

Carlos González Higuera

*Métodos Lineales Multipaso para la
Resolución de Ecuaciones Diferenciales*

Multistep Linear Methods for Solving Differential
Equations

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, Mes de Año

DIRIGIDO POR
Severiano González Pinto

Severiano González Pinto
Departamento de Análisis
Matemático
Universidad de La Laguna
38200 La Laguna, Tenerife

Agradecimientos

Quiero agradecer a la familia González Armas, y en especial, a Severiano González Pinto por guiarme en este trabajo y hacerlo posible. También quiero agradecer a mis padres y a mi hermano Marcos por ser un apoyo continuo a lo largo de mi trayectoria en la universidad de La Laguna.

Carlos González Higuera
La Laguna, 10 de septiembre de 2021

Resumen · Abstract

Resumen

Este trabajo está dedicado al estudio de métodos lineales multipaso, siguiendo la línea que presenta Ernst Hairer y coautores en dos sus textos clásicos "Solving Ordinary Differential Equations" Nonstiff and Stiff problems editados por Springer y que han sido ampliamente citados en la comunidad científica del análisis numérico de ecuaciones diferenciales. Se iniciará comentando algunos métodos clásicos desarrollados históricamente (Adams Bashforth, Adams Moulton, Nystrom, etc), y se continuará en los siguientes capítulos con los métodos lineales multipaso mas generales, centrándose en el estudio del orden, la estabilidad y las propiedades de convergencia. Se comentarán y demostrarán importantes resultados, así como dos de los Teoremas establecidos a lo largo del siglo XX por el gran matemático sueco Germund Dahlquist, los cuales han desempeñado un papel clave en el estudio de la estabilidad y convergencia de los métodos lineales multipaso.

Palabras clave: *Métodos Lineales Multipaso – Orden de Consistencia – 0-estabilidad – Convergencia – Dominio de estabilidad – Boundary Locus.*

Abstract

This paper is dedicated to the study of linear multistep methods by following the line presented by Ernst Hairer and coauthors in his classical books "Solving ordinary differential equations" non-stiff and stiff problems, edited by Springer. The textbooks have been widely quoted and constitute a cornerstone in the numerical analysis community. Our work will start with some classical multistep methods in the historical order they were developed by Adams Bashforth, Adams Moulton, Nystrom, etc. We will continue in the following chapters with general multistep methods focusing on the study of order, stability and convergence properties.

Important results will be discussed and demonstrated, as well as two of the theorems established throughout the 20th century by the great mathematician Germund Dahlquist, which play a crucial role in the study of stability and convergence of linear multistep methods.

Keywords: *Multistep Linear Methods - Order of Consistency - 0-stability - Convergence - Stability Region - Boundary Locus.*

Contenido

Agradecimientos	III
Resumen/Abstract	V
Introducción	IX
1. Métodos Clásicos	1
1.1. Métodos Explícitos de Adams	1
1.2. Métodos Implícitos de Adams	3
1.3. Relación de Recurrencia de los Coeficientes	4
1.3.1. Adams Explícito:	4
1.3.2. Adams Implícito:	5
1.4. Métodos Predictor-Corrector	6
1.5. Métodos Explícitos de Nyström	7
1.6. Métodos Implícitos de Milne-Simpson	7
1.7. Métodos BDF	8
1.8. Experimentos Numéricos	9
2. Error y Condiciones de Orden	11
2.1. Error Global y Error Local	11
2.2. Orden de los métodos multipaso	13
2.3. La constante de Error	15
2.4. Métodos Irreducibles	18
3. 0-Estabilidad y Primera barrera de Dalhquist	19
3.1. Máximo Orden Posible de Consistencia para Métodos Lineales Multipaso 0-Estables	20
4. Convergencia de los métodos multipaso	27
4.1. Formulación como método de un solo paso	31
4.2. Demostración de la convergencia	34

5. Teoría Estabilidad Lineal	37
5.1. La Región de Estabilidad	38
5.2. Ejemplos	40
5.2.1. Métodos de Adams	41
5.2.2. Métodos de Nyström	42
5.2.3. Métodos BDF	43
5.3. La Segunda Barrera Dahlquist	43
Bibliografía	49
Poster	51

Introducción

Los métodos multipaso fueron introducidos a mitad del siglo XIX alrededor de 1855 por Adams, quien auxilió a Bashforth (1883) que estaba investigando en problemas de capilaridad y quería determinar la forma de una gota de sangre (u otro líquido) siguiendo el modelo físico que conduce a un sistema de ecuaciones diferenciales no lineales que no pueden resolverse exactamente usando las técnicas analíticas standard. Adams le propuso usar métodos multipaso de tipo explícito o implícito basados en la integración numérica, los cuales hacen uso de la información obtenida en varios pasos anteriores, es decir, de la solución numérica obtenida en un número prefijado de pasos anteriores, digamos k , avanzando paso a paso la integración hasta el punto final del intervalo de integración. Esto representa una diferencia importante con los métodos de un paso, los cuales hacen uso solo de la información del paso anterior. Es curioso que con la excepción del método de Euler que es de un paso y orden 1 (es decir da errores globales en el punto final del tamaño del paso usado en la integración), los métodos multipaso fueran propuestos antes que los métodos de un paso. Esto se debe a que permiten obtener órdenes de convergencia más altos que los métodos de un paso (por ejemplo los métodos de tipo Runge-Kutta). Sin embargo, el inconveniente de los métodos de tipo multipaso con respecto a los de un paso, estriba en que necesitan valores adicionales de arranque, no provistos en el Problema de Valor Inicial (PVI), y a menudo hay que recurrir a métodos Runge-Kutta para proporcionar estos valores adicionales de arranque. Además los métodos de un paso tienen mejores propiedades de estabilidad, lo que permite avanzar la integración con tamaños de paso mayores, lo cual en muchos casos compensa el hecho de que los métodos multipaso alcancen mayor orden en un solo paso de integración. Digamos que los principales códigos usados en la actualidad (en el software corriente) para integrar PVI se basan o bien en métodos de un paso de tipo Runge-Kutta o en métodos lineales multipaso (métodos de Adams para problemas no stiff) o métodos BDF, para problemas de tipo stiff (solución analítica muy estable, en el sentido de que pequeñas perturbaciones en el valor inicial conducen rápidamente a la solución exacta con valor inicial sin perturbar). En

esta memoria abordaremos los aspectos de consistencia, 0-estabilidad, convergencia y estabilidad lineal de los métodos lineales multipaso.

Consideramos PVI para sistemas de ecuaciones diferenciales ordinarias de primer orden (EDOs),

$$y' = f(t, y), \quad y(t_0) = y_0, \quad t \in [t_0, t^*], \quad t_0 = 0, \quad y, f(t, y) \in \mathbf{R}^m.$$

Para su integración mediante métodos multipaso (de k pasos), vamos a usar una partición uniforme del intervalo de integración. Por tanto, para un tamaño de paso fijo $h = t^*/N$, donde N es el número de pasos que vamos a dar para alcanzar el punto final del intervalo t^* , definimos los nodos del intervalo en la forma $t_n = nh$, $n = 0, 1, 2, \dots, N$. La integración se avanzará paso a paso obteniendo así una solución discreta de la forma (t_n, y_n) , $n = 0, 1, \dots, N$. Los k primeros valores se denominan valores de arranque o valores iniciales (t_n, y_n) , $n = 0, 1, \dots, k-1$, y si $k \geq 2$, entonces los valores de arranque adicionales (t_j, y_j) , $j = 1, 2, \dots, k-1$, deben ser suministrados por algún algoritmo conveniente. Observe que (t_0, y_0) lo proporciona el valor inicial. Una vez obtenidos los valores de arranque de forma conveniente entonces se calculan todos los demás (t_n, y_n) , $n = k, k+1, \dots, N$, mediante el método lineal multipaso elegido.

Métodos Clásicos

En este capítulo iniciamos el estudio de los métodos lineales multipaso, mencionando algunas fórmulas de los métodos clásicos multipaso más destacadas a lo largo de la historia del análisis numérico, incluyendo relaciones de recurrencia de coeficientes tanto para métodos explícitos como implícitos.

Considerando Problemas de Valor Inicial

$$\begin{cases} y'(t) = f(t, y) \\ y(t_0) = y_0 \end{cases}, \quad t \in [t_0, t^*], \quad t_0 = 0$$

comenzaremos describiendo los métodos explícitos de Adams (también conocidos como métodos de Adams-Bashforth) de k pasos.

1.1. Métodos Explícitos de Adams

Si conocemos $y(t_n)$, el valor exacto para $y(t_{n+1})$ será

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} y'(t) dt = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt \quad (1.1)$$

Supongamos que conocemos las aproximaciones numéricas $y'_n, y'_{n-1}, \dots, y'_{n-k}$ a las derivadas de solución exacta $y'(t_n), \dots, y'(t_{n-k})$ ($\equiv t_n, \dots, t_{n-k}$), podemos expresar f de la forma siguiente

$$f_i = f(t_i, y_i) = y'_i, \quad \text{para } i = n - k, \dots, n. \quad (1.2)$$

Por tanto, podemos sustituir $y(t_n)$ por y_n y reemplazar la función $f(t, y(t))$ en (1.1) por el polinomio de interpolación a través de los puntos $(t_i, f_i) | i = n - k, \dots, n$. Utilizando la fórmula regresiva de Newton-Gregory podemos expresar este polinomio en la forma siguiente

$$p(t) = p(t_n + sh) = \sum_{j=0}^{k-1} (-1)^j \binom{-s}{j} \nabla^j f_n$$

donde

$$\nabla^0 f_n = f_n, \quad \nabla f_n = f_n - f_{n-1}, \quad \dots, \quad \nabla^{j+1} f_n = \nabla^j f_n - \nabla^{j-1} f_{n-1}, \quad j \geq 1$$

$$s = \frac{t - t_n}{h} \rightarrow t = t_n + sh.$$

Con este cambio estamos tomando t_n como origen de coordenadas y h como unidad de longitud. Por tanto,

$$\int_{t_n}^{t_{n+1}} y'(t) dt \simeq \int_{t_n}^{t_{n+1}} p(t) dt = h \int_{s=0}^1 \sum_{j=0}^{k-1} (-1)^j \binom{-s}{j} \nabla^j f_n ds = h \sum_{j=0}^{k-1} \gamma_j \nabla^j y'(t_n)$$

donde γ_j debe tener la siguiente forma

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

Para

$$\begin{aligned} j = 0 : \gamma_0 &= (-1)^0 \int_0^1 \binom{-s}{0} ds = \int_0^1 ds = 1 \\ j = 1 : \gamma_1 &= (-1)^1 \int_0^1 \binom{-s}{1} ds = \int_0^1 (-s) ds = \frac{1}{2} \\ &\vdots \end{aligned}$$

Si deseamos expresar las diferencias regresivas en términos de f_{n-j} , entonces obtenemos para los distintos valores de k

$$\begin{aligned} k = 1 : y_{n+1} - y_n &= h \sum_{j=0}^0 \gamma_j \nabla^j f_n = h f_n \\ k = 2 : y_{n+1} - y_n &= h \sum_{j=0}^1 \gamma_j \nabla^j f_n = h \left(f_n + \frac{1}{2} (f_n - f_{n-1}) \right) \\ &= h \left(\frac{3}{2} f_n - \frac{1}{2} f_{n-1} \right) \\ &\vdots \end{aligned}$$

1.2. Métodos Implícitos de Adams

En los métodos implícitos de Adams (también conocidos como métodos de Adams Moulton) se usa el polinomio interpolador a la derivada $y'(t)$ en los nodos $t_{n+1}, t_n, \dots, t_{n-(k-1)}$ a pesar de que no se conoce el valor $y'(t_{n+1})$, pero esto nos dará una fórmula suponiendo que ese valor se conozca

$$p^*(t) = p^*(t_n + sh) = p^*(t_{n+1} + (s-1)h)$$

llamamos $\mu = (s-1)$ y tenemos

$$p^*(t) = \sum_{j=0}^k (-1)^j \binom{\mu}{j} \nabla^j f_{n+1} = \sum_{j=0}^k (-1)^j \binom{-s+1}{j} \nabla^j f_{n+1}$$

considerando s como la variable $s = \frac{t-t_n}{h}$. Luego, tenemos que

$$\int_{t_n}^{t_{n+1}} y'(t) dt \simeq \int_{t_n}^{t_{n+1}} p^*(t) dt = \int_{s=0}^1 \sum_{j=0}^k (-1)^j \binom{-s+1}{j} \nabla^j f_{n+1} ds = h \sum_{j=0}^k \gamma_j^* \nabla^j f_{n+1}$$

donde

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

Para

$$\begin{aligned} j=0 : \gamma_0^* &= (-1)^0 \int_0^1 \binom{-s+1}{0} ds = \int_0^1 ds = 1 \\ j=1 : \gamma_1^* &= (-1)^1 \int_0^1 \binom{-s+1}{1} ds = \int_0^1 (-s) ds = -\frac{1}{2} \\ &\vdots \end{aligned}$$

Los primeros ejemplos son de la forma:

$$k=0 : y_{n+1} - y_n = h \sum_{j=0}^0 \gamma_j^* \nabla^j f_{n+1} = h f_n$$

$$k=1 : y_{n+1} - y_n = h \sum_{j=0}^1 \gamma_j^* \nabla^j f_{n+1} = h(f_{n+1} - \frac{1}{2}(f_{n+1} - f_n)) = [\frac{1}{2}f_{n+1} + \frac{1}{2}f_n]$$

\vdots

1.3. Relación de Recurrencia de los Coeficientes

1.3.1. Adams Explícito:

Veamos que pueden obtenerse, de forma sencilla, los coeficientes γ_j . Consideramos $G(t)$ función real que tiene γ_j como coeficientes en su desarrollo de Maclaurin, es decir

$$G(t) = \sum_{j=0}^{\infty} \gamma_j t^j$$

Con la definición de γ_j y suponiendo que la serie converge (entonces $|t| < 1$), tenemos

$$G(t) = \sum_{j=0}^{\infty} \left((-1)^j \int_0^1 \binom{-s}{j} ds \right) t^j = \sum_{j=0}^{\infty} (-t)^j \int_0^1 \binom{-s}{j} ds$$

como sabemos que la serie converge absolutamente para $|t| < 1$, y aplicando el teorema binomial se obtiene que

$$\begin{aligned} G(t) &= \int_0^1 \left(\sum_{j=0}^{\infty} \binom{-s}{j} (-t)^j \right) ds = \int_0^1 (1-t)^{-s} ds \\ &= \left. \frac{(1-t)^{-s}}{\ln(1-t)} \right|_0^1 = -\frac{(1-t)^{-1}}{\ln(1-t)} + \frac{1}{\ln(1-t)} = \frac{1 - \frac{1}{1-t}}{\ln(1-t)}. \end{aligned}$$

Esto nos queda que

$$G(t) = \sum_{j=0}^{\infty} \gamma_j t^j = \frac{-t}{(1-t) \ln(1-t)}$$

esta última expresión podemos reescribirla de la forma siguiente

$$-\frac{\ln(1-t)}{t} G(t) = \frac{1}{1-t} \quad (1.3)$$

Haciendo los cálculos pertinentes

$$\begin{aligned} (1-t)^{-1} &= 1 + t + t^2 + t^3 + \dots \\ -\ln(1-t) &= \int_0^t \frac{d\mu}{1-\mu} = \int_0^t (1 + \mu + \mu^2) d\mu = \mu + \frac{\mu^2}{2} + \frac{\mu^3}{3} + \dots \Big|_0^t = t + \frac{1}{2}t^2 + \frac{1}{3}t^3 + \dots \end{aligned}$$

De la expresión (1.3) deducimos que

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

$$\begin{aligned} \gamma_0 + (\gamma_1 + \frac{1}{2}\gamma_0)t + (\gamma_2 + \frac{1}{2}\gamma_1 + \frac{1}{3}\gamma_0)t^2 + \dots + (\gamma_n + \frac{1}{2}\gamma_{n-1} + \dots + \frac{1}{n+1}\gamma_0)t^n \\ = 1 + t + t^2 + \dots \end{aligned}$$

Comparando coeficientes de potencias respectivas de t en ambos miembros se obtiene la siguiente relación, la cual permite computar los coeficientes γ_j de forma recursiva,

$$\frac{\gamma_n}{1} + \frac{\gamma_{n-1}}{2} + \dots + \frac{\gamma_0}{n+1} = 1, \quad n = 0, 1, 2, \dots$$

1.3.2. Adams Implícito:

Igual que el caso explícito es posible dar una fórmula recurrente sencilla, para calcular los coeficientes γ_m^* . En efecto si ponemos

$$G^*(t) = \sum_{j=0}^{\infty} \gamma_j^* t^j$$

se tiene, bajo las mismas suposiciones de convergencia, que

$$G^*(t) = \sum_{j=0}^{\infty} \left((-1)^j \int_0^1 \binom{-s+1}{j} ds \right) t^j = \int_0^1 \left(\sum_{j=0}^{\infty} \binom{-s+1}{j} (-t)^j \right) ds$$

y aplicando ahora el teorema binomial

$$\int_0^1 \left(\sum_{j=0}^{\infty} \binom{-s+1}{j} (-t)^j \right) ds = \int_0^1 (1-t)^{-s+1} ds = \frac{(1-t)^{-s+1}}{\ln(1-t)} \Big|_0^1$$

tenemos que

$$G^*(t) = \sum_{j=0}^{\infty} \gamma_j^* t^j = \frac{-t}{\ln(1-t)}$$

Esta ultima expresión podemos reescribirla de la forma siguiente

$$- \frac{\ln(1-t)}{t} G(t) = 1 \tag{1.4}$$

Procediendo de forma análoga al caso explícito deducimos que

$$\gamma_0^* + (\gamma_1^* + \frac{1}{2}\gamma_0^*)t + (\gamma_2^* + \frac{1}{2}\gamma_1^* + \frac{1}{3}\gamma_0^*)t^2 + \dots + (\gamma_n^* + \frac{1}{2}\gamma_{n-1}^* + \dots + \frac{1}{n+1}\gamma_0^*)t^n + \dots = 1$$

Comparando coeficientes de potencias respectivas de t en ambos miembros se tienen que

$$\gamma_n^* + \frac{1}{2}\gamma_{n-1}^* + \frac{1}{3}\gamma_{n-2}^* + \dots + \frac{1}{n+1}\gamma_0^* = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \geq 1 \end{cases}$$

1.4. Métodos Predictor-Corrector

La combinación de un método multipaso explícito con uno implícito da lugar a un método de predicción y corrección, o bien “Método Predictor-Corrector”, donde el método explícito predice una aproximación de y_{n+1} , luego se evalúa la función con dicha aproximación y luego el método implícito la corrige. Los distintos modos PEC, PECE, PECEC, PECECE, aparecerán dependiendo del número de evaluaciones (E) que se hagan y del número de correcciones (C).

