - \mathbf{u}_n\| \leq M \left( \max_{0 \leq j \leq k-1} \{\|\mathbf{v}_j - \mathbf{u}_j\|\} + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right), \quad n = k, k+1, \dots, N. \quad (3.6)$$

Decimos que un método multipaso  $(\rho, \sigma)$  es estable si su operador en diferencias asociado es estable.

En la expresión (3.3), cada vez que calculamos  $\sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i})$  aparecen errores inevi-

tables de redondeo, es decir, una serie de perturbaciones, pues ningún ordenador puede calcular con precisión infinita. Ahora, para entender mejor el sentido de la definición anterior vamos a suponer que tenemos dos ordenadores: uno es capaz de calcular de forma exacta y el otro tiene una precisión finita. Las dos ecuaciones (3.5) representan esos dos procedimientos diferentes de cálculo. La primera ecuación se corresponde con el ordenador que opera de forma exacta (sería un ordenador ideal), luego no hay perturbaciones:  $\{\mathbf{u}_n\}_{n=0}^N$  representa la solución numérica *ideal* de la *ecuación en diferencias* (3.4); en la segunda ecuación de (3.5),  $\{\mathbf{v}_n\}_{n=0}^N$  representa la solución numérica de la ecuación perturbada, es la solución *real* de la ecuación en diferencias debida a la presencia *inevitable* de las *perturbaciones*  $\{\mathbf{w}_n\}_{n=0}^N$  que provocan los errores de redondeo.

Cuando el método es estable, lo que aseguramos es que las sucesiones  $\{\mathbf{u}_n\}_{n=0}^N$  y  $\{\mathbf{v}_n\}_{n=0}^N$  no serán muy diferentes, puesto que su diferencia se encuentra bajo control. Así, aunque calculemos  $\{\mathbf{v}_n\}_{n=0}^N$  en lugar de  $\{\mathbf{u}_n\}_{n=0}^N$ , entre ambas no habrá diferencias notables.

Sin embargo, eso no se puede asegurar cuando el método no es estable, de manera que podría ocurrir que las soluciones  $\{\mathbf{u}_n\}_{n=0}^N$  y  $\{\mathbf{v}_n\}_{n=0}^N$  estuviesen muy alejadas entre sí, es decir, los errores de redondeo podrían hacer que obtuviéramos una solución  $\{\mathbf{v}_n\}_{n=0}^N$  muy alejada de la esperada y por tanto no tendría sentido emplear el método numérico.

Este concepto de estabilidad también se denomina *0-estabilidad*, pues se controla la acumulación de errores cuando el paso  $h$  tiende a cero, o también *D-estabilidad*, en honor a Germund Dahlquist (véase [6, 11]). En esta sección veremos que la estabilidad del método depende únicamente de los coeficientes  $\{\alpha_0, \alpha_1, \dots, \alpha_k\}$ , esto es, del primer polinomio característico  $\rho$ . En otras palabras, podemos considerar que el polinomio  $\sigma$  es nulo, que es equivalente a la condición  $f = 0$ .

### 3.2.1. Formulación como un método de un paso

Es posible reformular un método multipaso en la forma de los métodos de un paso (véase [2, 11]). A modo de ejemplo, consideremos el método del punto medio (2.25)

$$y_n = y_{n-2} + 2h f_{n-1}.$$

Por simplicidad, supondremos que el problema es escalar. La ley de recurrencia se puede escribir en forma vectorial:

$$\begin{pmatrix} y_n \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_{n-1} \\ y_{n-2} \end{pmatrix} + 2h \begin{pmatrix} f_{n-1} \\ 0 \end{pmatrix}.$$

Se puede usar la misma idea incluso si el método es implícito; por ejemplo, el método de Adams-Moulton de 3 pasos, definido por

$$y_n = y_{n-1} + \frac{h}{24} (9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3}),$$



tambi3n se puede reescribir as3:

$$\begin{pmatrix} y_n \\ y_{n-1} \\ y_{n-2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_{n-1} \\ y_{n-2} \\ y_{n-3} \end{pmatrix} + \frac{h}{24} \begin{pmatrix} 9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3} \\ 0 \\ 0 \end{pmatrix}.$$

En general, la ley de recurrencia del m3todo multipaso (3.1), bajo la hip3tesis de normalizaci3n  $\alpha_k = 1$ , equivale a

$$\begin{pmatrix} \mathbf{y}_n \\ \mathbf{y}_{n-1} \\ \mathbf{y}_{n-2} \\ \vdots \\ \mathbf{y}_{n-k+1} \end{pmatrix} = \begin{pmatrix} -\alpha_{k-1} & -\alpha_{k-2} & \dots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{y}_{n-1} \\ \mathbf{y}_{n-2} \\ \mathbf{y}_{n-3} \\ \vdots \\ \mathbf{y}_{n-k} \end{pmatrix} + h \begin{pmatrix} \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}) \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}.$$

Si denotamos

$$\mathbf{Y}_n = \begin{pmatrix} \mathbf{y}_n \\ \mathbf{y}_{n-1} \\ \mathbf{y}_{n-2} \\ \vdots \\ \mathbf{y}_{n-k+1} \end{pmatrix}, \quad A = \begin{pmatrix} -\alpha_{k-1} & -\alpha_{k-2} & \dots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \quad (3.7)$$

y

$$\Phi(t_n, \mathbf{Y}_n, \mathbf{Y}_{n-1}, h; \mathbf{f}) = \left( \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}), \mathbf{0}, \mathbf{0}, \dots, \mathbf{0} \right)^t, \quad (3.8)$$

entonces podemos escribir

$$\begin{aligned} \mathbf{Y}_{k-1} &= (\mathbf{y}_{k-1}, \dots, \mathbf{y}_0)^t \\ \mathbf{Y}_n &= A \mathbf{Y}_{n-1} + h \Phi(t_n, \mathbf{Y}_n, \mathbf{Y}_{n-1}, h; \mathbf{f}), \quad \text{con } n \geq k. \end{aligned} \quad (3.9)$$

que es como se formulan en general los m3todos de un paso. Podemos darle una interpretaci3n a esta notaci3n: una sucesi3n  $\{\mathbf{y}_n\}_{n=0}^N$  que satisfaga la ley (3.1) queda determinada por los  $k$  t3rminos iniciales  $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}$ , puesto que para calcular un t3rmino de la sucesi3n se necesitan  $k$  t3rminos anteriores. Por ejemplo, para avanzar desde  $t_{n-1}$  hasta  $t_n$  necesitamos *recordar* los valores de las aproximaciones  $\mathbf{Y}_{n-1} = \{\mathbf{y}_{n-1}, \mathbf{y}_{n-2}, \dots, \mathbf{y}_{n-k}\}$ , pero para seguir desde  $t_n$  hasta  $t_{n+1}$  haremos intervenir a  $\mathbf{Y}_n = \{\mathbf{y}_n, \mathbf{y}_{n-1}, \dots, \mathbf{y}_{n-k+1}\}$ ; as3, hemos eliminado el *recuerdo m3s antiguo* y a3adido al principio uno *m3s reciente*.

La formulación (3.9) engloba a todos los métodos lineales multipaso que hemos construido en el capítulo anterior, y también a métodos multipaso más generales, como los métodos *predictor-corrector* (véase [7,11]). Debe exigirse siempre que la función  $\Phi$  satisfaga las hipótesis siguientes:

$$\mathbf{f} = \mathbf{0} \quad \Rightarrow \quad \Phi = \mathbf{0}, \quad (3.10)$$

y además que exista  $L_\Phi > 0$  tal que para cualesquiera  $\{\mathbf{Y}_{n-1}, \mathbf{Y}_n\}$  y  $\{\mathbf{Z}_{n-1}, \mathbf{Z}_n\}$  se verifique

$$\|\Phi(t_n, \mathbf{Y}_n, \mathbf{Y}_{n-1}, h; \mathbf{f}) - \Phi(t_n, \mathbf{Z}_n, \mathbf{Z}_{n-1}, h; \mathbf{f})\| \leq L_\Phi \sum_{i=0}^k \|\mathbf{y}_{n-i} - \mathbf{z}_{n-i}\|. \quad (3.11)$$

Estas dos hipótesis se cumplen en el caso de un método lineal multipaso  $(\rho, \sigma)$  si  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$ ; en efecto, la función  $\Phi$  está definida por la expresión (3.8) y por tanto es idénticamente nula si  $\mathbf{f} = \mathbf{0}$ . Además,

$$\begin{aligned} & \|\Phi(t_n, \mathbf{Y}_n, \mathbf{Y}_{n-1}, h; \mathbf{f}) - \Phi(t_n, \mathbf{Z}_n, \mathbf{Z}_{n-1}, h; \mathbf{f})\| \\ &= \left\| \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}) - \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i}) \right\| \\ &\leq \beta L \sum_{i=0}^k \|\mathbf{y}_{n-i} - \mathbf{z}_{n-i}\| \end{aligned}$$

con  $\beta = \max\{\beta_0, \beta_1, \dots, \beta_k\}$ .

Es conveniente adaptar la definición de estabilidad para la notación (3.9).

**Definición 3.5.** *Decimos que el método (3.9) es estable si existen  $h^* > 0$  y  $M > 0$  tales que para cualquier partición uniforme  $\mathcal{P} = \{t_0 < t_1 < \dots < t_N = t_0 + T\}$  con diámetro  $h \leq h^*$  y cualesquiera tres colecciones  $\{\mathbf{U}_n\}_{n=k-1}^N$ ,  $\{\mathbf{V}_n\}_{n=k-1}^N$  y  $\{\omega_n\}_{n=k-1}^N$  que para  $k \leq n \leq N$  verifiquen*

$$\begin{aligned} \mathbf{U}_n &= A \mathbf{U}_{n-1} + h \Phi(t_n, \mathbf{U}_n, \mathbf{U}_{n-1}, h; \mathbf{f}), \\ \mathbf{V}_n &= A \mathbf{V}_{n-1} + h \Phi(t_n, \mathbf{V}_n, \mathbf{V}_{n-1}, h; \mathbf{f}) + \omega_n, \end{aligned} \quad (3.12)$$

se cumple

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq M \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| + \sum_{i=k}^n \|\omega_i\| \right), \quad n = k, k+1, \dots, N. \quad (3.13)$$

Como hemos dicho anteriormente, nos proponemos probar que, si el método es estable para  $\mathbf{f} = \mathbf{0}$ , entonces lo es para cualquier  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$ . Obviamente, el recíproco también es cierto. Para  $\mathbf{f} = \mathbf{0}$ , en virtud de la hipótesis (3.10), la definición de estabilidad se enunciaría así: existen  $h^* > 0$  y  $M > 0$  tales que para cualquier partición uniforme  $\mathcal{P} = \{t_0 < t_1 < \dots < t_N = t_0 + T\}$  con diámetro  $h \leq h^*$  y cualesquiera tres colecciones  $\{\mathbf{U}_n\}_{n=k-1}^N$ ,  $\{\mathbf{V}_n\}_{n=k-1}^N$  y  $\{\omega_n\}_{n=k-1}^N$  que

para  $k \leq n \leq N$  verifiquen

$$\begin{aligned} \mathbf{U}_n &= A \mathbf{U}_{n-1}, \\ \mathbf{V}_n &= A \mathbf{V}_{n-1} + \boldsymbol{\omega}_n, \end{aligned} \quad (3.14)$$

se cumple

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq M \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right), \quad n = k, k+1, \dots, N. \quad (3.15)$$

Supondremos en lo que sigue que el problema de valor inicial es escalar, pero eso no resta generalidad a nuestra exposición (véase [2]).

**Teorema 3.1.** *El método (3.9) es estable para cualquier  $f \in \mathcal{C}_L(I \times \mathbb{R})$  si y sólo si lo es para  $f = 0$ .*

*Demostración.* Realmente, basta probar que si el método es estable para  $f = 0$  entonces también lo es para cualquier función  $f \in \mathcal{C}_L(I \times \mathbb{R})$ , ya que  $f = 0$  es una función continua y lipschitziana respecto del segundo argumento.

