



# **Métodos GRK**

para ecuaciones  
diferenciales ordinarias

Tesis presentada para optar al  
Grado de Doctor en  
Ciencias Matemáticas, por:  
**Jorge Álvarez López**

**Departamento de Matemática Aplicada a la Ingeniería**  
**Universidad de Valladolid**



Métodos GRK  
para ecuaciones  
diferenciales ordinarias

**Jorge Álvarez López**

Departamento de Matemática Aplicada a la Ingeniería  
E.T.S. de Ingenieros Industriales  
Universidad de Valladolid

Jorge Álvarez López

Métodos GRK  
para ecuaciones  
diferenciales ordinarias



**Tesis presentada para optar al  
Grado de Doctor en  
Ciencias Matemáticas, por:  
Jorge Álvarez López**





AUTORIZACIÓN DEL DIRECTOR DE TESIS  
(Art. 8.1 del R.D. 778/98 de 30 de abril)

D. Jesús Rojo García, con D.N.I. nº 13048651S, profesor del departamento de Matemática Aplicada a la Ingeniería, como Director de la Tesis Doctoral titulada 'Métodos GRK para ecuaciones diferenciales ordinarias', presentada por D. Jorge Álvarez López, alumno del programa Matemática Aplicada, impartido por el departamento de Matemática Aplicada a la Ingeniería, autoriza la presentación de la misma, considerando que es un trabajo original de investigación bien elaborado y de gran nivel científico.

Valladolid, 23 de abril de 2002  
El Director de la Tesis,

Fdo.: Jesús Rojo García



## AGRADECIMIENTOS

Quiero agradecer en primer lugar a mi director de tesis Jesús Rojo García su inestimable ayuda a la hora de sacar adelante esta Tesis Doctoral. Sin el tiempo que amablemente ha dedicado a orientarme, ayudarme y animarme, este trabajo difícilmente hubiera concluido satisfactoriamente.

A mi esposa María José e hijo Arturo por su paciencia, confianza y apoyo en todo momento.

A mis padres Trinidad y Félix y a mi hermano Roberto, por su ayuda y palabras de ánimo en los momentos difíciles.

A todos los demás amigos, familiares y compañeros del Departamento de Matemática Aplicada a la Ingeniería, que no nombro por miedo a olvidar injustamente a alguno.

Gracias a todos por su ayuda, paciencia, palabras de ánimo y buenos consejos.



# Prólogo.

La Tesis Doctoral que presenta Jorge Álvarez para optar al grado de Doctor en Ciencias Matemáticas posee un aspecto esencial que conviene señalar. Y es que, lejos de constituir una disquisición puramente teórica, de mucha importancia científica, pero de dudosa aplicación a tareas prácticas (al trabajo diario de cálculo, que diríamos mejor), pretende en todo momento mantenerse anclado a las necesidades de quien desea enfrentarse a la integración efectiva de problemas de valores iniciales. El espectro de científicos y técnicos a los que estas notas pueden interesar es muy amplio. Este hecho se debe a que, hoy, buena parte de los problemas prácticos a los que todos nos enfrentamos están ligados a la resolución, especialmente numérica, de problemas de valores iniciales formulados en términos de sistemas de ecuaciones diferenciales.

Quienes somos conscientes de la importancia del tema a que aludimos, estamos en condiciones de reconocer las ventajas que nos proporcionan los denominados Métodos GRK (por 'generalized Runge-Kutta methods'). Su orientación al tratamiento de los problemas rígidos (o 'stiff' como aparece frecuentemente en la literatura) marcha simultánea con la atención que en las últimas décadas se viene prestando a este tipo de comportamiento. Enfocados a este propósito, los métodos GRK permiten aumentar el orden manteniendo el número de etapas. Es cierto que el número de etapas viene constituyendo el criterio empleado en la valoración del 'coste' de un método, aunque hay que reconocer que, dependiendo del problema, pueden existir mayores ingredientes de 'gasto computacional'.

Pero, a esta ventaja que ya sería importante por ella misma, hay que añadir otras que no lo son menos. Contrariamente a la generalidad de los métodos diseñados para problemas rígidos, los métodos GRK no incluyen la necesidad de evaluar la (generalmente costosa) matriz

Jacobiana.

Y, dejada para el final por ser una de las características mayores, está la importante mejora de la estabilidad lineal sin perder el carácter explícito (o linealmente implícito) de los métodos.

No cabe duda de que las anteriores ventajas no pueden ser gratuitas. En efecto, el autor de la Tesis nos hace ver que estos métodos no pueden ser aplicados a cualquier tipo de sistemas. Pero nos muestra al mismo tiempo la fecundidad de los problemas a los que los métodos GRK sí que pueden enfrentarse. Justamente termina el trabajo con una exploración, ligera pero importante, de las técnicas para ecuaciones especiales de segundo orden (entre los que se encuentran importantes problemas orbitales) a las que sus métodos permiten tratar ventajosamente. Posiblemente será éste el campo en el que los derivados de los métodos GRK presentarán mayores ventajas de uso. El futuro tendrá, como siempre, la última palabra.

JESÚS ROJO, Valladolid 2002 .

# Contenido

<b>Prólogo.</b>	<b>xi</b>
<b>Contenido</b>	<b>xiii</b>
<b>1 Introducción.</b>	<b>1</b>
1.1 Breve reseña histórica. . . . .	2
1.2 Ventajas y limitaciones de los nuevos métodos. . . . .	5
1.3 Cómo surgen los nuevos métodos. . . . .	8
<b>2 Métodos de dos etapas.</b>	<b>15</b>
2.1 Expresión general de los métodos de dos etapas. . . . .	16
2.2 Orden de consistencia de los métodos de dos etapas. . . . .	22
2.3 Propiedades de estabilidad lineal de los métodos de dos etapas. . . . .	28
<b>3 Métodos de <math>p</math> etapas.</b>	<b>39</b>
3.1 Expresión general de los métodos de $p$ etapas. . . . .	39
3.2 Convergencia de los métodos de $p$ etapas. . . . .	45
3.3 Métodos de tipo polinomial. . . . .	54
3.4 Métodos de tipo racional. . . . .	60
3.5 Propiedades de estabilidad lineal de los métodos de $p$ etapas. . . . .	65
<b>4 Métodos de <math>p</math> etapas para sistemas.</b>	<b>75</b>
4.1 Expresión general de los métodos de $p$ etapas para sistemas. . . . .	76
4.2 Orden de consistencia de los métodos para sistemas. . . . .	89
4.3 Propiedades de estabilidad lineal de los métodos para sistemas. . . . .	100

<b>5</b>	<b>Métodos para ecuaciones especiales de segundo orden.</b>	<b>125</b>
5.1	Métodos de dos etapas para ecuaciones especiales de segundo orden. . . . .	126
	<b>Bibliografía</b>	<b>147</b>



# Capítulo 1

## Introducción.

En este capítulo, tras una breve reseña histórica, nos proponemos plantear el problema general del que nos ocuparemos, así como los métodos numéricos que construiremos y estudiaremos en posteriores capítulos. Asimismo introduciremos algunos conceptos y resultados que nos serán de utilidad en lo que sigue, en especial a la hora de comprender la estructura de los nuevos métodos.

Por el momento, el problema general vendrá dado por la ecuación escalar autónoma de primer orden  $y'(x) = f(y(x))$  con condición inicial  $y(x_0) = y_0$  (problema de Cauchy). Las razones que nos mueven a considerar únicamente el caso escalar autónomo en este capítulo introductorio, en lugar del más habitual caso vectorial, se pondrán de manifiesto más adelante, cuando motivemos la introducción de nuevos métodos numéricos y veamos las dificultades que plantea su extensión al caso vectorial.

En cuanto a los métodos numéricos que aplicaremos a la resolución numérica de nuestro problema de partida, podemos adelantar que, en un sentido que precisaremos posteriormente, constituyen una generalización de los conocidos métodos Runge-Kutta y asimismo se relacionan con los denominados Rosenbrock. Por esta razón nos referiremos a ellos en ocasiones con el término *métodos GRK* (abreviatura de "generalized Runge-Kutta methods").

Conviene señalar sin embargo que existen diferencias esenciales entre los métodos que construiremos y los anteriormente citados. De hecho, los esquemas numéricos de cuya construcción y estudio nos ocuparemos, no aparecen publicados en la literatura específica sobre el tema.

## 1.1 Breve reseña histórica.

En las últimas décadas, gran número de libros y artículos en revistas internacionales se han ocupado del estudio de métodos numéricos para la integración de ecuaciones diferenciales rígidas (stiff). Se han propuesto multitud de métodos buscando buenas propiedades de estabilidad lineal y no lineal. Casi todos estos métodos son implícitos en mayor o en menor medida. Una buena referencia en lo relativo a métodos para problemas rígidos es [20].

Los más usados en la práctica son aquellos basados en métodos lineales multipaso, especialmente los métodos BDF (véanse por ejemplo [27, 18]), por ser muy eficientes para un gran número de problemas de este tipo. No obstante, Dahlquist probó en [15] que ningún método lineal multipaso de orden mayor que dos puede ser A-estable, por lo que estas fórmulas no son adecuadas cuando se trata de integrar algunos problemas rígidos (por ejemplo, aquellos cuya matriz Jacobiana tiene algún autovalor con parte imaginaria grande).

También han sido utilizados con cierta frecuencia métodos Runge-Kutta implícitos [12], debido a sus buenas propiedades de estabilidad (A-estabilidad, L-estabilidad y B-estabilidad entre otras). Sin embargo, es necesario resolver un sistema no lineal de ecuaciones (algebraicas) en cada paso de la integración, lo que hace estas fórmulas poco competitivas especialmente cuando la dimensión del problema es grande. De hecho, al integrar numéricamente un sistema de ecuaciones diferenciales de dimensión  $N$  con un método Runge-Kutta implícito de  $m$  etapas, es necesario en general resolver en cada paso un sistema no lineal de ecuaciones de dimensión  $m \cdot N$ . Por otra parte, y con el fin de retener las buenas propiedades de estabilidad de estos métodos, el sistema algebraico ha de ser resuelto en general mediante un método del tipo Newton-Raphson. El método de Newton requiere de la evaluación de  $m$  matrices Jacobianas y de la resolución de un sistema lineal de dimensión  $m \cdot N$  (cuya factorización LU requiere a su vez del orden de  $m^3 \cdot N^3 / 3$  productos) en cada iteración. Con el fin de reducir el número de evaluaciones de la matriz Jacobiana, se suele recurrir a un método de Newton modificado (a la hora de resolver las ecuaciones algebraicas) consistente en sustituir en el método de Newton las  $m$  evaluaciones de la matriz Jacobiana, por una única evaluación de dicha matriz (perdiendo de este modo la convergencia cuadrática). De este

modo se reduce el coste computacional asociado a la evaluación de la matriz Jacobiana, pero sigue siendo necesario realizar la factorización LU de una matriz de dimensión  $m \cdot N$ .

Para reducir el coste computacional que requiere la solución del sistema algebraico mediante un método de Newton modificado al aplicar un método Runge-Kutta implícito, se han identificado subfamilias de métodos Runge-Kutta cuya implementación permite ahorrar operaciones a la hora de obtener la factorización LU. Así, para la clase de métodos Runge-Kutta diagonalmente implícitos (DIRK) de  $m$  etapas, el coste computacional asociado a la factorización LU se reduce a  $m \cdot N^3/3$  operaciones de multiplicación. Considerando la clase de métodos Runge-Kutta simplemente diagonalmente implícitos (SDIRK) o la clase de métodos Runge-Kutta simplemente implícitos (SIRK), es posible reducir aún más el coste de dicha factorización hasta las  $N^3/3$  operaciones de multiplicación (véanse [20, 12, 1] para más detalles).

En un intento de superar algunas de las dificultades mencionadas anteriormente, en los últimos años se han introducido clases de métodos más generales (véanse por ejemplo [10, 11, 8]) que combinan métodos lineales multipaso y Runge-Kutta.

Con el fin de reducir el coste computacional que se requiere en cada paso de la integración numérica, también se han considerado métodos linealmente implícitos, eliminando de este modo la necesidad de resolver sistemas no lineales de ecuaciones algebraicas por medio de fórmulas del tipo Newton (que requieren evaluaciones adicionales en cada iteración). Para implementar estas fórmulas, basta resolver un sistema lineal de ecuaciones algebraicas en cada paso, con las ventajas que esto supone sobre los métodos considerados con anterioridad.

De entre los muchos métodos de este tipo cabe destacar los conocidos como métodos Rosenbrock [35] así como los ROW-métodos (también conocidos como métodos Rosenbrock-Wanner o métodos Rosenbrock modificados) [32, 24, 25]. Estos métodos tienen el inconveniente de que es necesario evaluar la matriz Jacobiana en cada paso, lo que les hace poco competitivos cuando esta evaluación es costosa desde un punto de vista computacional. Por esta razón se han considerado extensiones de los métodos Rosenbrock para las cuales la matriz Jacobiana sólo ha de ser evaluada cada cierto número de pasos [45, 22, 43]. Más aún, se han desarrollado métodos de tipo Rosenbrock que no requieren de la evaluación de la matriz Jacobiana. Entre otros podemos

citar los W-métodos [41], los MROW-métodos [53] y los métodos de Runge-Kutta generalizados [44] (véase también [16]). Para una excelente exposición de algunos de estos métodos véase [20].

También se han propuesto otros métodos no lineales que prescindan del uso de la matriz Jacobiana, entre los que cabe destacar los métodos racionales especialmente diseñados para problemas cuya solución posee singularidades. Dichos métodos vienen descritos en [26] pp. 209–216 (para más detalles véanse también [28] y [38]).

Los métodos explícitos (linealmente implícitos) que propondremos en esta memoria, pueden ser considerados como una generalización de los clásicos métodos Runge-Kutta explícitos, que permite una sensible reducción del número de etapas necesario para obtener fórmulas de un orden determinado. Veremos asimismo cómo los métodos GRK propuestos permiten obtener fórmulas explícitas (linealmente implícitas) con buenas propiedades de estabilidad (como A-estabilidad y L-estabilidad) y generalizaremos los mismos para hacerlos aplicables a algunos problemas no escalares y no autónomos que surgen con cierta frecuencia en las aplicaciones. Conviene señalar en este punto que, a diferencia de lo que ocurre con otros métodos específicamente diseñados para problemas rígidos (y de los que hemos hablado con anterioridad), nuestros métodos no requieren del uso de la matriz Jacobiana ni en su diseño ni en su implementación. También propondremos una generalización de los mismos para problemas de segundo orden de un tipo especial.

Tras este primer capítulo justificativo e introductorio, nos ocuparemos en el segundo capítulo del estudio de los nuevos métodos de dos etapas para problemas escalares autónomos de primer orden. En el tercer capítulo obtendremos la expresión general de los métodos de  $p$  etapas para este tipo de problemas y estudiaremos la convergencia, el orden y las propiedades de estabilidad lineal de los mismos. Una generalización de los métodos considerados en los tres primeros capítulos, para hacerlos aplicables a problemas vectoriales de primer orden de un tipo especial, se considera en el capítulo cuatro. Finalmente, en el último capítulo estudiamos brevemente una generalización de los métodos que permite obtener fórmulas aplicables a ecuaciones especiales de segundo orden (que también pueden ser consideradas una generalización de los métodos Runge-Kutta-Nyström).

## 1.2 Ventajas y limitaciones de los nuevos métodos.

Como se hará enseguida obvio, apenas comencemos la descripción de los nuevos métodos GRK para ecuaciones escalares, la principal y muy notable ventaja de los mismos es la mejora del orden manteniendo un número reducido de evaluaciones. El único tributo a pagar es un aumento del coste computacional asociado, que se traduce en un mayor número de operaciones de suma y multiplicación (y en ocasiones en la necesidad de introducir también alguna división) que en los tradicionales métodos Runge-Kutta explícitos del mismo número de etapas.

Si dejamos momentáneamente aparte la importante limitación que supone la restricción al caso escalar autónomo, conviene señalar que obtenemos métodos explícitos, de orden muy alto y con notables propiedades de estabilidad lineal, hechos todos ellos imposibles con el diseño usual de los métodos de Runge-Kutta explícitos. A título de ejemplo elemental: obtener, con sólo dos evaluaciones de la función por paso, orden 3 y L-estabilidad justifica ya ampliamente la consideración de los métodos que presentamos. El método que consigue las ventajas anteriores tiene la simple formulación

$$\begin{aligned} k_1 &= f(y_n), & k_2 &= f\left(y_n + \frac{2}{3}hk_1\right), \\ s &= \frac{3(k_2 - k_1)}{2k_1}, \\ y_{n+1} &= y_n + h k_1 \left(\frac{6 - s}{6 - 4s + s^2}\right). \end{aligned} \quad (1.1)$$

Conviene señalar en este punto que, como veremos más adelante en esta memoria, es también posible construir fórmulas de dos etapas y orden tres (pertenecientes a la familia de métodos que introducimos) a partir de funciones de estabilidad lineal prefijadas. Esto puede ser de gran utilidad al tratar problemas perturbados para los que se conozca la solución exacta del problema no perturbado, pues permite construir de forma sencilla fórmulas específicamente diseñadas para integrar exactamente el problema no perturbado.

Cuando pasamos a abordar el problema de valor inicial para sistemas, nos encontramos sin embargo con el inconveniente de que la



pp. 349–443), (4.24) (véanse [36] pp. 27 y [34]) y (4.26) (tomado de [27], pp. 213) son ejemplos, por lo demás muy interesantes, de nuestra afirmación. Dando un paso más en esta línea de las ecuaciones escalares, hemos adaptado nuestros métodos a la integración de buena parte de los problemas de segundo orden, siempre de un tipo especial, con lo que tocamos uno de los problemas de mayor interés en el tratamiento del caso orbital.

En el caso de las ecuaciones en derivadas parciales, las ideas sobre semi-discretización que transforman estos problemas en sistemas de ecuaciones diferenciales ordinarias llevan, o pueden llevar, en muchos casos a sistemas "separados", para los que nuestros métodos son aplicables. Más adelante ponemos ejemplos de esta opción y, en particular, retomamos el ejemplo de [20] (pp. 349–443) sobre la ecuación de Burgers

$$u_t + u u_x = \nu u_{xx}, \quad \nu > 0, \quad (1.4)$$

que, tras la discretización por el Método de líneas (MOL) adquiere la forma

$$u'_i = -\frac{u_{i+1}^2 - u_{i-1}^2}{4\Delta x} + \nu \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2}, \quad i = 1, 2, \dots, N, \quad (1.5)$$

que resolvemos con uno de nuestros métodos comparando las soluciones obtenidas con las de [20] y relacionando el trabajo de integración empleado en ambos métodos de resolución numérica.

Resumiendo, esta sección ha querido dejar clara la importante limitación de los métodos GRK cuyo diseño sigue, en lo relativo a su ámbito de aplicabilidad. Objetivo esencial de lo anterior es no producir engaño alguno al lector rápido de esta memoria. Pero dicho esto, hemos tratado de dejar claramente establecido que los nuevos métodos se pueden usar en gran cantidad de problemas considerados como prácticos en la literatura seria ya conocida. Además, en aquellos problemas para los cuales estos métodos explícitos (o linealmente implícitos) se pueden aplicar, la ganancia en orden y sobre todo en lo relativo a las propiedades de estabilidad lineal para problemas rígidos ("stiff") es notable y justifica plenamente su introducción y su estudio.

Establecido lo que precede, estamos en condiciones de comenzar la descripción de los nuevos métodos, comenzando por dar una idea intuitiva de su origen. Este punto será el que nos ocupe el resto de este capítulo introductorio.

## 1.3 Cómo surgen los nuevos métodos.

**1.3.1 La ecuación escalar autónoma de primer orden.** Para empezar nos ocupamos de buscar métodos más eficaces para la integración de la ecuación escalar autónoma de primer orden, o más concretamente del correspondiente problema de Cauchy

$$y'(x) = f(y(x)), \quad y(x_0) = y_0, \quad (1.6)$$

donde  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

En ocasiones, nos restringiremos a estudiar el problema de valores iniciales en un intervalo  $I = [a, b]$  de  $\mathbb{R}$  y tomaremos  $a = x_0$ . Asimismo, dependiendo del contexto, supondremos que la función  $f$  verifica ciertas propiedades, que precisaremos en el momento oportuno.

Como ya hemos dicho, los esquemas numéricos que aportamos mejoran considerablemente el orden de consistencia y las propiedades de estabilidad lineal que se obtienen con los métodos clásicos de tipo Runge-Kutta (abreviadamente R-K) explícitos con el mismo número de evaluaciones.

**1.3.2 Una primera aproximación a los nuevos métodos.** Es bien conocido que muchos métodos numéricos para la integración de ecuaciones diferenciales ordinarias pueden ser obtenidos de manera análoga a como se obtienen las fórmulas de cuadratura de tipo interpolación. Veamos a continuación, a modo de ejemplo, cómo es posible obtener el método de Euler y alguno de los nuevos métodos que nos proponemos introducir, siguiendo la idea antes expuesta.

A partir del problema de valor inicial (1.6) se deduce la identidad

$$\frac{y'(x)}{f(y(x))} = 1, \quad \forall x \in \mathbb{R}, \quad (1.7)$$

supuesto que la función  $f$  no se anula. Integrando la identidad anterior en el intervalo  $[x_n, x_{n+1}]$  deducimos

$$\int_{x_n}^{x_{n+1}} \frac{y'(x)}{f(y(x))} dx = x_{n+1} - x_n, \quad (1.8)$$



o equivalentemente, tras un cambio de variable en el primer miembro

$$\int_{y(x_n)}^{y(x_{n+1})} \frac{dy}{f(y)} = x_{n+1} - x_n. \quad (1.9)$$

La igualdad anterior nos permite obtener aproximaciones numéricas a la solución del problema (1.6) sin más que aproximar la integral por una regla de cuadratura. Pasamos a considerar algunas posibles elecciones de fórmulas de cuadratura, para ver a qué tipo de métodos numéricos dan lugar.

Si utilizamos la regla del rectángulo, esto es

$$\int_a^b g(t) dt \approx (b - a) g(a), \quad (1.10)$$

se deduce de (1.9) que

$$x_{n+1} - x_n = \int_{y(x_n)}^{y(x_{n+1})} \frac{dy}{f(y)} \approx \frac{y(x_{n+1}) - y(x_n)}{f(y(x_n))}. \quad (1.11)$$

Si denotamos por  $h = x_{n+1} - x_n$  y definimos  $y_{n+1}$  e  $y_n$  como aproximaciones a  $y(x_{n+1})$  y a  $y(x_n)$  respectivamente (de modo que la igualdad aproximada (1.11) pase a ser una igualdad), obtenemos el método

$$h = \frac{y_{n+1} - y_n}{f(y_n)}, \quad (1.12)$$

que no es otro que el método de Euler explícito dado por

$$y_{n+1} = y_n + hf(y_n). \quad (1.13)$$

Análogamente, si consideramos la regla del trapecio dada por

$$\int_a^b g(t) dt \approx \frac{b - a}{2} (g(a) + g(b)), \quad (1.14)$$

se tiene que

$$\begin{aligned} x_{n+1} - x_n &= \int_{y(x_n)}^{y(x_{n+1})} \frac{dy}{f(y)} \\ &\approx \frac{y(x_{n+1}) - y(x_n)}{2} \left( \frac{1}{f(y(x_n))} + \frac{1}{f(y(x_{n+1}))} \right) \\ &= \frac{y(x_{n+1}) - y(x_n)}{2} \frac{f(y(x_n)) + f(y(x_{n+1}))}{f(y(x_n)) f(y(x_{n+1}))}, \end{aligned} \quad (1.15)$$

y, haciendo uso de las notaciones anteriores, la correspondiente fórmula viene dada por

$$y_{n+1} = y_n + 2h \frac{f(y_n) f(y_{n+1})}{f(y_n) + f(y_{n+1})}. \quad (1.16)$$

La fórmula anterior es implícita. Con el fin de obtener un método explícito, podemos sustituir en el segundo miembro de la igualdad anterior  $y_{n+1}$  por la aproximación dada por  $y_n + h f(y_n)$  (esto es, por la aproximación que nos proporciona el método de Euler explícito) con lo que se obtiene la fórmula

$$y_{n+1} = y_n + 2h \frac{f(y_n) f(y_n + h f(y_n))}{f(y_n) + f(y_n + h f(y_n))}, \quad (1.17)$$

que puede ser descrita de forma análoga a como se hace en el caso de los métodos de tipo Runge-Kutta, mediante la expresión

$$y_{n+1} = y_n + 2h \frac{k_1 k_2}{k_1 + k_2}, \quad (1.18)$$

donde

$$k_1 = f(y_n), \quad k_2 = f(y_n + h k_1). \quad (1.19)$$

Es fácil comprobar que la fórmula anterior proporciona un método de orden dos.