Por ejemplo, para obtener un método Predictor-Corrector basados en los métodos de Adams de dos pasos, debemos elegir el Adams-Bashforth de dos pasos como predictor y el Adams-Moulton de dos pasos como corrector.

Aplicamos primero (AB_2) y calculamos una estimación inicial de $y_{n+1}^{(0)}$

$$P : y_{n+1}^{(0)} = y_n^{(1)} + \frac{h}{2}(3f_n^{(0)} - f_{n+1}^{(0)})$$

A continuación, evaluamos la función en la aproximación calculada

$$E : f_{n+1}^{(0)} = f(t_{n+1}, y_{n+1}^{(0)})$$

Luego aplicamos el corrector (AM_2), para mejorar la aproximación obtenida

$$C : y_{n+1}^{(1)} = y_n^{(1)} + \frac{h}{12}(5f_{n+1}^{(0)} + 8f_n^{(0)} - f_{n-1}^{(0)})$$

El ejemplo visto es el (PEC), modo Predecir-Evaluar-Corregir, que es una de las diferentes variantes del método Predictor-Corrector. Existen otras posibilidades como el modo ($PECE$) :

$$\begin{aligned} P : y_{n+1}^{(0)} &= y_n^{(1)} + \frac{h}{2}(3f_n^{(1)} - f_{n-1}^{(1)}) \\ E : f_{n+1}^{(0)} &= f(t_{n+1}, y_{n+1}^{(0)}) \\ C : y_{n+1}^{(1)} &= y_n^{(1)} + \frac{h}{12}(5f_{n+1}^{(0)} + 8f_n^{(1)} - f_{n-1}^{(1)}) \\ E : f_{n+1}^{(1)} &= f(t_{n+1}, y_{n+1}^{(1)}) \\ &\vdots \end{aligned}$$

1.5. Métodos Explícitos de Nyström

Los métodos de Nyström se basan también en cuadraturas, pero usando el intervalo $[t_{n-1}, t_{n+1}]$ en lugar del $[t_n, t_{n+1}]$, como se usó en los métodos de Adams. Por tanto,

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

De la misma forma que en Adams explícito, Nyström reemplaza la función desconocida $f(t, y(t))$ por el polinomio interpolador tal que

$$p(t) = p(t_n + sh) = \sum_{j=0}^{k-1} (-1)^j \binom{-s}{j} \nabla^j f_n,$$

y obtenido así la fórmula siguiente

$$p(t) = p(t_n + sh) = \sum_{j=0}^{k-1} (-1)^j k_j \nabla^j f_n,$$

donde

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

1.6. Métodos Implícitos de Milne-Simpson

Consideramos la ecuación integral

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

Procediendo de forma similar al caso Adams Implícito, sustituimos el integrando por el polinomio interpolador $p^*(t)$ en los nodos $t_{n+1}, t_n, \dots, t_{n-(k+1)}$, tal que

$$p^*(t) = \sum_{j=0}^k (-1)^j \binom{-s+1}{j} \nabla^j f_{n+1},$$

y obtenemos la fórmula implícita

$$y_{n+1} = y_{n-1} + h \sum_{j=0}^k k_j^* \nabla^j f_{n+1},$$

donde

$$k_j^* = (-1)^j \int_{-1}^1 \binom{-s+1}{j} ds.$$

1.7. Métodos BDF

Hasta ahora, las fórmulas de los distintos métodos vistos se han obtenido partiendo de la forma integral de la ecuación diferencial y sustituyendo $y'(t)$ por el polinomio de interpolación a la derivada en varios nodos. Por tanto, se tratan de fórmulas basadas en la integración numérica.

Los métodos BDF son fórmulas basadas en la derivación numérica (backward differentiation formulas). Fueron introducidas independientemente por Curtiss-Hirschfelder (1952) y Mitchell-Craggs (1953) y popularizadas por Gear (1971) para la integración de problemas stiff.

Sea $p(t)$ el polinomio que interpola a (t_j, y_j) en $j = n, n-1, \dots, n-(k-1)$ y que pasa además por (t_{n+1}, y_{n+1}) . Exigimos ahora que este polinomio verifique la ED en el punto t_{n+1} ;

$$p'(t_{n+1}) = f(t_{n+1}, y_{n+1}). \quad (1.5)$$

Luego, tenemos

$$p(t) = \sum_{j=0}^k (-1)^j \binom{-s+1}{j} \nabla^j y_{n+1}, \quad s = \frac{t-t_n}{h} \rightarrow t = t_n + sh$$

Para obtener con fórmula para y_{n+1} , impondremos la condición (1.5), teniendo así:

$$\frac{dp}{dt}(t = t_{n+1}) = f(t_{n+1}, y_{n+1}) = f_{n+1}$$

Desarrollándolo tenemos que

$$\begin{aligned} \left. \frac{dp}{dt} \right|_{t=t_{n+1}} &= \frac{1}{h} \cdot \left. \frac{dp}{ds} \right|_{s=1} = \frac{1}{h} \cdot \frac{d}{ds} \left(\sum_{j=0}^k (-1)^j \binom{-s+1}{j} \nabla^j y_{n+1} \right) \Big|_{s=1} \\ &= \frac{1}{h} \sum_{j=0}^k (-1)^j (-1)^j \frac{1}{j} \nabla^j y_{n+1} = \frac{1}{h} \sum_{j=0}^k \frac{1}{j} \nabla^j y_{n+1} \end{aligned}$$

Teniendo en cuenta los desarrollos pertinentes ¹, llegamos a la siguiente igualdad

$$\sum_{j=0}^k \frac{1}{j} \nabla^j y_{n+1} = hf_{n+1} \quad \text{fórmula BDF de } k \text{ pasos.}$$

¹ $\binom{-s+1}{j} = \frac{(-s+1)(-s)(-s-1)\dots(-s-(j-2))}{j!}, \quad \left. \frac{d}{ds} \binom{-s+1}{j} \right|_{s=1} = (-1)^j \frac{1}{j}$

1.8. Experimentos Numéricos

A efectos de comprobar el orden de los métodos de forma experimental vamos a considerar la ecuación del calor con un término fuente, de modo que tenga solución exacta suave, conocida de antemano y fijaremos condiciones de frontera de tipo Dirichlet y una condición inicial.

$$\begin{aligned} u_t(x, t) &= u_{xx}(x, t) + g(x, t), \quad x \in (0, 1), \quad t \in [0, 2] \\ u(0, t) &= a(t) := \exp(-t), \quad u(1, t) = b(t) := 4 \exp(-t), \quad t \in [0, 2] \\ u(x, 0) &= (x + 1)^2. \quad x \in (0, 1). \end{aligned} \quad (1.6)$$

Elegimos $g(x, t) = -(x^2 + 2x + 3) \exp(-t)$, de modo que la solución exacta de la EDP es

$$u(x, t) = (x + 1)^2 \exp(-t). \quad (1.7)$$

Si discretizamos la ecuación en la variable espacial usando diferencias centrales de segundo orden

$$v''(x) = \frac{v(x+h) - 2v(x) + v(x-h)}{h^2} - \frac{h^2}{12} v^{(4)}(x + \theta_h h), \quad |\theta_h| < 1,$$

obtenemos, tras usar la partición $(x_j = jh, j = 1, \dots, N, h = \frac{1}{N+1})$ del intervalo $(0, 1)$ en N subintervalos, el siguiente sistema de ODEs en el intervalo $t \in [0, 2]$,

$$\begin{aligned} u_1'(t) &= h^{-2}(-2u_1 + u_1) + g(x_1, t) + h^{-2}a(t), \\ u_j'(t) &= h^{-2}(u_{j-1} - 2u_j + u_{j+1}) + g(x_j, t), \quad j = 2, 3, \dots, N-1, \\ u_N'(t) &= h^{-2}(-2u_N + u_{N-1}) + g(x_N, t) + h^{-2}b(t), \end{aligned} \quad (1.8)$$

donde $u_j(t) = u(x_j, t)$. No se cometen errores en la discretización espacial pues $u_{xxxx}(x, t) \equiv 0$. Usaremos $N = 4$ para ver los errores cometidos por los métodos de Adams y BDF en el punto final de integración $t^* = 2$ (en la norma euclídea ponderada) y confirmar el orden de convergencia de los métodos. En el capítulo 5 usaremos también este ejemplo para verificar los intervalos de estabilidad lineal de los métodos anteriormente mencionados.

En la Tabla 1.1 podemos apreciar que los métodos explícitos presentan problemas de convergencia si Δt no se reduce de forma adecuada. Cuanto mayor es el número de pasos k , mas reducción en Δt debe hacerse para que el método conveja. Esto se explica por las propiedades de estabilidad lineal que se verán en el capítulo 5. El orden de convergencia de los métodos p puede verse en dicha tabla de forma experimental cuando los métodos convergen. En el caso de los métodos implícitos de Adams puede confirmarse en la Tabla 1.2 su orden de convergencia $p = k + 1$. Además estos métodos presentan menos problemas de estabilidad que los explícitos, pues convergen para valores mayores de Δt . Finalmente para los métodos BDF, Tabla 1.3, puede apreciarse que no presentan ningún problema de estabilidad y los métodos convergen (con orden k).

k	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-9}$	$\Delta t = 2^{-10}$
2	8.6066e+11	1.3787e+23	6.3177e+35	3.0753e+41	1.2052e+16	3.1791e-10	3.9656e-11	4.9518e-12
	p	-2.2953e+01	3.6858e+01	7.9028e+01	2.0034e+00	2.0017e+00	2.0008e+00	6.7850e-01
3	8.6066e+11	1.3787e+23	6.3177e+35	3.0753e+41	1.2052e+16	3.1791e-10	3.9656e-11	4.9518e-12
	p	-3.7221e+01	-4.2059e+01	-1.8893e+01	8.4400e+01	8.4971e+01	3.0030e+00	3.0015e+00
4	4.5849e+10	2.8048e+23	8.5334e+39	4.2506e+55	5.9919e+56	1.3297e+10	7.2068e-14	4.5284e-15
	p	-4.2476e+01	-5.4756e+01	-5.2145e+01	-3.8173e+00	1.5498e+02	7.7288e+01	3.9923e+00

Tabla 1.1: Errores en el punto final $t^* = 2$ usando Adams-Bashforth de $k = 2, 3, 4$ pasos, usando los Δt indicados y valores de arranque adicionales exactos. Cuando el método converge el error de cada método puede leerse como $p \approx \log_2(\epsilon(\Delta t)/\epsilon(2^{-1}\Delta t))$, donde $\epsilon(\Delta t)$ representa el error global en el punto final en la norma euclidea ponderada al hacer la integración temporal con paso Δt .

k	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-9}$	$\Delta t = 2^{-10}$
1	4.6086e-06	1.1525e-06	2.8815e-07	7.2039e-08	1.8010e-08	4.5025e-09	1.1256e-09	3.5186e-09
	p	1.9981e+00	1.9995e+00	1.9999e+00	2.0000e+00	2.0000e+00	2.0000e+00	2.0000e+00
2	2.5068e-06	1.4746e-07	1.8218e-08	2.2642e-09	2.8221e-10	3.5226e-11	4.3994e-12	5.5070e-13
	p	4.0875e+00	3.0169e+00	3.0083e+00	3.0041e+00	3.0021e+00	3.0013e+00	2.9980e+00
3	1.4674e-05	5.4032e-06	3.6537e-10	2.2555e-11	1.4013e-12	8.7605e-14	5.1221e-15	1.9652e-15
	p	1.4414e+00	1.3852e+01	4.0178e+00	4.0086e+00	3.9996e+00	4.0962e+00	1.3821e+00

Tabla 1.2: Errores en el punto final $t^* = 2$ con Adams-Moulton de $k = 1, 2, 3$ pasos (orden $k + 1$), usando los Δt indicados y valores de arranque adicionales exactos. El error de cada método puede leerse como $p \approx \log_2(\epsilon(\Delta t)/\epsilon(2^{-1}\Delta t))$

k	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-9}$	$\Delta t = 2^{-10}$
2	8.1052e-05	1.9329e-05	4.7200e-06	1.1662e-06	2.8985e-07	7.2251e-08	1.8036e-08	4.5058e-09
	p	2.0680e+00	2.0339e+00	2.0169e+00	2.0085e+00	2.0042e+00	2.0021e+00	2.0011e+00
3	8.0479e-06	9.3223e-07	1.1220e-07	1.3764e-08	1.7044e-09	2.1205e-10	2.6444e-11	3.3008e-12
	p	3.1099e+00	3.0546e+00	3.0272e+00	3.0136e+00	3.0068e+00	3.0034e+00	3.0020e+00
4	8.5754e-07	4.7999e-08	2.8464e-09	1.7331e-10	1.0692e-11	6.6522e-13	4.2393e-14	8.2203e-15
	p	4.1591e+00	4.0758e+00	4.0377e+00	4.0188e+00	4.0065e+00	3.9719e+00	2.3666e+00

Tabla 1.3: Errores en el punto final $t^* = 2$ con métodos BDF de $k = 2, 3, 4$ pasos (orden k), usando los Δt indicados y valores de arranque adicionales exactos. El error de cada método puede leerse como $p \approx \log_2(\epsilon(\Delta t)/\epsilon(2^{-1}\Delta t))$.

Error y Condiciones de Orden

En la sección I del trabajo, “Métodos Lineales Multipaso Clásicos”, vimos que los métodos multipaso tenían en común que tanto las aproximaciones numéricas y_n como los valores $f_n = f(t_n, y_n)$ aparecen en combinaciones lineales.

Además, debemos saber que existen métodos multipaso aún más generales que los métodos vistos. Estos son de la forma

$$\sum_{s=0}^k \alpha_s y_{n+s} = h \sum_{j=0}^k \beta_j f_{n+j}, \quad \text{método lineal de } k \text{ pasos.} \quad (2.1)$$

Los coeficientes α_s, β_s son parámetros reales independientes de h y de variable n . A lo largo de esta sección asumiremos que

$$\alpha_k \neq 0, \quad |\alpha_0| + |\beta_0| > 0$$

de esta manera la expresión (2.1) es una fórmula de k pasos, la cuál se puede resolver siempre respecto a y_{n+k} para valores de h suficientemente pequeños, que son los de interés, pues la convergencia teórica se prueba siempre para $h \rightarrow 0$.

Los diferentes métodos numéricos (2.1) dependen exclusivamente de los coeficientes α_s y β_s . Además, distinguiremos entre métodos explícitos ($\beta_k = 0$) e implícitos ($\beta_k \neq 0$).