Supongamos entonces que el método es estable para  $f = 0$ . Sea  $f \in \mathcal{C}_L(I \times \mathbb{R})$  y sea  $L$  su constante de Lipschitz. Sean  $h^*$  y  $M > 0$  las constantes de la definición de estabilidad para  $f = 0$ .

Sean  $\{\mathbf{U}_n\}_{n=k-1}^N$ ,  $\{\mathbf{V}_n\}_{n=k-1}^N$  y  $\{\boldsymbol{\omega}_n\}_{n=k-1}^N$  que para  $k \leq n \leq N$  verifiquen (3.12). Restando miembro a miembro se tiene

$$\mathbf{V}_n - \mathbf{U}_n = A(\mathbf{V}_{n-1} - \mathbf{U}_{n-1}) + \boldsymbol{\psi}_n$$

con

$$\boldsymbol{\psi}_n = h \left( \Phi(t_n, \mathbf{V}_n, \mathbf{V}_{n-1}, h; f) - \Phi(t_n, \mathbf{U}_n, \mathbf{U}_{n-1}, h; f) \right) + \boldsymbol{\omega}_n,$$

Como el método es estable para  $f = 0$ , entonces

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq M \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| + \sum_{i=k}^n \|\boldsymbol{\psi}_i\| \right), \quad k \leq n \leq N.$$

Para acotar  $\|\boldsymbol{\psi}_i\|$  utilizaremos la hipótesis (3.11):

$$\begin{aligned} \|\boldsymbol{\psi}_i\| &\leq h \|\Phi(t_i, \mathbf{V}_i, \mathbf{V}_{i-1}, h; f) - \Phi(t_i, \mathbf{U}_i, \mathbf{U}_{i-1}, h; f)\| + \|\boldsymbol{\omega}_i\| \\ &\leq h \beta L \sum_{j=0}^k |v_{i-j} - u_{i-j}| + \|\boldsymbol{\omega}_i\| \\ &\leq h \beta L \|\mathbf{V}_i - \mathbf{U}_i\| + h \beta L \|\mathbf{V}_{i-1} - \mathbf{U}_{i-1}\| + \|\boldsymbol{\omega}_i\|. \end{aligned}$$

Si sustituimos en la expresión anterior, obtenemos

$$\begin{aligned} \|\mathbf{V}_n - \mathbf{U}_n\| &\leq M \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| \right. \\ &\quad \left. + h\beta L \sum_{i=k}^n \|\mathbf{V}_i - \mathbf{U}_i\| + h\beta L \sum_{i=k}^n \|\mathbf{V}_{i-1} - \mathbf{U}_{i-1}\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right) \\ &= M \left( (1 + h\beta L) \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| \right. \\ &\quad \left. + 2h\beta L \sum_{i=k}^{n-1} \|\mathbf{V}_i - \mathbf{U}_i\| + h\beta L \|\mathbf{V}_n - \mathbf{U}_n\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right), \quad k \leq n \leq N. \end{aligned}$$

Podemos suponer que las constantes  $h^*$  y  $M > 0$  se eligen de modo que  $h^* L \beta M < 1$ ; de no ser así, basta reducir el valor de  $h^*$ . Sea  $\widetilde{M} = M/(1 - h^* L \beta M)$  y sea  $h \in (0, h^*)$ ; si denotamos  $\mathbf{Z}_n = \mathbf{V}_n - \mathbf{U}_n$  y también  $\boldsymbol{\omega}_{k-1} = \mathbf{V}_{k-1} - \mathbf{U}_{k-1} = \mathbf{Z}_{k-1}$ , entonces

$$\begin{aligned} \|\mathbf{Z}_n\| &\leq \widetilde{M} \left( (1 + h\beta L) \|\mathbf{Z}_{k-1}\| + 2h\beta L \sum_{i=k}^{n-1} \|\mathbf{Z}_i\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right) \\ &\leq \widetilde{M} \left( 2h\beta L \sum_{i=k-1}^{n-1} \|\mathbf{Z}_i\| + \sum_{i=k-1}^n \|\boldsymbol{\omega}_i\| \right), \quad k \leq n \leq N. \end{aligned}$$

Ahora, por inducción (véase [2]) se puede probar

$$\|\mathbf{Z}_n\| \leq \widetilde{M} (1 + 2h\beta L \widetilde{M})^n \sum_{i=k-1}^n \|\boldsymbol{\omega}_i\| \quad k \leq n \leq N.$$

Finalmente, si tenemos en cuenta la desigualdad  $(1 + x) \leq e^x$  y que además  $nh \leq Nh = T$ , entonces podemos deducir

$$\|\mathbf{Z}_n\| \leq \widetilde{M} e^{2T\beta L \widetilde{M}} \sum_{i=k-1}^n \|\boldsymbol{\omega}_i\| \quad k \leq n \leq N.$$

Si deshacemos los cambios de variable y denotamos  $\widehat{M} = \widetilde{M} e^{2T\beta L \widetilde{M}}$ , hemos probado

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq \widehat{M} \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right), \quad k \leq n \leq N,$$

que concuerda con (3.13); por lo tanto, el método es estable.  $\square$

El resultado anterior establece que la estabilidad para  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$  se garantiza si y solo si el método es estable para  $\mathbf{f} = \mathbf{0}$ . El objetivo ahora va a ser caracterizar de forma sencilla cuándo podemos asegurar esto último. De nuevo nos restringiremos al caso escalar, por sencillez en la notación.

**Teorema 3.2.** *El método (3.9) es estable para  $f = 0$  si y sólo si existe una constante  $K > 0$  tal que*

$$\sup_n \|A^n\| \leq K. \quad (3.16)$$

*Demostración.* Supongamos que el método es estable para  $f = 0$ . Sean  $h^* > 0$  y  $M > 0$  las constantes de la definición de estabilidad y sea  $h \leq h^*$ . Sean  $\{\mathbf{U}_n\}_{n=k-1}^N$ ,  $\{\mathbf{V}_n\}_{n=k-1}^N$  y  $\{\boldsymbol{\omega}_n\}_{n=k-1}^N$  tres colecciones que verifiquen

$$\begin{aligned} \mathbf{U}_n &= A\mathbf{U}_{n-1}, \\ \mathbf{V}_n &= A\mathbf{V}_{n-1} + \boldsymbol{\omega}_n, \end{aligned}$$

para  $k \leq n \leq N$ . Si elegimos  $\boldsymbol{\omega}_n = \mathbf{0}$ , entonces

$$\mathbf{V}_n - \mathbf{U}_n = A^{n-k+1}(\mathbf{V}_{k-1} - \mathbf{U}_{k-1}), \quad k \leq n \leq N.$$

Como el método es estable, sabemos que se verifica

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq M \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\|, \quad k \leq n \leq N.$$

y por tanto se deduce (3.16) con  $K = M$ .

Recíprocamente, supongamos que existe  $K > 0$  tal que se verifica (3.16). Sean  $\{\mathbf{U}_n\}_{n=k-1}^N$ ,  $\{\mathbf{V}_n\}_{n=k-1}^N$  y  $\{\boldsymbol{\omega}_n\}_{n=k-1}^N$  que para  $k \leq n \leq N$  verifiquen

$$\begin{aligned} \mathbf{U}_n &= A\mathbf{U}_{n-1}, \\ \mathbf{V}_n &= A\mathbf{V}_{n-1} + \boldsymbol{\omega}_n. \end{aligned}$$

Entonces, por inducción, se prueba

$$\mathbf{V}_n - \mathbf{U}_n = A^{n-k+1}(\mathbf{V}_{k-1} - \mathbf{U}_{k-1}) + \sum_{i=k}^n A^{n-i} \boldsymbol{\omega}_i, \quad k \leq n \leq N$$

y, en virtud de (3.16), concluimos que también se verifica

$$\|\mathbf{V}_n - \mathbf{U}_n\| \leq K \left( \|\mathbf{V}_{k-1} - \mathbf{U}_{k-1}\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right).$$

Así, el método es estable para  $f = 0$ . □

En general, el resultado anterior es de difícil aplicación práctica. Veremos a continuación cómo expresar la condición (3.16) en otros términos. Recordemos (véase [5]) que, dada cualquier matriz

$A$ , real o compleja, existe una matriz  $C$  invertible tal que  $A$  puede factorizarse en la forma

$$A = C J C^{-1},$$

donde  $J$  es una matriz de Jordan (se dice que es la *forma canónica de Jordan* de  $A$ ) cuyos elementos en la diagonal son los valores propios de  $A$ . La matriz  $J$  es una matriz *diagonal por bloques*

$$J = \begin{pmatrix} J_{q_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{q_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{q_r}(\lambda_r) \end{pmatrix},$$

donde cada bloque  $J_{q_i}(\lambda_i)$  es una matriz cuadrada  $q_i \times q_i$  de la forma

$$J_q(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} = \lambda I_q + N_q.$$

La condición (3.16) equivale a  $\sup_n \|J^n\| \leq K$ , aunque no necesariamente con la misma constante  $K$ . Al ser la forma canónica de Jordan  $J$  una matriz diagonal por bloques, resulta que  $J^n$  también lo es y sus bloques son las potencias de los bloques de  $J$ . Ahora bien, si  $J_q(\lambda)$  es un bloque de tamaño  $q \times q$ , entonces

$$J_q^n(\lambda) = \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots & \binom{n}{q-2}\lambda^{n-q+2} & \binom{n}{q-1}\lambda^{n-q+1} \\ 0 & \lambda^n & \binom{n}{1}\lambda^{n-1} & \dots & \binom{n}{q-3}\lambda^{n-q+3} & \binom{n}{q-2}\lambda^{n-q+2} \\ 0 & 0 & \lambda^n & \dots & \binom{n}{q-4}\lambda^{n-q+4} & \binom{n}{q-3}\lambda^{n-q+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ 0 & 0 & 0 & \dots & 0 & \lambda^n \end{pmatrix}.$$

Las normas  $\|J^n\|$  estarán uniformemente acotadas si y solo si se cumplen las condiciones siguientes:

- $|\lambda| \leq 1$ , para cada valor propio  $\lambda$ ;

- Si  $|\lambda| = 1$ , entonces el valor propio  $\lambda$  no puede tener asociados bloques de tamaño  $q > 1$ , sino solo bloques de tamaño  $q = 1$  (bloques simples). Dicho de otro modo, la multiplicidad de  $\lambda$  como valor propio coincide con la dimensión del subespacio propio asociado a  $\lambda$ .

En el caso particular de un método lineal multipaso  $(\rho, \sigma)$ , recordemos que la matriz  $A$  tiene la expresión siguiente (véase (3.7)), bajo la hipótesis  $\alpha_k = 1$ :

$$A = \begin{pmatrix} -\alpha_{k-1} & -\alpha_{k-2} & \dots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

y, por tanto, el polinomio característico está definido por

$$p(\lambda) = \det(A - \lambda I_k) = (-1)^k (\lambda^k + \alpha_{k-1} \lambda^{k-1} + \dots + \alpha_1 \lambda + \alpha_0) = (-1)^k \rho(\lambda).$$

Además, el subespacio propio  $S(\lambda)$  asociado a un valor propio es siempre de dimensión 1, más concretamente, viene dado por

$$S(\lambda) = \langle \mathbf{v} \rangle, \quad \text{con } \mathbf{v} = (\lambda^{k-1}, \lambda^{k-2}, \dots, \lambda, 1).$$

Esto motiva la definición siguiente (véase [1])

**Definición 3.6** (Condición de las raíces). *Se dice que un polinomio  $P$  satisface la condición de las raíces si todas sus raíces se encuentran en el disco unidad cerrado del plano complejo y todas las raíces múltiples están en el disco unidad abierto, es decir, si se cumple*

$$P(\zeta) = 0 \quad \Rightarrow \quad \begin{cases} |\zeta| < 1, \\ \text{ó} \\ |\zeta| = 1 \text{ y } \zeta \text{ es una raíz simple.} \end{cases}$$

Los Teoremas 3.1 y (3.2), junto con los razonamientos anteriores y la Definición (3.6) (*condición de las raíces*), permiten enunciar el resultado fundamental siguiente sobre la estabilidad de los métodos lineales multipaso.