Finalmente, la regla de cuadratura de tipo Radau I de dos abscisas dada por

$$\int_a^b g(t) dt \approx \frac{b-a}{4} \left( g(a) + 3g\left(a + \frac{2}{3}(b-a)\right) \right), \quad (1.20)$$

da lugar a la siguiente aproximación

$$\begin{aligned} x_{n+1} - x_n &= \int_{y(x_n)}^{y(x_{n+1})} \frac{dy}{f(y)} \\ &\approx \frac{y(x_{n+1}) - y(x_n)}{4} \left( \frac{1}{f(y(x_n))} + \frac{3}{f\left(y(x_n) + \frac{2}{3}(y(x_{n+1}) - y(x_n))\right)} \right) \\ &= \frac{y(x_{n+1}) - y(x_n)}{4} \frac{3f(y(x_n)) + f\left(y(x_n) + \frac{2}{3}(y(x_{n+1}) - y(x_n))\right)}{f(y(x_n)) f\left(y(x_n) + \frac{2}{3}(y(x_{n+1}) - y(x_n))\right)}, \quad (1.21) \end{aligned}$$

que permite deducir la fórmula

$$y_{n+1} = y_n + 4h \frac{f(y_n) f\left(y_n + \frac{2}{3}(y_{n+1} - y_n)\right)}{3f(y_n) + f\left(y_n + \frac{2}{3}(y_{n+1} - y_n)\right)}. \quad (1.22)$$

De nuevo, la fórmula así obtenida es implícita. Sustituyendo en el segundo miembro de la igualdad anterior  $y_{n+1}$  por la aproximación dada por  $y_n + h f(y_n)$ , obtenemos el siguiente método explícito

$$y_{n+1} = y_n + 4h \frac{k_1 k_2}{3k_1 + k_2}, \quad (1.23)$$

donde

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3} h k_1\right). \quad (1.24)$$

Más adelante veremos que la fórmula anterior proporciona un método de orden tres y dos evaluaciones por paso.

Es fácil comprobar que la función de estabilidad lineal asociada a los métodos (1.18) y (1.23) viene dada por una función racional. Esto nos permite contemplar la posibilidad de construir métodos como los anteriormente descritos con buenas propiedades de estabilidad lineal (A-estabilidad y L-estabilidad). En posteriores secciones veremos que de hecho es posible construir métodos de dos etapas A-estables y L-estables de orden tres.

Podemos observar en los tres ejemplos anteriores que la función  $f$  interviene en los segundos miembros de las fórmulas a través de una función racional dada por el cociente de dos polinomios homogéneos (en las etapas  $k_1$  y  $k_2$ ), siendo el grado del numerador una unidad mayor que el grado del denominador. Esta observación nos va a permitir dar una expresión general de los nuevos métodos GRK de dos etapas, que más adelante generalizaremos al caso de más etapas. Pero antes veamos un modo alternativo de motivar la introducción de los nuevos métodos.

**1.3.3 Otra aproximación a los nuevos métodos.** También es posible llegar a nuestros métodos partiendo de un enfoque completamente distinto. Para ello supondremos que la función  $f$  en el problema de Cauchy (1.6) es lo suficientemente regular como para que todas las

derivadas que van a intervenir tengan sentido. Derivando la igualdad  $y'(x) = f(y(x))$  respecto de la variable  $x$  obtenemos

$$y''(x) = f_y(y(x)) y'(x) = f_y(y(x)) f(y(x)), \quad (1.25)$$

donde  $f_y$  denota la derivada de  $f$  respecto de su argumento y hemos aplicado la regla de la cadena en el segundo miembro de la igualdad. Podemos abreviar la expresión así obtenida y escribirla en la forma  $y'' = f_y(y) f(y)$ , o equivalentemente, haciendo uso de la ecuación de partida, como

$$\frac{d}{dx} f(y) = f_y(y) f(y), \quad (1.26)$$

que pone de manifiesto que, formalmente, el operador  $d/dx$  actuando sobre  $f \circ y$  se comporta como el producto de  $f_y \circ y$  por  $f \circ y$ . La sencilla propiedad que acabamos de observar y que, como veremos, se extiende sin dificultad a derivadas de mayor orden de  $f \circ y$  respecto de  $x$ , es la llave que nos permitirá llegar a la expresión de los nuevos métodos explícitos de un paso, de un modo completamente distinto al descrito en el apartado anterior.

Pasemos a estudiar con más detalle la propiedad anterior, su extensión a derivadas de orden superior y sus aplicaciones a la hora de reescribir el desarrollo en serie de Taylor de la solución, en una forma más conveniente para nuestros propósitos.

Partiendo de la última igualdad obtenida, derivamos sucesivas veces respecto de la variable  $x$  haciendo uso de la regla de la cadena, de la ecuación y de que derivamos a lo largo de una solución de la misma y obtenemos

$$\begin{aligned} \frac{d}{dx} f(y) &= f_y(y) f(y), \\ \frac{d^2}{dx^2} f(y) &= [f_{yy}(y) f(y) + (f_y(y))^2] f(y), \\ \frac{d^3}{dx^3} f(y) &= [f_{yyy}(y) (f(y))^2 + 4f_{yy}(y) f_y(y) f(y) + (f_y(y))^3] f(y). \end{aligned} \quad (1.27)$$

Es fácil ver que, en general, la derivada de cualquier orden admite una expresión análoga a las anteriores, esto es, en su segundo miembro aparece  $f(y)$  como factor (consecuencia inmediata de la regla de la cadena y de la ecuación).

Veamos en qué se traduce esto, a la hora de calcular el desarrollo limitado de Taylor de la solución exacta de nuestro problema de Cauchy.

Consideremos, por simplicidad, el desarrollo de Taylor de grado cuatro de la solución en un entorno del punto  $x$ . Se tendrá

$$y(x+h) = y(x) + y'(x)h + \frac{y^{(2)}(x)}{2}h^2 + \frac{y^{(3)}(x)}{6}h^3 + \frac{y^{(4)}(x)}{24}h^4 + O(h^5), \quad (1.28)$$

o equivalentemente, tras sustituir las derivadas por sus correspondientes expresiones en términos de la función  $f$ , tendremos

$$y(x+h) = y(x) + fh + \frac{f_y f}{2}h^2 + \frac{[f_{yy}f + f_y^2]f}{6}h^3 + \frac{[f_{yyy}f^2 + 4f_{yy}f_y f + f_y^3]f}{24}h^4 + O(h^5), \quad (1.29)$$

donde hemos obviado el punto de evaluación que en todos los casos es  $y(x)$ . Si ahora manipulamos ligeramente la igualdad anterior, ésta tomará la forma

$$\frac{y(x+h) - y(x)}{h} = \left( 1 + \frac{f_y}{2}h + \frac{f_{yy}f + f_y^2}{6}h^2 + \frac{f_{yyy}f^2 + 4f_{yy}f_y f + f_y^3}{24}h^3 \right) f + O(h^4), \quad (1.30)$$

que pone de manifiesto que el segundo miembro de la igualdad aproxima a la derivada  $y'(x) = f$ . Hagamos ahora una observación más importante. Si elevamos al cuadrado ambos miembros de la igualdad y dividimos la expresión así obtenida por  $f$ , suponiendo no nulo este valor, tendremos

$$\frac{(y(x+h) - y(x))^2}{fh^2} = \left( 1 + f_y h + \frac{4f_{yy}f + 7f_y^2}{12}h^2 + \frac{f_{yyy}f^2 + 6f_{yy}f_y f + 3f_y^3}{12}h^3 \right) f + O(h^4), \quad (1.31)$$

y es fácil observar que los términos que aparecen en el segundo miembro de la igualdad, son idénticos a los que observábamos en el segundo

miembro de la igualdad precedente, salvo en los coeficientes numéricos de los mismos. El proceso anterior se podría repetir sin dificultad (elevando ambos miembros a la potencia  $n$  y dividiendo por  $f^{n-1}$ ) y seguiríamos observando el mismo comportamiento en el segundo miembro.

Vamos a ver a continuación cómo se traslada esta propiedad de la solución  $y$  de nuestra ecuación autónoma de partida a la función  $f$ .

Consideremos el desarrollo en serie de Taylor hasta orden tres de la función  $f(y(x+h))$  (esto es, de  $f$  a lo largo de una solución de la ecuación escalar autónoma), en un entorno del punto  $x$ . Tendremos

$$f(y(x+h)) = f + f_y f h + \frac{f_{yy} f^2 + f_y^2 f}{2} h^2 + \frac{f_{yyy} f^3 + 4 f_{yy} f_y f^2 + f_y^3 f}{6} h^3 + O(h^4), \quad (1.32)$$

(en el segundo miembro de la igualdad anterior hemos omitido el punto de evaluación que es en todos los casos  $y(x)$ ). Equivalentemente se tiene

$$f(y(x+h)) = \left( 1 + f_y h + \frac{f_{yy} f + f_y^2}{2} h^2 + \frac{f_{yyy} f^2 + 4 f_{yy} f_y f + f_y^3}{6} h^3 \right) f + O(h^4). \quad (1.33)$$

Se puede comprobar de forma sencilla que, al igual que sucedía anteriormente, si elevamos la igualdad anterior a la  $n$ -ésima potencia y la dividimos por  $f^{n-1}$  (suponiendo que esta cantidad es no nula), los términos que aparecen en el segundo miembro de la igualdad resultante son idénticos a los que aparecían en la igualdad de partida, salvo en los coeficientes numéricos que, en general, habrán cambiado.

Partiendo de esta última observación, estamos en condiciones de motivar la introducción de los nuevos métodos numéricos y abordar un estudio preliminar de los mismos. Observaremos cómo estas propiedades aparentemente desprovistas de aplicaciones prácticas, nos permitirán disponer de más parámetros libres a la hora de establecer el orden de consistencia y las propiedades de estabilidad de los mismos, sin necesidad de aumentar el número de evaluaciones por paso con respecto a los esquemas de tipo Runge-Kutta clásicos.

## Capítulo 2

### Métodos de dos etapas.

En este capítulo nos ocuparemos del estudio de los métodos GRK de dos etapas. Las razones que nos mueven a empezar considerando este caso particular son las siguientes:

- La estructura de las fórmulas así obtenidas es notablemente más sencilla que en el caso general de  $p$  etapas.
- La familia de métodos resultante es lo suficientemente amplia como para ilustrar las diferentes ventajas que se pueden obtener sobre otros métodos clásicos.

De hecho, mostraremos cómo los nuevos métodos explícitos de dos etapas, aplicados al problema escalar autónomo, permiten obtener orden tres con tan sólo dos evaluaciones por paso. Nótese en este punto que los métodos Runge-Kutta explícitos de dos etapas sólo permiten obtener fórmulas de orden dos. Más aún, veremos que de hecho es posible obtener fórmulas explícitas  $A$ -estables y  $L$ -estables de orden tres (para el problema escalar autónomo). Sabemos que esto último es imposible cuando consideramos la familia de métodos Runge-Kutta explícitos.

Describiremos asimismo cómo obtener métodos que minimizan, en un sentido que precisaremos en su momento, la parte principal del error local de truncación. Por último veremos cómo es posible construir métodos de la nueva familia a partir de funciones de estabilidad prefijadas y realizaremos unos experimentos numéricos para poner de manifiesto las ventajas computacionales de nuestros métodos sobre otros bien conocidos.

## 2.1 Expresión general de los métodos de dos etapas.

**2.1.1 Descripción de los métodos.** La versión más general de los métodos que hemos propuesto informalmente, para el caso de dos evaluaciones por paso, es la siguiente: consideraremos métodos que admiten una expresión del tipo

$$y_{n+1} = y_n + hF(k_1, k_2), \quad (2.1)$$

con  $k_1$  y  $k_2$  dados por

$$k_1 = f(y_n), \quad k_2 = f(y_n + hc_2k_1), \quad (2.2)$$

y donde  $F(x, y)$  es una función homogénea de grado uno definida en un cierto cono  $A \subset \mathbb{R}^2$ .

Por tanto, la función  $F$  satisface la condición

$$F(\alpha x, \alpha y) = \alpha F(x, y), \quad \forall \alpha > 0, \forall (x, y) \in A \subset \mathbb{R}^2, \quad (2.3)$$

y está definida en un cono  $A \subset \mathbb{R}^2$ , esto es, en un subconjunto de  $\mathbb{R}^2$  que verifica

$$\alpha(x, y) \in A, \quad \forall \alpha > 0, \forall (x, y) \in A. \quad (2.4)$$

Conviene comentar en este punto que, en lo que sigue, muy frecuentemente la relación (2.3) será válida tomando  $A = \mathbb{R}^2$  y  $\alpha \in \mathbb{R}$ .

La familia de métodos que acabamos de introducir, generaliza en cierto sentido a los clásicos Runge-Kutta explícitos de dos etapas. Aplicados al problema (1.6), éstos pueden ser descritos completamente en términos de la subfamilia de los nuevos métodos que se obtiene tomando  $F(x, y) = b_1x + b_2y$  y  $c_2 = a_{21}$  en (2.1) y (2.2). Con la nueva formulación conseguiremos considerables ventajas que iremos detallando en lo que sigue, aunque a costa de perder la linealidad, lo que complicará el estudio de los métodos y las expresiones que los definen.

Prestaremos especial atención al caso en el que  $F(x, y)$  es una función racional dada por el cociente de dos polinomios homogéneos  $N(x, y)$  y  $D(x, y)$  de grados  $r + 1$  y  $r$  respectivamente (para cierto  $r \in \mathbb{N}$ ). Las funciones racionales  $F$  del tipo anterior son homogéneas de grado uno y pueden ser consideradas como una generalización de las funciones consideradas en los apartados 1.3.2 y 1.3.3 del capítulo anterior (véase especialmente la observación final en dicho capítulo).



**2.1.2 Una simplificación útil.** En adelante asumiremos que la constante  $c_2$  en (2.2) es no nula (cuando  $c_2 = 0$  el método podría ser considerado de una etapa).

Con el fin de simplificar el estudio de nuestros métodos, conviene describir el segundo miembro de la expresión (2.1) de una forma más adecuada. Con tal fin introducimos el nuevo término

$$s = \frac{k_2 - k_1}{c_2 k_1}, \quad (2.5)$$

donde  $k_1$  y  $k_2$  vienen dados como en (2.2).

Es fácil ver que  $s = h f_y(y_n) + O(h^2)$ , lo que pone de manifiesto que  $s$  aproxima numéricamente a  $h f_y(y_n)$  y nos será de utilidad cuando busquemos métodos con buenas propiedades de estabilidad lineal y cuando abordemos el estudio del orden de consistencia de los métodos.

En términos de  $s$ , cualquier método de la familia (2.1) con las condiciones ya expuestas, puede ser descrito mediante una fórmula del tipo

$$y_{n+1} = y_n + h k_1 G(s), \quad (2.6)$$

donde  $k_1$  y  $s$  vienen dados como en (2.2) y (2.5).

Es fácil comprobar que a cada método dado en términos de  $k_1$  y  $k_2$  mediante una fórmula del tipo (2.1), le corresponde un único método expresado en términos de  $k_1$  y  $s$  a través de una fórmula de la forma (2.6) y recíprocamente. Para ello, basta observar que de la definición de  $s$  (véase 2.5) se deduce que  $k_2/k_1 = 1 + c_2 s$ , y que la función  $G$  se puede obtener a partir de la función  $F$  (y recíprocamente  $F$  a partir de  $G$ ). En efecto, por ser  $F$  homogénea de grado uno se tiene

$$F(k_1, k_2) = k_1 F\left(1, \frac{k_2}{k_1}\right) = k_1 F(1, 1 + c_2 s), \quad (2.7)$$

y basta tomar

$$G(s) = F(1, 1 + c_2 s). \quad (2.8)$$

Análogamente, se puede obtener  $F$  a partir de  $G$  observando que

$$G(s) = G\left(\frac{k_2 - k_1}{c_2 k_1}\right), \quad (2.9)$$

y tomando

$$F(k_1, k_2) = k_1 G\left(\frac{k_2 - k_1}{c_2 k_1}\right). \quad (2.10)$$

De hecho, ambos métodos son en realidad el mismo, escrito de dos maneras diferentes.

Nótese que sólo damos esta notación para los métodos de dos evaluaciones por paso. Se podrá generalizar al caso de tres o más evaluaciones (mediante la introducción de varios términos  $s_i$ ), como veremos en capítulos posteriores.

Cuando  $k_1 = f(y_n) = 0$  (como consecuencia de lo cual  $k_2 = 0$ ) tomaremos  $y_{n+1} = y_n$  en vez de (2.6), por lo que la solución numérica pasará a ser constante (al igual que sucede en otros métodos de un paso). De este modo, aunque  $s$  no está definida cuando  $k_1 = 0$ , el método sí lo estará. Lo anterior habrá de ser tenido en cuenta especialmente a la hora de implementar los métodos.

La introducción del término  $s$  simplificará considerablemente las ecuaciones que surgen al estudiar el orden de consistencia de los métodos así construidos. Podemos adelantar en este punto que, al ser  $s = O(h)$  (como hemos visto anteriormente), se tendrá que  $s^k = O(h^k)$ , por lo que muchos de los parámetros que caracterizan los métodos sólo aparecerán en las ecuaciones relativas a las condiciones de orden suficientemente grande.

**2.1.3 Métodos de tipo polinomial.** Empezaremos considerando métodos GRK de dos etapas dados por (2.1–2.2) para los cuales la función homogénea de grado uno  $F$  viene dada por una expresión del tipo

$$F(x, y) = x \sum_{i=0}^r A_i \left( \frac{y}{x} \right)^i, \quad (2.11)$$

para cierto entero no negativo  $r$  y ciertas constantes reales  $A_i$ . Llamaremos métodos de dos etapas de tipo polinomial a tales métodos.

Nótese que cuando  $x = 0$  la función  $F$  en (2.11) puede no estar definida. Sin embargo, como ya iremos viendo, en la mayoría de los casos es posible definir la función  $F$  en los puntos de la forma  $(0, y)$  extendiéndola por continuidad (de hecho, basta hacerlo para el punto  $(0, 0)$ , pues, como ya hemos observado anteriormente  $k_1 = 0 \Rightarrow k_2 = 0$ ).

La familia de métodos de dos etapas de tipo polinomial sigue conteniendo todos los métodos Runge-Kutta explícitos de dos etapas (basta tomar  $r = 1$  en (2.11)).

Si expresamos los métodos de dos etapas de tipo polinomial en términos de  $s$  se tiene que la función  $F$  definida en (2.11) que caracteriza dichos métodos permite obtener una función polinómica  $G$  (de ahí el nombre que hemos dado a los métodos) definida a través de la relación (2.8) que en nuestro caso concreto toma la forma

$$G(s) = F(1, 1 + c_2 s) = \sum_{i=0}^r A_i (1 + c_2 s)^i. \quad (2.12)$$

Por tanto, la función  $G$  puede ser descrita mediante una expresión del tipo

$$G(s) = \sum_{i=0}^r a_i s^i. \quad (2.13)$$

Las razones que nos mueven a considerar esta familia de métodos de tipo polinomial se pondrán de manifiesto más adelante, cuando abordemos el estudio de las condiciones de orden para nuestros métodos. De hecho, podemos adelantar que las condiciones de orden para los métodos más generales que estamos considerando, se obtienen sin dificultad a partir de las condiciones de orden de los métodos de tipo polinomial, sin más que considerar el desarrollo de Taylor de la función  $G$  en términos de la variable  $s$  hasta un determinado orden.

**2.1.4 Métodos de tipo racional.** Como ya comentamos con anterioridad, prestaremos especial atención a los métodos dados por (2.1–2.2) para los cuales  $F(x, y)$  es una función racional dada por el cociente de dos polinomios homogéneos  $N(x, y)$  y  $D(x, y)$  de grados  $r + 1$  y  $r$  respectivamente (para cierto  $r \in \mathbb{N}$ ), esto es

$$F(x, y) = \frac{N(x, y)}{D(x, y)} = \frac{\sum_{i=0}^{r+1} N_i x^{r-i+1} y^i}{\sum_{i=0}^r D_i x^{r-i} y^i}. \quad (2.14)$$

Denominaremos a tales métodos GRK, métodos de tipo racional.

En la expresión que define a  $F(x, y)$  en (2.14), podríamos considerar  $r \in \mathbb{Z}$ , pero con ello no obtendríamos ningún método esencialmente distinto (pues si  $r$  es negativo, basta multiplicar el numerador y el denominador de  $F(x, y)$  por  $x^{-r} y^{-r}$  para recuperar un método equivalente perteneciente a la familia anterior).

Nótese que si deseamos evitar la posibilidad de que un cierto método de tipo racional, con  $r$  fijado, sea equivalente a un método del mismo tipo pero con un valor de  $r$  más pequeño, bastará imponer que  $N(x, y)$  y  $D(x, y)$  sean primos entre sí, y así lo haremos en adelante.

Por otra parte, queremos que los métodos, para  $h = 0$ , se reduzcan a  $y_{n+1} = y_n$ . Tomando  $h = 0$  en la expresión que define a estos métodos (véase (2.2)), se tiene que  $k_2 = k_1 = f(y_n)$  y por tanto

$$F(k_1, k_1) = k_1 \frac{N(1, 1)}{D(1, 1)} = k_1 \frac{\sum_{i=0}^{r+1} N_i}{\sum_{i=0}^r D_i}. \quad (2.15)$$

A la vista de la expresión anterior, y con el fin de que  $F$  esté definida cuando  $h = 0$ , deberemos imponer que  $D(1, 1) \neq 0$ . De hecho, sin pérdida de generalidad, podemos suponer en adelante que los métodos de tipo racional verifican la condición

$$D(1, 1) = \sum_{i=0}^r D_i = 1. \quad (2.16)$$

De este modo, obtenemos simultáneamente la unicidad en la representación de cada método.

Si reescribimos los métodos de tipo racional en términos de  $s$  haciendo uso de la relación (2.8) se obtiene que la función  $G$  asociada a la función  $F$  (véase (2.14)) viene dada por una expresión del tipo

$$G(s) = \frac{\sum_{i=0}^n n_i s^i}{1 + \sum_{i=1}^d d_i s^i}, \quad (2.17)$$

donde los valores de  $n$  y  $d$ , que dan cuenta respectivamente de los grados de los polinomios (en la variable  $s$ ) que definen el numerador y el denominador de  $G(s)$ , se pueden obtener a partir de (2.14) sin más que tomar  $n = \max\{i \in \mathbb{N}/N_i \neq 0\}$  y  $d = \max\{i \in \mathbb{N}/D_i \neq 0\}$ . Que  $d_0 = 1$  es una consecuencia inmediata de la condición (2.16).

Por ser  $s^k = O(h^k)$ , los parámetros  $n_k$  y  $d_k$  sólo aparecerán en las ecuaciones relativas a las condiciones de orden igual o superior a  $k + 1$ .

Conviene señalar en este punto que en [49] Wambecq introduce una familia de métodos no lineales de un paso para ecuaciones escalares de primer orden, contruidos a partir de aproximantes de Padé. Dichos métodos, a los que denomina métodos no lineales Runge-Kutta, son de tipo racional y Wambecq describe su construcción a partir de métodos Runge-Kutta clásicos aplicando formalmente un procedimiento de aceleración de la convergencia (mediante la construcción de una pseudo tabla de Padé). Los métodos contruidos de este modo pueden ser considerados casos particulares de la familia de métodos de tipo racional que consideramos en esta memoria (al menos formalmente). No obstante, ninguno de los métodos propuestos por Wambecq tiene un orden de convergencia superior al número de etapas considerado. Además en [49] el autor considera la ecuación escalar general, en tanto que nosotros restringimos nuestra atención al caso autónomo. Esto nos permite obtener fórmulas con órdenes de convergencia superiores al número de etapas consideradas (orden 3 con 2 etapas, orden 5 con 3 etapas, orden 7 con 4 etapas, etc) y evitar los problemas que presentan dichos métodos al ser generalizados para hacerlos aplicables a sistemas de EDOs de primer orden.

En [50], [51] y [52] Wambecq propone una generalización de los métodos introducidos en [49] para hacerlos aplicables a sistemas de EDOs, basada en la definición de un nuevo producto de vectores compatible con la inversa de Samelson de un vector y estudia las propiedades y las condiciones de orden de las fórmulas propuestas. Distintos autores como Hairer en [17] y Calvo en [14] se ocupan del estudio de algunas propiedades de estabilidad de los métodos de Wambecq en años posteriores. Liu en [29] pone de manifiesto una importante limitación de los métodos propuestos por Wambecq para sistemas. Concretamente en dicho artículo el autor muestra que en algunos problemas rígidos pequeñas perturbaciones en las condiciones iniciales se amplifican dramáticamente (al aplicar las fórmulas de Wambecq) hasta el punto de producir grandes errores relativos tras muy pocos pasos de integración.