2.1. Error Global y Error Local

Se entiende como **error global** del método multipaso (2.1) a la diferencia entre el valor $y(t)$ de la solución exacta del problema (PVI) y el valor obtenido por el método después de $n + k$ pasos, pues hemos considerado que el método necesita k valores adicionales de arranque y por tanto sería el error dado por el método tras n pasos de integración dados con el método considerado,

$$e_g(t_{n+k}) = y(t_{n+k}) - y_{n+k}. \quad (2.2)$$

El análisis de error local se hace asumiendo que los valores previos usados por el método son iguales a los valores de la solución exacta del (PVI), es decir

$$y_{n+j} = y(t_{n+j}), \quad \text{para } j = 0, 1, \dots, k-1.$$

Por lo tanto, para evitar confusión, denotaremos por \tilde{y}_{n+k} al valor obtenido, asumiendo $y_{n+j} = y(t_{n+j}) : j = 0, 1, \dots, k-1$, definiendo así el **error local** del método multipaso (2.1) mediante la diferencia siguiente

$$e_l(t_{n+k}) = y(t_{n+k}) - \tilde{y}_{n+k} \quad (2.3)$$

Con el objetivo de medir el error local, asociaremos a (2.1) el operador de diferencia lineal L definido por (observe que para la solución exacta de la ODE se tiene que $y'(t) = f(t, y(t))$),

$$L(y, t_n, h) = \sum_{j=0}^k \alpha_j y(t_n + jh) - h \sum_{j=0}^k \beta_j y'(t_n + jh), \quad (2.4)$$

donde $y(t_n)$ es una función diferenciable definida en un intervalo que contiene $t_n + jh, \forall i = 0, 1, \dots, k$. Haciendo uso de la definición (2.3), tenemos que el método lineal multipaso verifica que:

$$\alpha_k \tilde{y}_{n+k} + h \beta_k f(t_{n+k}, \tilde{y}_{n+k}) + \sum_{j=0}^{k-1} \alpha_j y(t_{n+j}) - h \sum_{j=0}^{k-1} \beta_j f(t_{n+j}, y'(t_{n+j})) = 0.$$

Realizando la diferencia entre la expresión anterior y la expresión (2.4) tenemos que

$$\begin{aligned} L(y, t_n, h) &= \alpha_k y(t_{n+k}) + h \beta_k f(t_{n+k}, \tilde{y}_{n+k}) - (\alpha_k \tilde{y}_{n+k} - h \beta_k f(t_{n+k}, \tilde{y}_{n+k})) \\ &= \alpha_k e_l(t_{n+k}) - h \beta_k [f(t_{n+k}, y(t_{n+k})) - f(t_{n+k}, \tilde{y}_{n+k})]. \end{aligned}$$

Aplicando el Teorema del valor medio; $\exists \eta_{n+k} \in (y(t_{n+k}), \tilde{y}_{n+k})$ tal que

$$\begin{aligned} L(y, t_n, h) &= \alpha_k e_l(t_{n+k}) - h \beta_k f_y(t_{n+k}, \eta_{n+k}) e_l(t_{n+k}) \\ &= [\alpha_k I - h \beta_k f_y(t_{n+k}, \eta_{n+k})] e_l(t_{n+k}). \end{aligned}$$

Mediantes los calculos anteriores, se deduce el siguiente Lema:

Lema 2.1. *Considerando la ecuación diferencial $y' = f(t, y), y(t_0) = y_0$ donde f es continuamente diferenciable y sea $y(t)$ su solución. Entonces el error local verifica*

$$e_l(t_{n+k}) = y(t_{n+k}) - \tilde{y}_{n+k} = (\alpha_k I - h \beta_k f_y(t_{n+k}, \eta_{n+k}))^{-1} L(y, t_n, h),$$

donde η es un valor entre $y(t_{n+k})$ y \tilde{y}_{n+k} .

Esto nos dice que prácticamente el error de discretización local y el error dado por el operador son de la misma magnitud, es decir

$$e_l(t_{n+k}) = (\alpha_k^{-1} + \mathcal{O}(h)) L(y_n, t_n, h).$$

2.2. Orden de los métodos multipaso

Como consecuencia de la acumulación del error local tras n pasos de tamaño h se espera que el error global tenga una potencia menos en h que la que tiene el error local, pues la suma de n errores locales es $n \mathcal{O}(h^{p+1}) = \mathcal{O}(h^p)$, ya que $n \cdot h = \hat{t}$ para $n \gg 1$, $\hat{t} > 0$. Luego, demostraremos más adelante que

$$e_l = \mathcal{O}(h^{p+1}) \Rightarrow e_g = \mathcal{O}(h^p) \quad (\text{bajo ciertas condiciones razonables})$$

Usaremos de nuevo el operador diferencial L (2.4), con el fin de explicar e introducir el concepto de orden para los métodos multipaso.

Un método tiene **orden de consistencia** $p \geq 1$ cuando verifica:

$$L(y, t_n, h) = \mathcal{O}(h^{p+1})$$

para todas las funciones $y(t)$ suficientemente regulares. Veamos que condiciones debe cumplir la expresión (2.4) para que la afirmación anterior sea cierta.

Supondremos por simplicidad que $y(t)$ es analítica en el intervalo correspondiente, entonces usando el desarrollo de Taylor para $y(t_{n+s}) = y(t_n + sh)$ y $y'(t_{n+s}) = y'(t_n + sh)$ alrededor de un punto $t = t_n$ se tiene que

$$\begin{aligned} y(t_{n+s}) &= y(t_n + sh) = y(t_n) + y'(t_n)sh + \dots + \frac{y^{(q)}(t_n)}{q!}(sh)^q + \dots \\ y'(t_{n+s}) &= y'(t_n + sh) = y'(t_n) + y''(t_n)sh + \dots + \frac{y^{(q+1)}(t_n)}{q!}(sh)^q + \dots \end{aligned}$$

Por tanto, podemos escribir

$$\begin{aligned} L(y, t_n, h) &= \sum_{s=0}^k \alpha_s y(t_{n+s}) - \sum_{s=0}^k \beta_s y'(t_{n+s}) \\ &= \sum_{s=0}^k \alpha_s \sum_{q=0}^{\infty} \frac{y^{(q)}(t_n)}{q!} (sh)^q - h \sum_{s=0}^k \beta_s \sum_{l=0}^{\infty} \frac{y^{(l+1)}(t_n)}{l!} (sh)^l \\ &= \sum_{q=0}^{\infty} \left(\sum_{s=0}^k \alpha_s s^q \right) \frac{y^{(q)}(t_n)}{q!} h^q - \sum_{l=0}^{\infty} \left(\sum_{s=0}^k \beta_s s^l \right) \frac{y^{(l+1)}(t_n)}{l!} h^{l+1} \end{aligned}$$

tomamos $q = l + 1$

$$\begin{aligned} &= \sum_{q=0}^{\infty} \left(\sum_{s=0}^k s^q \alpha_s \right) \frac{y^{(q)}(t_n)}{q!} h^q - \sum_{q=1}^{\infty} \left(\sum_{s=0}^k s^{q-1} \beta_s \right) \frac{y^{(q)}(t_n)}{(q-1)!} h^q \\ &= \sum_{s=0}^k \alpha_s y(t_n) + \sum_{q=1}^{\infty} \left(\sum_{s=0}^k (s^q \alpha_s - q s^{q-1} \beta_s) \right) \frac{y^{(q)}(t_n)}{q!} h^q \end{aligned}$$

A partir de estos cálculos se deduce el siguiente Teorema:

Teorema 2.1. *Los métodos multipaso*

$$\sum_{s=0}^k \alpha_s y_{n+s} = h \sum_{j=0}^k \beta_j f_{n+j}$$

tienen orden de consistencia p exactamente, sí y sólo si:

$$\begin{aligned} (i) \quad & \sum_{s=0}^k \alpha_s = 0 \\ (ii) \quad & \sum_{s=0}^k s^q \alpha_s = \sum_{s=0}^k q s^{q-1} \beta_s \quad \text{para } q = 1, 2, \dots, p \\ (iii) \quad & \sum_{s=0}^k s^{p+1} \alpha_s \neq \sum_{s=0}^k (p+1) s^p \beta_s \end{aligned}$$

Los **polinomios característicos** ó polinomios generadores del método lineal multipaso (MLM) (2.1) son:

$$\varrho(\zeta) = \sum_{s=0}^k \alpha_s \zeta^s, \quad \sigma(\zeta) = \sum_{s=0}^k \beta_s \zeta^s \quad (2.5)$$

La introducción de estos polinomios esta justificada por el siguiente argumento. Considerando sucesiones $Z_n = (z_0, z_1, \dots)$ y denotando con E el operador desplazamiento $Ez_n = z_{n+1}$ podemos escribir las ecuaciones (2.1) de la forma

$$\sum_{s=0}^k (\alpha_s E^s) y_n = h \sum_{s=0}^k (\beta_s E^s) f_n \quad \iff \quad \varrho(E) y_n = h \sigma(E) f_n$$

Teorema 2.2. *Los métodos multipaso*

$$\sum_{s=0}^k \alpha_s y_{n+s} = h \sum_{j=0}^k \beta_j f_{n+j}$$

tienen orden de consistencia p exactamente, sí y sólo si se cumple alguna de las condiciones siguientes

$$\begin{aligned} (i) \quad & \varrho(e^h) - h\sigma(e^h) = Ch^{p+1} + \mathcal{O}(h^{p+2}), \quad C \neq 0 \\ (ii) \quad & \frac{\varrho(\zeta)}{\log \zeta} - \sigma(\zeta) = C(\zeta - 1)^p + \mathcal{O}(\zeta - 1)^{p+1}, \quad C \neq 0 \end{aligned}$$

Observación: Las condiciones para consistencia de orden 1, también llamadas condiciones de consistencia, pueden ponerse en la forma siguiente

$$\varrho(1) = 0, \quad \varrho'(1) = \sigma(1)$$

2.3. La constante de Error

Hay métodos que tienen el mismo orden pero proporcionan diferentes errores globales y estos se distinguirán por su **constante de error**.

Se denomina coeficiente de error de los métodos (2.1) al primer coeficiente distinto de cero C_{p+1} del operador diferencial lineal L . En consecuencia, los métodos (2.1) tienen orden de consistencia p si

$$L(y, t_n, h) = C_{p+1}h^{p+1}y^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}) \quad (2.6)$$

siendo C_{p+1} el coeficiente de error, el cual viene dado por

$$C_{p+1} = \frac{1}{(p+1)!} \left(\sum_{s=0}^k s^{p+1} \alpha_s - (p+1) \sum_{s=0}^k s^p \beta_s \right) \quad (2.7)$$

El problema que tenemos con este coeficiente de error es que si multiplicamos la fórmula (2.1) por cualquier constantes no nula, el método lineal multipaso no varía (da las mismas soluciones), sin embargo la constante de error está afectada por la constante multiplicativa usada. Por ello C_{p+1} no es adecuada como medida de precisión del método. Como ya hemos mencionado con anterioridad, los errores globales son de orden p si el método tiene errores locales de orden $p+1$ y es convergente (lo cual asumiremos de momento, y se verá en capítulos posteriores).

Supongamos que arrancamos con valores iniciales exactos para el PVI

$$\begin{cases} y'(t) = f(t, y) \\ y(0) = y_0 \end{cases}$$

donde $y_j = y(t_j)$, $j = 0, 1, \dots, k-1$, $t_j = jh$ con $t \in [0, t^*]$.

Entonces asumiendo $e_g(t_n) = \varphi(t_n)h^p + \mathcal{O}(h^{p+1})$, (esto es cierto, pero no podemos afirmarlo en este momento), podemos escribir $e_g(t_n) = y(t_n) - y_n$. Por tanto, se puede considerar

$$\tilde{e}_g(t_n) = \frac{e_g(t_n)}{h^p} = \frac{y(t_n) - y_n}{h^p} = \varphi(t_n) + \mathcal{O}(h), \quad (h \rightarrow 0) \quad (2.8)$$

siendo este el error global escalado por h^p . Considerando ahora la solución exacta del P.V.I y (2.6) se obtiene la siguiente expresión:

$$\sum_{j=0}^k \alpha_j y(t_{n+j}) - h \sum_{j=0}^k \beta_j f(t_{n+j}, y(t_{n+j})) = h^{p+1} C_{p+1} y^{(p+1)}(t_n) + \mathcal{O}(h^{p+2}) = L(y, t_n, h)$$

(Escribiremos L_n en vez de $L(y, t_n, h)$ por comodidad de notación). Y además, tenemos

$$\sum_{j=0}^k \alpha_j y_{n+j} - h \sum_{j=0}^k \beta_j f(t_{n+j}, y_{n+j}) = 0$$

Realizando la diferencia de estas dos últimas ecuaciones deducimos que:

$$\sum_{j=0}^k \alpha_j e_g(t_{n+j}) - h \sum_{j=0}^k \beta_j (f(t_{n+j}, y_{n+j}) - f(t_{n+j}, y_{n+j} - e_g(t_{n+j}))) = L_n$$

Haciendo uso del Teorema del Valor Medio, deducimos que existe ($\sigma_{n+j} \in (0, 1)$ punto intermedio) tal que,

$$\sum_{j=0}^k \alpha_j e_g(t_{n+j}) - h \sum_{j=0}^k \beta_j f_y(t_{n+j}, y(t_{n+j}) - \sigma_{n+j} e_g(t_{n+j})) e_g(t_{n+j}) = L_n.$$

Deducimos ahora que

$$\begin{aligned} \sum_{j=0}^k \alpha_j e_g(t_{n+j}) - h \sum_{j=0}^k \beta_j f_y(t_{n+j}, y(t_{n+j})) e_l(t_{n+j}) \\ = h^{p+1} C_{p+1} y^{(p+1)}(t_{n+k}) + \mathcal{O}(h^{p+1}) + \mathcal{O}(h e_g(t_n)). \end{aligned} \quad (2.9)$$

Dividiendo ahora por h^p , deducimos (suponiendo que el método es convergente) que,

$$\begin{aligned} \sum_{j=0}^k \alpha_j \tilde{e}_g(t_{n+j}) - h \sum_{j=0}^k \beta_j f_y(t_{n+j}, y(t_{n+j})) \tilde{e}_l(t_{n+j}) \\ = h C_{p+1} y^{(p+1)}(t_{n+k}) + \mathcal{O}(h^2). \end{aligned} \quad (2.10)$$

Suponiendo consistencia se tiene que $\sum_{j=0}^k \alpha_j = 0$, $\sum_{j=0}^k \beta_j = \sum_{j=0}^k j \alpha_j$ y además,

para los polinomios fundamentales $\varrho(\zeta)$ y $\sigma(\zeta)$ tenemos $\sum_{j=0}^k \beta_j = \sigma(1)$.

De la expresión (2.10) se obtiene, al usar la notación $J_{n+j} = f_y(t_{n+j}, y(t_{n+j}))$,

$$\sum_{j=0}^k \alpha_j \tilde{e}_g(t_{n+j}) - h \sum_{j=0}^k \beta_j J_{n+j} \tilde{e}_g(t_{n+j}) = h \frac{C_{p+1}}{\sigma(1)} \left(\sum_{j=0}^k \beta_j y^{(p+1)}(t_{n+j}) \right) + \mathcal{O}(h^2) \quad (2.11)$$

Observe que la segunda parte de la igualdad se obtiene mediante el Teorema del Valor Medio al escribir,

$$\left(\sum_{j=0}^k \beta_j \right) y^{(p+1)}(t_{n+k}) = \sum_{j=0}^k \beta_j [y^{(p+1)}(t_{n+j}) + (y^{(p+1)}(t_{n+k}) - y^{(p+1)}(t_{n+j}))]$$

donde $y^{(p+1)}(t_{n+k}) - y^{(p+1)}(t_{n+j}) = (j-k)hy^{(p+2)}(t_{n+k}) + \mathcal{O}(h^2) = \mathcal{O}(h)$. Por tanto, nos queda

$$\left(\sum_{j=0}^k \beta_j \right) y^{(p+1)}(t_{n+k}) = \sum_{j=0}^k \beta_j y^{(p+1)}(t_n + jh) + \mathcal{O}(h).$$

Volviendo a la expresión (2.11) y despreciando los términos de orden h^2 nos queda

$$\sum_{j=0}^k \alpha_j \tilde{e}_g(t_{n+j}) = h \sum_{j=0}^k \beta_j \left(f_y(t_{n+j}, y(t_{n+j})) \tilde{e}_g(t_{n+j}) + \frac{C_{p+1}}{\sigma(1)} y^{(p+1)}(t_{n+j}) \right). \quad (2.12)$$

Suponiendo que los valores iniciales son exactos se tiene que $\tilde{e}_g(t_j) = 0$, $j = 0, 1, \dots, k-1$. Por otro lado, la fórmula (2.12) no es otra cosa que el resultado de aplicar el mismo MLM al PVI

$$\begin{cases} x'(t) &= A(t)x(t) + b(t), & t \in [0, t^*] \\ x(0) &= 0, \end{cases} \quad (2.13)$$

siendo $A(t) = f_y(t, y(t))$ y $b(t) = \frac{C_{p+1}}{\sigma(1)} y^{(p+1)}(t)$. Por tanto, si el método lineal multipaso (2.12) converge tenemos que

$$\lim_{h \rightarrow 0} \|\tilde{e}_g(t_n) - x(t_n)\| = 0.$$

Teniéndose además que

$$\tilde{e}_g(t_n) - x(t_n) = \mathcal{O}(h^p), \quad p \geq 1. \quad (2.14)$$

La solución del problema lineal de coeficientes variables de (2.13) nos queda de la forma:

$$x(t) = C \int_0^t R(t, s) y^{(p+1)}(s) ds, \quad C = \frac{C_{p+1}}{\sigma(1)}, \quad (2.15)$$

donde $R(t, s) = Y(t)Y(s)^{-1}$, siendo $Y(t)$ matriz fundamental de soluciones del sistema diferencial $x'(t) = A(t)x(t)$. Teniendo ahora en cuenta (2.14) y (2.15) se explica porqué la constante de error de un método lineal multipaso de orden p viene dada por $C = \frac{C_{p+1}}{\sigma(1)}$. Observe que los errores globales son proporcionales a esta constante si el método convergiera.