**Teorema 3.3.** *El método lineal multipaso  $(\rho, \sigma)$  es estable para cualquier  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$  si y sólo si el polinomio  $\rho$  satisface la condición de las raíces.*

### 3.2.2. Estabilidad de los métodos de Adams, de Nyström y de Milne-Simpson

En los métodos de Adams de  $k$  pasos, los coeficientes  $\{\alpha_0, \alpha_1, \dots, \alpha_k\}$  del primer polinomio característico son

$$\alpha_k = 1, \quad \alpha_{k-1} = -1, \quad \alpha_j = 0, \quad \text{si } j = 0, 1, 2, \dots, k-2.$$

Como consecuencia,  $\rho(\xi) = \xi^{k-1}(\xi - 1)$  y, por tanto,  $\rho$  tiene la raíz simple  $\xi_1 = 1$  y la raíz múltiple  $\xi_2 = 0$ . Por el Teorema 3.3, todos los métodos de Adams son estables.

Lo mismo podemos afirmar para los métodos de Nyström y los métodos de Milne-Simpson, donde  $\rho(\xi) = \xi^{k-2}(\xi^2 - 1)$ . Así,  $\rho$  tiene las raíces simples  $\xi_1 = 1$  y  $\xi_2 = -1$  y la raíz múltiple  $\xi_3 = 0$ ; todos estos métodos cumplen la condición de las raíces.

### 3.2.3. Estabilidad de los métodos BDF

Estudiar la estabilidad en los métodos BDF es algo más complicado. El polinomio  $\rho$  se obtiene a partir de la fórmula general (2.27)

$$\sum_{j=1}^k \frac{1}{j} \nabla^j y_n = h f_n.$$

En un método de  $k$  pasos, a cada sumando  $\nabla^j y_n$  le corresponde el polinomio

$$\binom{j}{0} \xi^k - \binom{j}{1} \xi^{k-1} + \binom{j}{2} \xi^{k-2} + \dots + (-1)^{k-j} \binom{j}{j} \xi^{k-j} = \xi^{k-j} (\xi - 1)^j.$$

Por tanto, el polinomio  $\rho$  de un método BDF de  $k$  pasos es de la forma

$$\rho(\xi) = \sum_{j=1}^k \frac{1}{j} \xi^{k-j} (\xi - 1)^j.$$

Claramente  $\xi = 1$  es una raíz simple de  $\rho$ . Para determinar la estabilidad del método es necesario localizar todas las raíces de  $\rho$ . En la Figura 3.1 se representa el disco unidad y la localización de las raíces de  $\rho$  para algunos valores de  $k$ . Para  $1 \leq k \leq 6$  el polinomio  $\rho$  tiene una raíz en  $z = 1$  y  $k - 1$  raíces en el interior del disco unidad. Si  $k = 7$ ,  $\rho$  tiene una raíz en  $z = 1$ , cuatro raíces en el interior del disco y otras dos son aproximadamente  $z = 0.0768 \pm 1.0193i$ , que quedan fuera del disco; lo mismo sucede si  $k = 8$ .

**Teorema 3.4.** *El método BDF de  $k$  pasos es estable si  $k \leq 6$  y es inestable si  $k \geq 7$ .*

La demostración de este resultado excede de los objetivos de este trabajo, pero puede consultarse en el libro de Hairer, Nørsett y Wanner [6]. La estabilidad para  $k \leq 6$ , que es un número finito de



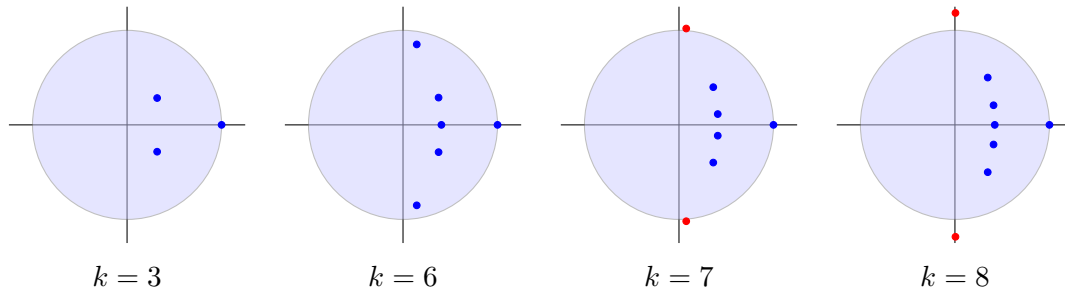


Figura 3.1: Localización de las raíces del polinomio  $\rho$  para algunas elecciones de  $k$ .

casos, fue probada por Mitchell y Craggs (1953). La segunda parte, la inestabilidad para  $k \geq 7$ , es más complicada. La primera demostración se debe a Cryer en 1971 y 1972 [4], pero también existen demostraciones alternativas debidas a Creedon y Miller (1975), Grigorieff (1977) y Hairer y Wanner (1983). En esta última, los autores prueban la inestabilidad para  $k \geq 12$ ; en el caso  $k < 6 < 12$  recurren a procedimientos numéricos porque se trata de un número finito de casos. Además, en lugar de estudiar las raíces de  $\rho$ , realizan el cambio de variable  $\xi = 1/(1 - z)$  y consideran el polinomio

$$p(z) = (1 - z)^k \rho\left(\frac{1}{1 - z}\right) = \sum_{j=1}^k \frac{z}{j}.$$

Con este cambio de variable, el interior del disco unidad se transforma en el exterior del disco centrado en  $z = 1$  y radio 1.

### 3.3. Consistencia y orden

La estabilidad está relacionada con la propagación de los errores de redondeo, es una propiedad de la ley de recurrencia. El concepto que estudiaremos en esta sección tiene que ver con la forma en la que el método aproxima a la solución de la ecuación diferencial.

En un método lineal de  $k$  pasos, la aproximación  $\mathbf{y}_n$  en  $t_n$  depende de las  $k$  aproximaciones anteriores  $\mathbf{y}_{n-1}, \mathbf{y}_{n-2}, \dots, \mathbf{y}_{n-k}$  a través de la relación (3.1). Es utópico pensar que si esas  $k$  aproximaciones anteriores coincidiesen con el valor exacto de la solución en los  $k$  instantes previos, es decir, si

$$\mathbf{y}_{n-1} = \mathbf{y}(t_{n-1}), \quad \mathbf{y}_{n-2} = \mathbf{y}(t_{n-2}), \quad \dots \quad \mathbf{y}_{n-k} = \mathbf{y}(t_{n-k}),$$

entonces el método debiera proporcionar un valor exacto en  $t_n$ , es decir,  $\mathbf{y}_n = \mathbf{y}(t_n)$ . Ahora bien, cabe pensar que cuanto menor sea la discrepancia entre el valor exacto y el valor aproximado calculado a partir de los datos exactos, mejores propiedades numéricas tendrá el método elegido. Sobre eso tratan los conceptos de consistencia y orden y a eso dedicaremos esta sección.

**Definición 3.7.** Para cada solución  $\mathbf{y}$  de la ecuación diferencial del problema (3.2), llamaremos

error local (o error de consistencia) en el instante  $t$  con paso  $h$  asociado a la solución  $\mathbf{y}$ , y lo denotaremos por  $\varepsilon(t, h, \mathbf{y})$ , a la diferencia entre  $\mathbf{y}(t)$  y el valor  $\mathbf{y}^*(t)$  proporcionado por el método a partir de los  $k$  valores exactos  $\mathbf{y}(t-h), \mathbf{y}(t-2h), \dots, \mathbf{y}(t-kh)$ , es decir,

$$\varepsilon(t, h, \mathbf{y}) = \mathbf{y}(t) - \mathbf{y}^*(t), \quad (3.17)$$

donde  $\mathbf{y}^*(t)$  está definido por

$$\alpha_k \mathbf{y}^*(t) + \sum_{i=1}^k \alpha_{k-i} \mathbf{y}(t-ih) = h \beta_k \mathbf{f}(t, \mathbf{y}^*(t)) + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t-ih, \mathbf{y}(t-ih)). \quad (3.18)$$

El error local es un error intrínseco del método, es la medida de lo que le falta al valor proporcionado por el método, si pudiese ser calculado a partir de los valores exactos en los instantes anteriores, para coincidir con el valor exacto en el instante actual, como se muestra en la Figura 3.2. De alguna forma, representa el precio que se paga por resolver el problema de forma aproximada.

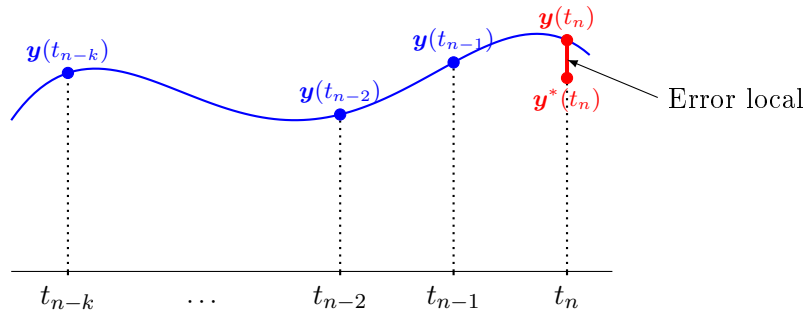


Figura 3.2: Ilustración del error local. La aproximación  $\mathbf{y}^*(t_n)$  se supone que se ha calculado utilizando los valores exactos en los  $k$  instantes anteriores.

**Definición 3.8.** Se dice que un método tiene orden al menos  $p$  para el problema (3.2) si para cualquier solución  $\mathbf{y} \in \mathcal{C}^{p+1}(I)$  de la ecuación diferencial del problema el error local satisface la condición

$$\sum_{n=0}^{N-1} \|\varepsilon(t_n, h, \mathbf{y})\| = \mathcal{O}(h^p). \quad (3.19)$$

Si el orden de un método es  $p \geq 1$ , se dice que el método es consistente.

Realmente, basta con probar  $\varepsilon(t_n, h, \mathbf{y}) = \mathcal{O}(h^{p+1})$ . En efecto, en esas condiciones existe  $C > 0$  tal que

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

ya que  $T$ , que representa la longitud del intervalo  $I$ , es igual a la suma de todos los pasos.

El mayor inconveniente de la definición del error local es debido a la presencia del término  $\mathbf{y}^*(t)$ . Veremos cómo podemos usar una expresión equivalente y de manejo más sencillo; además, en ella intervendrá el operador en diferencias que hemos introducido en el estudio de la estabilidad.