Veremos más adelante que la generalización de los métodos GRK que propondremos para hacerlos aplicables a algunos sistemas de EDOs, permite obtener órdenes de convergencia superiores al número de etapas consideradas (entre otras ventajas que detallaremos) y además no presentan los problemas observados en [29] para las fórmulas de Wambecq para sistemas.

## 2.2 Orden de consistencia de los métodos de dos etapas.

**2.2.1 Condiciones de orden para los métodos de dos etapas de tipo polinomial.** Veamos que es posible obtener métodos de dos etapas de tipo polinomial con orden de consistencia tres.

Para ello, como es habitual en otros métodos de un paso, empecaremos considerando el error local de truncación  $T_{n+1}$  asociado a un método de tipo polinomial en el punto  $x_n$ , que viene dado por una expresión del tipo

$$T_{n+1} = y(x_{n+1}) - y(x_n) - hk_1 G(s), \quad (2.18)$$

donde ahora  $k_1$ ,  $k_2$  y  $s$  son

$$k_1 = f(y(x_n)), \quad k_2 = f(y(x_n) + hc_2 k_1), \quad s = \frac{k_2 - k_1}{c_2 k_1}, \quad (2.19)$$

y  $G$  viene definida por

$$G(s) = \sum_{i=0}^r a_i s^i. \quad (2.20)$$

Supondremos que la función  $f$  es lo suficientemente regular como para garantizar que las derivadas que van a aparecer tengan sentido.

Para mostrar que es posible obtener fórmulas de orden tres, bastará considerar la subfamilia de métodos de tipo polinomial que se obtiene tomando  $r = 2$  en (2.20), por ser  $s = O(h)$ .

El desarrollo de Taylor de  $y(x_{n+1}) = y(x_n + h)$  es

$$y(x_{n+1}) = y(x_n) + hf + \frac{h^2}{2} f_y f + \frac{h^3}{6} [f_y^2 f + f_{yy} f^2] + O(h^4), \quad (2.21)$$

donde  $f$ ,  $f_y$  y  $f_{yy}$  se suponen evaluadas todas ellas en el punto  $y(x_n)$ . Si desarrollamos  $s = s(h)$  como función de  $h$  (nótese que  $s$  depende de  $c_2$ ), obtenemos

$$s = hf_y + \frac{h^2}{2} c_2 f_{yy} f + O(h^3), \quad (2.22)$$

que pone de manifiesto que  $s = O(h)$ . Ahora es fácil deducir que el desarrollo del error local de truncación viene dado por

$$\begin{aligned} T_{n+1} &= h(1 - a_0)f + \frac{h^2}{2}(1 - 2a_1)f_y f \\ &+ \frac{h^3}{6}[(1 - 3c_2a_1)f_{yy}f^2 + (1 - 6a_2)f_y^2 f] + O(h^4), \end{aligned} \quad (2.23)$$

y deducir las condiciones de orden:

$$a_0 = 1, \quad (2.24)$$

$$a_1 = 1/2, \quad (2.25)$$

$$c_2a_1 = 1/3, \quad (2.26)$$

$$a_2 = 1/6, \quad (2.27)$$

a partir de las cuales obtenemos la expresión general de los métodos de dos etapas de tipo polinomial y orden tres, que vendrán dados por

$$y_{n+1} = y_n + hk_1G(s), \quad (2.28)$$

donde  $k_1$ ,  $k_2$  y  $s$  vienen dados por

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3}hk_1\right), \quad s = \frac{3(k_2 - k_1)}{2k_1}, \quad (2.29)$$

y la función  $G$  admite la siguiente expresión

$$G(s) = 1 + \frac{1}{2}s + \frac{1}{6}s^2 + \sum_{i=3}^r a_i s^i, \quad (2.30)$$

esto es, basta tomar  $c_2 = 2/3$ ,  $a_0 = 1$ ,  $a_1 = 1/2$  y  $a_2 = 1/6$ . Los parámetros  $a_i$  con  $i \geq 3$  se pueden tomar de manera arbitraria (pues no afectan a las condiciones de orden  $\leq 3$ ). Por tanto, tomaremos  $a_i = 0$  para  $i \geq 3$  salvo en aquellos casos en los que sea posible obtener propiedades adicionales de interés que justifiquen no tomarlos así.

Las condiciones de orden adicionales para orden cuatro (ahora tomando  $r = 3$  en (2.20)) vienen dadas por

$$c_2^2 a_1 = 1/4, \quad (2.31)$$

$$c_2 a_2 = 1/6, \quad (2.32)$$

$$a_3 = 1/24, \quad (2.33)$$

y es fácil comprobar que ningún método de dos etapas de tipo polinomial tiene orden mayor que tres (las condiciones (2.31) y (2.32) no pueden ser satisfechas). Sin embargo, la condición (2.33) se puede satisfacer obteniendo de este modo métodos de orden tres que minimizan la parte principal del error local de truncación. Nótese en este punto que los otros términos de la parte principal del error local de truncación (aquellos asociados a las condiciones (2.31) y (2.32)) son los mismos para cualquier método de orden tres.

### 2.2.2 Condiciones de orden para los métodos de dos etapas de tipo racional.

A partir de las condiciones de orden de los métodos de tipo polinomial, es sencillo deducir las de los métodos de tipo racional. Para ello basta hacer el desarrollo de Taylor de la función  $G$  definida en (2.17) (para los métodos de tipo racional) en términos de  $s$  y comparar dicho desarrollo con el asociado a la función  $G$  en (2.20) que se obtiene para los métodos de tipo polinomial. Para ilustrar el procedimiento anteriormente descrito, pasamos a obtener las condiciones de orden para los métodos racionales a partir de las correspondientes condiciones de orden para los métodos polinomiales.

Para obtener todos los métodos de dos etapas de tipo racional de orden tres, es suficiente observar que de (2.25–2.26) se deduce que  $c_2 = 2/3$  y que ha de verificarse

$$G(s) = \frac{\sum_{i=0}^n n_i s^i}{1 + \sum_{i=1}^d d_i s^i} = 1 + \frac{1}{2}s + \frac{1}{6}s^2 + O(s^3). \quad (2.34)$$

Evidentemente, cuando buscamos métodos de orden tres sin otras propiedades adicionales bastará tomar  $n = d = 2$  en (2.34). Llegamos así a las siguientes relaciones

$$c_2 = 2/3, \quad (2.35)$$

$$n_0 = 1, \quad (2.36)$$

$$n_1 = \frac{1}{2} + d_1, \quad (2.37)$$

$$n_2 = \frac{1}{6} + \frac{1}{2}d_1 + d_2. \quad (2.38)$$



Si deseamos además minimizar la parte principal del error local de truncación, tendremos que tomar  $n = d = 3$  en (2.34) y desarrollar el segundo miembro hasta un orden superior, obteniendo de este modo la condición adicional

$$n_3 = \frac{1}{24} + \frac{1}{6}d_1 + \frac{1}{2}d_2 + d_3. \quad (2.39)$$

La expresión general de un método de dos etapas de tipo racional y orden tres viene dada por

$$y_{n+1} = y_n + hk_1 G(s), \quad (2.40)$$

donde  $k_1$ ,  $k_2$  y  $s$  vienen dados por

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3}hk_1\right), \quad s = \frac{3(k_2 - k_1)}{2k_1}, \quad (2.41)$$

y la función  $G$  toma la forma

$$G(s) = \frac{1 + \frac{1+2d_1}{2}s + \frac{1+3d_1+6d_2}{6}s^2 + \sum_{i=3}^n n_i s^i}{1 + d_1 s + d_2 s^2 + \sum_{i=3}^d d_i s^i}. \quad (2.42)$$

Posteriormente veremos cómo es posible explotar la estructura de nuestros métodos racionales, a la hora de obtener fórmulas con buenas propiedades de estabilidad lineal tales como la A-estabilidad y la L-estabilidad. Pero antes veamos un experimento numérico.

**2.2.3 Un primer ejemplo de método de orden tres.** Consideremos a modo de ejemplo sencillo el siguiente método de tipo polinomial

$$y_{n+1} = y_n + hk_1 \left( \frac{5}{8} + \frac{3}{8} \left( \frac{k_2}{k_1} \right)^2 \right), \quad (2.43)$$

donde

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3}hk_1\right), \quad (2.44)$$

Reescribiendo el mismo en términos de  $s$ , a través de la relación  $k_2/k_1 = 1 + (2/3)s$ , que se deduce de la notación introducida en (2.5), se obtiene la expresión

$$y_{n+1} = y_n + hk_1 \left( 1 + \frac{1}{2}s + \frac{1}{6}s^2 \right). \quad (2.45)$$

Estamos por lo tanto ante un primer ejemplo de método de dos evaluaciones por paso y orden tres (concretamente el que obteníamos en el apartado anterior tomando  $a_i = 0$  para  $i \geq 3$ ).

Conviene en este punto ilustrar lo anterior mediante algún ejemplo numérico. Para ello consideremos el siguiente problema de valores iniciales

$$y' = 1 - y^2, \quad y(0) = 0, \quad (2.46)$$

cuya solución exacta viene dada por  $y(x) = (e^{2x} - 1)/(e^{2x} + 1)$ . Integrando el problema anterior para distintos valores de  $x$  y con distintas amplitudes de paso  $h$  y utilizando el método citado, se obtienen soluciones numéricas al mismo cuyos errores (en módulo) recogemos en la tabla 2.1. Los resultados recogidos en la tabla confirman numérica-

**Tabla 2.1:** Errores para el método de tipo polinomial.

error	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$x = 1.0$	0.6267D-05	0.8245D-06	0.1057D-06	0.1338D-07
$x = 3.0$	0.5719D-05	0.6606D-06	0.7936D-07	0.9725D-08
$x = 5.0$	0.2464D-06	0.2846D-07	0.3419D-08	0.4189D-09
$x = 7.0$	0.7107D-08	0.8215D-09	0.9868D-10	0.1209D-10
$x = 9.0$	0.1776D-09	0.2054D-10	0.2468D-11	0.3022D-12

mente que el orden de consistencia de nuestro método es tres.

Como ya hemos comentado previamente, los correspondientes métodos de tipo Runge-Kutta explícito de dos etapas no tienen en ningún caso orden de consistencia superior a dos, ni tan siquiera aplicados al problema que nos ocupa. Por ejemplo, el método de Heun de dos etapas y orden dos, viene dado por

$$y_{n+1} = y_n + h \left( \frac{1}{4}k_1 + \frac{3}{4}k_2 \right), \quad (2.47)$$

donde

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3}hk_1\right), \quad (2.48)$$

y hace uso, exactamente, de las mismas evaluaciones y en los mismos puntos ( $k_1$  y  $k_2$  coinciden en ambos métodos). Aplicando este último esquema numérico al mismo problema de valores iniciales, se obtienen los resultados recogidos en la tabla 2.2, que confirman el ya conocido orden dos. Observando los errores obtenidos en uno y otro caso, po-

**Tabla 2.2:** Errores para el método de Heun de dos etapas.

error	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$x = 1.0$	$0.7298D-03$	$0.1745D-03$	$0.4267D-04$	$0.1055D-04$
$x = 3.0$	$0.1532D-03$	$0.3540D-04$	$0.8534D-05$	$0.2096D-05$
$x = 5.0$	$0.5758D-05$	$0.1309D-05$	$0.3142D-06$	$0.7706D-07$
$x = 7.0$	$0.1611D-06$	$0.3615D-07$	$0.8645D-08$	$0.2118D-08$
$x = 9.0$	$0.4002D-08$	$0.8866D-09$	$0.2114D-09$	$0.5175D-10$

demos concluir que el estudio de los métodos que nos ocupa, puede aportar ventajas considerables respecto de los clásicos. Más aún, si consideramos el método de Heun de tres etapas y orden tres dado por

$$y_{n+1} = y_n + h \left( \frac{1}{4}k_1 + \frac{3}{4}k_3 \right), \quad (2.49)$$

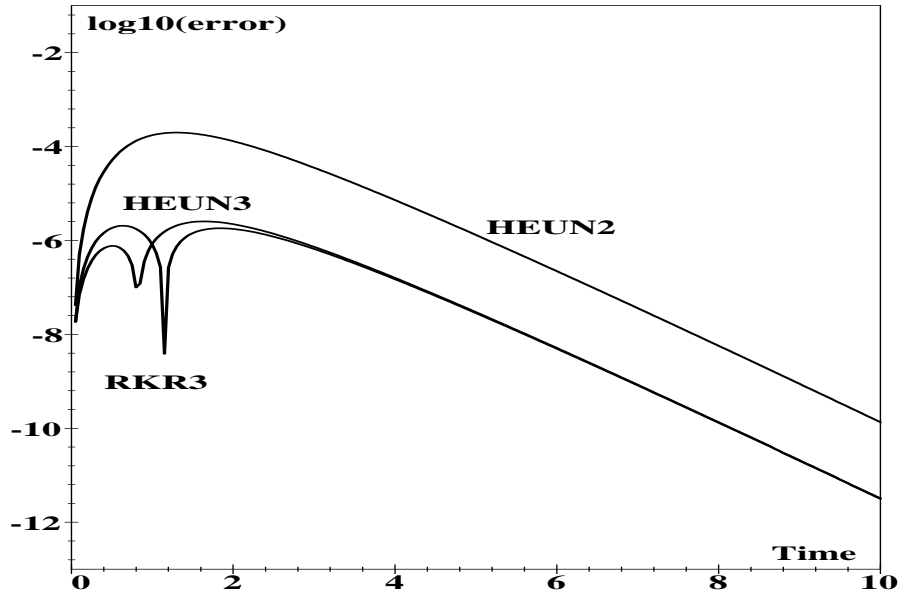
donde

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{1}{3}hk_1\right), \quad k_3 = f\left(y_n + \frac{2}{3}hk_2\right), \quad (2.50)$$

y lo aplicamos al problema que nos ocupa, obtenemos la tabla 2.3 que recoge los errores cometidos. Es evidente, a la vista de las tablas, que nuestro primer ejemplo de método de tipo polinomial de dos etapas se comporta, en cuanto al orden de consistencia y al error, de forma muy parecida al método de Heun de tres etapas y orden tres, y ello haciendo uso de una evaluación menos por paso. Las figuras 2.1 y 2.2 complementan las observaciones anteriores.

**Tabla 2.3:** Errores para el método de Heun de tres etapas.

error	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$x = 1.0$	$0.6910D-05$	$0.8471D-06$	$0.1045D-06$	$0.1298D-07$
$x = 3.0$	$0.6283D-05$	$0.7298D-06$	$0.8793D-07$	$0.1079D-07$
$x = 5.0$	$0.2568D-06$	$0.2975D-07$	$0.3578D-08$	$0.4387D-09$
$x = 7.0$	$0.7298D-08$	$0.8451D-09$	$0.1016D-09$	$0.1245D-10$
$x = 9.0$	$0.1811D-09$	$0.2097D-10$	$0.2521D-11$	$0.3090D-12$

**Figura 2.1:**  $y' = 1 - y^2$ ,  $y(0) = 0$ ,  $h = 0.05$ ,  $x \in [0, 10]$ .

## 2.3 Propiedades de estabilidad lineal de los métodos de dos etapas.

**2.3.1 Función de estabilidad lineal de los métodos de dos etapas.** Pasamos ahora a estudiar las propiedades de estabilidad lineal de nuestros métodos de dos etapas (en su formulación más general). Cuando aplicamos uno de estos métodos (expresado en términos de la

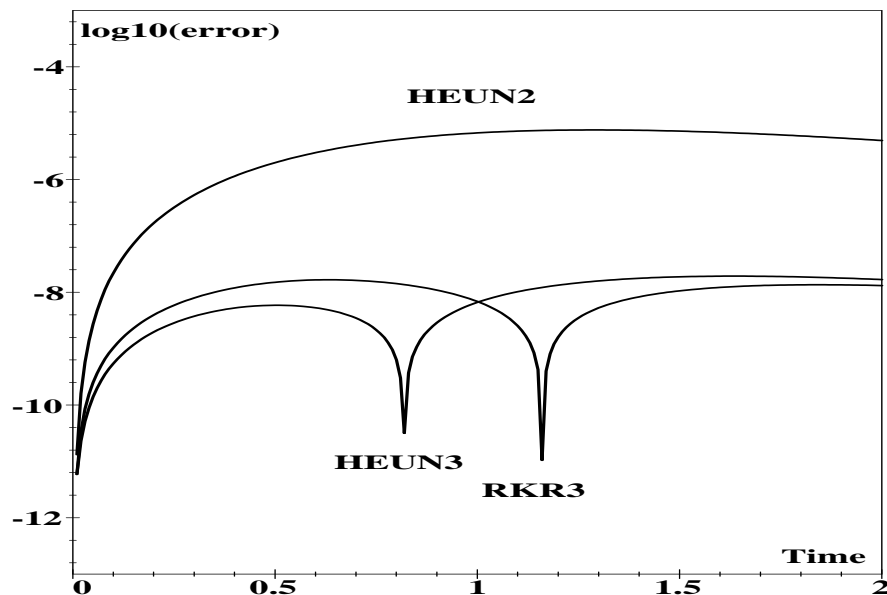


Figura 2.2:  $y' = 1 - y^2$ ,  $y(0) = 0$ ,  $h = 0.01$ ,  $x \in [0, 2]$ .

función  $G$ ) a la ecuación test dada por

$$y' = \lambda y, \quad \lambda \in \mathbb{C}, \quad (2.51)$$

obtenemos una relación del tipo

$$y_{n+1} = R(z) y_n, \quad (2.52)$$

donde  $R(z)$  denota la función de estabilidad lineal asociada al método considerado, con  $z = h\lambda$ . Más aún, de las expresiones que definen  $k_1$  y  $s$  obtenemos las relaciones

$$k_1 = \lambda y_n, \quad (2.53)$$

$$s = z, \quad (2.54)$$

que nos permiten deducir la siguiente expresión para la función de estabilidad lineal

$$R(z) = 1 + z G(z). \quad (2.55)$$

**2.3.2 Métodos de dos etapas A-estables y L-estables.** Es fácil deducir del apartado anterior que la función de estabilidad lineal de un método de tipo polinomial es siempre un polinomio. Por tanto, no es posible obtener fórmulas de tipo polinomial con buenas propiedades de estabilidad lineal tales como A-estabilidad y L-estabilidad. No obstante, también es claro que la función de estabilidad lineal asociada a nuestros métodos de tipo racional viene dada por una función racional, lo que permite albergar ciertas esperanzas a la hora de buscar fórmulas A-estables y L-estables. Nuestras esperanzas se ven confirmadas en lo que sigue y ello sin renunciar al orden máximo de consistencia, que para el caso de dos etapas que nos ocupa venía dado por tres.

Por ejemplo, tomando  $d_1 = -1/2$ ,  $d_2 = 1/12$  y  $n_i = d_i = 0$  para  $i \geq 3$  en (2.42) (recuérdese que  $c_2 = 2/3$ ) se obtiene un método de dos etapas de tipo racional y de orden tres cuya función de estabilidad asociada viene dada por

$$R(z) = \frac{12 + 6z + z^2}{12 - 6z + z^2}, \quad (2.56)$$

esto es, el (2, 2)-aproximante de PADÉ de la función  $e^z$ . De este modo obtenemos un primer método A-estable de dos etapas y orden tres, que admite la expresión

$$y_{n+1} = y_n + hk_1 \left( \frac{12}{12 - 6s + s^2} \right), \quad (2.57)$$

en términos de  $s$ , y en nuestra primera formulación vendría dado en la forma

$$y_{n+1} = y_n + h \left( \frac{16 k_1^3}{31 k_1^2 - 18 k_1 k_2 + 3 k_2^2} \right). \quad (2.58)$$

Es fácil comprobar que la condición (2.39) se satisface para los valores de los parámetros que hemos tomado, por lo que el método anterior minimiza la parte principal del error local de truncación.

Asimismo es sencillo obtener métodos L-estables de orden tres (dados por una función  $G(s)$  racional). Por ejemplo, para la elección de los parámetros dada por  $d_1 = -2/3$ ,  $d_2 = 1/6$  y  $n_i = d_i = 0$  para  $i \geq 3$  (y  $c_2 = 2/3$ ), se obtiene una fórmula de orden tres cuya función de estabilidad lineal es el (1, 2)-aproximante de PADÉ de la función  $e^z$

$$R(z) = \frac{6 + 2z}{6 - 4z + z^2}, \quad (2.59)$$

y es por tanto L-estable. En términos de  $s$  vendrá dada por la expresión

$$y_{n+1} = y_n + h k_1 \left( \frac{6 - s}{6 - 4s + s^2} \right), \quad (2.60)$$

y en nuestra formulación original toma la forma

$$y_{n+1} = y_n + h \left( \frac{10 k_1^3 - 2 k_1^2 k_2}{19 k_1^2 - 14 k_1 k_2 + 3 k_2^2} \right). \quad (2.61)$$

También es posible construir un método de tipo racional, orden tres y L-estable cuya función de estabilidad lineal asociada es el (1,3)-aproximante de PADÉ de la función  $e^z$

$$R(z) = \frac{24 + 6z}{24 - 18z + 6z^2 - z^3}, \quad (2.62)$$

y que minimiza la parte principal del error local de truncación (esto es, se satisface la condición (2.39)). Para ello basta tomar  $c_2 = 2/3$  y  $d_1 = -3/4$ ,  $d_2 = 1/4$ ,  $d_3 = -1/24$ ,  $n_i = 0$  para  $i \geq 3$  y  $d_i = 0$  para  $i \geq 4$  en (2.42), obteniendo de este modo la fórmula

$$y_{n+1} = y_n + h k_1 \left( \frac{24 - 6s + s^2}{24 - 18s + 6s^2 - s^3} \right), \quad (2.63)$$

esto es

$$y_{n+1} = y_n + h \left( \frac{94 k_1^4 - 36 k_1^3 k_2 + 6 k_1^2 k_2^2}{181 k_1^3 - 171 k_1^2 k_2 + 63 k_1 k_2^2 - 9 k_2^3} \right). \quad (2.64)$$

Nótese que los métodos de orden tres que minimizan la parte principal del error local de truncación exhiben orden cuatro cuando son aplicados a problemas lineales del tipo  $y' = \alpha y + \beta$  (con  $\alpha$  y  $\beta$  constantes).

Conviene observar también en este punto que las funciones  $F$  (homogéneas de grado uno) asociadas a los dos primeros métodos (véase (2.58) y (2.61)) están definidas en  $\mathbb{R}^2 - \{(0,0)\}$ . Es fácil comprobar que en ambos casos existe el límite de  $F(x,y)$  cuando  $(x,y) \rightarrow (0,0)$  y vale 0, por lo que si tomamos  $F(0,0) = 0$ , las funciones estarán definidas y serán continuas en todo  $\mathbb{R}^2$ .

En el tercer método (véase (2.64)) la situación es diferente, pues el denominador de la función asociada  $F$  se anula en una recta que pasa por el origen y además no se puede extender a dicha recta de modo que pase a ser continua. Esto habrá de ser tenido en cuenta a la hora de implementar dicho método.

**2.3.3 Experimentos numéricos con problemas rígidos.** El estudio teórico de las propiedades de estabilidad no lineal de las fórmulas que hemos construido es complicado (por la estructura no lineal que presentan) y no lo abordaremos en esta memoria. No obstante, realizaremos algunos experimentos numéricos con problemas no lineales y rígidos del tipo  $y'(x) = f(y(x))$ , en los cuales  $f$  tiene una constante de Lipschitz unilateral igual a 0. Sabemos que para este tipo de problemas las soluciones son contractivas, esto es, dadas dos soluciones  $u$  y  $v$  del problema se verifica que  $|u(x) - v(x)| \leq |u(x_0) - v(x_0)|$ ,  $\forall x \geq x_0$ .