2.4. Métodos Irreducibles

Sean $\varrho(\zeta)$ y $\sigma(\zeta)$ las fórmulas de los polinomios generadores del método (2.1) y supongamos que tengan un factor en común $\varphi(\zeta)$. Luego,

$$\varrho^*(\zeta) = \frac{\varrho(\zeta)}{\varphi(\zeta)} \quad \sigma^*(\zeta) = \frac{\sigma(\zeta)}{\varphi(\zeta)}$$

son los polinomios generadores de un nuevo método linear multipaso.

Considerando una sucesión $Z_n = (z_0, z_1, \dots)$ y denotando con E el operador desplazamiento $EZ_n = Z_{n+1}$ podemos escribir este nuevo método de la forma siguiente

$$\varrho^*(E)y_n = h\sigma^*(E)f_n$$

Si multiplicamos esta expresión por $\varphi(E)$ observamos que cualquier solución de este nuevo método es también solución del método

$$\varrho(E)y_n = h\sigma(E)f_n$$

Por ende, deducimos que estos dos métodos son equivalentes.

Denotando por L^* el nuevo operador diferencial asociado al nuevo método, y C_{p+1}^* a su nueva constante de error se tiene que

$$L(y, t, h) = \varphi(E)L^*(y, t, h) = C_{p+1}^*h^{p+1}\varphi(E)y^{(p+1)}(t) + \mathcal{O}(h^{p+2}) \quad (2.16)$$

Sabemos que φ es un polinomio de la forma $\varphi(\zeta) = a_0 + a_1\zeta + \dots + a_k\zeta^k$, el cual, puede ser expresado de la forma siguiente

$$\varphi(\zeta) = \varphi(1) + \varphi'(1)(\zeta - 1) + \frac{1}{2!}\varphi''(1)(\zeta - 1)^2 + \dots = \varphi(1) + \varphi'(1)\mathcal{O}(h) + \frac{1}{2!}\varphi''(1)\mathcal{O}(h^2) + \dots$$

Introduciendo el operador desplazamiento y una función regular $f(t)$ tenemos que

$$\varphi(E)f(t) = \varphi(I)f(t) + \varphi'(1)(E - I)f(t) + \frac{1}{2!}\varphi''(1)(E - I)^2f(t) + \dots, \quad (E - I) = \mathcal{O}(h)$$

Por lo que la expresión (2.16) es equivalente a

$$L(y, t, h) = \varphi(E)L^*(y, t, h) = C_{p+1}^*\varphi(1)h^{p+1}y^{(p+1)}(t) + \mathcal{O}(h^{p+2})$$

de lo que deducimos que $C_{p+1} = \varphi(1)C_{p+1}^*$, y consecuentemente,

$$\frac{C_{p+1}}{\sigma(1)} = \frac{C_{p+1}^*}{\sigma^*(1)}$$

Concluimos que ambos métodos tienen la misma constante de error.

Todos los cálculos anteriores nos muestran que aquellos métodos multipaso con polinomios generadores que posean factores en común no son interesantes. Por ello, asumiremos que $\varrho(\zeta)$ y $\sigma(\zeta)$ no tiene factores en común.

Llamaremos **métodos irreducibles** a los métodos que cumplan la propiedad anterior.

0-Estabilidad y Primera barrera de Dalhquist

Durante el siglo XX, matemáticos como Heinz Rustishauser o Germund Dalhquist observaron que la existencia de un error local pequeño y un alto orden para ciertos métodos lineales multipaso no implicaba que estos fueran de utilidad, pues la solución numérica puede presentar una considerable inestabilidad incluso para tamaños de paso h muy pequeños. Cuando un método sea estable en estas situaciones se dirá que no es 0-estable. Esta teoría y sus implicaciones en la convergencia es la que vamos a desarrollar en este capítulo.

Para el estudio de la 0-estabilidad de los métodos generales (2.1) nos centraremos en el comportamiento de la solución cuando ($h \rightarrow 0$), es decir,

$$\alpha_k y_{n+k} + \alpha_{k-1} y_{n+k-1} + \dots + \alpha_0 y_n = 0. \quad (3.1)$$

Tomando $y_i = \zeta^i$ en (3.1), y dividiendo entre ζ^n , obtenemos el polinomio característico de la ecuación en diferencias anterior

$$\varrho(\zeta) = \alpha_k \zeta^k + \alpha_{k-1} \zeta^{k-1} + \dots + \alpha_0 = 0. \quad (3.2)$$

La magnitud de las raíces del polinomio anterior caracterizarán las propiedades de 0-estabilidad del método.

Cuando las raíces del polinomio característico tienen multiplicidad $m > 1$, aparecen soluciones de (3.1) de la forma $y_n = n^{i-1} \zeta^n, \forall i = 1, \dots, m$. Tenemos así el siguiente resultado.

Lema 3.1. Sean ζ_1, \dots, ζ_l las raíces del polinomio generador $\varrho(\zeta) = 0$, con m_1, \dots, m_l sus respectivas multiplicidades. Entonces la solución general de (3.1) viene dada por

$$y_n = p_1(n) \zeta_1^n + \dots + p_l(n) \zeta_l^n,$$

donde los $p_i(n)$ son polinomios de grado $m_i - 1$

Definición 3.1. Los métodos multipaso (2.1) son **0-estables**, si el polinomio característico $\varrho(\zeta) = 0$ satisface la **condición de raíces** (root condition), es decir, que $\forall i = 1, \dots, l, |\zeta_i| \leq 1$ y que todas las raíces con módulo 1 sean simples.

Definición 3.2. *Importante destacar tres definiciones básicas:*

- I. Los métodos se dicen **fuertemente cero-estables** si la única raíz de módulo 1 es $\zeta = 1$ y esta es además simple.
- II. Los métodos son **débilmente cero-estables** si son 0-estables y posean alguna raíz simple con módulo 1.
- III. Los métodos que no cumplan la condición de las raíces son **inestables**.

Para los métodos explícitos e implícitos de Adams:

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

$$y_{n+k} = y_n + h(\beta_k f_{n+1} + \dots + \beta_0 f_{n-k+1})$$

tenemos, $\varrho(\zeta) = \zeta^k - \zeta^{k-1}$. Además de tener la raíz simple 1, hay $k - 1$ raíces en 0. Por lo tanto, los métodos de Adams son 0-estables (fuertemente 0-estables). Sin embargo, los métodos de Nyström explícitos y de Milne-Simpson, se tiene que $\varrho(\zeta) = \zeta^k - \zeta^{k-2}$, con lo cual aparece una nueva raíz $\zeta = -1$, de módulo 1 lo que los hace débilmente cero-estables.

El siguiente teorema resume las propiedades de estabilidad de las familias de fórmulas multipaso anteriormente vistas.

Teorema 3.1. *Los métodos de k -pasos Adams Bashfort, Adams-Moulton, Nyström y las fórmulas generalizas de Milne Simpson son 0-estables para todo $k \geq 1$. Los métodos de k -pasos BDF-fórmulas son 0-estables para $1 \leq k \leq 6$, pero inestables para $k \geq 7$.*

3.1. Máximo Orden Posible de Consistencia para Métodos Lineales Multipaso 0-Estables

Es habitual en el estudio de las propiedades de convergencia de los métodos, ¿por qué no usar simplemente los métodos multipaso de máximo orden de consistencia posible (puede conseguirse métodos de orden $2k$ con k pasos) y olvidar las clases especiales como los métodos Adams y Nyström?

La respuesta está en las propiedades de 0-estabilidad de los métodos (pues veremos que una condición necesaria para convergencia es 0-estabilidad) y aunque este hecho fue investigado por muchos autores durante la primera mitad del siglo XX para métodos particulares, fue Germund Dahlquist quien dio el resultado general que caracteriza el orden de consistencia máximo para métodos 0-estables de k pasos. Este resultado se conoce como la “**primera barrera de Dahlquist**”

Teorema 3.2. *El orden de consistencia p de un método lineal 0-estable de k -pasos debe satisfacer que*

$$p \leq \begin{cases} k + 2 & \text{si } k \text{ es par} \\ k + 1 & \text{si } k \text{ es impar} \\ k & \text{si el método es explícito} \end{cases}$$

A continuación veremos algunas notaciones y lemas que nos serán de gran utilidad para la demostración del Teorema anterior.

En primer lugar presentamos la transformación greco-romana (de Möbius):

$$\zeta = \frac{z + 1}{z - 1} \longleftrightarrow z = \frac{\zeta + 1}{\zeta - 1} \tag{3.3}$$

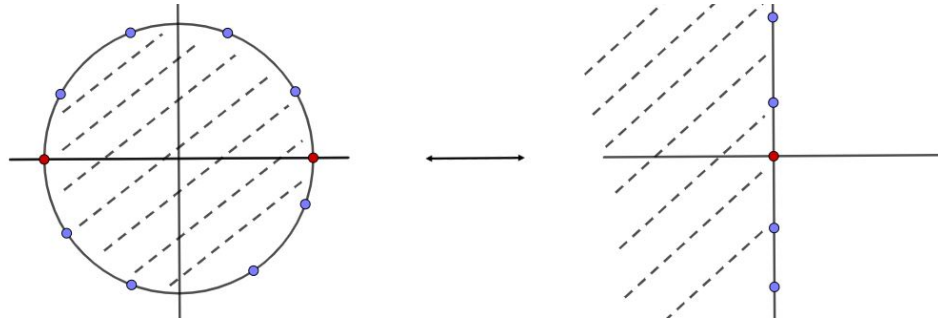


Figura 3.1: Transformación Greco-Romana

Esta transformación lleva de forma biunívoca cualquier punto del semiplano $\text{Re } z < 0$ dentro del disco $|\zeta| < 1$, el semiplano superior $\text{Im } z > 0$ sobre el semiplano inferior, los puntos del eje imaginario en el círculo $|\zeta| = 1$. Además $\zeta = \infty$ cuando $z = 1$ y $\zeta = 0$ para $z = -1$.

Consideramos los polinomios:

$$R(z) = \left(\frac{z - 1}{2}\right)^k \varrho(\zeta) = \sum_{j=0}^k a_j z^j$$

$$S(z) = \left(\frac{z - 1}{2}\right)^k \sigma(\zeta) = \sum_{j=0}^k b_j z^j,$$

donde $\varrho(\zeta)$ verifica condición de las raíces sí y solo si las raíces z_j de $R(z)$ tienen parte real negativa ó bien $\text{Re } z_j = 0$ y z_j es simple.

Lema 3.2. *Los métodos generales (2.1) que tienen orden de consistencia al menos cero y son 0-estables, verifican:*

- (i) $a_k = 0$ y $a_{k-1} = 2^{1-k} \varrho'(1) \neq 0$
- (ii) *Todos los coeficientes de $R(z)$ tienen el mismo signo*

Demostración. (i)

$$\begin{aligned}
 R(z) &= \left(\frac{z-1}{2} \right)^k (\alpha_0 + \alpha_1 \zeta^1 + \cdots + \alpha_k \zeta^k), \text{ hacemos sustitución de } \zeta. \\
 &= \left(\frac{z-1}{2} \right)^k \left(\alpha_0 + \alpha_1 \left(\frac{z+1}{z-1} \right) + \cdots + \alpha_k \left(\frac{z+1}{z-1} \right)^k \right) \\
 &= \frac{(z-1)^k}{2^k} \sum_{j=0}^k \alpha_j \left(\frac{z+1}{z-1} \right)^j = \frac{1}{2^k} \sum_{j=0}^k \alpha_j (z+j)^j (z-j)^{k-j} \\
 &= a_0 + a_1 z^1 + \cdots + a_k z^k
 \end{aligned}$$

Si $\zeta = 1 \leftrightarrow z = +\infty i$, luego dividiendo entre z^k y tomando límite:

$$\lim_{z \rightarrow \infty} \frac{R(z)}{z^k} = a_k = \frac{1}{2^k} \varrho(1) = 0$$

Luego, $R(z) = a_{k-1} z^{k-1} + \cdots + a_1 z = \left(\frac{z-1}{2} \right)^k \varrho(\zeta)$, volvemos a dividir y tomar límites:

$$\begin{aligned}
 \lim_{z \rightarrow \infty} \frac{R(z)}{z^{k-1}} &= a_{k-1} = \lim_{z \rightarrow \infty} \frac{(z-1)^k}{2^k z^{k-1}} \varrho(\zeta) = \frac{1}{2^k} \lim_{\zeta \rightarrow 1} \frac{(\zeta+1)\varrho(\zeta)}{(\zeta-1)} \\
 &= \frac{2}{2^k} \lim_{\zeta \rightarrow 1} \frac{\varrho(\zeta)}{\zeta-1} = 2^{-k+1} \varrho'(1).
 \end{aligned}$$

Demostración. (ii)

$$R(z) = a_{k-1} \prod_{j=1}^r (z+x_j) \prod_{j=1}^c ((z+u_j)^2 + v_j^2),$$

teniendo en cuenta la factorización anterior, donde $-x_j$ son las raíces reales y $-u_j \pm iv_j$ son los pares conjugados de raíces complejas.

Si un polinomio tiene, por ejemplo, raíces $-1, 3, 2 \pm 4i$ se tiene que

$$p(z) = k(z+1)(z-3)((z-2)^2 + 4^2), \quad \text{para } k \neq 0$$

Por tanto, para $R(z) = a_{k-1} \prod_{j=1}^r (z+x_j) \prod_{j=1}^c ((z+u_j)^2 + v_j^2)$ se tiene por estabilidad que $x_j \geq 0$, $u_j \geq 0$, entonces deducimos que todos los coeficientes tienen el mismo signo. \square

Lema 3.3. *Los mtodos (2.1) tienen orden de consistencia p s, y solo si*

$$R(z) \left(\log \frac{z+1}{z-1} \right)^{-1} - S(z) = C_{p+1} \left(\frac{2}{z} \right)^{p-k} + \mathcal{O} \left(\frac{2}{z} \right)^{p-k+1}, \quad \text{cuando } z \rightarrow \infty \quad (3.4)$$

Demostracin. Si (2.1) tiene orden de consistencia p si, y slo si

$$\frac{\varrho(\zeta)}{\log(\zeta)} - \sigma(\zeta) = C_{p+1}(\zeta - 1)^p + \mathcal{O}(\zeta - 1)^{p+1}, \quad \zeta \rightarrow 1$$

Luego, como $R(z) = \left(\frac{z-1}{2}\right)^k \varrho(\zeta) S(z) = \left(\frac{z-1}{2}\right)^k \sigma(\zeta)$ donde $\zeta = \frac{z+1}{z-1}$, se tiene que:

$$\begin{aligned} \frac{\left(\frac{2}{z-1}\right)^k R(z)}{\log\left(\frac{z+1}{z-1}\right)} - \left(\frac{2}{z-1}\right)^k S(z) &= C_{p+1} \left(\frac{2}{z-1}\right)^p + \mathcal{O}\left(\frac{2}{z-1}\right)^{p+1} \Rightarrow \\ \frac{R(z)}{\log\left(\frac{z+1}{z-1}\right)} - S(z) &= C_{p+1} \left(\frac{2}{z-1}\right)^{p-k} + \mathcal{O}\left(\frac{2}{z-1}\right)^{p-k+1} \Leftrightarrow \\ \frac{R(z)}{\log\left(\frac{z+1}{z-1}\right)} - S(z) &= C_{p+1}^* \left(\frac{1}{z}\right)^{p-k} + \mathcal{O}\left(\frac{1}{z}\right)^{p-k+1} \end{aligned}$$

Por tanto, $C_{p+1}^* = 2^{p-k} C_{p+1}$, cuando $z \rightarrow \infty$ □

Lema 3.4. *Los coeficientes de la serie de Laurent*

$$\log \left(\frac{z+1}{z-1} \right)^{-1} = \frac{z}{2} - \mu_1 z^{-1} - \mu_3 z^{-3} - \mu_5 z^{-5} - \dots \quad (3.5)$$

satisfacen que $\mu_{2j+1} > 0$, $\forall j \geq 0$

Demostracin. Teniendo en cuenta

$$\begin{aligned} \log(1-w) &= w + \frac{w^2}{2} + \frac{w^3}{3} + \dots + \frac{w^n}{n} + \dots, |w| < 1 \\ \log \frac{1+w}{1-w} &= 2 \left(w + \frac{w^3}{3} + \frac{w^5}{5} + \dots + \frac{w^{2n+1}}{2n+1} + \dots \right), |w| < 1 \end{aligned}$$

haciendo el cambio $w = \frac{1}{z} \Rightarrow |z| > 1 \Leftrightarrow |w| < 1$, luego

$$\begin{aligned} \log \frac{1 + \frac{1}{z}}{1 - \frac{1}{z}} &= \log \frac{z+1}{z-1} = 2z^{-1} \left(1 + z^{-2} \frac{1}{3} + z^{-4} \frac{1}{5} + \dots + z^{-2n} \frac{1}{2n+1} + \dots \right), |z| > 1 \\ \frac{1}{\log\left(\frac{z+1}{z-1}\right)} &= \frac{1}{\frac{2}{z} \left(1 + z^{-2} \frac{1}{3} + z^{-4} \frac{1}{5} + \dots \right)} = \frac{z}{2} \cdot \frac{1}{1 + z^{-2} \frac{1}{3} + z^{-4} \frac{1}{5} + \dots} \end{aligned}$$

$$= \frac{z}{2}(1 - \mu_1 z^{-2} - \mu_3 z^{-4} - \mu_5 z^{-6} \dots)$$

Veamos ahora que $\mu_{2j+1} > 0$, $\forall j \geq 0$

$$\begin{aligned} 1 &= \frac{z}{2} \log \left(\frac{z+1}{z-1} \right) (1 - \mu_1 z^{-2} - \mu_3 z^{-4} - \mu_5 z^{-6} - \dots) \\ &= \frac{z}{2} \cdot \frac{2}{z} \left(1 + z^{-2} \frac{1}{3} + z^{-4} \frac{1}{5} + \dots + z^{-2n} \frac{1}{2n+1} + \dots \right) (1 - \mu_1 z^{-2} - \mu_3 z^{-4} - \mu_5 z^{-6} - \dots) \\ &= 1 + z^{-2} \left(\frac{1}{3} - \mu_1 \right) + z^{-4} \left(\frac{1}{5} - \frac{\mu_1}{3} - \mu_3 \right) + z^{-6} \left(\frac{1}{7} - \frac{\mu_3}{3} - \frac{\mu_1}{5} - \mu_5 \right) + \dots \end{aligned}$$