Sin pérdida de generalidad, podemos suponer  $\alpha_k = 1$ ; además, en el caso de los métodos explícitos, sabemos que  $\beta_k = 0$ , luego

$$\mathbf{y}^*(t) = - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}(t - ih) = h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t - ih, \mathbf{y}(t - ih)).$$

y por tanto

$$\varepsilon(t, h, \mathbf{y}) = \mathbf{y}(t) - \mathbf{y}^*(t) = \sum_{i=0}^k \alpha_{k-i} \mathbf{y}(t - ih) - h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t - ih, \mathbf{y}(t - ih)).$$

Esta expresión del error local se puede escribir en forma más breve utilizando la definición del operador en diferencias  $\mathcal{L}_h$  del método que hemos definido en (3.3); se verifica

$$\varepsilon(t, h, \mathbf{y}) = \mathcal{L}_h(\mathbf{y}(t)).$$

Si el método es implícito, entonces  $\beta_k \neq 0$  y en ese caso  $\mathbf{y}^*(t)$  está definido de forma implícita a través de (3.18). En este caso, el error local y el operador en diferencias  $\mathcal{L}_h$  no coinciden, como vamos a ver:

$$\begin{aligned} \varepsilon(t, h, \mathbf{y}) &= \mathbf{y}(t) - \mathbf{y}^*(t) \\ &= \mathbf{y}(t) + \sum_{i=1}^k \alpha_{k-i} \mathbf{y}(t - ih) - h \beta_k \mathbf{f}(t, \mathbf{y}^*(t)) + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t - ih, \mathbf{y}(t - ih)) \\ &= \mathcal{L}_h(\mathbf{y}(t)) + h \beta_k (\mathbf{f}(t, \mathbf{y}(t)) - \mathbf{f}(t, \mathbf{y}^*(t))). \end{aligned}$$

Si  $\mathbf{f} \in \mathcal{C}_L(I \times \mathbb{R}^m)$ , entonces

$$\begin{aligned} \|\varepsilon(t, h, \mathbf{y})\| &\leq \|\mathcal{L}_h(\mathbf{y}(t))\| + h \beta_k \|\mathbf{f}(t, \mathbf{y}(t)) - \mathbf{f}(t, \mathbf{y}^*(t))\| \\ &\leq \|\mathcal{L}_h(\mathbf{y}(t))\| + h \beta_k L \|\mathbf{y}(t) - \mathbf{y}^*(t)\| \\ &= \|\mathcal{L}_h(\mathbf{y}(t))\| + h \beta_k L \|\varepsilon(t, h, \mathbf{y})\|, \end{aligned}$$

con  $L$  la constante de Lipschitz de  $\mathbf{f}$ . Igualmente, si escribimos

$$\mathcal{L}_h(\mathbf{y}(t)) = \varepsilon(t, h, \mathbf{y}) - h \beta_k (\mathbf{f}(t, \mathbf{y}(t)) - \mathbf{f}(t, \mathbf{y}^*(t))).$$

y utilizamos un razonamiento análogo, se obtiene

$$\|\mathcal{L}_h(\mathbf{y}(t))\| \leq \|\varepsilon(t, h, \mathbf{y})\| + h \beta_k L \|\varepsilon(t, h, \mathbf{y})\|.$$

Esta acotación y la anterior se pueden escribir de forma conjunta:

$$(1 - h \beta_k L) \|\varepsilon(t, h, \mathbf{y})\| \leq \|\mathcal{L}_h(\mathbf{y}(t))\| \leq (1 + h \beta_k L) \|\varepsilon(t, h, \mathbf{y})\|.$$

y por tanto, en norma,  $\varepsilon(t, h, \mathbf{y})$  y  $\mathcal{L}_h(\mathbf{y}(t))$  tienen el mismo comportamiento cuando  $h \rightarrow 0$ . Esto significa que en la expresión (3.19) de la definición de orden  $p$  podemos reemplazar  $\|\varepsilon(t, h, \mathbf{y})\|$  por  $\|\mathcal{L}_h(\mathbf{y}(t))\|$ . Fijémonos en que, si  $\mathbf{y}$  es solución de la ecuación diferencial  $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$ , entonces

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}(t)) &= \sum_{i=0}^k \alpha_{k-i} \mathbf{y}(t - ih) - h \sum_{i=0}^k \beta_{k-i} \mathbf{f}(t - ih, \mathbf{y}(t - ih)) \\ &= \sum_{i=0}^k \alpha_{k-i} \mathbf{y}(t - ih) - h \sum_{i=0}^k \beta_{k-i} \mathbf{y}'(t - ih) \end{aligned}$$

y su desarrollo en serie de potencias del paso  $h$  se reduce al desarrollo de funciones de una variable. De hecho, en esta última expresión no interviene la función  $\mathbf{f}$  del problema y, por tanto, basta suponer que  $\mathbf{y}$  es una función suficientemente regular.

### 3.3.1. Orden de los métodos de Adams, de Nyström y de Milne-Simpson

En los métodos de Adams es posible determinar el comportamiento asintótico de  $\mathcal{L}_h(\mathbf{y}(t_n))$  y por tanto deducir el orden de estos métodos (véase [10]). Por ejemplo, para  $k \geq 1$ , el método de Adams-Bashforth de  $k$  pasos está definido por

$$\mathbf{y}_n = \mathbf{y}_{n-1} + \int_{t_{n-1}}^{t_n} \mathbf{p}_{k-1}(t) dt,$$

donde  $\mathbf{p}_{k-1}$  es el polinomio de interpolación de Lagrange asociado a los  $k$  nodos  $\{t_{n-1}, \dots, t_{n-k}\}$  y valores  $\{\mathbf{f}(t_{n-1}, \mathbf{y}_{n-1}), \dots, \mathbf{f}(t_{n-k}, \mathbf{y}_{n-k})\}$ .

Así, el operador  $\mathcal{L}_h$  verifica

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}(t_n)) &= \mathbf{y}(t_n) - \mathbf{y}(t_{n-1}) - \int_{t_{n-1}}^{t_n} \mathbf{p}_{k-1}^*(t) dt \\ &= \int_{t_{n-1}}^{t_n} (\mathbf{y}'(t) - \mathbf{p}_{k-1}^*(t)) dt, \end{aligned}$$

donde  $\mathbf{p}_{k-1}^*(\cdot)$  es ahora el polinomio de interpolación asociado a los  $k$  nodos  $\{t_{n-1}, \dots, t_{n-k}\}$  y a los valores exactos  $\{\mathbf{f}(t_{n-1}, \mathbf{y}(t_{n-1})), \dots, \mathbf{f}(t_{n-k}, \mathbf{y}(t_{n-k}))\}$ , es decir,  $\mathbf{p}_{k-1}^*$  interpola a  $\mathbf{y}'$ . Usando la fórmula de representación del error en la interpolación polinómica, se deduce

$$\mathcal{L}_h(\mathbf{y}(t_n)) = \int_{t_{n-1}}^{t_n} \frac{1}{k!} \mathbf{y}^{(k)}(\theta_t) (t - t_{n-1}) \cdot (t - t_{n-2}) \cdot \dots \cdot (t - t_{n-k}) dt,$$

para algún  $\theta_t$  comprendido entre  $t$  y los nodos, es decir,  $\theta_t \in [t_{n-k}, t]$ .

En el intervalo de integración  $[t_{n-1}, t_n]$  el producto  $(t - t_{n-1}) \cdot (t - t_{n-2}) \cdot \dots \cdot (t - t_{n-k})$  que aparece en el integrando no cambia de signo; por tanto se verifica

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}(t_n)) &= \frac{1}{k!} \mathbf{y}^{(k)}(c) \int_{t_{n-1}}^{t_n} (t - t_{n-1}) \cdot (t - t_{n-2}) \cdot \dots \cdot (t - t_{n-k}) dt \\ &= \frac{1}{k!} \mathbf{y}^{(k)}(c) \int_0^1 h^{k+1} s \cdot (s+1) \cdot \dots \cdot (s+k-1) ds \\ &= \mathbf{y}^{(k)}(c) h^{k+1} (-1)^k \int_0^1 \binom{-s}{k} ds \\ &= \mathbf{y}^{(k)}(c) \gamma_k h^{k+1}, \end{aligned}$$

para algún  $c \in [t_{n-k}, t_n]$ . El método es de orden  $k$ , pues  $\mathcal{L}_h(\mathbf{y}(t_n)) = \mathcal{O}(h^{k+1})$ .

Razonando de forma análoga, se puede probar que el método de Adams-Moulton de  $k$  pasos, si  $k \geq 2$ , es de orden  $k+1$ . De los dos métodos de un sólo paso, el de Euler implícito es de orden 1 (se obtiene con  $k=0$ ) y el del trapecio (eligiendo  $k=1$ ) es de orden 2.

Con los métodos de Nyström y de Milne-Simpson los razonamientos son similares.

### 3.3.2. Análisis del orden

Como hemos dicho, para determinar el orden de un método lineal de  $k$  pasos, lo más fácil es desarrollar  $\mathcal{L}_h(\mathbf{y}(t))$  en potencias del paso  $h$ , con  $\mathbf{y}$  suficientemente regular. Haremos un desarrollo de Taylor de cada sumando  $\mathbf{y}(t - ih)$  eligiendo  $t^* = t - kh$  como centro del desarrollo; lo mismo haremos con  $\mathbf{y}'(t - ih)$ :

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}(t)) &= \sum_{i=0}^k \alpha_{k-i} \mathbf{y}(t - ih) - h \sum_{i=0}^k \beta_{k-i} \mathbf{y}'(t - ih) \\ &= \sum_{i=0}^k \alpha_i \mathbf{y}(t^* + ih) - h \sum_{i=0}^k \beta_i \mathbf{y}'(t^* + ih). \end{aligned}$$

Ahora basta desarrollar tanto  $\mathbf{y}(t^* + ih)$  como  $\mathbf{y}'(t^* + ih)$  en potencias crecientes del paso  $h$ , con  $\mathbf{y} \in \mathcal{C}^{p+1}(I)$ . Por un lado, se tiene

$$\mathbf{y}(t^* + ih) = \sum_{n=0}^p \frac{i^n h^n}{n!} \mathbf{y}^{(n)}(t^*) + \frac{i^{p+1} h^{p+1}}{(p+1)!} \mathbf{y}^{(p+1)}(\tau_i)$$

para algún  $\tau_i \in (t^*, t^* + ih)$ . Del mismo modo,

$$\mathbf{y}'(t^* + ih) = \sum_{n=0}^{p-1} \frac{i^n h^n}{n!} \mathbf{y}^{(n+1)}(t^*) + \frac{i^p h^p}{p!} \mathbf{y}^{(p+1)}(\sigma_i)$$

para algún  $\sigma_i \in (t^*, t^* + ih)$ . Entonces,

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}(t)) &= \underbrace{\sum_{i=0}^k \alpha_i \left( \sum_{n=0}^p \frac{i^n h^n}{n!} \mathbf{y}^{(n)}(t^*) \right) - h \sum_{i=0}^k \beta_i \left( \sum_{n=0}^{p-1} \frac{i^n h^n}{n!} \mathbf{y}^{(n+1)}(t^*) \right)}_A \\ &+ \underbrace{\sum_{i=0}^k \alpha_i \frac{i^{p+1} h^{p+1}}{(p+1)!} \mathbf{y}^{(p+1)}(\tau_i) - h \sum_{i=0}^k \beta_i \frac{i^p h^p}{p!} \mathbf{y}^{(p+1)}(\sigma_i)}_B \end{aligned}$$

Claramente  $B = \mathcal{O}(h^{p+1})$  si  $\mathbf{y} \in \mathcal{C}^{p+1}(I)$ . En cuanto al término  $A$ , si se reordena se puede escribir en la forma siguiente:

$$\begin{aligned} A &= \mathbf{y}(t^*) \sum_{i=0}^k \alpha_i + h \mathbf{y}'(t^*) \sum_{i=0}^k (i \alpha_i - \beta_i) + \frac{h^2}{2!} \mathbf{y}''(t^*) \sum_{i=0}^k (i^2 \alpha_i - 2i \beta_i) \\ &+ \frac{h^3}{3!} \mathbf{y}'''(t^*) \sum_{i=0}^k (i^3 \alpha_i - 3i^2 \beta_i) + \dots + \frac{h^p}{p!} \mathbf{y}^{(p)}(t^*) \sum_{i=0}^k (i^p \alpha_i - p i^{p-1} \beta_i). \end{aligned}$$

En lo sucesivo, denotaremos

$$\begin{aligned} c_0 &= \sum_{i=0}^k \alpha_i, \\ c_1 &= \sum_{i=0}^k (i \alpha_i - \beta_i), \\ c_n &= \sum_{i=0}^k (i^n \alpha_i - n i^{n-1} \beta_i), \quad n \geq 2, \end{aligned}$$

lo que nos permite escribir

$$\mathcal{L}_h(\mathbf{y}(t)) = c_0 \mathbf{y}(t^*) + \frac{c_1}{1!} h \mathbf{y}'(t^*) + \frac{c_2}{2!} h^2 \mathbf{y}''(t^*) + \dots + \frac{c_p}{p!} h^p \mathbf{y}^{(p)}(t^*) + \mathcal{O}(h^{p+1})$$

Por tanto, el orden del método es al menos  $p$  si y solo si

$$c_0 = c_1 = \dots = c_p = 0,$$

y es de orden  $p$  si además  $c_{p+1} \neq 0$ . El valor de  $\frac{c_{p+1}}{(p+1)!}$  se denomina coeficiente principal del error local. Se puede probar que estas condiciones y el valor del coeficiente principal no dependen del

punto elegido para hacer el desarrollo (véase [2,6]).