Con el fin de ilustrar el comportamiento de nuestros métodos A-estables y L-estables, consideraremos el siguiente problema

$$\begin{aligned} y' &= -by\sqrt{c^2 + y^2}, \\ y(0) &= a, \end{aligned} \quad (2.65)$$

que depende de los parámetros positivos  $a$ ,  $b$  y  $c$ , y cuya solución viene dada por

$$y(x) = \frac{ac}{c \operatorname{ch}(bcx) + \sqrt{a^2 + c^2} \operatorname{sh}(bcx)}. \quad (2.66)$$

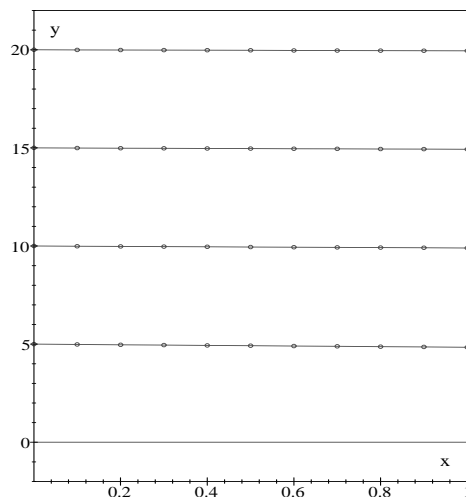
La derivada parcial respecto de  $y$  de la función  $f(y) = -by\sqrt{c^2 + y^2}$  en (2.65) es

$$f_y = -b \frac{c^2 + 2y^2}{\sqrt{c^2 + y^2}}, \quad (2.67)$$

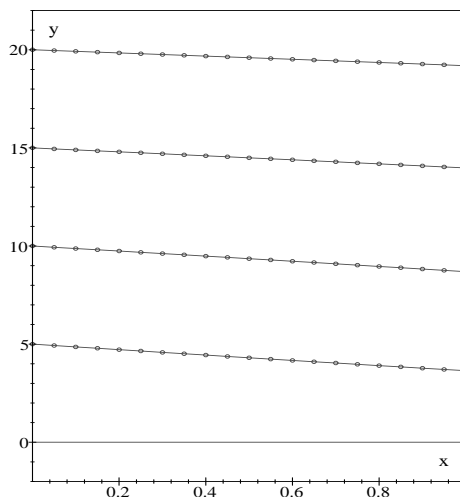
por lo que la función  $f$  tiene constante de Lipschitz unilateral igual a 0 ( $f_y < 0$ ). Como consecuencia de lo anterior, las soluciones exactas de este problema no lineal muestran un comportamiento contractivo. De hecho las soluciones, tras una fase transitoria, tienden a la solución estacionaria  $y(x) \equiv 0$  (esto es, a la solución que se obtiene tomando  $a = 0$  en (2.65)). Más aún, es fácil deducir de (2.67) y de nuestro comentario anterior que  $f_y \approx -bc$  a lo largo de la integración (al menos después de la fase transitoria).

Aplicando las fórmulas obtenidas en el apartado anterior al problema rígido que obtenemos tomando  $b = 5$  y  $c = 2000$  en (2.65), e integrando sobre el intervalo  $[0, 1]$  con condiciones iniciales  $a = 5, 10, 15, 20$ , se obtienen las figuras 2.3 y 2.4 para el método A-estable definido en (2.57) y las figuras 2.5 y 2.6 para el método L-estable definido en (2.60), que muestran el comportamiento cualitativo de las soluciones numéricas así obtenidas. Para el método L-estable definido

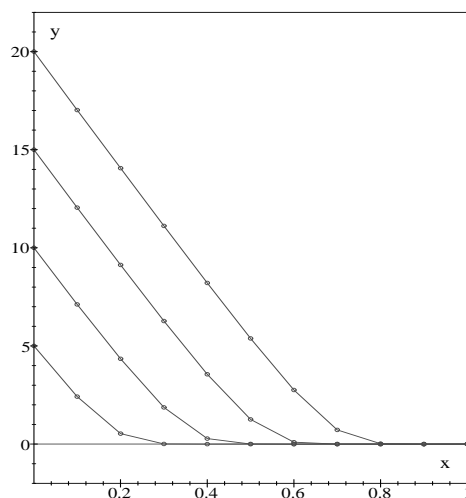




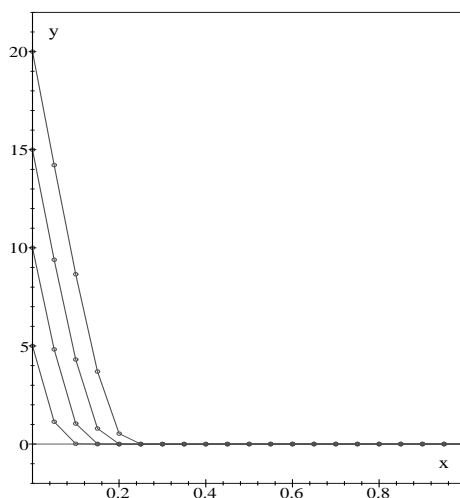
**Figura 2.3:** Paso fijo  
 $h = 0.1$



**Figura 2.4:** Paso fijo  
 $h = 0.05$



**Figura 2.5:** Paso fijo  
 $h = 0.1$



**Figura 2.6:** Paso fijo  
 $h = 0.05$

en (2.63) se obtienen figuras análogas (casi idénticas) a las correspondientes al método (2.60). El comportamiento de las soluciones numéricas es contractivo en todos los casos, si bien es evidente que los

métodos L-estables muestran un mejor comportamiento cualitativo.

Nótese que para el rango de valores de la condición inicial  $a$  que estamos considerando, se tiene que  $f_y \approx -10^4$  a lo largo de la integración. Es fácil comprobar que cuando aplicamos a este problema un método Runge-Kutta explícito la integración fracasa (se produce un "overflow") a menos que tomemos el paso fijo  $h$  lo suficientemente pequeño. Esto es consecuencia inmediata de las regiones de estabilidad absolutas que poseen dichos métodos.

Las soluciones numéricas que se obtienen al aplicar nuestras fórmulas al problema (2.65) con los parámetros  $a$ ,  $b$  y  $c$  fijados de modo que se verifique la condición  $bc \gg 1$  (que garantiza el carácter rígido del problema) y paso fijo  $h = 0.1$ , muestran que el comportamiento cualitativo de las soluciones no siempre es tan bueno. Por ejemplo, cuando  $f_{yy}$  dada por

$$f_{yy} = -by \frac{3c^2 + 2y^2}{(c^2 + y^2)^{3/2}}, \quad (2.68)$$

es grande (y por tanto  $f_y$  es una función que varía rápidamente) las soluciones numéricas tienden al estado estacionario muy lentamente. Sin embargo, disminuyendo el tamaño del paso  $h$  las cosas mejoran rápidamente. De hecho, el tamaño del paso que debemos tomar al proceder con la integración parece depender mucho más del carácter no lineal de  $f$  (mayor cuanto mayor es  $b$ ) que de la rigidez del problema (mayor cuanto mayor es  $bc$ ).

**2.3.4 Construcción de métodos a partir de funciones de estabilidad prefijadas.** Vamos a describir a continuación cómo es posible construir métodos de dos pasos de nuestra familia a partir de funciones de estabilidad lineal prefijadas.

Conviene recordar que, como vimos en el apartado 2.3.1, al aplicar un método de dos etapas descrito en términos de  $s$  a la ecuación test (2.51), obteníamos  $s = z$  (con  $z = \lambda h$ ) independientemente del valor del parámetro  $c_2$  considerado. Por otra parte, la función  $G$  que determina el método de dos etapas considerado y la función de estabilidad lineal  $R$  asociada al método están relacionadas a través de (2.55).

Dada una función  $R(z)$ , podemos construir una familia uniparamétrica de métodos de dos etapas cuya función de estabilidad lineal asociada venga dada por dicha función  $R(z)$ . Basta observar que de (2.55)

se deduce la relación

$$G(z) = \frac{R(z) - 1}{z}, \quad (2.69)$$

que permite dar la familia uniparamétrica de fórmulas

$$y_{n+1} = y_n + h k_1 G(s) = y_n + h k_1 \frac{R(s) - 1}{s}, \quad (2.70)$$

dependiente del parámetro  $c_2$  a través de  $k_1$  y  $s$  que vienen dados por

$$k_1 = f(y_n), \quad k_2 = f(y_n + h c_2 k_1), \quad s = \frac{k_2 - k_1}{c_2 k_1}. \quad (2.71)$$

El parámetro  $c_2$  puede ser determinado en ocasiones de modo que la fórmula resultante satisfaga alguna propiedad de interés.

Para ilustrar el procedimiento anteriormente descrito, consideremos la familia de funciones racionales

$$R(z) = \frac{1 + \frac{1}{2}(1 - \alpha)z + \frac{1}{4}(\beta - \alpha)z^2}{1 - \frac{1}{2}(1 + \alpha)z + \frac{1}{4}(\beta + \alpha)z^2}, \quad (2.72)$$

dependiente de los dos parámetros  $\alpha$  y  $\beta$ .

Es fácil comprobar que la función  $R(z)$  es una aproximación de orden dos (al menos) a la función exponencial  $e^z$  para cualesquiera valores de  $\alpha$  y  $\beta$ , de orden tres (al menos) si  $\alpha \neq 0$  y  $\beta = 1/3$ , y de orden cuatro si  $\alpha = 0$  y  $\beta = 1/3$ .

Un resultado de Liniger y Willoughby muestra que la función  $R(z)$  es A-aceptable si y sólo si se verifica que  $\alpha \geq 0$  y  $\beta \geq 0$ , y que es L-aceptable si y sólo si  $\alpha = \beta > 0$  (véase por ejemplo [27], pp. 233–234).

Para cada elección de los parámetros  $\alpha$  y  $\beta$  en (2.72), el procedimiento que acabamos de describir permite obtener una familia uniparamétrica de métodos de dos etapas (dependiente de  $c_2$ ), cuya función de estabilidad lineal asociada viene dada por (2.72) y que admiten la siguiente expresión

$$y_{n+1} = y_n + h k_1 \left( \frac{R(s) - 1}{s} \right) = y_n + h k_1 \frac{1 - \frac{1}{2}\alpha s}{1 - \frac{1}{2}(1 + \alpha)s + \frac{1}{4}(\beta + \alpha)s^2}. \quad (2.73)$$

donde  $k_1$  y  $s$  vienen dados como en (2.71).

De nuestros comentarios previos se deduce que los métodos así obtenidos son A-estables cuando tomamos  $\alpha \geq 0$  y  $\beta \geq 0$  y L-estables

si tomamos  $\alpha = \beta > 0$ . Más aún, es fácil comprobar a partir de las condiciones de orden (para orden dos son (2.36) y (2.37)) que cualquier método definido por (2.73) es de orden dos independientemente de los valores que se tomen para los parámetros  $\alpha$ ,  $\beta$  y  $c_2$ . Tomando  $\beta = 1/3$  y  $c_2 = 2/3$  se satisfacen también el resto de condiciones de orden para orden tres (esto es, las condiciones (2.35) y (2.38)). Por tanto, tomando  $\beta = 1/3$  y  $c_2 = 2/3$  obtenemos métodos de orden tres que son A-estables si además tomamos  $\alpha \geq 0$  y L-estables para el valor  $\alpha = 1/3$ .

Nótese que cuando la función  $R(z)$  es una aproximación a la función exponencial  $e^z$  de orden al menos tres, bastará tomar el valor  $c_2 = 2/3$  en las expresiones que definen  $k_2$  y  $s$  para poder garantizar que el método construido a partir de dicha función  $R(z)$  sea de orden tres.

**2.3.5 Un método de dos etapas de tipo exponencial.** Los métodos GRK de dos etapas que hemos venido construyendo hasta ahora eran de tipo polinomial o racional. Nos disponemos a dar un primer ejemplo de método de dos etapas para el cual la función  $G$  que lo caracteriza no es de tipo racional. Para ello, partiremos de la función  $R(z) = e^z$  y siguiendo el procedimiento descrito en el apartado anterior, construimos una familia uniparamétrica de métodos cuya función de estabilidad lineal viene dada por dicha función  $R(z)$  (con  $z = \lambda h$ ). Obtenemos así la familia de fórmulas

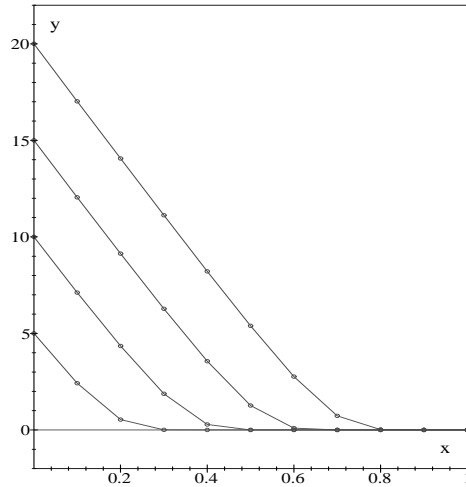
$$y_{n+1} = y_n + h k_1 \left( \frac{e^s - 1}{s} \right), \quad (2.74)$$

dependientes del parámetro  $c_2$  a través de (2.71). Los métodos así obtenidos aplicados a problemas lineales a coeficientes constantes ( $y' = \alpha y + \beta$ ), proporcionan la solución exacta (en ausencia de errores de redondeo). Por tanto, las fórmulas anteriores no son mejorables en lo que respecta a sus propiedades de estabilidad lineal.

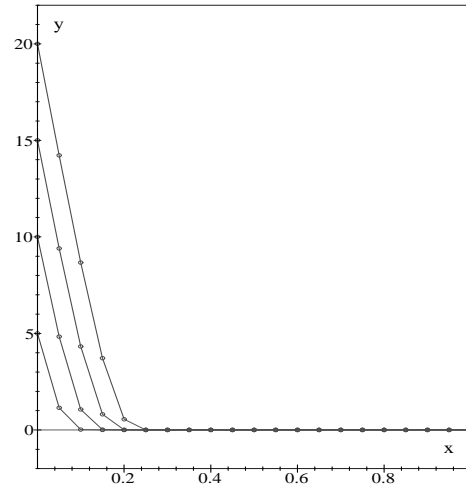
Si tomamos  $c_2 = 2/3$  obtenemos un método de orden tres para problemas escalares autónomos, que integra exactamente problemas lineales a coeficientes constantes.

Para ilustrar el comportamiento numérico de dicho método (el obtenido de (2.74) tomando  $c_2 = 2/3$  en (2.71)) lo aplicamos al problema no lineal rígido considerado en (2.65) tomando los mismos valores que

consideramos en el apartado 2.3.3. Se obtienen así las figuras 2.7 y 2.8. Las figuras son casi idénticas a las obtenidas para el método L-estable



**Figura 2.7:** Paso fijo  
 $h = 0.1$



**Figura 2.8:** Paso fijo  
 $h = 0.05$

del apartado 2.3.3. Conviene señalar sin embargo que, a la vista de los resultados numéricos, las soluciones aproximadas que proporciona este método exponencial tienden a más velocidad al estado estacionario una vez que están próximas a dicha solución estacionaria (aunque en las figuras no es posible apreciar esto).



## Capítulo 3

### Métodos de $p$ etapas.

Tras haber estudiado con cierto detalle los métodos GRK de dos etapas en el capítulo precedente, nos proponemos ahora abordar el estudio del caso general completando asimismo el análisis de algunos aspectos en los que no hemos incidido en el capítulo anterior tales como la cero estabilidad, la consistencia y la convergencia de los métodos.

Empezaremos el capítulo dando la formulación general de los métodos de  $p$  etapas e introduciendo reformulaciones de los mismos (en la línea de lo ya visto en el caso de dos etapas) que nos permitirán simplificar el estudio posterior. Pasaremos después al estudio de la convergencia y del orden de los métodos. Posteriormente, obtendremos la forma general de los métodos de tres etapas y orden cinco, centrándonos nuestra atención en algunos métodos concretos con propiedades específicas tales como A-estabilidad, L-estabilidad, minimización de la parte principal del error local de truncación, etc. Para terminar, realizaremos experimentos numéricos con las fórmulas obtenidas y comparaciones con otros métodos clásicos.

### 3.1 Expresión general de los métodos de $p$ etapas.

**3.1.1 Descripción de los métodos de  $p$  etapas.** Para el problema escalar autónomo dado por

$$y'(x) = f(y(x)), \quad y(x_0) = y_0, \quad (3.1)$$

consideraremos la familia de métodos de  $p$  etapas definidos por

$$y_{n+1} = y_n + hF_{p+1}(k_1, k_2, \dots, k_p), \quad (3.2)$$

donde

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hF_2(k_1)) \\ k_3 &= f(y_n + hF_3(k_1, k_2)) \\ &\vdots \\ k_p &= f(y_n + hF_p(k_1, k_2, \dots, k_{p-1})), \end{aligned} \quad (3.3)$$

y donde para cada  $i$  con  $2 \leq i \leq p+1$ , las funciones  $F_i$  son homogéneas de grado uno, esto es, verifican que para todo  $\alpha > 0$  y  $(x_1, x_2, \dots, x_{i-1})$  en un cono  $A_i \subset \mathbb{R}^{i-1}$  se tiene que

$$F_i(\alpha x_1, \alpha x_2, \dots, \alpha x_{i-1}) = \alpha F_i(x_1, x_2, \dots, x_{i-1}). \quad (3.4)$$

De nuevo conviene observar (como ya hicimos en el capítulo anterior) que para muchos de los métodos que consideraremos en lo que sigue, la relación (3.4) será válida tomando  $A_i = \mathbb{R}^{i-1}$  y  $\alpha \in \mathbb{R}$ . Frecuentemente las funciones  $F_i$  no estarán definidas en el punto  $(0, 0, \dots, 0)$ , pero se podrán redefinir por continuidad en dicho punto tomando  $F_i(0, 0, \dots, 0) = 0$ . En otros casos no será posible hacer lo anterior (véase el comentario que realizamos en el último párrafo del apartado 2.3.2 al estudiar el método (2.64)) y ello habrá de ser tenido en cuenta a la hora de implementar los métodos.

Como ya hemos comentado, la familia de métodos que acabamos de introducir generaliza la familia de métodos Runge-Kutta explícitos de  $p$  etapas. De hecho, tomando  $F_i(x_1, x_2, \dots, x_{i-1}) = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ii-1}x_{i-1}$  ( $2 \leq i \leq p$ ) en (3.3) y  $F_{p+1}(x_1, x_2, \dots, x_p) = b_1x_1 + b_2x_2 + \dots + b_px_p$  en (3.2), obtenemos todos los métodos Runge-Kutta explícitos de  $p$  etapas.

Más aún, nuestros métodos pueden ser considerados como una generalización de los métodos Runge-Kutta explícitos en la que los coeficientes pasan de ser constantes a depender de las etapas. Para mostrar que esto es así basta observar que por ser las funciones  $F_i$  homogéneas de grado uno se verificará por el teorema de Euler para funciones



homogéneas

$$F_i(k_1, k_2, \dots, k_{i-1}) = \sum_{j=1}^{i-1} \frac{\partial F_i}{\partial x_j}(k_1, k_2, \dots, k_{i-1}) k_j, \quad 2 \leq i \leq p+1, \quad (3.5)$$

y por tanto será suficiente tomar el tablero de Butcher

$$\begin{array}{c|cccc} & a_{21} & & & \\ & a_{31} & a_{32} & & \\ & \vdots & \vdots & \ddots & \\ & a_{p1} & a_{p2} & \cdots & a_{pp-1} \\ \hline & b_1 & b_2 & \cdots & b_{p-1} & b_p \end{array} \quad (3.6)$$

donde los parámetros  $a_{ij}$  y  $b_i$  vienen dados en términos de las etapas  $k_i$  y de las funciones  $F_i$  a través de las relaciones

$$\begin{aligned} a_{ij} &= \frac{\partial F_i}{\partial x_j}(k_1, k_2, \dots, k_{i-1}) \quad (1 \leq j < i \leq p), \\ b_i &= \frac{\partial F_{p+1}}{\partial x_i}(k_1, k_2, \dots, k_p) \quad (1 \leq i \leq p). \end{aligned} \quad (3.7)$$

Asumiremos en lo que sigue que para cada  $i$  con  $2 \leq i \leq p+1$  están definidas todas las derivadas parciales de las funciones  $F_i$  hasta un determinado orden  $q \geq 1$  y son continuas en un entorno del punto  $(1, 1, \dots, 1) \in \mathbb{R}^{i-1}$ . Como consecuencia de lo anterior, se tendrá en particular que los valores

$$c_i = F_i(1, 1, \dots, 1) \quad 2 \leq i \leq p+1, \quad (3.8)$$

están perfectamente definidos.

Como se pondrá de manifiesto más adelante, estos  $c_i$  juegan un papel similar al de aquellos asociados a los métodos Runge-Kutta. De hecho, las etapas  $k_i$  pueden ser vistas como aproximaciones a los valores  $y'(x_n + c_i h)$ .

Los nuevos métodos de  $p$  etapas que acabamos de describir son difíciles de estudiar debido al carácter no lineal que los mismos pueden presentar a través de las funciones  $F_i$  que los definen. Por razones que ya hemos comentado en el capítulo precedente, y que se pondrán de manifiesto más adelante al estudiar el orden de convergencia y las propiedades de estabilidad lineal de los métodos, damos a continuación una reformulación de los métodos que simplificará nuestra tarea considerablemente.

**3.1.2 Reformulación de los métodos de  $p$  etapas.** Asumiremos en adelante que los valores  $c_i = F_i(1, 1, \dots, 1)$  son no nulos. Aunque con esto se pierde algo de generalidad, es fácil ver que para un número de etapas fijado (al menos para  $p = 2, 3, 4$ ) el orden máximo se alcanza sólo cuando todos los  $c_i$  son no nulos, y por tanto todos los métodos de interés (al menos desde el punto de vista del orden de consistencia) seguirán estando entre los considerados. Además obtendremos una serie de ventajas adicionales que se irán poniendo de manifiesto en lo que sigue.

Definamos los términos

$$s_i = \frac{k_i - k_1}{c_i k_1}, \quad 2 \leq i \leq p, \quad (3.9)$$

donde las etapas  $k_i$  vienen dadas por (3.3). Cuando sea  $k_1 = 0$  no estarán definidos los términos  $s_i$ , y simplemente tomaremos  $y_{n+1} = y_n$  en línea con lo ya comentado con anterioridad.

Como demostraremos en el Lema 3.1, se tiene que  $s_i = O(h)$ , análogamente a como ocurría con el término  $s$  en el caso de dos etapas, lo que facilitará el estudio posterior. De hecho, los términos  $s_i$  pueden ser considerados como aproximaciones a  $hf_y(y_n)$ , pues se verifica que  $s_i = hf_y(y_n) + O(h^2)$  (si  $f$  es lo suficientemente regular), por lo que podemos obtener aproximaciones a la matriz Jacobiana  $f_y(y_n)$  sin más que tomar  $s_i/h$ .

Un proceso recursivo permite mostrar que las etapas  $k_i$  (con  $2 \leq i \leq p$ ) se pueden obtener a partir de los valores de  $k_1$  y  $s_j$  con  $2 \leq j \leq i - 1$ . Como consecuencia de lo anterior, es posible reescribir cualquier método de  $p$  etapas de los descritos anteriormente (ahora en términos de  $k_1$  y de los  $s_i$ ) en la forma

$$y_{n+1} = y_n + hk_1 G_{p+1}(s_2, s_3, \dots, s_p), \quad (3.10)$$

donde los  $s_i$  vienen dados por (3.9) en función de las etapas

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hk_1 G_2) \\ k_3 &= f(y_n + hk_1 G_3(s_2)) \\ &\vdots \\ k_p &= f(y_n + hk_1 G_p(s_2, s_3, \dots, s_{p-1})), \end{aligned} \quad (3.11)$$

y las funciones  $G_i$  se obtienen a partir de las funciones homogéneas de grado uno  $F_i$  (análogamente a como ocurría en el caso de dos etapas en el capítulo anterior) a través de las relaciones

$$\begin{aligned} F_i(k_1, k_2, \dots, k_{i-1}) &= k_1 F_i\left(1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1}\right) \\ &= k_1 F_i(1, 1 + c_2 s_2, \dots, 1 + c_{i-1} s_{i-1}), \end{aligned} \quad (3.12)$$

sin más que tomar

$$G_i(s_2, s_3, \dots, s_{i-1}) = F_i(1, 1 + c_2 s_2, \dots, 1 + c_{i-1} s_{i-1}), \quad (3.13)$$

para  $2 \leq i \leq p+1$ . Nótese que cuando  $i = 2$ ,  $G_2$  es simplemente la constante  $G_2 = F_2(1) = c_2$ .

De la relación (3.13) se deduce que para cada  $i$  con  $2 \leq i \leq p+1$  las derivadas parciales de las funciones  $G_i$  existen y son continuas hasta un determinado orden  $q \geq 1$  en un entorno del punto  $(0, 0, \dots, 0) \in \mathbb{R}^{i-2}$ , y por tanto se tendrá que  $c_i = G_i(0, 0, \dots, 0)$ .

A partir de las relaciones anteriores es posible reescribir cualquier método (3.2) en la forma (3.10).