Igualando coeficientes se tiene :

$$\mu_1 = \frac{1}{3}, \quad \mu_3 = \frac{1}{5} - \frac{\mu_1}{3} = \frac{4}{45}, \quad \mu_5 = \frac{1}{7} - \frac{4}{135} - \frac{1}{15} = \frac{44}{945}$$

Buscamos el coeficiente de z^{-2n} :

$$\frac{1}{2n+1} - \mu_1 \frac{1}{2n-1} - \mu_3 \frac{1}{2n-3} - \mu_5 \frac{1}{2n-5} - \dots - \mu_{2n-1} \frac{1}{1} = 0$$

\Rightarrow los coeficientes μ_j satisfacen la relación de recurrencia:

$$\mu_{2n-1} + \frac{1}{3} \mu_{2n-3} + \dots + \mu_3 \frac{1}{2n-3} + \mu_1 \frac{1}{2n-1} = \frac{1}{2n+1} \quad (3.6)$$

sustituyendo ahora $n+1$ por n tenemos:

$$\mu_{2n+1} + \frac{1}{3} \mu_{2n-1} + \dots + \mu_3 \frac{1}{2n-1} + \mu_1 \frac{1}{2n+1} = \frac{1}{2n+3} \quad (3.7)$$

Multiplicando (3.7) por $(2n+3)$ y restándole la expresión (3.6) multiplicada por $(2n+1)$ nos queda que :

$$(2n+3)\mu_{2n+1} + \sum_{j=0}^{n-1} \mu_{2i+1} \left(\frac{2n+3}{2(n-i)+1} - \frac{2n+1}{2(n-i)-1} \right)$$

Luego,

$$\mu_{2n+1} + \sum_{i=0}^n \mu_{2i+1} \left(\frac{1}{2i+1} - \frac{1}{2i-1} \right) = 0$$

Como $\frac{1}{2i+1} - \frac{1}{2i-1} = \frac{2}{4i^2-1} > 0 \geq 1 \Rightarrow \mu_1 > 0 \Rightarrow \mu_3 > 0 \Rightarrow \dots \Rightarrow \mu_{2j+1} > 0 \geq 0$ \square

En este punto ya poseemos las herramientas necesarias para proceder con la demostración del Teorema 3.2.

Demostracin (Teorema Dahlquist).

Insertando la serie (3.5) en la frmula (3.4) se obtiene que

$$\begin{aligned} R(z) \left(\log \frac{z+1}{z-1} \right)^{-1} - S(z) &= (a_0 + a_1z + \dots + a_{k-1}z^{k-1}) \left(\frac{z}{2} - \mu z^{-1} - \mu z^{-3} - \dots \right) - S(z) \\ &= P_k(z) + d_1z^{-1} + d_2z^{-2} + \mathcal{O}(z^{-3}) \end{aligned}$$

donde

$$\begin{aligned} d_1 &= - \sum_{j=0}^{k-1} \mu_{2j+1} a_{2j} = -\mu_1 a_0 - \mu_3 a_2 - \mu_5 a_4 \dots \\ d_2 &= - \sum_{j=1}^k \mu_{2j+1} a_{2j-1} = -\mu_3 a_1 - \mu_5 a_3 - \mu_7 a_5 \dots \end{aligned}$$

Todos los sumandos de d_1 y de d_2 tienen el mismo signo (por el Lema 3.3 y el Lema 3.4). Adems, por el Lema 3.2 se tiene que $a_{k-1} \neq 0$.

Por tanto, $d_2 \neq 0$ si k es par y $d_1 \neq 0$ si k es impar. Se deduce, por la frmula (3.4), que si k par entonces $p - k \leq 2$ y si k es impar se tiene que $p - k \leq 1$.

Las dos primeras afirmaciones del teorema han sido demostradas, veamos ahora que $p \leq k$ para $\frac{\beta_k}{\alpha_k} \leq 0$.

Asumimos, por contradiccin, que $p > k$. Entonces si $p - k > 0$ se tiene que

$$P_k(z) = z \left(\frac{a_0}{2} - a_2\mu_1 - a_4\mu_3 - \dots \right) + \left(\frac{a_1}{2} - a_3\mu_1 - a_5\mu_3 - \dots \right) + \dots = S(z).$$

Para $\mu_j = 0$ con j par, se cumple la siguiente igualdad:

$$S(z) = R(z) \left(\frac{z}{2} - \sum_{j=1}^{k-1} \mu_j z^{-j} \right) + \sum_{j=1}^{k-1} \left(\sum_{s=j}^{k-1} \mu_s a_{s-j} \right) z^{-j} \quad (3.8)$$

tomando $z = 1$ obtenemos

$$\frac{S(1)}{R(1)} = \left(\frac{1}{2} - \sum_{j=1}^{k-1} \mu_j \right) + \sum_{j=1}^{k-1} \left(\sum_{s=j}^{k-1} \mu_s a_{s-j} \right) \frac{1}{R(1)},$$

que teniendo en cuenta que

$$S(z) = \frac{1}{2^k} \sum_{j=0}^k \beta_j (z+1)^j (z-1)^{k-j} \Rightarrow S(1) = \frac{1}{2^k} \beta_k 2^k = \beta_k$$

$$R(z) = \frac{1}{2^k} \sum_{j=0}^k \alpha_j (z+1)^j (z-1)^{k-j} = \sum_{j=0}^k a_j z^k \Rightarrow S(1) = \alpha_k = \sum_{j=0}^{k-1} a_j$$

y que todos los a_j tienen el mismo signo que a_{k-1} , se concluye que $\frac{\beta_k}{\alpha_k} > 0$, lo cual es una contradicción.

Por el lema 3.4, $z \rightarrow 1$, se tiene que

$$\sum_{j=1}^{\infty} \mu_j = \frac{1}{2}.$$

Luego el primer sumando de (3.8) es estrictamente positivo. Utilizando de forma aditiva el lema 3.2 se demuestra de forma similar la no negatividad del segundo sumando.

□

Los métodos multipaso 0-estables que alcanzan el orden más alto posible $k+2$ tienen una estructura muy especial.

Se denominan **métodos simétricos** a aquellos que coinciden con su adjunto. Es decir, los que verifican

$$\alpha_j = -\alpha_{k-j}, \quad \beta_j = \beta_{k-j}, \quad j = 0, 1, \dots, k.$$

La regla Trapezoidal es un método simétrico (observe que se llega el mismo valor si se da un paso de tamaño h y luego un paso hacia atrás con paso $-h$):

$$\begin{aligned} y_{n+1} - y_n &= \frac{h}{2}[f_n + f_{n+1}], \quad \text{que aplicando la R.T. con paso } -h \text{ tenemos:} \\ y_{n-1} - y_n &= -\frac{h}{2}[f_n + f_{n+1}] \Leftrightarrow y_n - y_{n+1} = -\frac{h}{2}[f_{n+1} + f_n] \\ &\Leftrightarrow y_{n+1} - y_n = \frac{h}{2}[f_n + f_{n+1}] \end{aligned}$$

Teorema 3.3. *Los métodos multipaso 0-estables de orden $k+2$ son simétricos.*

Convergencia de los métodos multipaso

Dahlquist fue el primero en presentar una demostración de la convergencia general para los métodos lineales (ρ, σ) , dando condiciones necesarias y suficientes de la misma. Se introdujo posteriormente, usando las ideas de Butcher, una forma más elegante de probar los resultados de convergencia previamente obtenidos por Dahlquist. Las ideas de Butcher se basaban en plantear las fórmulas multipaso como fórmulas de un paso en un espacio dimensional superior. Esta formulación permite dar caracterizaciones algebraicas elegantes de estabilidad y además es idóneo porque se generaliza fácilmente para el análisis de los métodos multipaso cuando el paso de integración es variable.

No se puede esperar una convergencia razonable de los métodos numéricos, si el problema de la ecuación diferencial

$$y' = f(t, y), \quad y(t_0) = y_0 \quad (4.1)$$

no posee una solución única. Por lo tanto, asumimos como hemos hecho desde el principio las siguientes hipótesis,

$$f \text{ es continua en } D = \{(t, y); t \in [t_0, t^*], \|y(t) - y\| \leq b\} \quad (4.2)$$

donde $y(t)$ es la solución exacta de (4.1) y b algún número positivo. Además, asumimos que f satisface la condición de Lipschitz, es decir,

$$\|f(t, y) - f(t, z)\| \leq L \|y - z\| \quad \text{para } (t, y), (t, z) \in D \quad (4.3)$$

Si aplicamos los métodos multipaso (2.1) con tamaño de paso h al problema (4.1) obtenemos una secuencia de soluciones y_i . Además, dados t y h tales que $(t - t_0)/h = n$ es un número entero, usaremos la siguiente notación para la solución numérica sobre el punto t :

$$y_h(t) = y_n, \quad \text{si } t - t_0 = nh. \quad (4.4)$$

Definición 4.1. Presentamos las definiciones de convergencia y de convergencia de orden p :

i) Los métodos lineales multipaso (2.1) son **convergentes**, si para todos los (PVI) (4.1) satisfacen (4.2) y (4.3),

$$y(t) - y_h(t) \rightarrow 0, \quad h \rightarrow 0, \quad t \in [t_0, t^*]$$

siempre que los valores iniciales satisfagan

$$y(t_0 + ih) - h_h(t_0 + ih) \rightarrow 0, \quad h \rightarrow 0, \quad i = 0, 1, \dots, k - 1$$

ii) El método (2.1) es **convergente de orden p** , si para cualquier problema (4.1) con f suficientemente diferenciable, existe un h_0 positivo tal que

$$\|y(t) - y_h(t)\| \leq Ch^p, \quad \text{para } h \leq h_0$$

siempre que los valores iniciales satisfagan

$$\|y(t_0 + ih) - y_h(t_0 + ih)\| \leq C_0 h^p, \quad \text{para } h \leq h_0, \quad i = 0, 1, \dots, k - 1$$

Asumimos claramente mediante esta definición que existe una solución de (4.1) en $[t_0, t^*]$.

El objetivo de esta sección es demostrar que la estabilidad junto con la consistencia son condiciones necesarias y suficientes para la convergencia de los métodos multipaso. Esto se expresa en el famoso resultado

$$\text{convergencia} = \text{estabilidad} + \text{consistencia}$$

Teorema 4.1. Si el método multipaso (2.1) es convergente, entonces necesariamente presenta

i) 0-estabilidad

ii) Consistencia (consistencia de orden 1) $\iff \varrho(1) = 0 \wedge \varrho'(1) = \sigma(1)$

Teniendo en cuenta la estructura de los métodos (2.1)¹, procedemos a la demostración del teorema.

Demostración. i) Convergencia \implies 0-estabilidad

Aplicamos los métodos (2.1) para el PVI

¹ $\varrho(\zeta) = \sum_{j=0}^k \alpha_j \zeta^j$; $\sigma(\zeta) = \sum_{j=0}^k \beta_j \zeta^j$; $\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j f_{n+j}$, donde $f_{n+j} = f(t_{n+j}, y_{n+j})$

$$\begin{cases} y' = 0 \\ y(0) = 0 \end{cases}$$

donde se tiene que $h = \frac{t^*}{N}$, ($h \rightarrow 0 \Leftrightarrow N \rightarrow \infty$).

$$\text{Si } f \equiv 0 \Rightarrow \sum_{j=0}^k \alpha_j y_{n+j} = 0, \quad n = 0, 1, \dots, N - k. \quad (4.5)$$

Elegiremos ahora valores iniciales convenientes $y_0^{(h)}, y_1^{(h)}, \dots, y_{k-1}^{(h)}$ de modo que $\lim_{h \rightarrow 0} y_i^{(h)} = 0$ ($i = 0, 1, \dots, k - 1$). Recordando que las soluciones de (4.5) vienen dadas mediante combinaciones lineales de las potencias de las raíces del polinomio característico:

$$\varrho(\zeta) = 0 \Rightarrow \sum_{j=0}^k \alpha_j \zeta^j = 0 \quad (4.6)$$

Sea ζ_1 una raíz arbitraria de $\varrho(\zeta)$, veamos que $|\zeta_1| < 1$ o bien si se tiene $|\zeta_1| = 1$ entonces ζ_1 es simple:

Si $\exists |\zeta_1| > 1 \Rightarrow y_n = h\zeta_1^n$ cumple la ecuación (4.5) y la condición impuesta a los valores iniciales, pues si $h \rightarrow 0$ entonces $y_i = h\zeta_1^i \rightarrow 0$ para $i = 0, 1, \dots, k - 1$. Por otro lado, tenemos que

$$y_N = y_h(t^*) = h(\zeta_1)^N = h(\zeta_1)^{t^*/h} \quad \text{de donde deducimos que}$$

$$|y_N| = h|\zeta_1|^{t^*/h} \rightarrow \infty, \quad \text{por lo tanto } |\zeta_1|^{t^*/h} > 1, \text{ llegando así a una contradicción.}$$

Si $|\zeta_1| = 1$ y tiene multiplicidad mayor que 1, entonces

$$y_n = \sqrt{h} n \zeta_1^n$$

es una solución particular de (4.5) que además verifica $y_i(h) \rightarrow 0$ cuando $h \rightarrow 0$ para $i = 0, 1, \dots, k - 1$.

Por otro lado, $Y_N = y_n(t^*) = N\sqrt{h}(\zeta_1)^N$, y tomando ahora $|\cdot|$, se tiene:

$$|y_N| = N\sqrt{h}|\zeta_1|^N = N\sqrt{h} = \frac{t^*}{h}\sqrt{h} = \frac{t^*}{\sqrt{h}} \rightarrow \infty, \quad h \rightarrow 0.$$

Para ambos casos se obtienen soluciones que no convergen a la solución exacta $y = 0$ en el punto $t = t^*$. Esto nos lleva a una contradicción con la hipótesis de convergencia. Por tanto, el método (2.1) debe ser 0-estable. \square

Demostración. ii) Convergencia \Rightarrow Consistencia.

Tomando el PVI :

$$\begin{cases} y' = 0 \\ y(0) = 1 \end{cases}$$

cuya solución es $y(t) = 1$ para $t \in [0, 1]$. Aplicamos el método siguiente :

$$\begin{cases} \sum_{j=0}^k \alpha_j y_{n+j} = 0, & n = 0, 1, \dots, N - k \\ y_n = 1, & n = 0, 1, \dots, k - 1 \end{cases} \quad (4.7)$$

Si escribimos $n = N - k$, con $N = \frac{t^*}{h}$, tenemos:

$$\sum_{j=0}^k \alpha_j y_{N-k+j} = \alpha_0 y_{N-k} + \alpha_1 y_{N-k+1} + \dots + \alpha_k y_N = 0 \Rightarrow \sum_{j=0}^k \alpha_j y_h(t^* + (j-k)h) = 0$$

Tomando $h \rightarrow 0$ en (4.7), se tiene que $\lim_{h \rightarrow 0} y_h(t^* + (j-k)h) = \lim_{h \rightarrow 0} y_n(t^*) = 1 \Rightarrow \sum_{j=0}^k \alpha_j = \varrho(1) = 0$. (Primera condición de consistencia).

Para probar la segunda condición de consistencia elegimos el PVI $\begin{cases} y' = 1 \\ y(0) = 0 \end{cases}$

con solución $y(t) = t \Rightarrow y(t^*) = t^*$, ($t^* = Nh \Leftrightarrow h = \frac{t^*}{N}$, $N \rightarrow \infty \Leftrightarrow h \rightarrow 0$).