Según la Definición 3.19, un método es consistente si su orden es al menos 1, es decir, si  $\varepsilon(t_n, h, \mathbf{y}) = \mathcal{O}(h^2)$ . Por lo anterior, eso es equivalente a exigir  $c_0 = c_1 = 0$ . Estas condiciones se pueden relacionar con los polinomios  $\rho$  y  $\sigma$  del método. En efecto,

$$c_0 = \sum_{i=0}^k \alpha_i = \rho(1), \quad \text{y} \quad c_1 = \sum_{i=0}^k (i \alpha_i - \beta_i) = \rho'(1) - \sigma(1).$$

Por tanto, el método (3.1) es consistente si y solo si  $\rho(1) = 0$  y  $\rho'(1) = \sigma(1)$ . Estas dos condiciones también se denominan *condiciones de consistencia*.

**Ejemplo 3.1.** *El método de Adams Moulton de 3 pasos es*

$$y_n = y_{n-1} + h \left( \frac{9}{24} f_n + \frac{19}{24} f_{n-1} - \frac{5}{24} f_{n-2} + \frac{1}{24} f_{n-3} \right),$$

*luego los coeficientes del método son*

$$\alpha_3 = 1, \quad \alpha_2 = 1, \quad \alpha_1 = 0, \quad \alpha_0 = 0, \quad \beta_3 = \frac{9}{24}, \quad \beta_2 = \frac{19}{24}, \quad \beta_1 = -\frac{5}{24}, \quad \beta_0 = \frac{1}{24}.$$

*Así, obtenemos*

$$c_0 = c_1 = c_2 = c_3 = c_4 = 0, c_5 \neq 0,$$

*luego podemos concluir que el método es de orden 4. Se podría haber llegado a la misma conclusión siguiendo el procedimiento descrito en la sección 3.3.1.*

**Ejemplo 3.2.** *El método de Nyström de tres pasos es*

$$y_n = y_{n-2} + h \left( \frac{7}{3} f_{n-1} - \frac{2}{3} f_{n-2} + \frac{1}{3} f_{n-3} \right),$$

*luego los coeficientes del método son*

$$\alpha_3 = 1, \quad \alpha_2 = 0, \quad \alpha_1 = -1, \quad \alpha_0 = 0, \quad \beta_3 = 0, \quad \beta_2 = \frac{7}{3}, \quad \beta_1 = -\frac{2}{3}, \quad \beta_0 = \frac{1}{3}.$$

*Así obtenemos*

$$c_0 = c_1 = c_2 = c_3 = 0, c_4 = 8,$$

*luego el método es de orden 3 y el coeficiente principal del error local es 8.*

## 3.4. Análisis de la convergencia

En esta sección estudiaremos condiciones que nos ayuden a determinar si un método es convergente. El estudio de la estabilidad y el orden del método que hemos llevado a cabo en las

secciones anteriores serán de gran ayuda, como veremos a continuación.

**Teorema 3.5.** *Si un método lineal multipaso (3.1) es estable y consistente, entonces es convergente.*

*Demostración.* Supongamos que el método multipaso (3.1) es estable y consistente, y escribamos el esquema del método en forma matricial

$$\begin{aligned} \mathbf{Y}_{k-1} &= (\mathbf{y}_{k-1}, \dots, \mathbf{y}_0)^t \\ \mathbf{Y}_n &= A \mathbf{Y}_{n-1} + h \Phi(t_n, \mathbf{Y}_n, \mathbf{Y}_{n-1}, h; \mathbf{f}), \quad \text{con } n \geq k, \end{aligned} \quad (3.20)$$

donde  $\mathbf{Y}_n$  representa la colección de soluciones aproximadas. Por otro lado, consideramos

$$\bar{\mathbf{Y}}_n = \begin{pmatrix} \mathbf{y}(t_n) \\ \mathbf{y}(t_{n-1}) \\ \mathbf{y}(t_{n-2}) \\ \vdots \\ \mathbf{y}(t_{n-k+1}) \end{pmatrix}.$$

Entonces se tiene

$$\bar{\mathbf{Y}}_n = A \bar{\mathbf{Y}}_{n-1} + h \Phi(t_n, \bar{\mathbf{Y}}_n, \bar{\mathbf{Y}}_{n-1}, h; \mathbf{f}) + \boldsymbol{\omega}_n,$$

donde

$$\boldsymbol{\omega}_n = \begin{pmatrix} \varepsilon(t_n, h, \mathbf{y}) \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

con  $t_n \in \mathcal{P}$ ,  $k \leq n \leq N$ . Como el método es estable, existe  $M > 0$  tal que

$$\begin{aligned} \|\bar{\mathbf{Y}}_n - \mathbf{Y}_n\| &\leq M \left( \|\bar{\mathbf{Y}}_{k-1} - \mathbf{Y}_{k-1}\| + \sum_{i=k}^n \|\boldsymbol{\omega}_i\| \right) \\ &= M \left( \|\bar{\mathbf{Y}}_{k-1} - \mathbf{Y}_{k-1}\| + \sum_{i=k}^n \|\varepsilon(t_i, h, \mathbf{y})\| \right) \quad n = k, k+1, \dots, N. \end{aligned}$$

Como el método es consistente, se tiene

$$\lim_{h \rightarrow 0^+} \sum_{i=k}^N \|\varepsilon(t_i, h, \mathbf{y})\| = 0.$$



Por tanto, si también se cumple

$$\lim_{h \rightarrow 0^+} \max\{\|\bar{\mathbf{Y}}_n - \mathbf{Y}_n\| : t_n \in \mathcal{P} \text{ con } 0 \leq n \leq k-1\} = 0,$$

entonces se obtiene

$$\lim_{h \rightarrow 0^+} \max\{\|\bar{\mathbf{Y}}_n - \mathbf{Y}_n\| : n \geq k\} = 0.$$

y, por lo tanto, el método es convergente.  $\square$

**Teorema 3.6.** *Si un método lineal multipaso (3.1) es convergente, entonces es estable y consistente.*

*Demostración.* Veamos primero que el método es estable. Para estudiar la estabilidad, consideremos el problema de valor inicial escalar

$$\begin{cases} y' = 0, & \text{en } [0, 1], \\ y(0) = 0, \end{cases}$$

cuya solución es la función nula ( $y(t) = 0$  para cualquier  $t \in [0, 1]$ ). Para probar que el método es estable es suficiente ver que el polinomio  $\rho$  cumple la condición de las raíces, por el Teorema 3.3. Supongamos que el polinomio  $\rho$  tiene alguna raíz  $\xi \neq 0$  (si 0 fuese la única raíz de  $\rho$ , claramente se cumpliría la condición de las raíces y el método sería estable).

Para cada  $N \in \mathbb{N}$ , sea  $h = \frac{1}{N}$  y sea  $\mathcal{P}_N = \{t_0, t_1, \dots, t_N\}$  una partición uniforme del intervalo  $[0, 1]$ , de diámetro  $h$ . Como  $f = 0$  y el problema es escalar, la ley de recurrencia del método (3.1) es:

$$\alpha_k y_n + \alpha_{k-1} y_{n-1} + \dots + \alpha_1 y_{n-k+1} + \alpha_0 y_{n-k} = 0, \quad n = k, k+1, \dots, N.$$

Si tomamos como valores de arranque  $y_j = h \xi^j$ , para  $j = 0, 1, \dots, k-1$ , entonces se prueba por inducción que  $y_n = h \xi^n$ , para  $n = k, k+1, \dots, N$ . Como el método es convergente y los valores de arranque satisfacen

$$\lim_{h \rightarrow 0^+} \max\{|y(t_n) - y_n| : t_n \in \mathcal{P}_N \text{ con } 0 \leq n \leq k-1\} = \lim_{h \rightarrow 0^+} h \max\{1, |\xi|, |\xi|^2, \dots, |\xi|^{k-1}\} = 0,$$

entonces se verifica

$$\lim_{h \rightarrow 0^+} \max\{|y(t_n) - y_n| : t_n \in \mathcal{P}_N \text{ con } k \leq n \leq N\} = 0.$$

Ahora bien, como  $y(t_n) = 0$  para cualquier  $n$  y además  $y_n = h \xi^n$ , entonces eso es equivalente a

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \max\{|\xi^n| : k \leq n \leq N\} = 0.$$

Así, necesariamente,  $|\xi| \leq 1$ , es decir, o bien  $|\xi| < 1$  o bien  $|\xi| = 1$ . Si todas las raíces son tales

que  $|\xi| < 1$ , por la *condición de las raíces*, el método ya sería estable, luego solo nos falta ver qué ocurre si existe alguna raíz con  $|\xi| = 1$ . En este caso, para que el método satisfaga la condición de las raíces, debemos probar que  $\xi$  es una raíz simple. Por reducción al absurdo, supongamos que no lo es, esto es, supongamos que  $\rho$  tiene una raíz  $\xi$  de multiplicidad al menos 2, tal que  $|\xi| = 1$ . De esta forma  $\xi$  es raíz de  $\rho$  y de  $\rho'$ , luego  $\rho(\xi) + \xi\rho'(\xi) = 0$ . Entonces escogemos como valores de arranque  $y_j = h(j+1)\xi^j$  y por inducción probamos que  $y_n = h(n+1)\xi^n$ , para  $n = k, k+1, \dots, N$ . Como el método es convergente, se tiene que cumplir

$$\lim_{h \rightarrow 0^+} \max\{|y_n| : k \leq n \leq N\} = 0$$

Pero aquí llegamos a una contradicción, ya que si  $|\xi| = 1$  entonces

$$|y_n| = |h(n+1)\xi^n| = h(n+1) = \frac{n+1}{N},$$

y por tanto se obtiene

$$\lim_{h \rightarrow 0^+} \max\{|y_n| : k \leq n \leq N\} = \lim_{N \rightarrow +\infty} \frac{1}{N} \max\{(n+1) : k \leq n \leq N\} = \lim_{N \rightarrow +\infty} \frac{N+1}{N} = 1.$$

Luego, necesariamente  $\xi$  es una raíz simple. En cualquiera de los casos, hemos probado que se cumple la *condición de las raíces* y que, por lo tanto, el método es estable.

Probaremos ahora la consistencia del método. Para ello, basta probar que se cumple  $\rho(1) = 0$  y  $\rho'(1) = \sigma(1)$ , porque esto equivale a  $c_0 = c_1 = 0$ . Para lo primero, consideraremos el problema

$$\begin{cases} y' = 0, & \text{en } [0,1], \\ y(0) = 1, \end{cases}$$

cuya solución es  $y(t) = 1$  para cualquier  $t \in [0,1]$ . Si elegimos como valores de arranque los valores exactos  $y_0 = \dots = y_{k-1} = 1$ , que son independientes de  $h$ , entonces  $y_n$  también será independiente de  $h$ . Como el método es convergente, entonces

$$\lim_{h \rightarrow 0^+} \max\{|y(t_n) - y_n| : t_n \in \mathcal{P}_N \text{ con } k \leq n \leq N\} = 0.$$

es decir,

$$\lim_{h \rightarrow 0^+} \max\{|1 - y_n| : t_n \in \mathcal{P}_N \text{ con } k \leq n \leq N\} = 0,$$

luego necesariamente  $y_n = 1$  para cualquier  $n$  y, por tanto,  $\rho(1) = 0$ .