Recíprocamente, también es posible obtener las funciones homogéneas de grado uno  $F_i$  a partir de las funciones  $G_i$ , esto es, obtener a partir de cualquier método descrito en términos de  $k_1$  y de los  $s_i$ , su formulación en términos de los  $k_i$ . Para ello, basta tomar

$$F_i(k_1, \dots, k_{i-1}) = k_1 G_i\left(\frac{k_2 - k_1}{c_2 k_1}, \dots, \frac{k_{i-1} - k_1}{c_{i-1} k_1}\right), \quad 2 \leq i \leq p+1. \quad (3.14)$$

Cuando  $i = 2$ , tomaremos  $F_2(k_1) = k_1 G_2$ .

Antes de pasar al estudio de la consistencia y convergencia de los métodos, veamos el siguiente lema que garantiza (bajo hipótesis poco restrictivas) que  $s_i = O(h)$ .

**Lema 3.1** *Supongamos que  $f : \mathbb{R} \rightarrow \mathbb{R}$  en (3.1) es continua y Lipschitziana con constante de Lipschitz  $L$ , esto es:  $|f(y) - f(y^*)| \leq L|y - y^*|$ ,  $\forall y, y^* \in \mathbb{R}$ . Supongamos también que para cada  $j$  con  $2 \leq j \leq p+1$  las funciones  $G_j$  en (3.13) son continuas en un entorno  $W_{j-2}$  del punto  $(0, 0, \dots, 0) \in \mathbb{R}^{j-2}$  y que los términos  $s_i$  (con  $2 \leq i \leq p$ ) vienen dados como en (3.9) en función de las etapas  $k_i$*

de (3.11) con  $y$  en lugar de  $y_n$ . Entonces, para cada  $\epsilon > 0$  existe un  $h_1 > 0$  tal que se verifica

$$\begin{aligned} |s_i| &\leq hL \frac{|c_i| + \epsilon}{|c_i|}, \quad 2 \leq i \leq p, \\ |G_j(s_2, s_3, \dots, s_{j-1})| &\leq |c_j| + \epsilon, \quad 2 \leq j \leq p+1, \end{aligned} \quad (3.15)$$

para todo  $h$  con  $0 < h \leq h_1$  y para todo  $y \in \mathbb{R}$  tal que  $f(y) \neq 0$  (esto es  $k_1 \neq 0$ ).

Recordemos que  $c_i = G_i(0, 0, \dots, 0)$ , como hemos apuntado anteriormente.

*Demostración*

La demostración se sigue de la condición de Lipschitz que satisface la función  $f$  y de que las funciones  $G_j$  son continuas en un entorno  $W_{j-2}$  del punto  $(0, 0, \dots, 0) \in \mathbb{R}^{j-2}$ , por un procedimiento recursivo.

De la continuidad de las funciones  $G_j$  en  $W_{j-2}$  se sigue que para cada  $\epsilon > 0$  existe un  $\delta > 0$  de modo que si  $\max_{1 \leq i \leq p-1} |x_i| \leq \delta$  se verifica que  $|G_j(x_1, x_2, \dots, x_{j-2})| \leq |c_j| + \epsilon$  para  $3 \leq j \leq p+1$ . Cuando  $j = 2$  se tiene que  $G_2$  es constante y por tanto también se verifica en este caso  $|G_2| = |c_2| \leq |c_2| + \epsilon$ .

Ahora, de la condición de Lipschitz que satisface la función  $f$ , tomando  $h_1 = \delta \min_{2 \leq i \leq p} (|c_i| / (\epsilon + |c_i|) L)$  se obtiene recursivamente

$$\begin{aligned} |s_i| &= \frac{|k_i - k_1|}{|c_i| |k_1|} = \frac{|f(y + hk_1 G_i(s_2, s_3, \dots, s_{i-1})) - f(y)|}{|c_i| |k_1|} \\ &\leq \frac{hL}{|c_i|} |G_i(s_2, s_3, \dots, s_{i-1})| \leq hL \frac{|c_i| + \epsilon}{|c_i|}, \\ |G_{i+1}(s_2, s_3, \dots, s_i)| &\leq |c_{i+1}| + \epsilon, \end{aligned} \quad (3.16)$$

para cada  $i$  con  $2 \leq i \leq p$ ,  $0 < h \leq h_1$  y para cada  $y \in \mathbb{R}$ .  $\square$

Nótese que para los valores de  $y_n$  tales que  $f(y_n) = 0$  (y por tanto  $k_1 = 0$ ) tomábamos  $y_{n+1} = y_n$ , por lo que es irrelevante que en este caso los  $s_i$  no estén definidos.

Obsérvese asimismo que cuando  $k_1 = f(y) = 0$ , supuesto que el punto  $y$  es un cero aislado de la función  $f$ , podríamos haber definido los  $s_i$  a través del límite (cuando exista)

$$s_i = \lim_{k_1 \rightarrow 0} \frac{k_i - k_1}{c_i k_1}. \quad (3.17)$$

En este caso, consideremos una sucesión  $(y^{(n)})_{n \in \mathbb{N}}$  con  $y^{(n)} \rightarrow y$  tal que para todo  $n \in \mathbb{N}$   $f(y^{(n)}) \neq 0$ . Obviamente se verifican las desigualdades (3.16) para estos valores  $y^{(n)}$ , por lo que en el límite cuando  $y^{(n)} \rightarrow y$  también han de verificarse. De este modo, las conclusiones del lema siguen siendo válidas también cuando  $k_1 = 0$ .

## 3.2 Convergencia de los métodos de $p$ etapas.

**3.2.1 Consistencia y orden de los métodos de  $p$  etapas.** En lo que sigue, asumiremos que la función  $f : \mathbb{R} \rightarrow \mathbb{R}$  en (3.1) es continua y que existe una constante  $L$  (constante de Lipschitz) tal que se verifica

$$|f(y) - f(y^*)| \leq L|y - y^*|, \quad (3.18)$$

para cualesquiera  $y, y^* \in \mathbb{R}$  (condición de Lipschitz).

Es bien conocido que, en las condiciones anteriores, para cada  $y_0 \in \mathbb{R}$  existe una solución única  $y(x)$  al problema de valores iniciales (3.1) (en cualquier intervalo  $[x_0, b]$ ) que es continua y diferenciable.

Estudiaremos a continuación la consistencia de los métodos de  $p$  etapas descritos por (3.2). Para ello, supondremos únicamente que  $y'(x)$  existe y es continua en un cierto intervalo  $[x_0, b]$ .

Las demostraciones las realizaremos suponiendo en todos los casos que las funciones  $F_i$  que definen los métodos son homogéneas de grado uno en  $\mathbb{R}^{i-1}$ .

Usando las notaciones de Henrici para métodos de un paso, nuestros métodos pueden expresarse en la forma  $y_{n+1} = y_n + h\Phi(x_n, y_n, h)$  donde la función  $\Phi$  (que en el caso que nos ocupa no depende de  $x$  explícitamente por ser el problema (3.1) autónomo) viene dada por

$$\Phi(y, h) = F_{p+1}(k_1, k_2, \dots, k_p), \quad (3.19)$$

en términos de las etapas  $k_i$

$$\begin{aligned} k_1 &= f(y) \\ k_2 &= f(y + hF_2(k_1)) \\ k_3 &= f(y + hF_3(k_1, k_2)) \\ &\vdots \\ k_p &= f(y + hF_p(k_1, k_2, \dots, k_{p-1})). \end{aligned} \quad (3.20)$$

Definimos el error local de truncación  $T(x, h)$  asociado a cada uno de los métodos por

$$\begin{aligned} T(x, h) &= y(x+h) - y(x) - h\Phi(y, h) \\ &= y(x+h) - y(x) - hF_{p+1}(k_1, k_2, \dots, k_p), \quad x \in [x_0, b], \end{aligned} \quad (3.21)$$

donde  $h > 0$  y las etapas  $k_i$  vienen dadas por (3.20), donde ahora  $y(x)$  es la solución exacta de (3.1).

**Definición 3.2** Diremos que el método

$$y_{n+1} = y_n + hF_{p+1}(k_1, k_2, \dots, k_p), \quad (3.22)$$

es consistente (con el problema (3.1) satisfaciendo las condiciones impuestas) si

$$\lim_{h \rightarrow 0} \frac{T(x, h)}{h} = 0, \quad (3.23)$$

uniformemente para  $x \in [x_0, b]$ .

Como es bien sabido, la consistencia es una condición necesaria para la convergencia, por lo que deseamos obtener la condición que han de satisfacer nuestros métodos para ser consistentes. Estamos en condiciones de dar el siguiente teorema:

**Teorema 3.3** El método

$$y_{n+1} = y_n + hF_{p+1}(k_1, k_2, \dots, k_p), \quad (3.24)$$

es consistente con el problema (3.1) (bajo las anteriores hipótesis) si y sólo si se verifica que

$$c_{p+1} = F_{p+1}(1, 1, \dots, 1) = 1. \quad (3.25)$$

*Demostración*

Para cada  $x \in [x_0, b]$  y cada  $h > 0$  se tiene por el teorema del valor medio que

$$y(x+h) - y(x) = hy'(\alpha_x), \quad (3.26)$$

para cierto  $\alpha_x \in (x, x+h)$ . Se sigue de (3.21) que

$$\frac{T(x, h)}{h} = y'(\alpha_x) - F_{p+1}(k_1, k_2, \dots, k_p). \quad (3.27)$$

En el límite cuando  $h \rightarrow 0$  se tiene que  $y'(\alpha_x) \rightarrow y'(x)$  uniformemente para  $x \in [x_0, b]$ . Además  $k_i \rightarrow k_1$  para cada  $i$  con  $2 \leq i \leq p$ . Finalmente, por ser  $F_{p+1}$  continua y homogénea de grado uno se verificará

$$\lim_{h \rightarrow 0} \frac{T(x, h)}{h} = (1 - F_{p+1}(1, 1, \dots, 1))f(y(x)), \quad (3.28)$$

(ya que  $y'(x) = f(y(x))$  y  $k_1 = f(y(x))$ ) y es fácil deducir la condición de consistencia (3.25).  $\square$

Nótese en este punto que, usando las notaciones de Henrici para métodos de un paso, la condición de consistencia (3.25) podría darse en la forma equivalente  $\Phi(y, 0) = f(y)$ .

Se comprueba sin dificultad que la condición de consistencia para los métodos expresados usando la nueva formulación descrita en el apartado 3.1.2 viene dada por  $G_{p+1}(0, 0, \dots, 0) = 1$ .

Definimos ahora la consistencia de orden  $q$  análogamente a como se hace para otros métodos de un paso.

**Definición 3.4** Diremos que el método (3.2) es consistente de orden  $q$  (con el problema (3.1)), si  $q$  es el mayor entero para el que existen un  $N \geq 0$  y un  $h_0 > 0$  tal que se verifica que  $\sup_{x_0 \leq x \leq b} |T(x, h)| \leq N h^{q+1}$  para todo  $h \in (0, h_0]$ .

Si todas las derivadas parciales de  $f(y)$  hasta orden  $q$  existen y son continuas, entonces la consistencia se sigue de la consistencia de orden  $q \geq 1$ .

**3.2.2 Convergencia de los métodos de  $p$  etapas.** Ahora estamos en condiciones de abordar el estudio de la convergencia de los métodos GRK, para lo cual consideraremos el error global, esto es, el error en la solución numérica al cabo de varios pasos. En lo que sigue, se considerará el tamaño de paso  $h$  fijado con el fin de simplificar el estudio, pero todos los resultados que obtendremos seguirán siendo válidos con ligeras modificaciones en otros casos. Nuestra tarea se reduce a estimar el error global y esto se puede llevar a cabo de dos maneras distintas:

- Considerando los errores locales y cómo se propagan a lo largo de las curvas solución (exactas).

- Estudiando cómo se propagan los errores locales de truncación a lo largo de las soluciones numéricas.

Para más detalles véanse [18] y [12]. La primera estrategia es más sencilla y puede proporcionar cotas muy ajustadas del error global cuando se dispone de estimaciones precisas del error propagado para las soluciones exactas. No obstante, nosotros seguiremos la segunda estrategia por ser la más frecuentemente utilizada. Esto es debido fundamentalmente a que esta segunda estrategia es fácilmente generalizable al caso de métodos multipaso y puede ser una herramienta importante en el estudio de la existencia de desarrollos asintóticos.

Análogamente a como ocurre con otros métodos de un paso, necesitamos mostrar que la función  $\Phi(y, h) = F_{p+1}(k_1, k_2, \dots, k_p)$  satisface una condición de Lipschitz en la variable  $y$  para valores del paso  $h$  suficientemente pequeños. Esta propiedad se deduce sin dificultad para la mayoría de los métodos de un paso, a partir de la condición de Lipschitz que satisface la función  $f$ . Para nuestros métodos, la tarea es bastante más complicada debido a que las fórmulas muestran en general una dependencia no lineal de las etapas, a diferencia de lo que ocurre con los métodos Runge-Kutta, por poner un ejemplo. Antes de nada necesitamos algunos resultados previos.

**Lema 3.5** *Si todas las derivadas parciales de las funciones  $F_i$  hasta orden  $q \geq 1$  existen y son continuas en un entorno convexo  $W_{i-1}$  del punto  $(1, 1, \dots, 1) \in \mathbb{R}^{i-1}$ , entonces existe una constante  $\Lambda$  tal que para todo  $h$  con  $0 < h \leq h_0$  y para cualesquiera  $y, y^* \in \mathbb{R}$  se verifica*

$$|\Phi(y, h) - \Phi(y^*, h)| \leq \Lambda |y - y^*|. \quad (3.29)$$

*Demostración*

Empezamos observando que para cada  $i$  con  $2 \leq i \leq p+1$  y para cualesquiera  $y, y^* \in \mathbb{R}$  tales que  $f(y) \neq 0$  y  $f(y^*) \neq 0$  (y por tanto  $k_1 \neq 0$  y  $k_1^* \neq 0$ ) se verifica que

$$\begin{aligned} F_i(k_1, \dots, k_{i-1}) - F_i(k_1^*, \dots, k_{i-1}^*) & \quad (3.30) \\ &= k_1 F_i\left(1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1}\right) - k_1^* F_i\left(1, \frac{k_2^*}{k_1^*}, \dots, \frac{k_{i-1}^*}{k_1^*}\right) \\ &= \frac{k_1 - k_1^*}{2} \left( F_i\left(1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1}\right) + F_i\left(1, \frac{k_2^*}{k_1^*}, \dots, \frac{k_{i-1}^*}{k_1^*}\right) \right) \end{aligned}$$



$$+ \frac{k_1 + k_1^*}{2} \left( F_i \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1} \right) - F_i \left( 1, \frac{k_2^*}{k_1^*}, \dots, \frac{k_{i-1}^*}{k_1^*} \right) \right),$$

donde los  $k_i$  y los  $k_i^*$  vienen dados a través de las relaciones (3.20) (en el segundo caso con  $y^*$  en lugar de  $y$ ). Al final de la demostración comentamos como proceder en el caso en el que  $f(y) = 0$  y/o  $f(y^*) = 0$ .

Ahora probaremos algunas desigualdades que necesitaremos más adelante. A partir de la condición de Lipschitz que satisface la función  $f$  se deduce que

$$\begin{aligned} |k_1 - k_1^*| &\leq L|y - y^*|, \\ |k_i - k_i^*| &\leq L|y - y^*| + hL|F_i(k_1, \dots, k_{i-1}) - F_i(k_1^*, \dots, k_{i-1}^*)|, \end{aligned} \quad (3.31)$$

para cada  $i$  con  $2 \leq i \leq p$ .

De la continuidad de la función  $f$  se deduce que para cada  $\epsilon > 0$  existe un  $h_1 > 0$  tal que para cualesquiera  $y, y^* \in \mathbb{R}$  se verifica que  $|(k_i/k_1) - 1| < \epsilon$  y  $|(k_i^*/k_1^*) - 1| < \epsilon$  ( $2 \leq i \leq p$ ) para los valores de  $h$  con  $0 < h \leq h_1$ . Tomando  $\epsilon > 0$  lo suficientemente pequeño, podemos asegurar de este modo que los puntos  $(1, k_2/k_1, \dots, k_{i-1}/k_1)$  y  $(1, k_2^*/k_1^*, \dots, k_{i-1}^*/k_1^*)$  pertenecen al entorno  $W_{i-1}$  para los valores de  $h$  con  $0 < h \leq h_1$ . Por tanto, a partir del teorema del valor medio se deduce que para los valores de  $h$  con  $0 < h \leq h_1$  ha de verificarse que

$$F_i \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1} \right) - F_i \left( 1, \frac{k_2^*}{k_1^*}, \dots, \frac{k_{i-1}^*}{k_1^*} \right) = \sum_{j=2}^{i-1} \left( \frac{k_j}{k_1} - \frac{k_j^*}{k_1^*} \right) \frac{\partial F_i}{\partial x_j}(\xi_i), \quad (3.32)$$

para cada  $i$  con  $2 \leq i \leq p+1$ , para ciertos puntos intermedios  $\xi_i$  en el segmento contenido en  $W_{i-1}$  que une los puntos  $(1, k_2/k_1, \dots, k_{i-1}/k_1)$  y  $(1, k_2^*/k_1^*, \dots, k_{i-1}^*/k_1^*)$ .

De la continuidad de las derivadas parciales de primer orden de las funciones  $F_i$  en  $W_{i-1}$  se deduce que para cada  $\epsilon > 0$  existe un  $h_2 > 0$  tal que para los valores de  $h$  con  $0 < h \leq h_2$  y para cada  $y \in \mathbb{R}$  se verifica que

$$\begin{aligned} \left| F_i \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1} \right) \right| &\leq |c_i| + \epsilon, \quad 2 \leq i \leq p+1, \\ \left| \frac{\partial F_i}{\partial x_j} \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1} \right) \right| &\leq |c_{i,j}| + \epsilon, \quad 2 \leq i \leq p+1, \quad 2 \leq j \leq i-1, \end{aligned} \quad (3.33)$$

donde

$$c_i = F_i(1, \dots, 1) \quad \text{y} \quad c_{i,j} = \frac{\partial F_i}{\partial x_j}(1, \dots, 1).$$

Es fácil ver que tomando  $h_3 = \min(h_1, h_2)$  se verifica para los valores de  $h$  con  $0 < h \leq h_3$  la relación (3.32) donde ahora

$$\left| \frac{\partial F_i}{\partial x_j}(\xi_i) \right| \leq |c_{i,j}| + \epsilon. \quad (3.34)$$

Necesitaremos también la siguiente identidad

$$(k_1 + k_1^*) \left( \frac{k_i}{k_1} - \frac{k_i^*}{k_1^*} \right) = (k_1^* - k_1) \left( \frac{k_i}{k_1} + \frac{k_i^*}{k_1^*} \right) + 2(k_i - k_i^*), \quad (3.35)$$

fácil de comprobar, de la que se deduce a partir de las desigualdades (3.31) que ha de verificarse

$$\left| (k_1 + k_1^*) \left( \frac{k_i}{k_1} - \frac{k_i^*}{k_1^*} \right) \right| \leq 2(2 + \epsilon)L|y - y^*| + 2hL\nu_i, \quad (3.36)$$

donde

$$\nu_i = |F_i(k_1, \dots, k_{i-1}) - F_i(k_1^*, \dots, k_{i-1}^*)|, \quad 2 \leq i \leq p+1. \quad (3.37)$$

Nótese que  $\nu_{p+1} = |\Phi(y, h) - \Phi(y^*, h)|$ .

Ahora de las desigualdades anteriores se obtiene la siguiente fórmula recursiva

$$\begin{aligned} \nu_i \leq & \left( |c_i| + \epsilon + (2 + \epsilon) \sum_{j=2}^{i-1} (|c_{i,j}| + \epsilon) \right) L|y - y^*| \\ & + hL \sum_{j=2}^{i-1} (|c_{i,j}| + \epsilon) \nu_j, \end{aligned} \quad (3.38)$$

válida para todo  $i$  con  $2 \leq i \leq p+1$ .

Introduciendo las notaciones  $\nu = (\nu_2, \nu_3, \dots, \nu_{p+1})^T$ ,  $C^\epsilon = (\epsilon + |c_2|, \epsilon + |c_3|, \dots, \epsilon + |c_{p+1}|)^T$  y

$$C_{\partial}^{\epsilon} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ c_{3,1}^{\epsilon} & 0 & 0 & \cdots & 0 & 0 & 0 \\ c_{4,1}^{\epsilon} & c_{4,2}^{\epsilon} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ c_{p-1,1}^{\epsilon} & c_{p-1,2}^{\epsilon} & c_{p-1,3}^{\epsilon} & \ddots & 0 & 0 & 0 \\ c_{p,1}^{\epsilon} & c_{p,2}^{\epsilon} & c_{p,3}^{\epsilon} & \cdots & c_{p,p-2}^{\epsilon} & 0 & 0 \\ c_{p+1,1}^{\epsilon} & c_{p+1,2}^{\epsilon} & c_{p+1,3}^{\epsilon} & \cdots & c_{p+1,p-2}^{\epsilon} & c_{p+1,p-1}^{\epsilon} & 0 \end{pmatrix} \quad (3.39)$$

donde los  $c_{i,j}^{\epsilon}$  vienen dados por  $c_{i,j}^{\epsilon} = \epsilon + |c_{i,j}|$ , podemos reescribir las desigualdades (3.38) en la forma

$$\nu \leq (C^\epsilon + (2 + \epsilon) C_{\partial}^{\epsilon} \mathbb{1}) L |y - y^*| + h L C_{\partial}^{\epsilon} \nu, \quad (3.40)$$

donde las desigualdades han de considerarse componente a componente y  $\mathbb{1}$  viene dado por  $\mathbb{1} = (1, 1, \dots, 1)^T$ . Ahora es fácil deducir que se verifica

$$(I - h L C_{\partial}^{\epsilon}) \nu \leq (C^\epsilon + (2 + \epsilon) C_{\partial}^{\epsilon} \mathbb{1}) L |y - y^*|. \quad (3.41)$$

Tomemos ahora  $h_0 > 0$  con  $h_0 < h_3$  lo suficientemente pequeño como para que se satisfaga que  $h L \|C_{\partial}^{\epsilon}\| < 1$  para todo  $h$  con  $0 < h < h_0$  (donde  $\|\cdots\|$  es cualquier norma matricial). De lo anterior se sigue que existe la matriz inversa  $(I - h L C_{\partial}^{\epsilon})^{-1}$  y que todos sus elementos son no negativos. Por tanto, obtenemos

$$\nu \leq (I - h L C_{\partial}^{\epsilon})^{-1} (C^\epsilon + (2 + \epsilon) C_{\partial}^{\epsilon} \mathbb{1}) L |y - y^*|, \quad (3.42)$$

para los valores de  $h$  tales que  $0 < h < h_0$  y se deduce sin dificultad que ha de verificarse

$$\nu_{p+1} \leq e_p (I - h L C_{\partial}^{\epsilon})^{-1} (C^\epsilon + (2 + \epsilon) C_{\partial}^{\epsilon} \mathbb{1}) L |y - y^*|, \quad (3.43)$$

con  $e_p = (0, \dots, 0, 1) \in \mathbb{R}^p$ . Finalmente, teniendo en cuenta que  $\nu_{p+1} = |\Phi(y, h) - \Phi(y^*, h)|$ , se sigue el resultado deseado para los

valores de  $h$  verificando que  $0 < h < h_0$  y para cualesquiera  $y, y^* \in \mathbb{R}$  (tales que  $f(y) \neq 0$  y  $f(y^*) \neq 0$ ), sin más que tomar como constante de Lipschitz

$$\Lambda = e_p(I - h_0 L C_\partial^\epsilon)^{-1} (C^\epsilon + (2 + \epsilon) C_\partial^\epsilon \mathbb{1}) L. \quad (3.44)$$

En el caso en el que sea  $f(y) = 0$  y/o  $f(y^*) = 0$  la demostración ha de ser modificada como vemos a continuación. Supongamos que  $f(y^*) = 0$  y  $f(y) \neq 0$ , esto es,  $k_1^* = 0$  y  $k_1 \neq 0$ . En este caso se verifican las relaciones

$$\begin{aligned} |k_1| &= |k_1 - k_1^*| \leq L|y - y^*|, & (3.45) \\ \left| F_{p+1} \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_p}{k_1} \right) \right| &\leq |c_{p+1}| + \epsilon, \end{aligned}$$

(la segunda para  $h$  suficientemente pequeño). Observando que en este caso es  $\Phi(y^*, h) = 0$  se deduce que

$$\begin{aligned} |\Phi(y, h) - \Phi(y^*, h)| &= |F_{p+1}(k_1, k_2, \dots, k_p)| \\ &= |k_1| \left| F_{p+1} \left( 1, \frac{k_2}{k_1}, \dots, \frac{k_p}{k_1} \right) \right| \\ &\leq (|c_{p+1}| + \epsilon)L|y - y^*|, & (3.46) \end{aligned}$$

tomando  $h$  suficientemente pequeño. Por tanto, bastaría tomar  $\Lambda = (|c_{p+1}| + \epsilon)L$  en este caso y tras compararlo con el obtenido en (3.44) quedarnos con el mayor. En el resto de los casos se hace análogamente.  $\square$