Aplicamos el método multipaso

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j f_{n+j} = h \sum_{j=0}^k \beta_j = h\sigma(1)$$

Buscamos una solución del método de la forma $y_n = Ct_n$, con $C = \frac{\sigma(1)}{\varrho'(1)}$, $\varrho'(1) \neq 0$, además se tiene que $C = 1$ por convergencia. Por tanto, se tiene:

$$\begin{aligned} \sum_{j=0}^k \alpha_j y_{n+j} = h\sigma(1) &\Rightarrow \sum_{j=0}^k \alpha_j C(n+j)h = h\sigma(1) \Leftrightarrow hC \sum_{j=0}^k (n+j)\alpha_j = h\sigma(1) \Leftrightarrow \\ &\Leftrightarrow Cn \sum_{j=0}^k \alpha_j + C \sum_{j=0}^k j\alpha_j = \sigma(1) \Leftrightarrow C\varrho'(1) = \sigma(1) \Leftrightarrow C = \frac{\sigma(1)}{\varrho'(1)} = 1 \end{aligned}$$

□

Es bastante notable y sorprendente que las condiciones i) y ii) del Teorema anterior, obtenidas de exigir convergencia sobre ecuaciones diferenciales muy simples $y' = 0$ y $y' = 1$, resulten no solo necesarias sino también suficientes para la convergencia. Esto lo veremos en las próximas sesiones de este capítulo.

4.1. Formulación como método de un solo paso

Estamos en el punto en que es útil reescribir un método multipaso como un método de un paso en un espacio dimensional superior.

Los métodos (2.1) pueden ser escritos de la forma

$$y_{n+k} = - \sum_{j=0}^{k-1} \frac{\alpha_j}{\alpha_k} y_{n+j} + \sum_{j=0}^k \frac{\beta_j}{\alpha_k} f_{n+j} \quad (4.8)$$

además, trabajaremos con el supervector

$$Y_n^\top = (y_{n+k-1}^\top, \dots, y_{n+1}^\top, y_n^\top) \in \mathbb{R}^{mk} : y_{n+j} \in \mathbb{R}^m$$

donde m es la dimensión de la ecuación diferencial.

El método multipaso (4.8) puede ser escrito de forma compacta como

$$Y_{n+1} = (A \otimes I)Y_n + h\Phi(t_n, Y_n, h) \quad (4.9)$$

siendo,

$$Y_{n+1}^\top = (y_{n+k}^\top, y_{n+k-1}^\top, \dots, y_{n+1}^\top), \quad A = \begin{pmatrix} -\frac{\alpha_{k-1}}{\alpha_k} & -\frac{\alpha_{k-2}}{\alpha_k} & \dots & -\frac{\alpha_1}{\alpha_k} & -\frac{\alpha_0}{\alpha_k} \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

donde \otimes denota el producto de Kronecker $A \otimes B = (a_{ij}B)$. En este caso $A \otimes I$ es matriz de bloques de dimensión mk con bloques (m, m) . Para una mayor sencillez de presentación supondremos a veces que $m = 1$ y $A \otimes I = A$.

La primera componente de (4.9) sería de la forma

$$y_{n+k} = -\frac{\alpha_{k-1}}{\alpha_k} y_{n+k-1} - \frac{\alpha_{k-2}}{\alpha_k} y_{n+k-2} \dots - \frac{\alpha_0}{\alpha_k} y_n + h\Phi(t_n, Y_n, h)$$

donde $\Phi_{n+k}(t_n, Y_n, h) = \sum_{j=0}^k \frac{\beta_j}{\alpha_k} f(t_{n+j}, y_{n+j})$, deduciendo así $\Phi_i(t_n, Y_n, h) = 0$, para $i = n+k-1, \dots, n+1$. Por tanto,

$$\Phi(t_n, Y_n, h) = (e_1 \otimes I)\Psi(t_n, Y_n, h), \quad \text{donde } e_1 = (1, 0, \dots, 0)^\top$$

y

$$\Psi \equiv \Psi(t_n, Y_n, h) = \sum_{j=0}^{k-1} \beta'_j f(t_n + jh, y_{n+j}) + h\beta'_k f\left(t_n + kh, h\Psi - \sum_{j=0}^{k-1} \alpha'_j y_{n+j}\right) \quad (4.10)$$

con

$$\beta'_j = \frac{\beta_j}{\alpha_k}, \quad \alpha'_j = \frac{\alpha_j}{\alpha_k}$$

La fórmula implícita (4.10) se deduce de la expresión (4.8):

$$y_{n+k} = - \sum_{j=0}^{k-1} \alpha'_j y_{n+j} + h\Psi \quad (4.11)$$

donde

$$\begin{aligned} \Psi &= \sum_{j=0}^k \beta'_j f_{n+j} = \sum_{j=0}^k \beta'_j f(t_{n+j}, y_{n+j}) \\ &= \sum_{j=0}^{k-1} \beta'_j f(t_n + jh, y_{n+j}) + \beta'_k f(t_{n+k}, y_{n+k}) \quad ; y_{n+k} \notin Y_n \end{aligned}$$

sustituyendo ahora (4.11) en f_{n+k} se obtiene (4.10).

Los siguientes lemas expresan los conceptos de orden y estabilidad en esta nueva notación.

Lema 4.1. *Sea $y(t)$ la solución del problema (4.1). Para $n = 0, 1, 2, \dots$ definimos el vector \hat{Y}_{n+1} como la solución numérica de un paso tal que*

$$\hat{Y}_{n+1} = (A \otimes I)\hat{Y}_n + h\Phi(t_n, Y_n, h)$$

con valores iniciales exactos

$$Y(t_n)^\top = (y(t_{n+k-1})^\top, y(t_{n+k-2})^\top, \dots, y(t_n)^\top)$$

i) Si el método (2.1) es de orden 1 y satisface (4.2), entonces existe $h_0 \geq 0$ tal que para $h \leq h_0$,

$$\| Y(t_{n+1}) - \hat{Y}_{n+1} \| \leq hw(h), \quad 0 \leq n \leq \frac{\hat{t}}{h} - k$$

donde $w(h) \rightarrow 0$, cuando $h \rightarrow 0$.

ii) Si el método multipaso (2.1) es de orden p y f es suficientemente diferenciable,, entonces existe una constante M tal que para un h suficientemente pequeño

$$\| Y(t_{n+1}) - \hat{Y}_{n+1} \| \leq Mh^{p+1}, \quad 0 \leq n \leq \frac{\hat{t}}{h} - k$$

Demostración trivial mediante las definiciones de error local y orden de consistencia ya vistas en los capítulos anteriores.

Lema 4.2. *Suponiendo que el método multipaso (2.1) es 0-estable. Entonces existe una norma vectorial (en \mathbb{R}^{mk}) tal que la matriz A de (4.9) satisface*

$$\| A \otimes I \| \leq 1$$

en la norma matricial subordinada.

Demostración. Si el método multipaso (2.1) es 0-estable \Rightarrow los autovalores de módulo 1 son simples y solo pueden tener multiplicidad algebraica mayor que 1 ($\mu_a(\lambda) > 1$) cuando $|\lambda| < 1$.

El polinomio característico de la matriz A es el polinomio característico del método (2.1), siendo $\varrho(\zeta) = \sum_{j=0}^k \alpha_j \zeta^j$ el que nos interesa para la 0-estabilidad.

Por la forma de Jordan $\exists T \in \mathbb{C}^{k,k}$ tal que

$$T^{-1}AT = J = \text{blockdiag}(J_1, J_2, \dots, J_r)$$

es decir

$$T^{-1}AT = \begin{pmatrix} (J_1) & & & \\ & (J_2) & & \\ & & \ddots & \\ & & & (J_r) \end{pmatrix}, \quad \text{donde } J_l = \begin{pmatrix} \lambda_l & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_l \end{pmatrix}, \quad \dim J_l \geq 1 \quad (4.12)$$

En la matriz (4.12) se pueden ordenar los bloques de Jordan, es decir, los primeros bloques pueden elegirse de modo que tengan los autovalores simples (en particular los de módulo 1) y después colocamos el resto de bloques de dimensiones superiores.

Considerando ahora la matriz

$$J_l = \begin{pmatrix} \lambda_l & 1 & \cdots & 0 \\ 0 & \lambda_l & & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \cdots & & \lambda_l \end{pmatrix}, \quad \dim J_l = m$$

la premultiplicamos por

$$D_l^{-1}(\epsilon) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ & \epsilon^{-1} & & \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \epsilon^{m-1} \end{pmatrix}$$

y lo postmultiplicamos por $D_l(\epsilon)$, se tiene que

$$D_l^{-1}(\epsilon) \cdot J_l \cdot D_l(\epsilon) = \begin{pmatrix} \lambda_l & \epsilon & \cdots & 0 \\ 0 & \lambda_l & \ddots & \vdots \\ \vdots & & \ddots & \epsilon \\ 0 & \cdots & & \lambda_l \end{pmatrix} = B_l(\epsilon)$$

Tomando $\|\cdot\|_\infty$, se observa que $\|B_l(\epsilon)\|_\infty = |\lambda_l| + \epsilon \Rightarrow \|B(\epsilon)\| \leq 1$ si el método (2.1) es 0-estable. Consideramos un $\epsilon > 0$ tal que $\epsilon + \text{Max}|\lambda| \leq 1$ donde $|\lambda_l| < 1$ y λ_l es autovalor de A. Por tanto,

$$\text{Si } T^{-1}AT = J \quad \rightarrow \quad B^{-1}(\epsilon)T^{-1}ATB(\epsilon) = J_\epsilon$$

donde

$$B(\epsilon) = \begin{pmatrix} B_1(\epsilon) & & \\ & \ddots & \\ & & B_r(\epsilon) \end{pmatrix}$$

Llamamos $T_\epsilon = TB(\epsilon)$, $T_\epsilon^{-1} = B^{-1}(\epsilon)T^{-1}$,

$$T_\epsilon^{-1}AT_\epsilon = J_\epsilon \tag{4.13}$$

donde

$$J_\epsilon = \begin{pmatrix} J_1(\epsilon) & & \\ & \ddots & \\ & & J_r(\epsilon) \end{pmatrix}, \quad \text{para } J_l(\epsilon) = \begin{pmatrix} \lambda_l & \epsilon & & \\ & \ddots & \ddots & \epsilon \\ & & & \lambda_l \end{pmatrix}$$

Deducimos así que $\|J_\epsilon\|_\infty = 1$, por la 0-estabilidad. (Recuérdese que las normas son equivalentes en espacios de dimensión finita).

□

4.2. Demostración de la convergencia

Ahora se puede establecer el teorema de convergencia para métodos multipaso.

Teorema 4.2. *Si el método multipaso (2.1) es 0-estable y de orden 1 entonces es convergente. Si el método (2.1) es 0-estable y de orden p entonces es convergente y de orden p .*

Demostración. Probaremos solo la segunda parte del Teorema, pues la primera consiste en repetir prácticamente los mismos argumentos que veremos a continuación.

Supongamos sin pérdida de generalidad que $f(t, y)$ está definida para todo $y \in \mathbb{R}^m$, $[t_0, t^*]$ y satisface la condición de Lipschitz

$$\| f(t, y) - f(t, z) \| \leq L \| y - z \|,$$

para $z \in \mathbb{R}^m$. Además, $f \in C^p([t_0, t^*] \times \mathbb{R}^m)$. Veamos que la función Ψ esta definida de la forma

$$\Psi \equiv \Psi(t_n, Y_n, h) = \sum_{j=0}^{k-1} \beta'_j f(t_n + jh, y_{n+j}) + h \beta'_k f \left(t_n + kh, h\Psi - \sum_{j=0}^{k-1} \alpha'_j y_{n+j} \right), \quad \Psi \in \mathbb{R}^{mk}$$

es Lipschitz respecto al argumento central $Y^\top = (y_{n+k-1}^\top, \dots, y_n^\top)$. Llamando $\bar{Y}_n^\top = (\bar{y}_{n+k-1}^\top, \dots, \bar{y}_n^\top)$ y $\bar{\Psi} = \Psi(t_n, \bar{Y}_n, h)$ comprobamos que

$$\| \Psi(t_n, Y_n, h) - \Psi(t_n, \bar{Y}_n, h) \| \leq L^* \| Y - \bar{Y} \|, \quad \text{para } \forall |h| \leq h_0$$

siendo L^* la constante de Lipschitz. Estudiando la primera parte de la desigualdad se tiene

$$\begin{aligned} \Psi - \bar{\Psi} &= \sum_{j=0}^{k-1} \beta'_j (f(t_{n+j}, y_{n+j}) - f(t_{n+j}, \bar{y}_{n+j})) + \\ &\quad \beta'_k [f(t_{n+k}, h\Psi - \sum_{j=0}^{k-1} \beta'_j y_{n+j}) - f(t_{n+k}, h\bar{\Psi} - \sum_{j=0}^{k-1} \beta'_j \bar{y}_{n+j})], \end{aligned}$$

tomando $\| \cdot \|$, y por desigualdad triangular tenemos que

$$\begin{aligned} \| \Psi - \bar{\Psi} \| &\leq \sum_{j=0}^{k-1} |\beta'_j| L \| y_{n+j} - \bar{y}_{n+j} \| + |\beta'_k| L \| h(\Psi - \bar{\Psi}) - \sum_{j=0}^{k-1} \beta'_j (y_{n+j} - \bar{y}_{n+j}) \| \\ \| \Psi - \bar{\Psi} \| &\leq \sum_{j=0}^{k-1} |\beta'_j| \| y_{n+j} - \bar{y}_{n+j} \| + hL |\beta'_k| \| \Psi - \bar{\Psi} \| + L |\beta'_k| \sum_{j=0}^{k-1} |\beta'_j| \| y_{n+j} - \bar{y}_{n+j} \| \\ (1 - hL |\beta'_k|) \| \Psi - \bar{\Psi} \| &\leq L \sum_{j=0}^{k-1} (|\beta'_j| + |\beta'_k| |\beta'_j|) \| y_{n+j} - \bar{y}_{n+j} \|, \end{aligned}$$

Escribiendo $\gamma_j = |\beta'_j| + |\beta'_k| \cdot |\beta'_j|$, y tomando un $h_0 \geq |h|$ tal que $1 - h_0 L |\beta'_k| < 1$ se obtiene la siguiente desigualdad

$$\| \Psi - \bar{\Psi} \| \leq \frac{L}{(1 - h_0 L |\beta'_k|)} \sum_{j=0}^{k-1} \gamma_j \| y_{n+j} - \bar{y}_{n+j} \|.$$

Definimos ahora una nueva norma;

$$|||Y||| = \max_{0 \leq j \leq k-1} \|y_{n+j}\|.$$

Por tanto,

$$|||\Psi - \hat{\Psi}||| \leq C \cdot |||Y - \hat{Y}|||,$$

donde

$$C = \frac{L}{1 - h_o L |\beta'_k|} \sum_{j=0}^{k-1} \gamma_j. \quad \text{para } \forall |h| \leq h_0$$

El resto de la demostración procede ahora de la misma manera que para los métodos de un solo paso. Teniendo en cuenta los procesos

$$\begin{aligned} Y_{n+1} &= (A \otimes I)Y_n + h\phi(t_n, Y_n, h) \\ \hat{Y}_{n+1} &= (A \otimes I)\hat{Y}_n + h\phi(t_n, \hat{Y}_n, h) + \ell_n, \quad \text{siendo } \ell_n \text{ el error local} \end{aligned}$$

donde \hat{Y}_n es la solución exacta. Definimos el error global $\epsilon_n := \hat{Y}_n - Y_n$, entonces:

$$\epsilon_{n+1} = (A \otimes I)\epsilon_n + h(\phi(t_n, \hat{Y}_n, h) - \phi(t_n, Y_n, h)) + \ell_n$$

tomando norma ahora, se tiene

$$\begin{aligned} \|\epsilon_{n+1}\| &\leq \|A \otimes I\| \cdot \|\epsilon_n\| + hL^* \|\epsilon_n\| + \|\ell_n\| \leq (1 + hL^*) \|\epsilon_n\| + \|\ell_n\| \\ &\leq e^{hL^*} \|\epsilon_n\| + \|\ell_n\| \leq e^{hL^*} (e^{hL^*} \|\epsilon_{n-1}\| + \|\ell_{n-1}\|) + \|\ell_n\| \leq \dots \\ &\leq e^{(n+1)hL^*} \|\epsilon_0\| + (\|\ell_n\| + e^{hL^*} \|\ell_{n-1}\| + e^{2hL^*} \|\ell_{n-2}\| + \dots + e^{nhL^*} \|\ell_0\|) \\ &= e^{L^*t_{n+1}} \|\epsilon_0\| + \ell^* (1 + e^{hL^*} + e^{2hL^*} + \dots + e^{nhL^*}), \end{aligned}$$

donde

$$\ell^* = \max_{0 \leq j \leq n} \|\ell_j\| \leq Ch^{p+1}$$

Luego, se tiene

$$\begin{aligned} \|\epsilon_{n+1}\| &\leq e^{L^*t_{n+1}} \|\epsilon_0\| + \ell^* \frac{e^{nhL^*} e^{hL^*} - 1}{e^{hL^*} - 1} \\ \|\epsilon_{n+1}\| &\leq e^{L^*t_{n+1}} \|\epsilon_0\| + \ell^* \frac{e^{t_{n+1}L^*} - 1}{e^{hL^*} - 1}, \end{aligned}$$

y como $e^{hL^*} > 1 + hL^*$ para $h > 0$, podemos escribir

$$\|\epsilon_{n+1}\| \leq e^{L^*t_{n+1}} \|\epsilon_0\| + \ell^* \frac{e^{t_{n+1}L^*} - 1}{hL^*} \leq e^{L^*t_{n+1}} \|\epsilon_0\| + Ch^p \frac{e^{t_{n+1}L^*} - 1}{L^*}.$$

Por tanto, si $\|\epsilon_0\| = \mathcal{O}(h^p) \Rightarrow \|\epsilon_n\| \leq \mathcal{O}(h^p)$

□

Teoría Estabilidad Lineal

En este capítulo estudiamos que restricciones se imponen al tamaño de paso “ h ” para realizar integraciones estables (obtener soluciones acotadas) en sistemas lineales $y'(t) = Jy(t)$ para intervalos largos (en realidad semi-infinitos, $t \in I = [t_0, \infty)$) en caso del que el sistema sea estable Lyapunov, $\text{Re } \lambda(J) < 0$.