Ahora probaremos que  $\rho'(1) = \sigma(1)$ . Como  $\rho(1) = 0$  y el método es estable, entonces 1 es raíz simple de  $\rho$ , por la condición de las raíces. Luego  $\rho(1)' \neq 0$  (de lo contrario, 1 sería raíz doble de

$\rho$ ). Vamos a probar que  $\frac{\sigma(1)}{\rho'(1)} = 1$ . Consideraremos el problema

$$\begin{cases} y' = 1, & \text{en } [0,1], \\ y(0) = 0, \end{cases}$$

cuya solución es  $y(t) = t$ . Elegimos como valores de arranque  $y_j = j h \frac{\sigma(1)}{\rho'(1)}$ . Por inducción se prueba que  $y_n = n h \frac{\sigma(1)}{\rho'(1)}$ , para  $n \geq k$ . Como el método es convergente, entonces

$$\lim_{h \rightarrow 0^+} \max\{|y(t_n) - y_n| : t_n \in \mathcal{P}_N \text{ con } k \leq n \leq N\} = 0.$$

Si sustituimos  $y(t_n)$  e  $y_n$  por sus expresiones, resulta

$$\lim_{h \rightarrow 0^+} \max\{|n h - n h \frac{\sigma(1)}{\rho'(1)}| : t_n \in \mathcal{P}_N \text{ con } k \leq n \leq N\} = 0.$$

y eso solo es posible si  $\frac{\sigma(1)}{\rho'(1)} = 1$ . □

Según los dos teoremas anteriores, la estabilidad y la consistencia no son solo condiciones suficientes para la convergencia, sino que además son necesarias. Por tanto, si un método no es estable o si no es consistente, es claro que no va a ser convergente. Como además hemos caracterizado la estabilidad en términos de la condición de las raíces y la consistencia equivale a  $\rho(1) = 0$  y  $\rho'(1) = \sigma(1)$ , resulta sencillo comprobar si un método es o no es convergente.

### 3.5. Métodos Predictor-Corrector

Como hemos visto, un método lineal multipaso puede ser explícito o implícito. Si  $\alpha_k = 1$ , podemos escribir la ley de recurrencia (3.1) en la forma

$$\mathbf{y}_n = h \beta_k \mathbf{f}(t_n, \mathbf{y}_n) - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}, \quad k \leq n \leq N,$$

donde hemos denotado  $\mathbf{f}_j = \mathbf{f}(t_j, \mathbf{y}_j)$  para referirnos a la evaluación de  $\mathbf{f}$  en los instantes previos y así aligerar la expresión.

Si el método es explícito, es decir, si  $\beta_k = 0$ , entonces la aproximación  $\mathbf{y}_n$  está definida de forma explícita; una vez calculada solo se necesita realizar la evaluación  $\mathbf{f}_n = \mathbf{f}(t_n, \mathbf{y}_n)$  porque se utilizará en el paso siguiente, pues el resto de evaluaciones ya se conoce de las iteraciones anteriores. Desde el punto de vista computacional, este es un coste pequeño siempre que el número de pasos no sea grande: esa situación puede ocurrir cuando se resuelven numéricamente problemas *rígidos* (suelen denominarse *stiff*); en esos casos los métodos explícitos son poco utilizados.

En los métodos implícitos, es decir, cuando  $\beta_k \neq 0$ , la aproximación  $\mathbf{y}_n$  está definida de forma implícita mediante una ecuación que en general será no lineal; la primera cuestión a la que debemos responder es si realmente  $\mathbf{y}_n$  está bien definido. Observemos que  $\mathbf{y}_n$  es un punto fijo de

$$\mathbf{G}_n(\mathbf{u}) = h \beta_k \mathbf{f}(t_n, \mathbf{u}) - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i},$$

Si  $\mathbf{f}$  es lipschitziana respecto de su segundo argumento, entonces  $\mathbf{G}_n$  también es una función lipschitziana:

$$\|\mathbf{G}_n(\mathbf{u}) - \mathbf{G}_n(\mathbf{v})\| = h |\beta_k| \|\mathbf{f}(t_n, \mathbf{u}) - \mathbf{f}(t_n, \mathbf{v})\| \leq h |\beta_k| L \|\mathbf{u} - \mathbf{v}\|.$$

Si el paso  $h$  se elige de forma que  $h |\beta_k| L < 1$ , entonces  $\mathbf{G}_n$  es contractiva y, como consecuencia,  $\mathbf{y}_n$  existe y es único. Que esté bien definido no significa que sea fácil de calcular; usualmente, para calcular  $\mathbf{y}_n$  se utilizará un algoritmo iterativo: iteración de punto fijo, método de Newton... Elijamos por ejemplo el método de iteración de punto fijo: a partir de una aproximación inicial  $\mathbf{y}_n^{(0)}$ , se construye una sucesión  $\{\mathbf{y}_n^{(\nu)}\}_\nu$  en la forma siguiente:

$$\mathbf{y}_n^{(\nu)} = \mathbf{G}_n(\mathbf{y}_n^{(\nu-1)}) = h \beta_k \mathbf{f}(t_n, \mathbf{y}_n^{(\nu-1)}) - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i},$$

para  $\nu \geq 1$ . Cuando el algoritmo se programe en un ordenador, se iterará hasta satisfacer un cierto criterio de tolerancia y usaremos el último término calculado  $\mathbf{y}_n^{(\nu)}$  como aproximación del valor teórico  $\mathbf{y}_n$ . Es decir, realmente no calcularemos  $\mathbf{y}_n$  (¡necesitaríamos infinitas iteraciones!) y por tanto no sabemos cómo puede afectar eso al orden del método. Por otro lado, es evidente que el coste computacional puede ser enorme si en cada paso  $n$  se requiere un número de iteraciones de punto fijo muy elevado. En ese caso, el uso de métodos implícitos no estaría justificado (véase [8]).

Habitualmente, los métodos multipaso implícitos se implementan mediante el proceso de *predicción-corrección*. Este proceso consiste en lo siguiente:

(P) en primer lugar, se calcula una *predicción* o estimación inicial  $\mathbf{y}_n^{(0)}$  mediante un método explícito auxiliar de coeficientes  $\{\alpha_1^*, \dots, \alpha_k^*\}$  y  $\{\beta_1^*, \dots, \beta_k^*\}$ ; es decir:

$$\mathbf{y}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i};$$

(E) a continuación, se evalúa  $\tilde{\mathbf{f}}_n = \mathbf{f}(t_n, \mathbf{y}_n^{(0)})$ ;

(C) luego, se *corrige* la estimación inicial: se define

$$\mathbf{y}_n = h \beta_k \tilde{\mathbf{f}}_n - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}),$$

(E) se hace una *evaluación* final:  $\mathbf{f}_n = \mathbf{f}(t_n, \mathbf{y}_n)$ .

Este proceso se denota *PECE* y equivale a realizar una iteración de punto fijo para aproximar  $\mathbf{y}_n$ . También se puede obviar la evaluación final: el proceso se denotaría *PEC* y en ese caso en el paso siguiente se usaría  $\tilde{\mathbf{f}}_n$  en lugar de  $\mathbf{f}_n$ . Existen otras posibilidades, como  $P(EC)^2E$  o  $P(EC)^2$ , que equivalen a dos iteraciones de punto fijo con evaluación final o sin ella.

En general, según el número  $\mu$  de iteraciones de punto fijo que hagamos (y siempre el mismo número), la forma del método será  $P(EC)^\mu E$  o  $P(EC)^\mu$ , dependiendo de si tras la última corrección decidimos o no actualizar el valor de  $\mathbf{f}_n$  mediante una evaluación más. El método explícito se conoce como método *predictor* y el implícito como *corrector*, y ambos forman el método *predictor-corrector*, que denotaremos por *PC*. Así, existen dos formas distintas de métodos *PC* que podemos resumir por  $P(EC)^\mu E^{1-q}$  con  $\mu \geq 1$  y  $q \in \{0, 1\}$ , donde *P* denota *Predicción*, *E* representa la *Evaluación* y *C* es la *Corrección*. Lo habitual es elegir  $\mu = 1$  o  $\mu = 2$ , como justificaremos más adelante.

El algoritmo de un método  $P(EC)^\mu E^{1-q}$  es el siguiente:

$$\begin{aligned} P : & \quad \mathbf{y}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i}^{(\mu-q)} \\ (EC)^\mu : & \quad \text{Para } \nu = 0, 1, \dots, \mu-1 \\ & \quad \mathbf{f}_n^{(\nu)} = \mathbf{f}(t_n, \mathbf{y}_n^{(\nu)}) \\ & \quad \mathbf{y}_n^{(\nu+1)} = h \beta_k \mathbf{f}_n^{(\nu)} - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu-q)} \\ E^{1-q} : & \quad \text{Si } q = 0 \quad \mathbf{f}_n^{(\mu)} = \mathbf{f}(t_n, \mathbf{y}_n^{(\mu)}) \end{aligned}$$

Claramente, un método predictor-corrector es un método multipaso, pero no es un método *lineal* multipaso. En el esquema anterior hemos supuesto que el número de pasos del método predictor coincide con el del método corrector; en esa situación, ese número común  $k$  es el número de pasos del par *PC*. En realidad, el método predictor y el método corrector podrían tener números de pasos  $k^*$  y  $k$  diferentes, en cuyo caso el número de pasos del par *PC* sería el máximo de  $k^*$  y  $k$ .

Si el número de pasos del corrector es menor que el del predictor, es decir, si  $k < k^*$ , siempre es posible considerar que el corrector es un método de  $k^*$  pasos si se reenumeran sus coeficientes y se elimina la condición  $|\alpha_0| + |\beta_0| \neq 0$ . Lo mismo se puede hacer en caso contrario, esto es, cuando el predictor es de menos pasos que el corrector. Por esa razón, en el desarrollo anterior

hemos supuesto que ambos métodos tienen el mismo número de pasos y lo hemos denotado  $k$ .

### 3.5.1. Orden de un método predictor-corrector

Consideremos un par predictor-corrector  $P(EC)^\mu E^{1-q}$ ; supondremos que el método explícito es de orden  $p^*$  y tiene asociado un operador en diferencias  $\mathcal{L}_h^*$ , mientras que el método implícito es de orden  $p$  y su operador en diferencias lo denotamos por  $\mathcal{L}_h$ . Supondremos además que las soluciones de la ecuación diferencial son suficientemente regulares (al menos continuidad en las derivadas hasta el mayor de  $p^*$  y  $p$ ).