**Definición 3.6** *Diremos que el método (3.2) es convergente si, para todo  $x \in [x_0, b]$ , en el límite cuando  $h \rightarrow 0$  (con  $x_n = x$  fijado) se tiene que  $y_n \rightarrow y(x)$ .*

Ahora estamos en condiciones de probar la convergencia de los métodos GRK

**Teorema 3.7** *Bajo las hipótesis anteriores, cualquier método (3.2) consistente es convergente.*

*Demostración*

De la condición de Lipschitz (3.29) que verifica la función  $\Phi$ , se deduce que para  $h$  suficientemente pequeño ha de verificarse

$$|y_{i+1} - y_{i+1}^*| \leq (1 + h\Lambda) |y_i - y_i^*| \leq \exp(h\Lambda) |y_i - y_i^*|. \quad (3.47)$$

Consideremos la solución numérica para un punto  $x > x_0$ , obtenida al cabo de un número determinado de pasos de tamaño fijo  $h = x_{i+1} - x_i$  ( $0 \leq i \leq n-1$ ) de modo que  $x_n = x$ . Nuestra tarea se reduce a estimar el error global  $E = |y(x_n) - y_n|$  en el punto  $x > x_0$  y para ello seguiremos la segunda estrategia anteriormente descrita. Para el error global en el punto  $x_n = x$  se verifica

$$E = |y(x_n) - y_n| \leq \sum_{i=1}^n |y_{n,i} - y_{n,i-1}|, \quad (3.48)$$

donde con  $y_{n,i}$  denotamos la aproximación obtenida al cabo de  $n-i$  pasos ( $y_{j,i}$   $j = i+1, i+2, \dots, n$ ), partiendo del valor de la solución exacta  $y_{i,i} = y(x_i)$  (como condición inicial en el punto  $x_i$ ) y utilizando el método (3.2). Nótese que con nuestras notaciones se tiene que  $y_{n,n} = y(x_n)$  y que  $y_{n,0} = y_n$ . Pasamos a acotar cada sumando del segundo miembro de (3.48). Para  $h$  suficientemente pequeño se verifica a partir de la desigualdad (3.47) que

$$\begin{aligned} |y_{n,i} - y_{n,i-1}| &\leq (1 + h\Lambda)^{n-i} |y_{i,i} - y_{i,i-1}| \\ &\leq \exp((n-i)h\Lambda) |y_{i,i} - y_{i,i-1}|, \quad 1 \leq i \leq n. \end{aligned} \quad (3.49)$$

Nótese que  $x_i = x_0 + ih$  y por tanto  $(n-i)h = x_n - x_i$ . Conviene observar que  $l_i = y_{i,i} - y_{i,i-1} = T(x_{i-1}, h)$  es el error local de truncación asociado al paso  $i$ -ésimo en el punto  $x_i$ . Por tanto, se deduce de (3.48) y de (3.49) que

$$\begin{aligned} E = |y(x_n) - y_n| &\leq l \frac{\exp(\Lambda(x_n - x_0)) - 1}{\exp(\Lambda h) - 1} \\ &\leq \frac{l}{\Lambda h} (\exp(\Lambda(x_n - x_0)) - 1), \end{aligned} \quad (3.50)$$

donde  $l = \max_{1 \leq i \leq n} |l_i|$ . Finalmente, cuando  $h \rightarrow 0$  (con  $nh = x - x_0$  fijado) se tiene por la consistencia que  $l/h \rightarrow 0$  y la convergencia se sigue trivialmente de la anterior desigualdad.  $\square$

**Definición 3.8** Diremos que el método (3.2) es convergente de orden  $q$  si para todo  $x_n \in [x_0, b]$  (con  $x_n = x$  fijado) y para todo  $h \in (0, h_0]$  existe una constante  $M \geq 0$  tal que se verifica  $|y(x_n) - y_n| \leq M h^q$ .

Veamos que cualquier método consistente de orden  $q$  es convergente de orden  $q$ .

**Teorema 3.9** Bajo las hipótesis anteriores, cualquier método (3.2) consistente de orden  $q$  es convergente de orden  $q$ .

*Demostración*

La demostración es consecuencia inmediata de (3.50), observando que ahora  $l \leq N h^{q+1}$  para  $h$  suficientemente pequeño y que por tanto se verifica

$$\begin{aligned} E = |y(x_n) - y_n| &\leq N h^{q+1} \frac{\exp(\Lambda(x_n - x_0)) - 1}{\exp(\Lambda h) - 1} \\ &\leq \frac{N}{\Lambda} h^q (\exp(\Lambda(x_n - x_0)) - 1), \end{aligned} \quad (3.51)$$

por lo que basta tomar  $M = (N/\Lambda) (\exp(\Lambda(b - x_0)) - 1)$ .  $\square$

Conviene advertir en este punto que los errores de redondeo que se producen en cada paso de la integración no han sido tenidos en cuenta en el estudio anterior. A este respecto sólo comentaremos que como ocurre con otros métodos, existe un tamaño mínimo de paso  $h$  (que depende de la precisión utilizada) por debajo del cual los errores debidos al redondeo son mayores que los errores de truncación.

### 3.3 Métodos de tipo polinomial.

**3.3.1 Expresión de los métodos de  $p$  etapas de tipo polinomial.** Análogamente a como hicimos en el capítulo precedente cuando estudiábamos los métodos de dos etapas, empezaremos prestando atención a la subfamilia de métodos de  $p$  etapas de tipo polinomial, esto es, a los métodos de  $p$  etapas dados a través de las expresiones (3.2–3.3) donde todas las funciones  $F_i$  (con  $2 \leq i \leq p+1$ ) son homogéneas de grado uno que admiten una expresión del tipo

$$F_i(x_1, x_2, \dots, x_{i-1}) = \sum_{j_2+j_3+\dots+j_{i-1}=0}^{r_i} A_{j_2 j_3 \dots j_{i-1}} x_1 \left(\frac{x_2}{x_1}\right)^{j_2} \left(\frac{x_3}{x_1}\right)^{j_3} \dots \left(\frac{x_{i-1}}{x_1}\right)^{j_{i-1}}, \quad (3.52)$$

para ciertos enteros no negativos  $r_i$ .

La anterior subfamilia de fórmulas de tipo polinomial contiene todos los métodos Runge-Kutta explícitos de  $p$  etapas. De hecho, basta tomar  $r_i = 1$  en (3.52) para obtenerlos de (3.2–3.3).

Como veremos en lo que sigue, restringimos nuestra atención a la anterior subfamilia de métodos porque, análogamente a como ocurría en el caso de dos etapas estudiado en el capítulo anterior, podemos reescribirlos en términos de unas funciones  $G_i$  de tipo polinomial. Esto simplificará considerablemente el estudio de las condiciones de orden, y además permitirá obtener de manera sencilla las condiciones de orden para los métodos más generales a partir de las condiciones de orden de los métodos de tipo polinomial.

Recordemos que con la nueva formulación descrita en el apartado 3.1.2 se tiene que

$$s_i = \frac{k_i - k_1}{c_i k_1}, \quad 2 \leq i \leq p, \quad (3.53)$$

donde los términos  $s_i$  verifican que  $s_i = hf_y(y_n) + O(h^2)$ , lo que será de utilidad cuando nos ocupemos del estudio de las propiedades de estabilidad lineal de los métodos y simplificará considerablemente el estudio de las condiciones de orden de los mismos.

En términos de  $k_1$  y de los  $s_i$ , el método precedente (de tipo polinomial) toma la forma (3.10–3.11), donde las funciones  $G_i$  vienen dadas en términos de los  $s_i$  y de las funciones  $F_i$  (homogéneas de grado uno) definidas en (3.52) a través de las relaciones

$$\begin{aligned} F_i(k_1, k_2, \dots, k_{i-1}) &= k_1 F_i\left(1, \frac{k_2}{k_1}, \dots, \frac{k_{i-1}}{k_1}\right) \\ &= k_1 F_i(1, 1 + c_2 s_2, \dots, 1 + c_{i-1} s_{i-1}), \end{aligned} \quad (3.54)$$

válidas para  $2 \leq i \leq p + 1$ , tomando

$$G_i(s_2, s_3, \dots, s_{i-1}) = F_i(1, 1 + c_2 s_2, \dots, 1 + c_{i-1} s_{i-1}), \quad 2 \leq i \leq p + 1. \quad (3.55)$$

Por tanto, las funciones  $G_i$  vendrán dadas por una expresión del tipo

$$G_i(z_2, z_3, \dots, z_{i-1}) = c_i \left( 1 + \sum_{j_2+j_3+\dots+j_{i-1}=1}^{r_i} a_{j_2 j_3 \dots j_{i-1}} z_2^{j_2} z_3^{j_3} \dots z_{i-1}^{j_{i-1}} \right). \quad (3.56)$$

Nótese que para esta subfamilia de métodos todas las  $G_i$  son funciones polinomiales de los  $s_i$  y por ello denominamos a tales fórmulas métodos de tipo polinomial.

Ahora pasamos a estudiar el orden máximo que es posible alcanzar, en términos del número de etapas considerado, para los métodos de tipo polinomial. En adelante supondremos que la función  $f$  es lo suficientemente regular como para poder asegurar que todas las derivadas parciales que aparecerán en los desarrollos de Taylor del error local de truncación tengan sentido.

**3.3.2 Orden de los métodos de dos etapas de tipo polinomial.** En el capítulo anterior, si bien con notaciones ligeramente distintas, ya obtuvimos las condiciones de orden para los métodos de dos etapas de tipo polinomial y vimos que el orden máximo alcanzable era tres. Ahora simplemente recogemos de forma resumida los resultados allí obtenidos con las notaciones que hemos introducido para el caso general de  $p$  etapas.

Las condiciones de orden para orden tres (con las nuevas notaciones) toman la forma

$$c_3 = 1, \quad (3.57)$$

$$c_3 a_1 = 1/2, \quad (3.58)$$

$$c_3 c_2 a_1 = 1/3, \quad (3.59)$$

$$c_3 a_2 = 1/6, \quad (3.60)$$

por lo que cualquier método de orden tres de dos etapas de tipo polinomial admite una expresión del tipo

$$y_{n+1} = y_n + h k_1 G_3(s_2), \quad (3.61)$$

donde las etapas  $k_1$ ,  $k_2$  y el término  $s_2$  vienen dados por

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3} h k_1\right), \quad s_2 = \frac{3(k_2 - k_1)}{2k_1}, \quad (3.62)$$

y la función  $G_3$  es de la forma

$$G_3(s_2) = 1 + \frac{1}{2} s_2 + \frac{1}{6} s_2^2 + \sum_{i=3}^{r_3} a_i s_2^i, \quad (3.63)$$



donde los parámetros libres  $a_i$  (con  $i \geq 3$ ) pueden tomarse arbitrariamente.

Las condiciones adicionales para orden cuatro vienen dadas por

$$c_3 c_2^2 a_1 = 1/4, \quad (3.64)$$

$$c_3 c_2 a_2 = 1/6, \quad (3.65)$$

$$c_3 a_3 = 1/24, \quad (3.66)$$

y aunque, como ya comentamos en el capítulo anterior, no es posible obtener métodos de orden cuatro de dos etapas, se puede minimizar la parte principal del error local de truncación tomando  $a_3 = 1/24$  (con lo que se satisface la condición (3.66)). De este modo, los métodos resultantes exhiben orden cuatro cuando los aplicamos a problemas lineales con coeficientes constantes.

**3.3.3 Orden de los métodos de tres etapas de tipo polinomial.** Veamos ahora que es posible obtener una familia de métodos de tres etapas y orden cinco de tipo polinomial. Como ya ocurría para los métodos de dos etapas, sólo algunos de los parámetros libres de que disponemos aparecerán en las condiciones de orden. Esto es consecuencia de que, como ya hemos comentado con anterioridad, los términos  $s_2$  y  $s_3$  son  $O(h)$ . No obstante, las condiciones de orden resultantes son todavía relativamente complicadas, por lo que definiremos un nuevo término  $\tilde{s}_3 = s_3 - s_2$  que simplificará aun más nuestro estudio. El nuevo término  $\tilde{s}_3$ , que sustituirá al término  $s_3$ , verifica que  $\tilde{s}_3 = O(h^2)$  (por ser  $s_i = hf_y(y_n) + O(h^2)$  para  $i = 2, 3$ ). Esto permite asegurar que el número de parámetros involucrados en las condiciones de orden será inferior. Cualquier método de tres etapas de tipo polinomial puede ser descrito mediante una expresión del tipo

$$y_{n+1} = y_n + hk_1 \tilde{G}_4(s_2, \tilde{s}_3), \quad (3.67)$$

donde  $\tilde{s}_3 = s_3 - s_2$ , y  $s_2$  y  $s_3$  vienen dados por

$$s_i = \frac{k_i - k_1}{c_i k_1}, \quad 2 \leq i \leq 3, \quad (3.68)$$

en términos de las etapas  $k_i$  (con  $1 \leq i \leq 3$ )

$$k_1 = f(y_n), \quad k_2 = f(y_n + hk_1 G_2), \quad k_3 = f(y_n + hk_1 G_3(s_2)), \quad (3.69)$$

y donde las funciones  $G_2$ ,  $G_3$  y  $\tilde{G}_4$  vienen descritas por

$$\begin{aligned} G_2 &= c_2, \\ G_3(s_2) &= c_3 \left( 1 + \sum_{i=1}^{r_3} a_i s_2^i \right), \\ \tilde{G}_4(s_2, \tilde{s}_3) &= c_4 \left( 1 + \sum_{i+2j=1}^{\tilde{r}_4} a_{ij} s_2^i \tilde{s}_3^j \right). \end{aligned} \quad (3.70)$$

Se comprueba sin dificultad que para orden cinco es suficiente considerar los parámetros en (3.70) que se obtienen tomando  $r_3 = 3$  y  $\tilde{r}_4 = 4$ , esto es, las condiciones de orden (para orden cinco) quedan completamente determinadas en función de los parámetros  $c_i$  ( $2 \leq i \leq 4$ ),  $a_i$  ( $1 \leq i \leq 3$ ), y  $a_{ij}$  ( $1 \leq i+2j \leq 4$ ). Esto es consecuencia inmediata de que  $s_2 = O(h)$  y de que  $\tilde{s}_3 = O(h^2)$ .

Las condiciones de orden que ha de satisfacer un método de tres etapas de tipo polinomial para tener orden cinco vienen dadas por

$$c_4 = 1, \quad (3.71)$$

$$c_4 a_{10} = 1/2, \quad (3.72)$$

$$c_4 (c_2(a_{10} - a_{01}) + c_3 a_{01}) = 1/3, \quad (3.73)$$

$$c_4 (a_{20} + a_1 a_{01}) = 1/6, \quad (3.74)$$

$$c_4 (c_2^2(a_{10} - a_{01}) + c_3^2 a_{01}) = 1/4, \quad (3.75)$$

$$c_4 (c_2(2a_{20} - a_{11} + a_1 a_{01}) + c_3(a_{11} + 2a_1 a_{01})) = 1/3, \quad (3.76)$$

$$c_4 (a_{30} + a_1 a_{11} + a_2 a_{01}) = 1/24, \quad (3.77)$$

$$c_4 (c_2^3(a_{10} - a_{01}) + c_3^3 a_{01}) = 1/5, \quad (3.78)$$

$$c_4 (c_2^2(2a_{20} - a_{11} + a_1 a_{01}) + c_3^2(a_{11} + 3a_1 a_{01})) = 7/20, \quad (3.79)$$

$$c_4 (c_2^2(a_{20} - a_{11} + a_{02}) + c_2 c_3(a_{11} - 2a_{02} + 2a_1 a_{01}) + c_3^2 a_{02}) = 2/15, \quad (3.80)$$

$$\begin{aligned} c_4 (c_2(3a_{30} - a_{21} + 2a_1(a_{11} - a_{02}) + 2a_2 a_{01}) \\ + c_3(a_{21} + 2a_1(a_{11} + a_{02}) + (a_1^2 + 2a_2)a_{01})) = 11/60, \end{aligned} \quad (3.81)$$

$$c_4 (a_{40} + a_1 a_{21} + a_1^2 a_{02} + a_2 a_{11} + a_3 a_{01}) = 1/120. \quad (3.82)$$

Es sencillo comprobar que el sistema anterior admite infinitas soluciones. De hecho, existen dos familias doblemente infinitas de soluciones

que se obtienen en términos de los parámetros libres  $a_2$  y  $a_3$ . Los coeficientes restantes se calculan del modo siguiente:

*Paso 1.* Obviamente  $c_4 = 1$  por (3.71). Los coeficientes  $a_{10}$ ,  $a_{01}$ ,  $c_2$  y  $c_3$  se calculan resolviendo el sistema formado por las ecuaciones (3.72), (3.73), (3.75) y (3.78). Se obtienen así un par de soluciones.

*Paso 2.* Ahora, para cada una de las soluciones obtenidas, podemos obtener  $a_1$ ,  $a_{20}$ ,  $a_{11}$  y  $a_{02}$  sin más que resolver el sistema lineal dado por las ecuaciones (3.74), (3.76), (3.79) y (3.80).

*Paso 3.* Finalmente, se obtienen los coeficientes  $a_{40}$ ,  $a_{30}$  y  $a_{21}$  resolviendo el sistema lineal dado por las ecuaciones (3.77), (3.81) y (3.82) (en términos de los parámetros  $a_2$  y  $a_3$ ).

Se obtienen así

$$\begin{aligned} c_2 &= \frac{6 \mp \sqrt{6}}{10}, & c_3 &= \frac{6 \pm \sqrt{6}}{10}, & c_4 &= 1, & a_1 &= \frac{-3 \pm 2\sqrt{6}}{5}, \\ a_{10} &= \frac{1}{2}, & a_{01} &= \frac{9 \pm \sqrt{6}}{36}, & a_{20} &= \frac{3 \mp \sqrt{6}}{12}, & a_{11} &= \frac{3 \pm 2\sqrt{6}}{72}, \\ a_{02} &= \frac{1 \pm 4\sqrt{6}}{72}, & a_{30} &= \frac{-9 \mp \sqrt{6}}{36} a_2, \\ a_{21} &= \frac{-(3 + 11a_2) \pm (1 - 4a_2)\sqrt{6}}{24}, \\ a_{40} &= \frac{-3(3 - 10a_2 + 30a_3) \pm (3 + 20a_2 - 10a_3)\sqrt{6}}{360}. \end{aligned} \quad (3.83)$$

El resto de parámetros se pueden tomar de manera arbitraria.

Para orden seis, hay que añadir a las condiciones (3.71–3.82) las condiciones adicionales (ahora es suficiente tomar  $r_3 = 4$  y  $\tilde{r}_4 = 5$  en (3.70))

$$c_4 \left( c_2^4(a_{10} - a_{01}) + c_3^4 a_{01} \right) = 1/6, \quad (3.84)$$

$$c_4 \left( c_2^3(2a_{20} - a_{11} + a_1 a_{01}) + c_3^3(a_{11} + 4a_1 a_{01}) \right) = 11/30, \quad (3.85)$$

$$\begin{aligned} c_4 \left( 2c_2^3(a_{20} - a_{11} + a_{02}) + c_2^2 c_3(a_{11} - 2a_{02} + 2a_1 a_{01}) \right. \\ \left. + c_2 c_3^2(a_{11} - 2a_{02} + 3a_1 a_{01}) + 2c_3^3 a_{02} \right) = 1/4, \end{aligned} \quad (3.86)$$

$$\begin{aligned} c_4 \left( c_2^2(3a_{30} - a_{21} + 2a_1(a_{11} - a_{02}) + 2a_2 a_{01}) \right. \\ \left. + c_3^2(a_{21} + a_1(3a_{11} + 2a_{02}) + 3(a_1^2 + a_2)a_{01}) \right) = 4/15, \end{aligned} \quad (3.87)$$

$$\begin{aligned}
& c_4 \left( c_2^2(3a_{30} - 2a_{21} + a_{12} + a_1(a_{11} - 2a_{02}) + a_2a_{01}) \right. \\
& + 2c_2c_3(a_{21} - a_{12} + a_1(2a_{11} - a_{02}) + (a_1^2 + 2a_2)a_{01}) \\
& \left. + c_3^2(a_{12} + 4a_1a_{02}) \right) = 17/90, \quad (3.88)
\end{aligned}$$

$$\begin{aligned}
& c_4 \left( c_2(4a_{40} - a_{31} + a_1(3a_{21} - 2a_{12}) + 2(a_1^2 - a_2)a_{02}) \right. \\
& \quad + 3a_2a_{11} + 3a_3a_{01}) + c_3(a_{31} + 2(a_1(a_{21} + a_{12}) \\
& \quad + (2a_1^2 + a_2)a_{02} + (a_1a_2 + a_3)a_{01}) + (a_1^2 + 2a_2)a_{11}) \left. \right) = 13/180, \quad (3.89)
\end{aligned}$$

$$c_4(a_{50} + a_1(a_{31} + a_1a_{12} + 2a_2a_{02}) + a_2a_{21} + a_3a_{11} + a_4a_{01}) = 1/720. \quad (3.90)$$

Es fácil comprobar que no existe ningún método de tres etapas de tipo polinomial y orden seis, sin más que observar que las condiciones (3.83) (para orden cinco) y (3.84–3.90) (para orden seis) no pueden satisfacerse simultáneamente. De hecho, las condiciones (3.84–3.86) no pueden ser satisfechas y los coeficientes de los términos asociados a dichas condiciones en la parte principal del error local de truncación son los mismos para cualquier método de orden cinco. Sin embargo, sí es posible satisfacer las condiciones (3.87–3.90) tomando

$$\begin{aligned}
a_2 &= \frac{-519 \pm 226\sqrt{6}}{300}, \quad a_{12} = \frac{-103 \pm 42\sqrt{6}}{864}, \\
a_{31} &= \frac{-3(119 + 660a_3) \pm 5(25 - 144a_3)\sqrt{6}}{4320}, \\
a_{50} &= \frac{(22597 + 2400a_3 - 7200a_4) \mp 8(1143 - 200a_3 + 100a_4)\sqrt{6}}{28800}, \quad (3.91)
\end{aligned}$$

y obteniendo de este modo métodos de tres etapas y orden cinco que minimizan la parte principal del error local de truncación. Nótese que ahora los coeficientes  $a_{30}$ ,  $a_{21}$  y  $a_{40}$  vienen dados en (3.83) por

$$\begin{aligned}
a_{30} &= \frac{221 \mp 101\sqrt{6}}{720}, \quad a_{21} = \frac{-123 \mp 22\sqrt{6}}{1440}, \\
a_{40} &= \frac{(59 - 180a_3) \mp 2(9 + 10a_3)\sqrt{6}}{720}, \quad (3.92)
\end{aligned}$$

como se comprueba sin más que sustituir los valores (3.91) en las expresiones que definían dichos coeficientes en (3.83).

### 3.4 Métodos de tipo racional.

**3.4.1 Expresión de los métodos de  $p$  etapas de tipo racional.** A continuación estudiaremos la familia de métodos de  $p$  etapas dados por (3.2–3.3), donde ahora todas las  $F_i$  (con  $2 \leq i \leq p+1$ ) son funciones homogéneas de grado uno de tipo racional. Más precisamente, consideraremos funciones  $F_i$  dadas a través de un cociente de dos polinomios homogéneos  $\tilde{N}_i(x_1, x_2, \dots, x_{i-1})$  y  $\tilde{D}_i(x_1, x_2, \dots, x_{i-1})$  de grados  $r_i + 1$  y  $r_i$  respectivamente para cierto entero no negativo  $r_i$ , esto es

$$\begin{aligned} F_i(x_1, x_2, \dots, x_{i-1}) &= \frac{\tilde{N}_i(x_1, x_2, \dots, x_{i-1})}{\tilde{D}_i(x_1, x_2, \dots, x_{i-1})} \\ &= \frac{\sum_{j_1+j_2+\dots+j_{i-1}=r_i+1} N_{j_1 j_2 \dots j_{i-1}} x_1^{j_1} x_2^{j_2} \dots x_{i-1}^{j_{i-1}}}{\sum_{j_1+j_2+\dots+j_{i-1}=r_i} D_{j_1 j_2 \dots j_{i-1}} x_1^{j_1} x_2^{j_2} \dots x_{i-1}^{j_{i-1}}}. \end{aligned} \quad (3.93)$$

La nueva familia de métodos contiene todos los métodos de tipo polinomial. Asumiendo que todos los  $c_i$  son no nulos, obtenemos que en términos de los  $s_i$  definidos en (3.53), haciendo uso de las relaciones (3.55), cualquier método de  $p$  etapas de tipo racional puede reescribirse en la forma (3.10–3.11) con las funciones  $G_i$  dadas por

$$G_i(z_2, z_3, \dots, z_{i-1}) = c_i \left( \frac{1 + \sum_{j_2+j_3+\dots+j_{i-1}=1}^{n_i^*} n_{j_2 j_3 \dots j_{i-1}} z_2^{j_2} z_3^{j_3} \dots z_{i-1}^{j_{i-1}}}{1 + \sum_{j_2+j_3+\dots+j_{i-1}=1}^{d_i^*} d_{j_2 j_3 \dots j_{i-1}} z_2^{j_2} z_3^{j_3} \dots z_{i-1}^{j_{i-1}}} \right). \quad (3.94)$$

Más aún, los valores de  $n_i^*$  y  $d_i^*$  en (3.94) se obtienen fácilmente a partir de las expresiones de las funciones  $F_i$  definidas en (3.93) del método asociado, a través de las relaciones

$$\begin{aligned} n_i^* &= \max\{j_2 + j_3 + \dots + j_{i-1} / N_{j_1 j_2 \dots j_{i-1}} \neq 0\} \\ d_i^* &= \max\{j_2 + j_3 + \dots + j_{i-1} / D_{j_1 j_2 \dots j_{i-1}} \neq 0\}, \end{aligned} \quad (3.95)$$

y por tanto, son siempre menores o iguales que  $r_i + 1$  y  $r_i$  respectivamente.