Aplicando un MLM (ρ, σ) (métodos (2.1)) de k -pasos al sistema autónomo con coeficientes constantes $y' = Jy$, se obtiene

$$\sum_{j=0}^k \alpha_j y_{n+j} = hJ \sum_{j=0}^k \beta_j y_{n+j} \quad (5.1)$$

y suponiendo ahora que J diagonalizable, es decir $J = SDS^{-1}$ con D matriz diagonal, podemos escribir (5.1) en la forma

$$\sum_{j=0}^k \alpha_j y_{n+j} = hS \sum_{j=0}^k \beta_j DS^{-1} y_{n+j}, \quad (5.2)$$

observe que en este caso S es una matriz regular cuyos vectores columna son vectores propios de J y D es la matriz diagonal formada por los valores propios de J , ordenados acorde a los autovectores correspondientes (columnas de S). Expresando el vector $\vec{y}_n = (y_{n,i})_{i=1}^m$ como combinación lineal de los autovectores, $x_n = S^{-1}y_n \Leftrightarrow Sx_n = y_n$, y premultiplicando por S^{-1} en (5.2) obtenemos las siguientes expresiones

$$\begin{aligned} \sum_{j=0}^k x_{n+j} &= h \sum_{j=0}^k \beta_j D x_{n+j} \iff \sum_{j=0}^k x_{n+j,i} = h \sum_{j=0}^k \beta_j \lambda_i x_{n+j,i} \quad i=1,2,\dots,m \\ &\iff \sum_{j=0}^k (\alpha_j - \mu_i \beta_j) x_{n+j,i} = 0, \quad \mu_i = h\lambda_i \end{aligned}$$

La última expresión es equivalente a que apliquemos el al M.L.M (ϱ, σ) al problema (test de Dahlquist)

$$y'(t) = \lambda, \quad \lambda \in \mathbb{C},$$

obteniendo de este modo la ecuación siguiente

$$\sum_{j=0}^k (\alpha_j - \mu\beta_j)y_{n+j} = 0 \quad (5.3)$$

la cual utilizaremos para el estudio de la estabilidad lineal.

5.1. La Región de Estabilidad

Analizando el polinomio característico de (5.3) se obtiene

$$\pi(\zeta; \mu) = \varrho(\zeta) - \mu\sigma(\zeta) = \sum_{j=0}^k (\alpha_j - \mu\beta_j)\zeta^j = 0 \quad (5.4)$$

que depende del parámetro complejo μ , y denotaremos por $\zeta_i(\mu)$, $i = 1, \dots, k$ a las k raíces del polinomio. Las soluciones de (5.3) están uniformemente acotadas si y sólo si $\pi(\zeta; \mu)$ verifica la condición de las raíces (C.R).

Así pues, se denomina **dominio de estabilidad absoluta** S del método lineal (ϱ, σ) al conjunto

$$S = \{\mu \in \mathbb{C} \mid \pi(\zeta; \mu) \text{ verifica la C.R y tiene grado exacto } k\}$$

Definición 5.1. *El MLM (ϱ, σ) presenta **A-estabilidad** si $\mathbb{C}^- \subset S$.*

A veces es desesable considerar S como un subconjunto del plano complejo ampliado (la esfera de Riemann \mathbb{C}_∞), en este caso definiremos el dominio de estabilidad por

$$S = \{\mu \in \mathbb{C} \cup \{\infty\} \mid \pi(\zeta; \mu) \text{ verifica la C.R y tiene grado exacto } k\}$$

El punto del infinito pertenecerá a S si y sólo si el polinomio σ verifica la C.R, es decir, $\pi(\zeta; \mu) = 0$. Esto equivale a

$$\frac{1}{\mu}\varrho(\zeta) - \sigma(\zeta) = 0$$

y para $\mu = \infty$ las raíces de $\pi(\zeta; \mu)$ son las de $\sigma(\zeta)$.

Para el caso de $\mu = 0$, $\pi(\zeta; 0) = \varrho(\zeta)$ y la condición de estabilidad se reduce al capítulo III de la 0-estabilidad. Esta claro que para todo método 0-estable $0 \in S$.

Teorema 5.1. *Todas las soluciones del M.L.M (p, σ) están acotada para $y' = Jy$ con J diagonalizable si y sólo si $h\lambda \in S, \forall \lambda \in \sigma(J)$*

Ilustraremos el dominio de estabilidad usando como ejemplo particular el método explícito de Adams de orden 4

$$y_{n+4} = y_{n+3} + \frac{h}{24}(55f_{n+3} - 59f_{n+2} + 37f_{n+1} - 9f_n),$$

siendo su polinomio característico $\pi(\zeta; \mu) = 0$ de la forma

$$\zeta^4 - (1 + \frac{55}{24}\mu)\zeta^3 + \frac{59}{24}\mu\zeta^2 - \frac{37}{24}\mu\zeta + \frac{9}{24}\mu = 0. \quad (5.5)$$

Observación: Debe recordarse que todas las funciones meromorfas de variable compleja $\zeta = f(z)$ son conformes (conservan los ángulos y la orientación) en un dominio D con tal que $f'(z) \neq 0, \forall z \in D$.

Sea $\zeta_1 \equiv \zeta_1(\mu)$ raíz del polinomio (5.7) tal que $\zeta_1(0) = 1$. Aplicando el desarrollo de Taylor se tiene que $\zeta_1(\mu) = 1 + \zeta_1'(0)\mu + O(\mu^2)$, por tanto, derivando $\pi(\zeta_1; \mu)$ con respecto de μ se tiene

$$\varrho'(\zeta_1)\zeta_1'(\mu) - \sigma(\zeta_1) - \mu\sigma'(\zeta_1)\zeta_1'(\mu) = 0,$$

y sustituyendo en $\mu = 0$

$$\varrho'(1)\zeta_1'(1) - \sigma(1) = 0 \Rightarrow \zeta_1'(0) = \frac{\sigma(1)}{\varrho'(1)} = 1.$$

Teniendo en cuenta las condiciones de consistencia ($\varrho(1) = 0, \varrho'(1) \neq 0, \sigma(1) = \varrho'(1)$), llegamos a que

$$\zeta_1(\mu) = 1 + \mu + O(\mu^2)$$

Esta expresión nos dice que si μ se desplaza localmente al semiplano complejo izquierdo la raíz $|\zeta_1(\mu)| < 1$, es decir, $\zeta_1(\mu)$ se mueve dentro del disco unidad.

The Boundary Locus. La localización de la frontera para el dominio de estabilidad S consiste en encontrar aquellos valores de μ para que alguna raíz $\zeta_i(\mu)$ tenga módulo 1.

Luego, si $|\zeta| = 1 \Rightarrow \zeta = e^{\theta i} = \cos \theta + i \sin \theta$ para algún $\theta \in (-\pi, \pi]$, se tiene que

$$\varrho(e^{\theta i}) - \mu\sigma(e^{\theta i}) = 0 \Rightarrow \mu = \frac{\varrho(e^{\theta i})}{\sigma(e^{\theta i})}$$

Por tanto, podemos definir el Boundary Locus como el conjunto

$$BL = \{\mu \in \mathbb{C} \mid \mu = \frac{\varrho(e^{i\theta})}{\sigma(e^{i\theta})}, \theta \in (-\pi, \pi]\} \quad (5.6)$$

A continuación damos el siguiente resultado que relaciona el BL con la frontera del dominio de estabilidad ∂S . La demostración usa conceptos topológicos básicos y la omitiremos por brevedad.

Teorema 5.2. *Para un MLM (ϱ, σ) , el Boundary Locus (5.6) contiene a la frontera del Dominio de Estabilidad lineal ∂S . Además el BL divide al plano complejo en un número finito de componentes simplemente conexas abiertas S_i , y se tiene que si un punto $z \in S_i \cap \text{Int}(S)$, entonces $S_i \subset \text{Int}(S)$. Y si $z \in S_i$ verifica $z \notin \text{Int}(S)$ entonces $S_i \cap \text{Int}(S) = \emptyset$.*

5.2. Ejemplos

La idea para obtener el BL de un método multipaso es sustituir y_n por ζ^n , siendo $\zeta = e^{i\theta}$, y despejar μ .

Considerando el método

$$\sum_{j=0}^k \alpha_j y_{n+j} - h \sum_{j=0}^k \beta_j f_{n+j} = 0$$

aplicado al test de Dahlquist $y' = \lambda y$, se obtiene

$$\sum_{j=0}^k \alpha_j y_{n+j} - \mu \sum_{j=0}^k \beta_j y_{n+j} = 0, \quad \mu = h\lambda$$

sustituyendo ahora por $y_n = \zeta^n$ y dividiendo por ζ^n

$$\sum_{j=0}^k \alpha_j \zeta^j = \mu \sum_{j=0}^k \beta_j \zeta^j \Rightarrow \mu = \frac{\sum_{j=0}^k \alpha_j \zeta^j}{\sum_{j=0}^k \beta_j \zeta^j} = \frac{\varrho(\zeta)}{\sigma(\zeta)}$$

5.2.1. Métodos de Adams

Los métodos explícitos de Adams (veáse sección 1.1) aplicados al problema $y' = \lambda y$ nos dan

$$y_{n+1} = y_n + \mu \sum_{j=0}^{k-1} \gamma_j \nabla^j y_n, \quad \text{con } \gamma_j = (-1)^j \int_0^1 \binom{-s}{j} ds.$$

Sustituyendo ahora por $y_n = \zeta^n$ y dividiendo por ζ^n

$$\zeta - 1 = \mu \sum_{j=0}^{k-1} \gamma_j (1 - \zeta^{-1})^j \Rightarrow \mu = \frac{\zeta - 1}{\sum_{j=0}^{k-1} \gamma_j (1 - \zeta^{-1})^j}$$

obteniéndose así el BL , para el anterior valor de μ , de la forma

$$BL = \{\mu(\theta) \in \mathbb{C} : \zeta = e^{\theta i}, \theta \in [0, 2\pi)\}$$

Para el caso de los métodos de Adams explícitos y $k = 1$, se obtiene que $BL_1 = \{e^{\theta i} : \theta \in [0, 2\pi)\}$ (círculo del método de Euler). Se puede observar en la figura 5.1 como el dominio de estabilidad decrece a medida que se incrementa el orden.

De forma muy similar se obtien el BL para los métodos implícitos de Adams (sección 1.2)

$$y_{n+1} = y_n + \mu \sum_{j=0}^{k-1} \gamma_j^* \nabla^j y_{n+1}, \quad \text{con } \gamma_j^* = (-1)^j \int_0^1 \binom{-s+1}{j} ds.$$

Sustituyendo $y_n = \zeta^n$ y dividiendo por ζ^{n+1}

$$1 - \zeta^{-1} = \mu \sum_{j=0}^{k-1} \gamma_j^* (1 - \zeta^{-1})^j \Rightarrow \mu = \frac{1 - \zeta^{-1}}{\sum_{j=0}^{k-1} \gamma_j^* (1 - \zeta^{-1})^j}.$$

Para $k = 1$, se obtiene la regla trapezoidal la cual es A-estable. Para $k = 2, 3, 4, 5$, los dominios de estabilidad son mucho más grandes que los de los métodos explícitos (veáse figura 5.2).

En la figura 5.2, los dominios de estabilidad de los métodos de Adams incluyen las componentes conexas del BL que tienen para real negativa y que están acotadas.

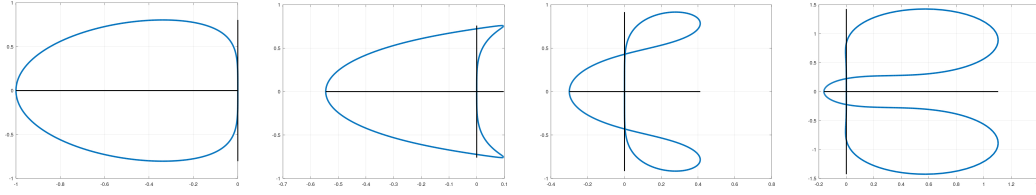


Figura 5.1: BL de Adams Bashforth de orden $k = 2, 3, 4, 5$

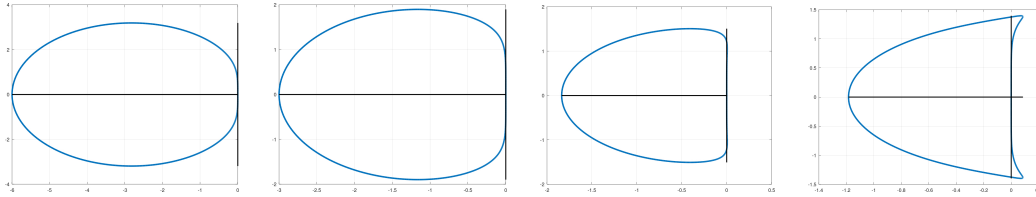


Figura 5.2: BL de Adams Moulton de orden $k = 2, 3, 4, 5$

5.2.2. Métodos de Nyström

Utilizaremos el ejemplo particular de Nyström explícito para $k = 1$, la regla del punto medio:

$$y_{n+1} = y_{n-1} + 2hf_n$$

obteniendo la ecuación $\pi(\zeta; \mu)$

$$\zeta^2 - 1 - 2\mu\zeta = 0 \Rightarrow \mu = \frac{\zeta - \zeta^{-1}}{2}$$

sustituyendo por $\zeta = e^{\theta i}$ se obtiene que $\mu = i \text{ sen } \theta$.

La curva del BL se mueve hacia arriba y abajo del eje imaginario entre $\pm i$. Todos los autovalores en el interior del semiplano negativo conducen a inestabilidades computacionales. Esto es causado por la segunda raíz de $\varrho(\zeta)$, la cual se desplaza fuera del círculo unidad cuando μ se mueve hacia la izquierda. Por ejemplo, si tomamos $\mu = -1$

$$\zeta^2 - 1 + 2\zeta = 0 \Rightarrow \begin{cases} \zeta_1 = -1 + \sqrt{2} \\ \zeta_2 = -1 - \sqrt{2} \Rightarrow |\zeta_2| > 1 \end{cases}$$

Este fenómeno es conocido como *estabilidad débil de la regla del punto medio* y fue el punto de partida de los estudios de estabilidad lineal emprendidos por Dahlquist desde mediados de los años cincuenta del siglo XX.

5.2.3. Métodos BDF

The backward differentiation formulas (véase sección 1.7) presenta su *BL* de la forma

$$\mu = \sum_{j=0}^k \frac{1}{j} (1 - \zeta^{-1})^j, \quad \zeta = e^{\theta i} : \sigma \in [0, 2\pi).$$

Se observa en la figura 5.3 que para $k = 3, 4, 5$ los métodos pierden cada vez más estabilidad en una parte del eje imaginario. En la figura 5.3, las regiones de estabilidad de los BDF son las que quedan fuera de las regiones acotadas. Solo presentan A-estabilidad (L-estabilidad) los casos $k = 1, 2$.

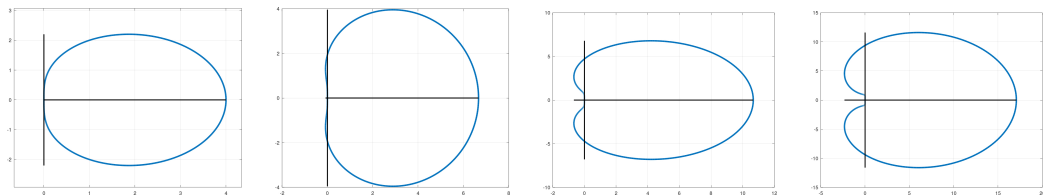


Figura 5.3: BL de Adams Moulton de orden $k = 2, 3, 4, 5$

5.3. La Segunda Barrera Dahlquist

Desde los años veinte hasta los años sesenta del siglo XX, muchos analistas numéricos buscaron en vano la obtención de métodos lineales multipaso A-estables para integrar sistemas lineales estables Lyapunov en intervalos largos. Fué Dahlquist en 1963, quien probó por primera vez el teorema conocido como segunda barrera de Dahlquist. Antes de probar ese resultado necesitamos el siguiente teorema previo.