En el instante  $t_n$ , se utiliza el método predictor para calcular una estimación inicial  $\mathbf{y}_n^{(0)}$  definida por

$$\mathbf{y}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i}^{(\mu-q)}.$$

Para determinar el orden del método predictor-corrector, lo que haremos será determinar el término principal del desarrollo del error local, como se hace en el libro de J.D. Lambert [8]. Recordemos que el error local es el error propio del método, es el error cometido tras un paso de integración cuando se parte de los valores exactos. Por tanto, vamos a calcular la aproximación en  $t_n$  partiendo de los valores exactos en  $t_{n-1}, t_{n-2}, \dots, t_{n-k}$ , es decir,

$$\mathbf{y}_{n-i}^{(\mu)} = \mathbf{y}(t_{n-i}), \quad i = 1, 2, \dots, k.$$

Denotaremos por  $\tilde{\mathbf{y}}_n^{(\nu)}$  a las aproximaciones proporcionadas por el método predictor y el método corrector a partir de los valores exactos. Con esta notación, resulta

$$\tilde{\mathbf{y}}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}(t_{n-i}) + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}(t_{n-i}, \mathbf{y}(t_{n-i})).$$

Por otra parte, si usamos la expresión (3.3) del operador en diferencias y tenemos en cuenta que el método predictor es explícito ( $\beta_k^* = 0$ ), se tiene

$$\mathbf{y}(t_n) = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}(t_{n-i}) + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}(t_{n-i}, \mathbf{y}(t_{n-i})) + \mathcal{L}_h^*(\mathbf{y}(t_n)).$$

Podemos restar ambas expresiones y recordar el análisis del orden que hicimos en la Sección 3.3.2; así, deducimos

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(0)} = \mathcal{L}_h^*(\mathbf{y}(t_n)) = \frac{c_{p^*+1}^*}{(p^*+1)!} h^{p^*+1} \mathbf{y}^{(p^*+1)}(t_n) + \mathcal{O}(h^{p^*+2}). \quad (3.21)$$

Ahora fijaremos nuestra atención en el método corrector. Nuevamente, de la definición del ope-

radar en diferencias, podemos escribir

$$\mathbf{y}(t_n) + \sum_{i=1}^k \alpha_{k-i} \mathbf{y}(t_{n-i}) = h \beta_k \mathbf{f}(t_n, \mathbf{y}(t_n)) + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}(t_{n-i})) + \mathcal{L}_h(\mathbf{y}(t_n)).$$

y

$$\begin{aligned} \tilde{\mathbf{y}}_n^{(\nu+1)} + \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} &= h \beta_k \mathbf{f}(t_n, \tilde{\mathbf{y}}_n^{(\nu)}) + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu-q)} \\ &= h \beta_k \mathbf{f}(t_n, \tilde{\mathbf{y}}_n^{(\nu)}) + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}(t_{n-i})) \quad \nu = 0, 1, \dots, \mu-1. \end{aligned}$$

De nuevo, si restamos ambas expresiones obtenemos

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu+1)} = h \beta_k (\mathbf{f}(t_n, \mathbf{y}(t_n)) - \mathbf{f}(t_n, \tilde{\mathbf{y}}_n^{(\nu)})) + \mathcal{L}_h(\mathbf{y}(t_n)), \quad \nu = 0, 1, \dots, \mu-1. \quad (3.22)$$

Podemos aplicar el Teorema del Valor Medio a cada función coordenada  $f_j$  de  $\mathbf{f}$ :

$$f_j(t_n, \mathbf{y}(t_n)) - f_j(t_n, \tilde{\mathbf{y}}_n^{(\nu)}) = \nabla f_j(\boldsymbol{\zeta}_j^{(\nu)}) \cdot (\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu)}),$$

donde  $\nabla f_j(\boldsymbol{\zeta}_j^{(\nu)})$  representa el vector gradiente de  $f_j$  respecto de las variables  $\{y_1, \dots, y_m\}$  en algún punto  $\boldsymbol{\zeta}_j^{(\nu)}$  del segmento de extremos  $\mathbf{y}(t_n)$  y  $\tilde{\mathbf{y}}_n^{(\nu)}$ . Este punto no tiene por qué ser el mismo para todas las funciones coordenadas. Denotemos

$$\boldsymbol{\zeta}^{(\nu)} = \begin{pmatrix} \boldsymbol{\zeta}_1^{(\nu)} \\ \vdots \\ \boldsymbol{\zeta}_m^{(\nu)} \end{pmatrix} \quad \text{y} \quad \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(\nu)}) = \begin{pmatrix} \nabla f_1(\boldsymbol{\zeta}_1^{(\nu)}) \\ \vdots \\ \nabla f_m(\boldsymbol{\zeta}_m^{(\nu)}) \end{pmatrix}.$$

Con la notación  $\bar{\mathbf{J}}_{\mathbf{f}}$  queremos destacar que cada vector gradiente aparece evaluado en un punto diferente para cada función coordenada de  $\mathbf{f}$ .

Así, la expresión (3.22) se puede escribir de esta forma equivalente

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu+1)} = h \beta_k \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(\nu)}) \cdot (\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu)}) + \mathcal{L}_h(\mathbf{y}(t_n)), \quad \nu = 0, 1, \dots, \mu-1.$$

Ahora bien, si el método corrector es de orden  $p$ , entonces

$$\begin{aligned} \mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu+1)} &= h \beta_k \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(\nu)}) \cdot (\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\nu)}) \\ &\quad + \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}), \quad \nu = 0, 1, \dots, \mu-1. \end{aligned} \quad (3.23)$$

Vamos a analizar qué sucede según sean  $p^*$  y  $p$ . En primer lugar, vamos considerar el caso  $p^* \geq p$ ,

es decir, el orden del método predictor es mayor o igual que el del corrector. Podemos utilizar la expresión (3.21) del predictor en (3.23) para  $\nu = 0$ , es decir, para la primera corrección; como  $p^* \geq p$ , entonces

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(1)} = \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}).$$

De nuevo, podemos usar esta expresión otra vez en (3.23) para  $\nu = 1$  y obtendremos

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(2)} = \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}).$$

Si reiteramos el proceso, para la última corrección (para  $\nu = \mu$ ) llegaremos a

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\mu)} = \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}).$$

Se deduce que el término principal del error del método predictor corrector coincide con el del método predictor.

En segundo lugar, consideremos el caso  $p^* = p-1$ , es decir, el orden del predictor es una unidad menos que el del corrector. Procedemos como antes: sustituimos la expresión (3.21) del predictor en (3.23) para  $\nu = 0$ ; en este caso obtenemos

$$\begin{aligned} \mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(1)} &= h \beta_k \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(0)}) \cdot (\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(0)}) + \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}) \\ &= \left( \beta_k \frac{c_p^*}{p!} \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(0)}) \cdot \mathbf{y}^{(p)}(t_n) + \frac{c_{p+1}}{(p+1)!} \mathbf{y}^{(p+1)}(t_n) \right) h^{p+1} + \mathcal{O}(h^{p+2}). \end{aligned}$$

Por tanto, si el método predictor corrector fuese del tipo  $P(EC)E^{1-q}$ , es decir, si  $\mu = 1$ , entonces el método predictor corrector y el método corrector tendrían el mismo orden  $p$  pero diferente término principal del error local.

En la misma situación  $p^* = p-1$  pero con  $\mu \geq 2$ , si razonamos como en el caso  $p^* \geq p$  y seguimos substituyendo esta última expresión en (3.23), entonces se deduce

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(\mu)} = \frac{c_{p+1}}{(p+1)!} h^{p+1} \mathbf{y}^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}).$$

y, por lo tanto, el método predictor corrector y el método corrector tienen el mismo orden  $p$  y el mismo término principal en el error local.

Finalmente, supongamos  $p^* = p-2$ , esto es, el orden del predictor es dos unidades menos que el del corrector. Como en los casos anteriores, comenzamos utilizando la expresión (3.21) del predictor en (3.23) para  $\nu = 0$ ; resulta lo siguiente:

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(1)} = \beta_k \frac{c_{p-1}^*}{(p-1)!} h^p \bar{\mathbf{J}}_{\mathbf{f}}(\boldsymbol{\zeta}^{(0)}) \cdot \mathbf{y}^{(p-1)}(t_n) + \mathcal{O}(h^{p+1})$$



Si además  $\mu = 1$ , o sea, si el método es  $P(EC)E^{1-q}$ , entonces su orden es  $p - 1$ . En cambio, si  $\mu = 2$  podemos utilizar esta expresión de nuevo en (3.23) para  $\nu = 1$  y de ese modo se deduce

$$\mathbf{y}(t_n) - \tilde{\mathbf{y}}_n^{(2)} = \left( \frac{c_{p-1}^*}{(p-1)!} (\beta_k \bar{\mathbf{J}}_f)^2 (\boldsymbol{\zeta}^{(0)}) \cdot \mathbf{y}^{(p-1)}(t_n) + \frac{c_{p+1}}{(p+1)!} \mathbf{y}^{(p+1)}(t_n) \right) h^{p+1} + \mathcal{O}(h^{p+2}).$$

El método alcanza el mismo orden  $p$  que el corrector, pero tiene diferente término principal del error local. Si fuese  $\mu \geq 3$ , entonces repitiendo los razonamientos anteriores se prueba que el orden del predictor corrector es  $p$  y su término principal del error local también coincide con el del corrector.

Si fuese  $p^* = p-3$ , la forma de proceder sería análoga, y lo mismo en el caso más general  $p^* = p-s$ , con  $s \geq 3$ .

En resumen, el orden y el término principal del error local del método  $P(EC)E^{1-q}$  dependen de la diferencia entre los órdenes  $p$  y  $p^*$  de cada uno de los dos métodos que forman el par, así como del número  $\mu$  de correcciones. Los resultados se pueden resumir así (véase [8]):

1. si  $p^* \geq p$ , o si  $p^* < p$  y  $\mu > p - p^*$ , el método  $PC$  y el corrector tienen el mismo orden  $p$  y mismo término principal del error local,
2. si  $p^* < p$  y  $\mu = p - p^*$ , el método  $PC$  y el corrector tienen el mismo orden  $p$  pero distinto término principal del error local, y
3. si  $p^* < p$  y  $\mu \leq p - p^* - 1$ , el orden del método  $PC$  es  $p^* + \mu$ , que siempre es inferior a  $p$ .

### 3.5.2. Estabilidad de un método predictor-corrector

Como ya hemos comentado, los métodos de tipo predictor-corrector que estamos estudiando son métodos multipaso, pero no son métodos *lineales* multipaso porque en ellos aparecen reevaluaciones de la función  $\mathbf{f}$ . Además, son métodos explícitos. Denotaremos  $k$  al número de pasos del par  $PC$ .

Como hemos visto, el algoritmo del método  $P(EC)^\mu E^{1-q}$  es el siguiente:

$$\begin{aligned}
 P : \quad & \mathbf{y}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i}^{(\mu-q)} \\
 (EC)^\mu : \quad & \text{Para } \nu = 0, 1, \dots, \mu-1 \\
 & \mathbf{f}_n^{(\nu)} = \mathbf{f}(t_n, \mathbf{y}_n^{(\nu)}) \\
 & \mathbf{y}_n^{(\nu+1)} = h \beta_k \mathbf{f}_n^{(\nu)} - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu-q)} \\
 E^{1-q} : \quad & \text{Si } q = 0 \quad \mathbf{f}_n^{(\mu)} = \mathbf{f}(t_n, \mathbf{y}_n^{(\mu)})
 \end{aligned}$$

Por simplicidad para la exposición, consideraremos el caso  $q = 0$ ; así, en cada instante  $t_n$  el método *PC* finaliza siempre con la evaluación  $\mathbf{f}_n^{(\mu)} = \mathbf{f}(t_n, \mathbf{y}_n^{(\mu)})$ .

Es posible escribir el método en la forma que hemos considerado en la Sección 3.2.1. En efecto, a partir de la *predicción*  $\mathbf{y}_n^{(0)}$  se hace la primera evaluación (para  $\nu = 0$ )

$$\mathbf{f}_n^{(0)} = \mathbf{f}(t_n, \mathbf{y}_n^{(0)}) = \mathbf{f}\left(t_n, -\sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i}^{(\mu)}\right)$$

y se utiliza esta evaluación para obtener la primera *corrección*  $\mathbf{y}_n^{(1)}$ :

$$\begin{aligned} \mathbf{y}_n^{(1)} &= h \beta_k \mathbf{f}_n^{(0)} - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu)} \\ &= h \beta_k \mathbf{f}\left(t_n, -\sum_{i=1}^k \alpha_{k-i}^* \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i}^* \mathbf{f}_{n-i}^{(\mu)}\right) - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu)} \end{aligned}$$

Si se reitera el proceso, se comprueba que en la expresión de cada corrección  $\mathbf{y}_n^{(\nu)}$  siempre aparecerá el término  $-\sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)}$ , en donde intervienen los coeficientes del polinomio  $\rho$  del método corrector.