Ahora podemos obtener las condiciones de orden para estos métodos de tipo racional a partir de las correspondientes condiciones de

orden de los métodos de tipo polinomial (que ya hemos estudiado con anterioridad). Para ello, como ya hemos adelantado, bastará considerar los desarrollos de Taylor de las funciones  $G_i(s_2, \dots, s_{i-1})$  definidas en (3.94) (como funciones de los  $s_j$  con  $2 \leq j \leq i-1$ ), y a continuación comparar éstos con los desarrollos asociados a los métodos de tipo polinomial (con el mismo número de etapas). Para ilustrar lo anterior, obtendremos las condiciones de orden para los métodos de tres etapas de tipo racional a partir de las condiciones correspondientes a los métodos de tres etapas de tipo polinomial. Conviene hacer notar en este punto que también es posible obtener las condiciones de orden para métodos del tipo más general (esto es, para métodos dados por (3.10–3.11) con las funciones  $G_i$  de cualquier tipo) del mismo modo, pues el procedimiento anteriormente descrito sólo involucra los desarrollos de las funciones  $G_i$  en términos de los  $s_j$ .

### 3.4.2 Orden de los métodos de dos etapas de tipo racional.

Para obtener todos los métodos de dos etapas y orden tres (con las nuevas notaciones introducidas) basta observar que de (3.62) y (3.63) se deduce que  $G_2 = c_2 = 2/3$  y que ha de satisfacerse

$$G_3(s_2) = c_3 \left( \frac{1 + \sum_{i=1}^{n_3^*} n_i s_2^i}{d_3^*} \right) = 1 + \frac{1}{2} s_2 + \frac{1}{6} s_2^2 + O(s_2^3). \quad (3.96)$$

Evidentemente, basta tomar  $n_3^* = d_3^* = 2$  en (3.96) y obtenemos las siguientes condiciones de orden (para orden tres)

$$c_2 = 2/3, \quad (3.97)$$

$$c_3 = 1, \quad (3.98)$$

$$n_1 = 1/2 + d_1, \quad (3.99)$$

$$n_2 = 1/6 + (1/2)d_1 + d_2. \quad (3.100)$$

Si deseamos minimizar la parte principal del error local de truncación, bastará tomar  $n_3^* = d_3^* = 3$  en (3.96) y desarrollar el segundo miembro un poco más, obteniendo de este modo la condición adicional

$$n_3 = 1/24 + (1/6)d_1 + (1/2)d_2 + d_3. \quad (3.101)$$

La expresión general de un método de dos etapas y orden tres de tipo racional es

$$y_{n+1} = y_n + hk_1 G_3(s_2), \quad (3.102)$$

donde  $k_1$ ,  $k_2$  y  $s_2$  vienen dados por

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3}hk_1\right), \quad s_2 = \frac{3(k_2 - k_1)}{2k_1}, \quad (3.103)$$

y la función  $G_3$  es

$$G_3(s_2) = \frac{1 + \frac{1+2d_1}{2}s_2 + \frac{1+3d_1+6d_2}{6}s_2^2 + \sum_{i=3}^{n_3^*} n_i s_2^i}{1 + d_1 s_2 + d_2 s_2^2 + \sum_{i=3}^{d_3^*} d_i s_2^i}. \quad (3.104)$$

Como veremos más adelante, cuando extendamos nuestros métodos de modo que éstos sean aplicables a ciertos sistemas de EDOs, es deseable considerar fórmulas en las cuales el denominador en (3.104) viene dado en la forma  $(1 - as_2)^\alpha$  ( $\alpha \in \mathbb{N}$ ) de modo que sólo sea necesario realizar una factorización LU en cada paso de la integración. Más aún, buscaremos métodos de dos etapas y orden tres de tipo racional de la forma anterior, que sean L-estables.

### 3.4.3 Orden de los métodos de tres etapas de tipo racional.

Ahora, siguiendo nuestras notaciones, podemos describir cualquier método de tres etapas de tipo racional por

$$y_{n+1} = y_n + hk_1 \tilde{G}_4(s_2, \tilde{s}_3), \quad (3.105)$$

donde  $\tilde{s}_3 = s_3 - s_2$ , y los términos  $s_2$  y  $s_3$  vienen dados a través de

$$s_i = \frac{k_i - k_1}{c_i k_1}, \quad 2 \leq i \leq 3, \quad (3.106)$$

en función de las etapas  $k_i$  ( $1 \leq i \leq 3$ )

$$k_1 = f(y_n), \quad k_2 = f(y_n + hk_1 G_2), \quad k_3 = f(y_n + hk_1 G_3(s_2)). \quad (3.107)$$

Las funciones racionales  $G_i$  vienen dadas por

$$\begin{aligned}
 G_2 &= c_2, \\
 G_3(s_2) &= c_3 \left( \frac{1 + \sum_{i=1}^{n_3^*} n_i s_2^i}{d_3^*} \right), \\
 \tilde{G}_4(s_2, \tilde{s}_3) &= c_4 \left( \frac{1 + \sum_{i+2j=1}^{\tilde{n}_4^*} n_{ij} s_2^i \tilde{s}_3^j}{\tilde{d}_4^*} \right). \tag{3.108}
 \end{aligned}$$

Cuando buscamos métodos de orden cinco, es suficiente considerar los parámetros en (3.108) con  $n_3^* = d_3^* = 3$  y  $\tilde{n}_4^* = \tilde{d}_4^* = 4$ , esto es, las condiciones de orden (menor o igual que cinco) quedan completamente determinadas en términos de los parámetros  $c_i$  (con  $2 \leq i \leq 4$ ),  $n_i$  y  $d_i$  (con  $1 \leq i \leq 3$ ), y  $n_{ij}$  y  $d_{ij}$  (con  $1 \leq i + 2j \leq 4$ ). De hecho, podemos obtener las condiciones para orden cinco comparando (como en el caso de dos etapas) los desarrollos de las funciones (3.108) con los de las funciones (3.70) que obtenemos tomando los valores de los parámetros (3.83). De este modo se obtiene que los  $c_i$  vienen dados como en (3.83). Los  $n_i$  vienen dados en términos de  $a_1$  como en (3.83) y los parámetros  $a_2, a_3, d_1, d_2$  y  $d_3$ , a través de las relaciones

$$n_i = \sum_{j=0}^i a_j d_{i-j}, \quad 1 \leq i \leq 3, \tag{3.109}$$

donde evidentemente tomamos  $a_0 = d_0 = 1$ . Los términos  $n_{ij}$  vienen dados por

$$n_{i_1 i_2} = \sum_{j_k=0}^{i_k} a_{j_1 j_2} d_{i_1 - j_1 i_2 - j_2}, \quad 1 \leq i_1 + 2i_2 \leq 4, \tag{3.110}$$

con  $a_{00} = d_{00} = 1$ , donde los parámetros  $a_{ij}$  son aquellos considerados en (3.70). Los valores  $d_{ij}$  se pueden tomar de manera arbitraria.

Para minimizar la parte principal del error local de truncación, basta tomar  $n_3^* = d_3^* = 4$  y  $\tilde{n}_4^* = \tilde{d}_4^* = 5$  en (3.108), y se obtienen de



este modo las condiciones adicionales

$$n_4 = \sum_{j=0}^4 a_j d_{4-j}, \quad n_{i_1 i_2} = \sum_{j_k=0}^{i_k} a_{j_1 j_2} d_{i_1 - j_1 i_2 - j_2}, \quad i_1 + 2i_2 = 5, \quad (3.111)$$

donde los  $a_i$  (con  $1 \leq i \leq 2$ ) y los  $a_{ij}$  (con  $1 \leq i+2j \leq 5$ ) vienen dados como en (3.83), (3.91) y (3.92), y el resto de parámetros se pueden tomar de forma arbitraria. Como antes, tomamos  $a_0 = d_0 = a_{00} = d_{00} = 1$ .

## 3.5 Propiedades de estabilidad lineal de los métodos de $p$ etapas.

### 3.5.1 Función de estabilidad de los métodos de $p$ etapas.

Pasamos ahora a estudiar las propiedades de estabilidad lineal de los métodos. Cuando aplicamos un método de  $p$  etapas descrito en términos de los  $s_i$  a la ecuación escalar test

$$y' = \lambda y, \quad \lambda \in \mathbb{C}, \quad (3.112)$$

obtenemos

$$y_{n+1} = R(z) y_n, \quad (3.113)$$

donde  $R(z)$  es la función de estabilidad lineal asociada con  $z = h\lambda$ . Más aún, de

$$s_i = \frac{k_i - k_1}{c_i k_1}, \quad 2 \leq i \leq p, \quad (3.114)$$

donde las etapas  $k_i$  vienen dadas por

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hk_1 G_2) \\ k_3 &= f(y_n + hk_1 G_3(s_2)) \\ &\vdots \\ k_p &= f(y_n + hk_1 G_p(s_2, s_3, \dots, s_{p-1})), \end{aligned} \quad (3.115)$$

se obtienen recursivamente

$$\begin{aligned}
 k_1 &= \lambda y_n \\
 s_2 &= z \\
 s_3 &= z G_3(z) \\
 &\vdots \\
 s_p &= z G_p(z, z G_3(z), \dots, z G_{p-1}(z, \dots, z G_{p-2}(\dots))). \quad (3.116)
 \end{aligned}$$

Ahora, recordando que el método venía dado en la forma

$$y_{n+1} = y_n + h k_1 G_{p+1}(s_2, s_3, \dots, s_p), \quad (3.117)$$

la función de estabilidad lineal es finalmente

$$R(z) = 1 + z G_{p+1}(s_2, s_3, \dots, s_p), \quad (3.118)$$

en términos de  $z$  a través de las relaciones (3.116).

**3.5.2 Métodos A-estables de dos y tres etapas.** Del apartado anterior se deduce que la función de estabilidad lineal asociada a un método de tipo polinomial es siempre un polinomio. Por tanto, no es posible obtener fórmulas de tipo polinomial con buenas propiedades de estabilidad lineal tales como la A-estabilidad o la L-estabilidad. Sin embargo, también es claro que podemos obtener métodos GRK de tipo racional cuya función de estabilidad lineal asociada viene dada por una función racional. Más aún, veremos que es posible obtener fórmulas A-estables y L-estables de tipo racional, sin renunciar a que las mismas sean del orden más alto posible para el número de etapas considerado.

Como hemos comentado anteriormente, es posible obtener métodos GRK L-estables de dos etapas y orden tres (de tipo racional), de modo que el denominador de la función de estabilidad lineal asociada es de la forma  $(1 - a s_2)^\alpha$  (con  $\alpha \in \mathbb{N}$ ), y de este modo sólo es necesaria una factorización LU por cada paso de integración. Para ello, basta tomar  $\alpha = 3$  y los parámetros en (3.104) satisfaciendo

$$d_1 = -3a, \quad d_2 = 3a^2, \quad d_3 = -a^3, \quad (3.119)$$

donde  $a$  es la raíz de  $6x^3 - 18x^2 + 9x - 1 = 0$  dada por

$$a = 1 + \frac{\sqrt{6}}{2} \sin\left(\frac{1}{3} \arctan\left(\frac{\sqrt{2}}{4}\right)\right) - \frac{\sqrt{2}}{2} \cos\left(\frac{1}{3} \arctan\left(\frac{\sqrt{2}}{4}\right)\right) \\ \approx 0.435866521508459, \quad (3.120)$$

y tomamos el resto de parámetros  $n_i = 0$  para  $i \geq 3$ , y  $d_i = 0$  para  $i \geq 4$ .

Ahora daremos algunos métodos A-estables (o L-estables) de tres etapas y orden cinco. Para todos estos métodos los  $c_i$  vienen dados como en (3.83) (tomando el signo superior), esto es

$$c_2 = \frac{6 - \sqrt{6}}{10}, \quad c_3 = \frac{6 + \sqrt{6}}{10}, \quad c_4 = 1. \quad (3.121)$$

Tomaremos asimismo  $d_i = 0$  ( $i \geq 1$ ) en todos los casos, obteniendo de este modo que  $n_i = a_i$ . Lógicamente no podemos anular los  $d_{ij}$  puesto que buscamos fórmulas cuya función de estabilidad lineal sea racional. Por ejemplo, tomando

$$d_{10} = \frac{-3}{5}, \quad d_{01} = \frac{3 - 7\sqrt{6}}{30}, \quad d_{20} = \frac{77 - 18\sqrt{6}}{100}, \\ d_{11} = \frac{153 + 29\sqrt{6}}{360}, \quad d_{30} = \frac{27 - 73\sqrt{6}}{600}, \quad d_{21} = \frac{-44 + 3\sqrt{6}}{120}, \\ d_{40} = \frac{-168 + 97\sqrt{6}}{600}, \quad n_1 = \frac{-3 + 2\sqrt{6}}{5}, \quad n_{10} = \frac{-1}{10}, \\ n_{01} = \frac{63 - 37\sqrt{6}}{180}, \quad n_{20} = \frac{216 - 79\sqrt{6}}{300}, \quad n_{11} = \frac{44 - 3\sqrt{6}}{120}, \\ n_{30} = \frac{168 - 97\sqrt{6}}{600}, \quad (3.122)$$

y el resto de parámetros iguales a cero, se obtiene de (3.105–3.108) un método de tres etapas y orden cinco cuya función de estabilidad lineal viene dada por el (2, 3) aproximante de Padé a la función exponencial, por lo que el método así obtenido es L-estable.

Tomando

$$d_{10} = \frac{-2}{3}, \quad d_{01} = \frac{3 - 7\sqrt{6}}{30}, \quad d_{20} = \frac{41 - 9\sqrt{6}}{50},$$

$$\begin{aligned}
d_{11} &= \frac{431 - 59\sqrt{6}}{600}, & d_{30} &= \frac{1396 - 619\sqrt{6}}{750}, & d_{21} &= \frac{1436 - 709\sqrt{6}}{3600}, \\
d_{40} &= \frac{432353 - 178017\sqrt{6}}{180000}, & d_{31} &= \frac{-20769 + 7966\sqrt{6}}{21600}, \\
d_{50} &= \frac{127698 - 38147\sqrt{6}}{1080000}, & d_{60} &= \frac{-7193669 + 2942716\sqrt{6}}{2160000}, \\
n_1 &= \frac{-3 + 2\sqrt{6}}{5}, & n_2 &= \frac{-519 + 226\sqrt{6}}{300}, & n_{10} &= \frac{-1}{6}, \\
n_{01} &= \frac{63 - 37\sqrt{6}}{180}, & n_{20} &= \frac{221 - 79\sqrt{6}}{300}, & n_{11} &= \frac{3474 - 1111\sqrt{6}}{5400}, \\
n_{30} &= \frac{43409 - 18001\sqrt{6}}{18000}, & n_{21} &= \frac{20769 - 7966\sqrt{6}}{21600}, \\
n_{40} &= \frac{1892669 - 781091\sqrt{6}}{540000}, & n_{50} &= \frac{7193669 - 2942716\sqrt{6}}{2160000}, \quad (3.123)
\end{aligned}$$

(el resto de parámetros nulos) obtenemos otro método L-estable que minimiza la parte principal del error local de truncación. La función de estabilidad asociada a dicho método viene dada por el (2, 4) aproximante de Padé a la función exponencial.

Finalmente, si tomamos

$$\begin{aligned}
d_{10} &= \frac{-1}{2}, & d_{01} &= \frac{3 - 7\sqrt{6}}{30}, & d_{20} &= \frac{36 - 9\sqrt{6}}{50}, \\
d_{11} &= \frac{1323 - 247\sqrt{6}}{1800}, & d_{30} &= \frac{5969 - 2566\sqrt{6}}{3000}, \\
d_{21} &= \frac{1159 - 486\sqrt{6}}{2400}, & d_{40} &= \frac{480158 - 199037\sqrt{6}}{180000}, \\
d_{31} &= \frac{-3729 + 1411\sqrt{6}}{3600}, & d_{50} &= \frac{135777 - 46528\sqrt{6}}{720000}, \\
d_{60} &= \frac{-1282889 + 525021\sqrt{6}}{360000}, & n_1 &= \frac{-3 + 2\sqrt{6}}{5}, \\
n_2 &= \frac{-519 + 226\sqrt{6}}{300}, & n_{01} &= \frac{63 - 37\sqrt{6}}{180}, \\
n_{20} &= \frac{216 - 79\sqrt{6}}{300}, & n_{11} &= \frac{421 - 144\sqrt{6}}{600}, \\
n_{30} &= \frac{45569 - 18791\sqrt{6}}{18000}, & n_{21} &= \frac{3729 - 1411\sqrt{6}}{3600}, \quad (3.124)
\end{aligned}$$

$$n_{40} = \frac{694953 - 286792\sqrt{6}}{180000}, \quad n_{50} = \frac{1282889 - 525021\sqrt{6}}{360000},$$

el método resultante es A-estable, y minimiza la parte principal del error local de truncación. La función de estabilidad asociada para esta fórmula viene dada por el (3, 3) aproximante de Padé a la función exponencial.

**3.5.3 Experimentos numéricos.** Para ilustrar el comportamiento cuando  $h \rightarrow 0$  de los métodos de esta sección, consideraremos el siguiente problema (tomado de [12], pp. 134)

$$\begin{aligned} y'(x) &= \frac{y(x)(1 - y(x))}{2y(x) - 1}, \\ y(0) &= \frac{5}{6}, \end{aligned} \tag{3.125}$$

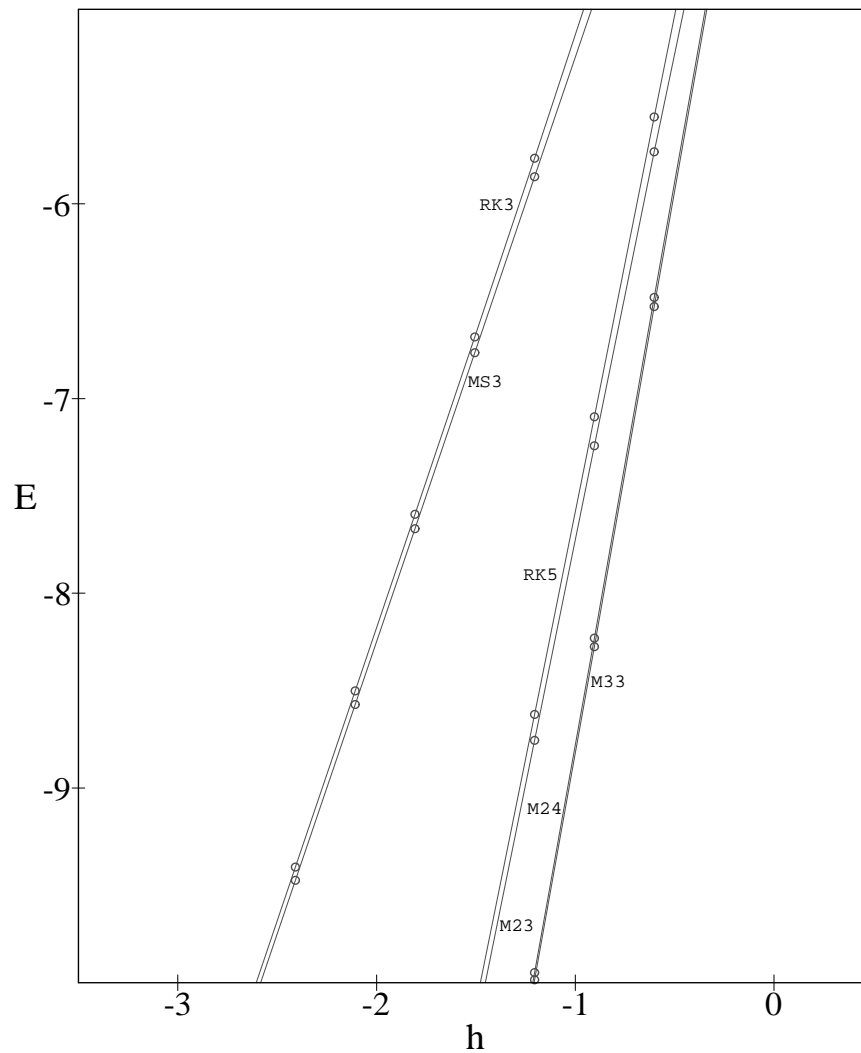
cuya solución viene dada por

$$y(x) = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{5}{36} e^{-x}}. \tag{3.126}$$

Con paso fijo  $h = 2^{-n}$  para los valores de  $n = 1, 2, 3, \dots, 9$  sobre  $2^n$  pasos, hemos obtenido aproximaciones al valor  $y(1)$  usando los métodos anteriores así como dos métodos Runge-Kutta.

La figura 3.1 muestra el error  $E$  obtenido usando diferentes tamaños de paso  $h$  (en escala logarítmica) al integrar el problema con cada uno de los métodos. El método de dos etapas y orden tres que obteníamos tomando los parámetros no nulos dados por (3.119) y (3.120), aparece representado por *MS3* en la figura. Los métodos de orden cinco y tres etapas, con funciones de estabilidad lineal dadas por los (2, 3), (2, 4) y (3, 3) aproximantes de Padé a la función exponencial aparecen representados por *M23*, *M24* y *M33* respectivamente. A efectos comparativos también incluimos en la figura un método Runge-Kutta explícito de tres etapas y orden tres (véase [12], pp. 134) que denotamos por *RK3*, y un método Runge-Kutta explícito de seis etapas y orden cinco (véase [12], pp. 202) que denotamos por *RK5*.

En la escala logarítmica que usamos en la figura, el error para cada método viene representado bastante bien por una línea recta



**Figura 3.1:** Errores frente al tamaño de paso (en doble escala logarítmica) para varios métodos.

cuya pendiente coincide con el orden del método. Para los métodos  $MS3$  y  $M23$  las pendientes son muy próximas a 3 y 5 respectivamente (como cabía esperar). Se puede comprobar que los métodos  $M24$  y  $M33$  proporcionan resultados muy parecidos cuando se aplican a este problema y que la pendiente de la recta que los aproxima es

mayor que 5 (de hecho  $\approx 5.5$ ). Es fácil explicar este comportamiento observando que ambos métodos comparten la propiedad de minimizar la parte principal del error local de truncación. A la vista de la gráfica podemos concluir que los métodos obtenidos minimizando la parte principal del error local de truncación (esto es, los métodos  $M24$  y  $M33$ ) mejoran considerablemente la aproximación proporcionada por el método  $M23$ .

También se desprende de la figura que nuestros métodos GRK se comportan mejor que los Runge-Kutta  $RK3$  y  $RK5$  del mismo orden, y ello a pesar de utilizar un número menor de evaluaciones de la función por paso.

Nuestros métodos, prescindiendo del coste computacional asociado a las evaluaciones de la función  $f$ , requieren de un mayor número de operaciones aritméticas por paso que los métodos Runge-Kutta explícitos del mismo número de etapas. Esto se debe a que en nuestras fórmulas la dependencia respecto de las etapas es no lineal en general, en tanto que en los Runge-Kutta explícitos siempre es lineal. Sin embargo, este coste computacional añadido queda sobradamente compensado en la mayoría de los problemas por el incremento en el orden alcanzable para un mismo número de etapas (y por tanto de evaluaciones de la función  $f$  por paso) y por las mejores propiedades de estabilidad lineal que se pueden obtener con nuestras fórmulas. De hecho, si la función a evaluar  $f$  tiene asociado un gran coste computacional y/o da lugar a un problema rígido, las diferencias tienden a ser muy acusadas en favor de los métodos que proponemos.