Teorema 5.3. *Sea un método lineal multipaso (ϱ, σ) irreducible, entonces*

$$A\text{-estable} \Leftrightarrow \operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} \right) > 0 \quad \forall |\zeta| > 1 \tag{5.7}$$

Demostración. Veamos primero que $A\text{-estabilidad} \Rightarrow \operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} \right) > 0, \forall |\zeta| > 1$. A-estabilidad implica que $\mathbb{C}^- \subset S(\text{dominio de estabilidad})$, es decir, para $\operatorname{Re} \mu \leq 0$ dado, las raíces del polinomio $\varrho(\zeta) - \mu\sigma(\zeta) = 0$ cumplen la condición de las raíces. Luego,

$$\operatorname{Re} \mu \leq 0 \Leftrightarrow \left(\operatorname{Re} \frac{\varrho(\zeta)}{\sigma(\zeta)} \leq 0 \Rightarrow |\zeta| \leq 1 \right),$$

esto es equivalente a

$$|\zeta| > 1 \Rightarrow \operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} \right) > 0$$

Probemos ahora que $\operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} \right) > 0, \forall |\zeta| > 1 \Rightarrow A\text{-estabilidad}$.

Fijamos μ_0 con $\operatorname{Re} \mu_0 \leq 0$ y sea ζ_0 una raíz de $\varrho(\zeta_0) - \mu_0 \sigma(\zeta_0) = 0$. Por tanto, nos bastará con probar que ζ_0 verifica la condición de las raíces. Se tiene que $\sigma(\zeta_0) \neq 0$, pues en caso contrario $\sigma(\zeta_0) = 0 \Rightarrow \varrho(\zeta_0) = 0 \Rightarrow \{\varrho, \sigma\}$ tienen una raíz en común, llegando así a la contradicción de que el método sería reducible. Por lo tanto, $\mu_0 = \frac{\varrho(\zeta_0)}{\sigma(\zeta_0)}$

$$\mu_0 = \frac{\varrho(\zeta_0)}{\sigma(\zeta_0)}$$

Si $|\zeta_0| > 1 \Rightarrow \operatorname{Re} \left(\frac{\varrho(\zeta_0)}{\sigma(\zeta_0)} \right) > 0$ por hipótesis, llegando nuevamente a una contradicción pues $\operatorname{Re} \mu_0 \leq 0$. Deducimos pues que, $|\zeta_0| \leq 1$. Falta por demostrar que ζ_0 es una raíz simple si $|\zeta_0| = 1$.

Por un argumento de continuidad de (5.7) se sigue que $|\zeta_0| = 1$ y $\operatorname{Re} \mu_0 < 0$ es absurdo. Por tanto, $|\zeta_0| = 1$ implica que $\operatorname{Re} \mu_0 = 0$. Desarrollando en un entorno de ζ_0 , se obtiene que

$$\frac{\varrho(\zeta)}{\sigma(\zeta)} - \mu_0 = C_1(\zeta - \zeta_0) + C_2(\zeta - \zeta_0)^2 + \dots$$

Si $C_1 = 0$, se tendría que

$$\operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} - \mu_0 \right) = \operatorname{Re} (C_2(\zeta - \zeta_0)^2 + \dots) = \operatorname{Re} \left(\frac{\varrho(\zeta)}{\sigma(\zeta)} \right),$$

llegando a una contradicción, pues $\operatorname{Re} (\zeta - \zeta_0)^2 < 0$ para algún $|\zeta| > 1$ con ζ tan próximo como se quiera a ζ_0 (Observe que $(\zeta - \zeta_0)^2$ gira da una vuelta completa cuando $(\zeta - \zeta_0)$ gira media vuelta). Luego, $C_1 \neq 0$ para

$$\frac{\varrho(\zeta)}{\sigma(\zeta)} - \mu_0 = (\zeta - \zeta_0)(C_1 + C_2(\zeta - \zeta_0) + \dots)$$

apreciándose así, que ζ_0 es raíz simple. \square

Teorema 5.4 (Segunda Barrera Dahlquist). *Sea un MLM (ϱ, σ) A-estable e irreducible, entonces su orden de consistencia*

$$i) p \leq 2$$

$$ii) p = 2 \Rightarrow \text{la constante de error satisface que } C \leq -\frac{1}{12}$$

siendo $-\frac{1}{12} = C_T$ la constante de error de la regla trapezoidal.

Demostración. Orden de consistencia p implica que

$$\varrho(e^h) - h\sigma(e^h) = C_{p+1}h^{p+1} + \mathcal{O}(h^{p+2}) \quad h \rightarrow 0 \quad (5.8)$$

Sabiendo que $\varrho(e^h) = \varrho\left(1 + h + \frac{h^2}{2!} + \dots\right)$ y llamando ahora $b = 1 + h + \frac{h^2}{2!} + \dots$;

$$\varrho(e^h) = \varrho(1) + \varrho'(1)b + \frac{1}{2!}\varrho''(1)b^2 + \dots = 0 + \sigma(1)h + \mathcal{O}(h^2), \quad h \rightarrow 0, \quad \sigma(1) \neq 0$$

Dividiendo la expresión (5.8) entre $h\varrho(e^h)$, se obtiene

$$\frac{1}{h} - \frac{\sigma(e^h)}{\varrho(e^h)} = Ch^{p-1} + \mathcal{O}(h^p), \quad h \rightarrow 0, \quad C = \frac{C_{p+1}}{\sigma(1)} \text{ constante de error}$$

Realizando el cambio de variable $e^h = \zeta \Leftrightarrow h = \ln \zeta$ ($h \rightarrow 0 \Leftrightarrow \zeta \rightarrow 1$), y teniendo en cuenta que $\ln \zeta = (\zeta - 1) + \mathcal{O}(\zeta - 1)^2$, $\zeta \rightarrow 1$, se tiene que

$$\frac{1}{\ln \zeta} - \frac{\sigma(\zeta)}{\varrho(\zeta)} = C(\zeta - 1)^{p-1} + \mathcal{O}(\zeta - 1)^p, \quad \zeta \rightarrow 1 \quad (5.9)$$

Para la regla trapezoidal (A-estable), se tiene entonces que

$$\frac{1}{\ln \zeta} - \frac{\sigma_T(\zeta)}{\varrho_T(\zeta)} = -\frac{1}{12}(\zeta - 1) + \mathcal{O}(\zeta - 1)^2, \quad \zeta \rightarrow 1 \quad (5.10)$$

Restando (5.10) y (5.9), deducimos que

$$d(\zeta) := \frac{\sigma(\zeta)}{\varrho(\zeta)} - \frac{\sigma_T(\zeta)}{\varrho_T(\zeta)} = \left(-C - \frac{1}{12}\right)(\zeta - 1) + \mathcal{O}(\zeta - 1)^2. \quad (5.11)$$

Luego, si el método tiene orden $p = 2$ entonces C es la constante de error del método, pero si $p > 2$ se tiene que $C = 0$. Teniendo en cuenta estas afirmaciones junto con $\operatorname{Re}\left(\frac{\sigma_T(\zeta)}{\varrho_T(\zeta)}\right) = 0$ si $|\zeta| = 1$ y que A-estabilidad implica $\operatorname{Re}\left(\frac{\sigma(\zeta)}{\varrho(\zeta)}\right) > 0$ si $|\zeta| > 1$, concluimos que

$$\lim_{\zeta \rightarrow \zeta_0} d(\zeta) \geq 0, \quad |\zeta| > 1, \quad |\zeta_0| = 1.$$

Por tanto, $d(\zeta)$ es una función analítica en $|\zeta| > 1$ y en $\zeta = \infty$. Por tanto, $\operatorname{Re} d(\zeta)$ alcanza su máximo y su mínimo en la frontera $|\zeta| = 1$. Entonces,

$$\text{si } \operatorname{Re} d(\zeta) \geq 0 \text{ en } |\zeta| = 1 \Rightarrow \operatorname{Re} d(\zeta) \geq 0 \text{ en } |\zeta| > 1$$

Eligiendo ahora $\zeta = 1 + \epsilon$, con $\operatorname{Re} \epsilon > 0$ y $|\epsilon| \lll 1$, y teniendo en cuenta (5.11) tenemos que

$$0 \leq \operatorname{Re} d(\zeta) = \left(-C - \frac{1}{12}\right)\epsilon + \mathcal{O}(\epsilon^2), \quad (5.12)$$

pero esto es imposible si $C = 0$, por lo que $C \neq 0 \Rightarrow p = 2$, demostrando así la primera parte del teorema. Para que se cumpla (5.12) se debe dar $\operatorname{Re}\left(-C - \frac{1}{12}\right) \geq 0 \Rightarrow C \leq -\frac{1}{12}$, concluyendo así la demostración. \square

A continuación ilustramos la teoría de la estabilidad lineal usando el problema dado en el capítulo 1 (1.6)-(1.8), con valor de $N = 19$ lo que implica $h = \Delta x = 1/20$. En este caso todos los autovalores de la matriz D son negativos y están distribuidos en el intervalo $(-800, 0)$. Los métodos BDF_k con $k \leq 6$ no tienen problema alguno para integrar este problema pues sus regiones de estabilidad contienen al semieje real negativo. Esto puede apreciarse en la Tabla 5.1, donde hemos usado el BDF_4 , el cual da buenos resultados incluso para valores Δt grandes. Sin embargo, los métodos implícitos de Adams (los explícitos tienen mayores problemas de estabilidad) deben verificar, para realizar integraciones estables, que $-800\Delta t \in S_k$, siendo S_k el dominio del Adams implícito de k pasos. Para ilustrar estos hechos hemos usado el método implícito de Adams con $k = 3$, el cual nos lleva a una restricción en el tamaño de paso de

$$\Delta t < \tau := \frac{3}{800} = 3.75 \cdot 10^{-3}.$$

En la Tabla 5.2 puede verse que si $\Delta t > \tau$ las integraciones son inestables, mientras que cuando $\Delta t \leq \tau$ estas no presentan ningún problema de convergencia.

BDF_4	$\Delta t = 2^0$	$\Delta t = 2^{-1}$	$\Delta t = 2^{-2}$	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$
	4.5993e-02	1.3438e-04	7.5314e-06	4.2889e-07	2.4002e-08	1.4233e-09	8.6660e-11	5.3457e-12
p		8.4189e+00	4.1573e+00	4.1342e+00	4.1594e+00	4.0758e+00	4.0377e+00	4.0189e+00

Tabla 5.1: Errores en el punto final $t^* = 2$ con el método BDF de $k = 4$ pasos (orden 4), usando los Δt indicados y valores de arranque adicionales exactos. El orden del método puede leerse como $p \approx \log_2(\epsilon(\Delta t)/\epsilon(2^{-1}\Delta t))$. Hemos usado en este caso $N = 19$.

AM_3	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	$\Delta t = 2^{-9}$	$\Delta t = 2^{-10}$	$\Delta t = 2^{-11}$	$\Delta t = 2^{-12}$	$\Delta t = 2^{-13}$
	3.7571e+25	1.8701e+51	3.3731e+68	2.9201e-07	1.1952e-15	2.2714e-15	2.4964e-15	1.0355e-14

Tabla 5.2: Errores en el punto final $t^* = 2$ con el método Adams Moulton de $k = 3$ pasos (orden 4), usando los Δt indicados y valores de arranque adicionales exactos. Hemos usado en este caso $N = 19$. Vemos que aparecen inestabilidades para valores de $\Delta t > 3/800$.

Conclusiones

“Ninguna investigación humana puede ser llamada ciencia real si no puede demostrarse matemáticamente”. Leonardo da Vinci

Este trabajo de fin de grado se apoya en el entendimiento y manejo de las siguientes áreas:

- Los Métodos Numéricos son el pilar de conocimiento necesario para poder realizar y comprender la base de este trabajo, al igual que el uso de programas tales como Matlab para poder computar ejemplos y ver sus comportamientos.
- Todas las asignaturas de la rama del Análisis Matemático son de vital importancia para el desarrollo de las demostraciones y fórmulas planteadas en el trabajo.
- Tener nociones básicas de Topología, en especial, la Teoría de Conjuntos es de gran ayuda para facilitar el entendimiento de algunos conceptos.

El gran desarrollo teórico y el planteamiento de algunos ejemplos nos han permitido obtener algunas conclusiones muy claras y destacables:

- El proceso de crear un método numérico para la resolución de EDOs no es trivial. Se requiere conocer conceptos como el orden de consistencia, la estabilidad, etc.
- El planteamiento de las fórmulas multipaso como fórmulas de un paso en un espacio dimensional superior permite dar caracterizaciones algebraicas elegantes de estabilidad, facilitando así, el estudio de los métodos multipaso cuando el paso de integración es variable.
- La importancia de la elección del tamaño de paso “h” para la realización de integraciones estables.

Bibliografía

- [1] E.Hairer S.P Nørsett G.Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Second Revised Edition.
- [2] E.Hairer G.Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Second Revised Edition.
- [3] Manuel. Calvo Pinilla Juan Ignacio Montijano Torcal; L Rández 1985. *Curso de analisis numérico : Métodos lineales multipaso para la resolución numérica de ecuaciones diferenciales ordinarias*. Disponible en la biblioteca de la Universidad de La Laguna: <https://absysnetweb.bbt.k.u11.es/cgi-bin/abnetopac/07173/ID30a63774?ACC=161>.
- [4] Wikipedia, 2021. Disponible en: https://es.wikipedia.org/wiki/M%C3%A9todo_lineal_multipaso.
- [5] Apuntes online, 2021. Disponible en: <https://sites.lsa.umich.edu/shixumeng/wp-content/uploads/sites/629/2020/04/Lec-5.10.pdf>.
- [6] GNU Octave *Programa y lenguaje de programación matemática para realizar cálculos numéricos*. <https://www.gnu.org/software/octave/index>.

Multistep Linear Methods for Solving Differential Equations

Abstract

This paper is dedicated to the study of multistep linear methods following the line presented by Ernst Hairer, along with other mathematicians, in two of his books. It will begin by commenting on some classic methods developed during history (Adams Bashforth, Adams Moulton,...), and will continue in the following chapters with general multistep methods focusing on the study of order, stability and convergence properties. Important results will be discussed and demonstrated, as well as two of the theorems established throughout the 20th century by the great mathematician Dahlquist, which take a crucial role in the study of stability.

1. Classical Linear Multistep Methods

There are families of classical multistep methods developed throughout the history of numerical analysis in the 20th century. A very famous example is the explicit k-step Adams method, also known as Adams-Bashforth methods.

$$y_{n+1} = y_n + h \sum_{j=0}^{k-1} \gamma_j \nabla^j f_n,$$

where

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

2. Error and order conditions

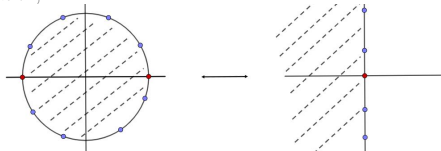
General formulas for multistep linear methods are shown along with local and global error definitions.

$$\sum_{s=0}^k \alpha_k y_{n+s} = h \sum_{j=0}^k \beta_j f_{n+s}.$$

These general methods are also expressed in the form of a differential operator in order to see what conditions are necessary for a method to have order of consistency p. There are methods that have the same order but provide different global errors and these will be distinguished by their error constant. In addition, the characteristic polynomials are shown, which take a crucial role throughout the paper.

3. Zero-Stability and the First Dahlquist Barrier

Basic definitions of zero stability are presented such as strongly zero-stable, weakly zero-stable, or unstable. A multitude of lemmas and necessary applications are shown, such as the use of the Greco-Roman transformation,



in order to demonstrate an important result: the Dahlquist First Barrier Theorem

4. Convergence of Multistep Methods

The objective of this section is to demonstrate that stability together with consistency are necessary and sufficient conditions for the convergence of the multistep methods. This is expressed in the famous result

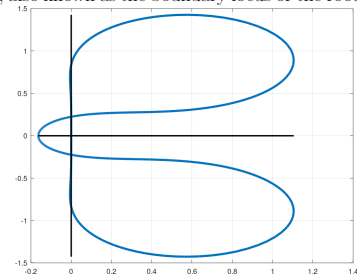
$$\text{convergence} = \text{stability} + \text{consistency}$$

Butcher's ideas introduced a higher quality to the proofs, who wrote the multistep formulas as one-step formulas in higher dimensional space.

$$Y_{n+1} = (A \otimes I)Y_n + h\Phi(t_n, Y_n, h)$$

5. Linear Stability Theory

Important definitions are presented such as the stability domain and its boundary, also known as the boundary locus or the root locus curve



In addition, another great Dahlquist result is shown, the Dahlquist second barrier theorem.

References

- [1] E.Hairer S.P. Nørsett G.Wanner. Solving Ordinary Differential Equations I. Nonstiff Problems. Second Revised Edition.
- [2] E.Hairer G.Wanner. Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems. Second Revised Edition.
- [3] Manuel. Calvo Pinilla Juan Ignacio Montijano Torcal; L Rández 1985. Curso de análisis numérico : Métodos lineales multipaso para la resolución numérica de ecuaciones diferenciales ordinarias. Disponible en la biblioteca de la Universidad de La Laguna: <https://absysnetweb.bbt.ull.es/cgi-bin/abnetopac/O7173/ID30a63774?ACC=161>.