Por otra parte, si denotamos

$$\mathbf{Y}_{n-1}^{(\mu)} = \begin{pmatrix} \mathbf{y}_{n-1}^{(\mu)} \\ \mathbf{y}_{n-2}^{(\mu)} \\ \vdots \\ \mathbf{y}_{n-k}^{(\mu)} \end{pmatrix}, \quad \Phi(t_n, \mathbf{Y}_{n-1}^{(\mu)}, h; \mathbf{f}) = \begin{pmatrix} h \beta_k \mathbf{f}_n^{(\mu-1)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}_{n-i}^{(\mu)} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix},$$

entonces el método se puede escribir en la forma (3.9), es decir,

$$\begin{aligned} \mathbf{Y}_{k-1}^{(\mu)} &= (\mathbf{y}_{k-1}, \dots, \mathbf{y}_0)^t \\ \mathbf{Y}_n^{(\mu)} &= A \mathbf{Y}_{n-1}^{(\mu)} + h \Phi(t_n, \mathbf{Y}_{n-1}^{(\mu)}, h; \mathbf{f}), \quad \text{con } n \geq k. \end{aligned}$$

La matriz  $A$ , que es del tipo definido en (3.7), se construye con los coeficientes del método corrector, como ya hemos dicho.

Estudiamos ahora la estabilidad del método. Claramente la función  $\Phi$  satisface la hipótesis (3.10), pues

$$\mathbf{f} = \mathbf{0} \quad \Rightarrow \quad \Phi(t_n, \mathbf{Y}_{n-1}^{(\mu)}, h; \mathbf{0}) = \mathbf{0}.$$

Veamos ahora que  $\Phi$  también cumple la condición de Lipschitz (3.11). Sean

$$\mathbf{Y}_{n-1}^{(\mu)} = (\mathbf{y}_{n-1}^{(\mu)}, \dots, \mathbf{y}_{n-k}^{(\mu)})^t \quad \text{y} \quad \mathbf{Z}_{n-1}^{(\mu)} = (\mathbf{z}_{n-1}^{(\mu)}, \dots, \mathbf{z}_{n-k}^{(\mu)})^t.$$

Puesto que

$$\|\Phi(t_n, \mathbf{Y}_{n-1}^{(\mu)}, h; \mathbf{f}) - \Phi(t_n, \mathbf{Z}_{n-1}^{(\mu)}, h; \mathbf{f})\| = h \beta_k \|\mathbf{f}(t_n, \mathbf{y}_n^{(\mu-1)}) - \mathbf{f}(t_n, \mathbf{z}_n^{(\mu-1)})\|$$

bastará acotar las diferencias  $\|\mathbf{f}(t_n, \mathbf{y}_n^{(\nu)}) - \mathbf{f}(t_n, \mathbf{z}_n^{(\nu)})\|$  para  $\nu = 0, 1, \dots, \mu - 1$ .

Para  $\nu = 0$ , se tiene

$$\mathbf{y}_n^{(0)} - \mathbf{z}_n^{(0)} = - \sum_{i=1}^k \alpha_{k-i}^* (\mathbf{y}_{n-i}^{(\mu)} - \mathbf{z}_{n-i}^{(\mu)}) + h \sum_{i=1}^k \beta_{k-i}^* (\mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}^{(\mu)}) - \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i}^{(\mu)})),$$

y por lo tanto, por ser  $\mathbf{f}$  una función lipschitziana, resulta

$$\begin{aligned} \|\mathbf{y}_n^{(0)} - \mathbf{z}_n^{(0)}\| &\leq \sum_{i=1}^k |\alpha_{k-i}^*| \|\mathbf{y}_{n-i}^{(\mu)} - \mathbf{z}_{n-i}^{(\mu)}\| + h L \sum_{i=1}^k |\beta_{k-i}^*| \|\mathbf{y}_{n-i}^{(\mu)} - \mathbf{z}_{n-i}^{(\mu)}\| \\ &= \sum_{i=1}^k (|\alpha_{k-i}^*| + h L |\beta_{k-i}^*|) \|\mathbf{y}_{n-i}^{(\mu)} - \mathbf{z}_{n-i}^{(\mu)}\| \\ &\leq (\text{máx}\{|\alpha_{k-i}^*|\} + h L \text{máx}\{|\beta_{k-i}^*|\}) \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\| \\ &= \Lambda^* \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\| \end{aligned}$$

donde hemos denotado

$$\Lambda^* = \text{máx}\{|\alpha_{k-i}^*|\} + h L \text{máx}\{|\beta_{k-i}^*|\},$$

y así también

$$\|\mathbf{f}(t_n, \mathbf{y}_n^{(0)}) - \mathbf{f}(t_n, \mathbf{z}_n^{(0)})\| \leq L \Lambda^* \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\|.$$

Repetimos el proceso para  $\nu = 1$ :

$$\begin{aligned} \mathbf{y}_n^{(1)} &= h \beta_k \mathbf{f}(t_n, \mathbf{y}_n^{(0)}) - \sum_{i=1}^k \alpha_{k-i} \mathbf{y}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{y}_{n-i}^{(\mu)}), \\ \mathbf{z}_n^{(1)} &= h \beta_k \mathbf{f}(t_n, \mathbf{z}_n^{(0)}) - \sum_{i=1}^k \alpha_{k-i} \mathbf{z}_{n-i}^{(\mu)} + h \sum_{i=1}^k \beta_{k-i} \mathbf{f}(t_{n-i}, \mathbf{z}_{n-i}^{(\mu)}), \end{aligned}$$

de modo que, si denotamos ahora

$$\Lambda = \text{máx}\{|\alpha_{k-i}|\} + h L \text{máx}\{|\beta_{k-i}|\},$$

podemos escribir

$$\begin{aligned}\|\mathbf{y}_n^{(1)} - \mathbf{z}_n^{(1)}\| &\leq h |\beta_k| L \Lambda^* \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\| + \Lambda \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\| \\ &= (h |\beta_k| L \Lambda^* + \Lambda) \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\|\end{aligned}$$

y también

$$\|\mathbf{f}(t_n, \mathbf{y}_n^{(1)}) - \mathbf{f}(t_n, \mathbf{z}_n^{(1)})\| \leq L (h |\beta_k| L \Lambda^* + \Lambda) \|\mathbf{Y}_{n-1}^{(\mu)} - \mathbf{Z}_{n-1}^{(\mu)}\|.$$

Si se reitera este proceso hasta  $\nu = \mu - 1$ , se prueba que  $\Phi$  cumple la condición de Lipschitz requerida.

De ese modo, la estabilidad del método *PC* depende solo de los coeficientes de la matriz  $A$  (Teorema 3.2), esto es, del polinomio  $\rho$  del método corrector. Podemos asegurar que el método predictor-corrector es estable si y solo si el polinomio  $\rho$  cumple la condición de las raíces 3.6, o equivalentemente, si y solo si el método corrector es estable.

**Ejemplo 3.3.** Consideremos el método *PECE* que utiliza como predictor el método de Adams-Bashforth de 2 pasos y como corrector el método de Adams-Moulton de 1 paso (método del trapecio). El algoritmo es el siguiente:

- en el instante  $t_n$  se supone que las aproximaciones  $\mathbf{y}_{n-2}^{(1)}$  e  $\mathbf{y}_{n-1}^{(1)}$  y sus evaluaciones

$$\mathbf{f}_{n-2}^{(1)} = \mathbf{f}(t_{n-2}, \mathbf{y}_{n-2}^{(1)}) \quad \text{y} \quad \mathbf{f}_{n-1}^{(1)} = \mathbf{f}(t_{n-1}, \mathbf{y}_{n-1}^{(1)})$$

son conocidas del instante anterior;

- se calcula la predicción en  $t_n$  con el método Adams-Bashforth de 2 pasos:

$$\mathbf{y}_n^{(0)} = \mathbf{y}_{n-1}^{(1)} + h \left( \frac{3}{2} \mathbf{f}_{n-1}^{(1)} - \frac{1}{2} \mathbf{f}_{n-2}^{(1)} \right). \quad (3.24)$$

- se hace la evaluación

$$\mathbf{f}_n^{(0)} = \mathbf{f}(t_n, \mathbf{y}_n^{(0)}),$$

- se corrige la aproximación en  $t_n$  mediante el método del trapecio:

$$\mathbf{y}_n^{(1)} = \mathbf{y}_{n-1}^{(1)} + h \left( \frac{1}{2} \mathbf{f}_n^{(0)} + \frac{1}{2} \mathbf{f}_{n-1}^{(1)} \right). \quad (3.25)$$

- se hace la evaluación final:

$$\mathbf{f}_n^{(1)} = \mathbf{f}(t_n, \mathbf{y}_n^{(1)}).$$

En este caso, tanto el método predictor como el corrector son de orden 2. Luego, por las reglas mencionadas anteriormente, el método *PC* y el corrector tienen el mismo orden y mismo término

principal del error local. Como el método del trapecio es estable, entonces podemos concluir que el método predictor-corrector es convergente de orden 2.

### 3.5.3. Implementación en Matlab

MATLAB es “una plataforma de programación y computación numérica utilizada por millones de ingenieros y científicos para analizar datos, desarrollar algoritmos y crear modelos”, según su propia definición. Matlab dispone de varios comandos que permiten resolver un problema de valor inicial mediante un método numérico. Algunos de estos comandos utilizan métodos multipaso predictor-corrector (véase [9]):

- `ode113` es un método de paso variable y orden variable (VSVO, por sus siglas en inglés: variable-step, variable-order) que se basa en métodos *PECE* Adams-Bashforth-Moulton de órdenes 1 a 13. El orden máximo utilizado es 12, pero se usa una fórmula de orden 13 para la estimación del error junto con técnicas de extrapolación local.
- `ode15s` es también un método de paso variable y orden variable (VSVO) y puede utilizar, de forma opcional, métodos BDF.
- `ode15i` es un método de paso variable y orden variable (VSVO) basado en las fórmulas BDF de órdenes 1 a 5.



# Bibliografía

- [1] John Charles Butcher. *The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods*. John Wiley & Sons, Chichester, 1987.
- [2] Manuel Calvo, Juan Ignacio Montijano, and Luis Rández. *Curso de análisis numérico: Métodos lineales multipaso para la resolución numérica de ecuaciones diferenciales ordinarias*. Servicio de Publicaciones de la Universidad de Zaragoza, 1986.
- [3] Michel Crouzeix and Alain Mignot. *Analyse numérique des équations différentielles*. Masson, deuxième édition, 1992.
- [4] Colin W Cryer. On the instability of high order backward-difference multistep methods. *BIT Numerical Mathematics*, 12(1):17–25, 1972.
- [5] Stanley Grossman and José Job Flores Godoy. *Álgebra lineal*. McGraw-Hill/Interamericana Editores, 2012.
- [6] Ernst Hairer, Syvert P Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, Berlin, second revised edition, 1993.
- [7] Eugene Isaacson and Herbert Bishop Keller. *Analysis of numerical methods*. Dover Publications, 1994.
- [8] John Denholm Lambert. *Numerical Methods for Ordinary Differential Systems. The Initial Value Problem*. John Wiley & Sons, Chichester, 1991.
- [9] Mathworks. Choose an ODE Solver. <https://es.mathworks.com/help/matlab/math/choose-an-ode-solver.html>. Fecha de consulta: 2021-06-07.
- [10] Pablo Pérez Riera. Notas sobre la resolución numérica de ecuaciones diferenciales ordinarias. Publicado en formato PDF, Oviedo, 2020.
- [11] Lawrence F Shampine. *Numerical solution of ordinary differential equations*. Chapman & Hall, New York, 1994.