Para mostrar el buen comportamiento de los métodos en cuanto a sus propiedades de estabilidad lineal, consideraremos también el siguiente problema (ya utilizado en el apartado 2.3.3 para ilustrar el comportamiento numérico de los métodos de dos etapas)

$$\begin{aligned} y'(x) &= -b y(x) \sqrt{c^2 + y^2(x)}, \\ y(0) &= a, \end{aligned} \tag{3.127}$$

dependiente de los tres parámetros positivos  $a$ ,  $b$  y  $c$ , cuya solución viene dada por

$$y(x) = \frac{a c}{c \operatorname{ch}(b c x) + \sqrt{a^2 + c^2} \operatorname{sh}(b c x)}. \tag{3.128}$$

La derivada parcial respecto de  $y$  de la función  $f(y) = -by\sqrt{c^2 + y^2}$  es

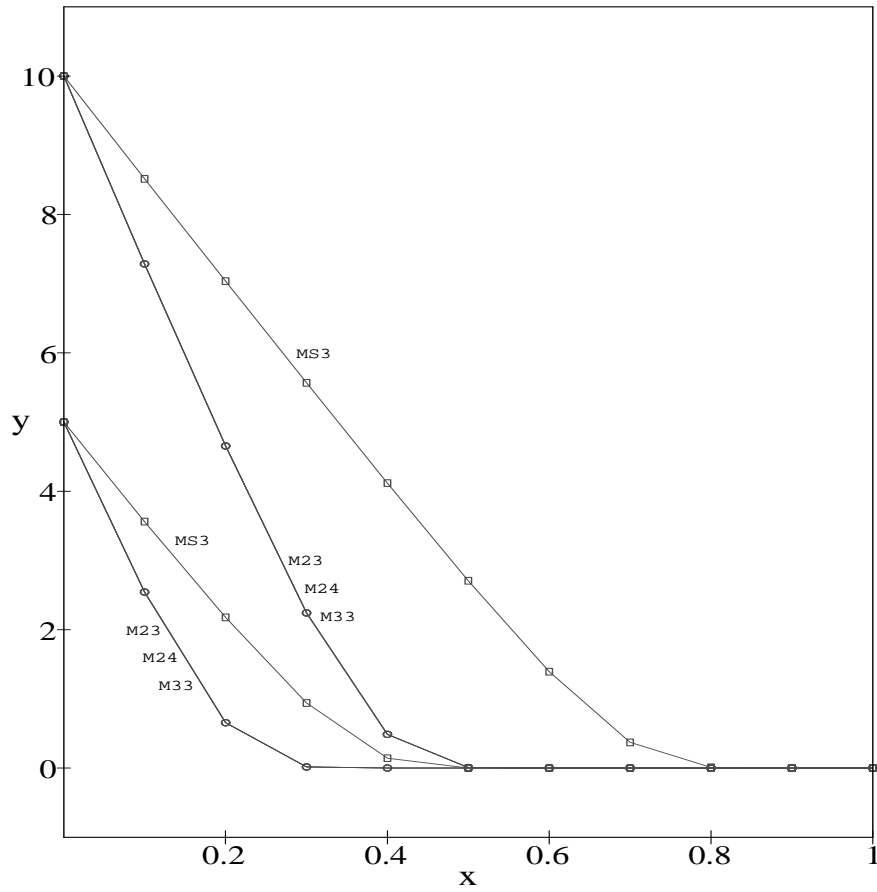
$$f_y = -b \frac{c^2 + 2y^2}{\sqrt{c^2 + y^2}}, \quad (3.129)$$

y por tanto se tiene que  $f_y < 0$ , y la función  $f$  admite una constante de Lipschitz unilateral dada por 0. Como consecuencia de lo anterior, las soluciones exactas de este problema no lineal muestran un comportamiento contractivo. De hecho, tras una corta fase transitoria, las soluciones tienden rápidamente a la solución estacionaria  $y(x) \equiv 0$ , esto es, a la solución del problema (3.127) que se obtiene tomando como condición inicial  $a = 0$ . También se deduce de (3.129) y de nuestro comentario previo que  $f_y \approx -bc$  a lo largo de la integración (al menos tras la fase transitoria).

Para ilustrar el comportamiento de los métodos representados por *MS3*, *M23*, *M24* y *M33* cuando son aplicados a un problema no lineal y rígido, tomamos  $b = 10$  y  $c = 3000$  en (3.127) e integramos este problema sobre el intervalo  $[0, 1]$  con paso fijo  $h = 0.1$  y tomando como condiciones iniciales  $a = 5, 10$ . La figura 3.2 muestra el buen comportamiento cualitativo de las soluciones numéricas que obtenemos de este modo. Los tres métodos de orden cinco se comportan de manera muy parecida para este problema. Nótese que para el rango de valores que estamos considerando para la condición inicial  $a$  tenemos que  $f_y \approx -30000$  a lo largo de la integración y por tanto, los dos métodos Runge-Kutta explícitos representados por *RK3* y *RK5* dan lugar a un overflow cuando son aplicados a este problema con paso fijo  $h \geq 0.0001$ .

Las soluciones numéricas que se obtienen a partir de nuestros métodos cuando son aplicados con paso fijo  $h = 0.1$  al problema (3.127) para distintos valores de los parámetros  $a$ ,  $b$  y  $c$ , con  $bc \gg 1$  (para retener el carácter rígido del problema), muestran que el comportamiento cualitativo no es siempre tan bueno. Por ejemplo, cuando  $f_{yy}$  es grande (esto es,  $f_y$  es una función que varía con rapidez) las soluciones numéricas tienden al estado estacionario muy lentamente. No obstante, tomando tamaños de paso más pequeños la situación mejora rápidamente. De hecho, el tamaño de paso que hemos de utilizar en cada caso parece depender mucho más del carácter no lineal de la función  $f$  que del grado de rigidez del problema.





**Figura 3.2:** Varias soluciones numéricas para  $a = 5, 10$ ,  $b = 10$ ,  $c = 3000$ , con  $h = 0.1$ ,  $x \in [0, 1]$ .



## Capítulo 4

### Métodos de $p$ etapas para sistemas.

Pasamos a continuación a estudiar una ligera modificación de nuestros métodos GRK que nos permitirá aplicarlos al tipo de sistemas de EDOs (autónomos y no autónomos) al que hacíamos referencia en el apartado 1.2 del capítulo introductorio. Tales sistemas, a los que denominamos sistemas "separados", vienen dados en términos de una función  $f$  del tipo definido en (1.4). En muchos casos podemos retener todas las ventajas observadas en el caso escalar autónomo; en otros, conservaremos una gran parte de dichas ventajas. La familia de sistemas "separados" es lo suficientemente amplia como para contener muchos problemas de interés. A modo de ejemplo, podemos adelantar que las fórmulas que proponemos serán aplicables a problemas que surgen al discretizar algunas EDPs mediante el Método de líneas (MOL) con diferencias finitas.

En este capítulo, tras describir brevemente el tipo de problemas al que serán aplicables las fórmulas y el modo en que éstas han de ser modificadas para su aplicación a sistemas, pasaremos a dar la expresión general de los métodos GRK de  $p$  etapas para sistemas separados. A continuación veremos un primer ejemplo de fórmula de dos etapas y orden tres y realizaremos diversos experimentos numéricos. Uno de los problemas considerados al realizar dichos experimentos numéricos surge al discretizar la ecuación de Burguer's mediante el Método de líneas (MOL) con diferencias finitas, lo que da lugar a un sistema separado de gran tamaño. Posteriormente estudiaremos las condiciones de orden y las propiedades de estabilidad lineal de los métodos y obtendremos distintas fórmulas de dos y tres etapas (algunas de las cuales minimizan la parte principal del error local de truncación).

Finalmente exhibimos algunos experimentos numéricos adicionales.

## 4.1 Expresión general de los métodos de $p$ etapas para sistemas.

**4.1.1 Forma general del sistema de EDOs.** Consideraremos en lo que sigue sistemas autónomos de EDOs de la forma

$$\begin{aligned} y'_{(1)} &= f_{11}(y_{(1)}) + f_{12}(y_{(2)}) + \dots + f_{1m}(y_{(m)}), \\ y'_{(2)} &= f_{21}(y_{(1)}) + f_{22}(y_{(2)}) + \dots + f_{2m}(y_{(m)}), \\ &\vdots \\ y'_{(m)} &= f_{m1}(y_{(1)}) + f_{m2}(y_{(2)}) + \dots + f_{mm}(y_{(m)}), \end{aligned} \quad (4.1)$$

es decir, sistemas de la forma  $y' = f(y)$  para los cuales la función  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$  viene dada por  $f(y) = F(y)\mathbb{1}$ , con  $y = (y_{(1)}, y_{(2)}, \dots, y_{(m)})$ ,  $\mathbb{1} = (1, 1, \dots, 1)^T$  y donde  $F$  es la matriz

$$F(y) = \begin{pmatrix} f_{11}(y_{(1)}) & f_{12}(y_{(2)}) & \dots & f_{1m}(y_{(m)}) \\ f_{21}(y_{(1)}) & f_{22}(y_{(2)}) & \dots & f_{2m}(y_{(m)}) \\ \vdots & \vdots & \dots & \vdots \\ f_{m1}(y_{(1)}) & f_{m2}(y_{(2)}) & \dots & f_{mm}(y_{(m)}) \end{pmatrix}, \quad (4.2)$$

con componentes  $f_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ . Nótese el uso que hacemos de los paréntesis en los subíndices para denotar las componentes de los vectores  $y$  e  $y'$ . Luego, en definitiva, tenemos el sistema  $y' = f(y)$  de dimensión  $m$  cuyas componentes vienen dadas por

$$y'_{(i)} = f_i(y_{(1)}, y_{(2)}, \dots, y_{(m)}) = \sum_{j=1}^m f_{ij}(y_{(j)}), \quad 1 \leq i \leq m. \quad (4.3)$$

En ocasiones llamaremos a este tipo de sistemas "separados". Sistemas de este tipo surgen por ejemplo al tratar de resolver algunas ecuaciones en derivadas parciales de tipo parabólico mediante el método de líneas (con diferencias finitas). Como ya hemos comentado, más adelante aplicaremos los métodos GRK a varios sistemas separados que surgen al aplicar el método de líneas a la ecuación de Burgers. Éstos serán algunos de los principales ejemplos que utilizaremos en

este capítulo a la hora de ilustrar el comportamiento numérico de las fórmulas obtenidas.

Nótese asimismo que aunque estamos considerando únicamente sistemas autónomos del tipo anteriormente descrito, los sistemas no autónomos de la forma  $y'(x) = f(y(x)) + g(x)$ , como es bien conocido, se pueden reescribir como un sistema autónomo de una dimensión más con sólo añadir la ecuación  $x' = 1$ . Si  $f$  es de la forma descrita anteriormente, entonces el sistema autónomo resultante también, y todo lo que veremos en adelante será perfectamente aplicable. Además, aunque el sistema resultante es de dimensión  $m + 1$ , los métodos pueden ser de hecho implementados sin incrementar dicha dimensión.

#### 4.1.2 Descripción de los métodos de $p$ etapas para sistemas.

Pasamos ahora a describir los métodos GRK que serán aplicables a sistemas del tipo (4.1). Lo haremos utilizando la reformulación de los métodos de  $p$  etapas vista en el caso escalar (véase (3.1.2)), si bien la modificaremos ligeramente para hacerla aplicable a la nueva situación.

Para el problema (4.1) con condición inicial  $y(x_0) = y_0$  ( $y_0 \in \mathbb{R}^m$ ) consideramos la familia de métodos de  $p$  etapas definida por

$$y_{n+1} = y_n + hG_{p+1}(S_2, S_3, \dots, S_p) k_1, \quad (4.4)$$

donde ahora  $y_n$ ,  $y_{n+1}$  y  $k_1$  son vectores de  $\mathbb{R}^m$  y tanto los  $S_i$  como los  $G_i(S_2, S_3, \dots, S_{i-1})$  son matrices cuadradas, esto es,  $S_i \in M_{\mathbb{R}}(m)$  y  $G_i(S_2, S_3, \dots, S_{i-1}) \in M_{\mathbb{R}}(m)$  (con  $M_{\mathbb{R}}(m)$  denotamos el espacio de matrices cuadradas de dimensión  $m$  con coeficientes reales). Las etapas  $k_i$  (con  $1 \leq i \leq p$ ) vienen dadas en la forma

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hG_2 k_1) \\ k_3 &= f(y_n + hG_3(S_2) k_1) \\ &\vdots \\ k_p &= f(y_n + hG_p(S_2, S_3, \dots, S_{p-1}) k_1), \end{aligned} \quad (4.5)$$

en términos de las matrices cuadradas  $S_i$  (con  $2 \leq i \leq p$ ) definidas

ahora por

$$\begin{pmatrix} \frac{f_{11}(y_{n(1)} + \alpha_1) - f_{11}(y_{n(1)})}{e_1 G_i(S_2, S_3, \dots, S_{i-1}) k_1} & \cdots & \frac{f_{1m}(y_{n(m)} + \alpha_m) - f_{1m}(y_{n(m)})}{e_m G_i(S_2, S_3, \dots, S_{i-1}) k_1} \\ \vdots & & \vdots \\ \frac{f_{m1}(y_{n(1)} + \alpha_1) - f_{m1}(y_{n(1)})}{e_1 G_i(S_2, S_3, \dots, S_{i-1}) k_1} & \cdots & \frac{f_{mm}(y_{n(m)} + \alpha_m) - f_{mm}(y_{n(m)})}{e_m G_i(S_2, S_3, \dots, S_{i-1}) k_1} \end{pmatrix}, \quad (4.6)$$

donde  $\alpha_j = h e_j G_i(S_2, S_3, \dots, S_{i-1}) k_1$  (con  $1 \leq j \leq m$ ) y con  $e_j$  denotamos el vector de  $\mathbb{R}^m$  que tiene todas sus componentes nulas a excepción de la que ocupa la componente  $j$ -ésima que vale 1. Nótese que el elemento que ocupa la fila  $p$  y columna  $q$  (con  $1 \leq p, q \leq m$ ) en la matriz  $S_i$  viene dado por

$$\frac{f_{pq}(y_{n(q)} + h e_q G_i(S_2, S_3, \dots, S_{i-1}) k_1) - f_{pq}(y_{n(q)})}{e_q G_i(S_2, S_3, \dots, S_{i-1}) k_1}. \quad (4.7)$$

Evidentemente, con  $G_i k_1$  representamos el producto matriz por vector.

Las etapas  $k_i$  pueden expresarse en términos de la función  $F$  (véase (4.2)) a través de la relación

$$k_i = F(y_n + h G_i(S_2, S_3, \dots, S_{i-1}) k_1) \mathbb{1}, \quad 1 \leq i \leq p, \quad (4.8)$$

donde  $F(y_n + h G_i(S_2, S_3, \dots, S_{i-1}) k_1)$  es la matriz cuadrada de dimensión  $m$  dada por

$$\begin{pmatrix} f_{11}(y_{n(1)} + h e_1 G_i(*) k_1) & \cdots & f_{1m}(y_{n(m)} + h e_m G_i(*) k_1) \\ \vdots & & \vdots \\ f_{m1}(y_{n(1)} + h e_1 G_i(*) k_1) & \cdots & f_{mm}(y_{n(m)} + h e_m G_i(*) k_1) \end{pmatrix}, \quad (4.9)$$

donde con  $(*)$  denotamos el punto  $(S_2, S_3, \dots, S_{p-1})$  y el elemento que ocupa la fila  $p$  y columna  $q$  (con  $1 \leq p, q \leq m$ ) en dicha matriz es

$$f_{pq}(y_{n(q)} + h e_q G_i(S_2, S_3, \dots, S_{i-1}) k_1). \quad (4.10)$$

De los comentarios anteriores se deduce que las matrices  $S_i$  son una generalización de los términos  $s_i$ , si bien ligeramente modificados. De hecho es fácil comprobar que los métodos descritos en capítulos precedentes para el problema escalar podrían haber sido formulados en

la forma que acabamos de describir, pero no lo vamos a hacer para no complicar innecesariamente el estudio y las expresiones de las fórmulas involucradas en dicho caso escalar.

Observemos en este punto que las matrices  $S_i$  para este tipo de sistemas nos proporcionan aproximaciones a  $h f_y(y_n)$ , esto es, a la matriz Jacobiana asociada a la función  $f$  en el punto  $y_n$  (multiplicada por el paso  $h$ ). Para verlo, basta observar que (4.7) puede ser considerado una aproximación a la derivada parcial de la componente  $p$ -ésima de  $f$  respecto de la variable  $q$ -ésima en el punto  $y_n$ . Esto es así para este tipo de sistemas (véase (4.3)) por verificarse que

$$\frac{\partial f_p}{\partial y_{(q)}}(y_n) = \frac{\partial f_{pq}}{\partial y_{(q)}}(y_{n(q)}) \approx \frac{f_{pq}(y_{n(q)} + h\delta_q) - f_{pq}(y_{n(q)})}{h\delta_q}, \quad (4.11)$$

donde en nuestro caso es  $\delta_q = e_q G_i(S_2, S_3, \dots, S_{i-1}) k_1$ .

De la observación anterior podemos concluir que los métodos GRK que acabamos de describir para este tipo especial de sistemas se parecen a los métodos de tipo Rosenbrock y especialmente a sus variantes tales como los W-métodos [41], los MROW-métodos [53] y los métodos de Runge-Kutta generalizados [44] para los cuales no es necesario calcular la matriz Jacobiana de forma exacta a la hora de implementarlos. En [20] se describen algunos de dichos métodos.

La principal novedad que introducen nuestros métodos respecto de los citados anteriormente radica en que, para las fórmulas que hemos introducido, las aproximaciones a la matriz Jacobiana se obtienen sin nuevas evaluaciones de la función  $f$ , pues las matrices  $S_i$  se construyen a partir de la información contenida en las propias etapas (si bien esto obliga a trabajar con matrices cuadradas de dimensión  $m$  en algunas fases de la evaluación).

**4.1.3 Un primer ejemplo de método de dos etapas para sistemas.** Empezaremos viendo un ejemplo de método de dos etapas antes de comenzar con un estudio sistemático de los mismos. Nótese que la función  $G_2$  que aparece en (4.5) es en realidad una constante (que denotaremos por  $c_2$ ) que multiplica a la matriz identidad de dimensión  $m$ , esto es, se tiene que  $G_2 = c_2 I$  (donde  $I$  es la matriz identidad de  $M_{\mathbb{R}}(m)$ ). Como consecuencia de lo anterior, un método de dos etapas para sistemas del tipo (4.1) vendrá dado en la forma

$$y_{n+1} = y_n + hG_3(S_2) k_1, \quad (4.12)$$

donde las etapas son

$$k_1 = f(y_n), \quad k_2 = f(y_n + hc_2k_1), \quad (4.13)$$

y  $S_2$  es la matriz

$$\begin{pmatrix} \frac{f_{11}(y_{n(1)}+hc_2e_1k_1)-f_{11}(y_{n(1)})}{c_2e_1k_1} & \dots & \frac{f_{1m}(y_{n(m)}+hc_2e_mk_1)-f_{1m}(y_{n(m)})}{c_2e_mk_1} \\ \vdots & & \vdots \\ \frac{f_{m1}(y_{n(1)}+hc_2e_1k_1)-f_{m1}(y_{n(1)})}{c_2e_1k_1} & \dots & \frac{f_{mm}(y_{n(m)}+hc_2e_mk_1)-f_{mm}(y_{n(m)})}{c_2e_mk_1} \end{pmatrix} \quad (4.14)$$

El método queda completamente determinado una vez fijada la constante  $c_2$  y la función  $G_3$ .

A partir de la teoría de Butcher sabemos que un método Runge-Kutta de orden  $q$  para problemas escalares puede mostrar un orden menor que  $q$  cuando es aplicado a sistemas de EDOs. No obstante, para  $q \leq 3$  cualquier método Runge-Kutta de orden  $q$  para problemas escalares autónomos muestra el mismo orden cuando se le aplica a sistemas de EDOs (para más detalles, véase [27] pp. 173-175). Lo mismo ocurre al considerar nuestros métodos, o sea para  $q \leq 3$ , cualquier método GRK de orden  $q$  para el caso escalar autónomo tiene orden  $q$  al aplicarlo a sistemas separados.

Nuestro primer ejemplo consistirá en dar un ejemplo de método L-estable. Para ello tomaremos la función  $G_3$  de tipo racional, si bien como el argumento  $S_2$  de la función es una matriz cuadrada de dimensión  $m$ , los denominadores habrán de interpretarse como inversas. Concretamente buscamos que el denominador de la función  $G_3(s)$  sea de la forma  $(1 - as)^\alpha$ , esto es  $G_3(S_2)$  en (4.12) tome la forma  $(I - aS_2)^{-\alpha} N(S_2)$  para un cierto  $\alpha \in \mathbb{N}$ , una cierta constante  $a \in \mathbb{R}$  y una cierta función polinomial  $N$  (con  $I$  denotamos la matriz identidad de dimensión  $m$ ). La razón que nos mueve a tomar  $G_3$  así no es otra que la de posibilitar que sólo sea necesaria una factorización LU por paso de integración. Tomando los valores  $\alpha = 3$ ,  $c_2 = 2/3$ ,  $a$  como la raíz del polinomio  $6x^3 - 18x^2 + 9x - 1 = 0$  dada por

$$\begin{aligned} a &= 1 + \frac{\sqrt{6}}{2} \sin\left(\frac{1}{3} \arctan\left(\frac{\sqrt{2}}{4}\right)\right) - \frac{\sqrt{2}}{2} \cos\left(\frac{1}{3} \arctan\left(\frac{\sqrt{2}}{4}\right)\right) \\ &\approx 0.435866521508459, \end{aligned} \quad (4.15)$$



y la función  $G_3$  de (4.12) de la forma

$$G_3(S_2) = (I - a S_2)^{-3} \left( I + \frac{1 - 6a}{2} S_2 + \frac{1 - 9a + 18a^2}{6} S_2^2 \right), \quad (4.16)$$

se obtiene un método L-estable de dos etapas y orden tres para sistemas del tipo considerado, cuya función de estabilidad lineal asociada es

$$R(z) = \frac{2 + 2(1 - 3a)z + (1 - 6a + 6a^2)z^2}{2(1 - az)^3}, \quad (4.17)$$

con  $a$  dado por (4.15). Para más detalles relativos a la L-estabilidad que proporciona la anterior función  $R(z)$ , [20], pp. 96–98 es una buena referencia. En [5] se describe este método con notaciones ligeramente distintas.

**4.1.4 Un primer experimento numérico con sistemas.** Estamos en condiciones de realizar un primer experimento numérico para mostrar el comportamiento del método introducido en el apartado anterior. Para ello consideramos la ecuación de Burgers dada por

$$u_t + u u_x = \nu u_{xx} \quad \text{o equivalentemente} \quad u_t + \left( \frac{u^2}{2} \right)_x = \nu u_{xx}, \quad \nu > 0, \quad (4.18)$$

donde  $u = u(x, t)$  y tomamos el rectángulo definido por  $0 \leq x \leq 1$  y  $0 \leq t \leq 1$ . Las condiciones iniciales y de frontera de Dirichlet que tomaremos para nuestro ejemplo son

$$u(x, 0) = (\text{sen}(3\pi x))^2 \cdot (1 - x)^{3/2}, \quad u(0, t) = u(1, t) = 0. \quad (4.19)$$

Este problema no lineal parabólico está tomado de [20] (pp. 349–443) y fue diseñado en un principio por Burgers (1948) como modelo matemático para describir la teoría de turbulencias. Hoy en día sigue siendo un ejemplo interesante de ecuación no lineal que recuerda en cierto modo a la ecuación de Navier-Stokes (en Dinámica de Fluidos) y que posee para valores pequeños de  $\nu$  "shock waves" y para  $\nu \rightarrow 0$  soluciones discontinuas.

Ahora aplicaremos el conocido método de líneas a la ecuación en derivadas parciales anterior como sigue. Discretizaremos a lo largo del eje  $x$  tomando una red uniforme y reemplazando las derivadas

espaciales que aparecen en la ecuación de la derecha de (4.18) por las correspondientes aproximaciones mediante las fórmulas de diferencias centradas. Tomando

$$\Delta x = \frac{1}{N+1}, \quad u_i(t) = u(i\Delta x, t) \quad i = 0, 1, \dots, N+1, \quad (4.20)$$

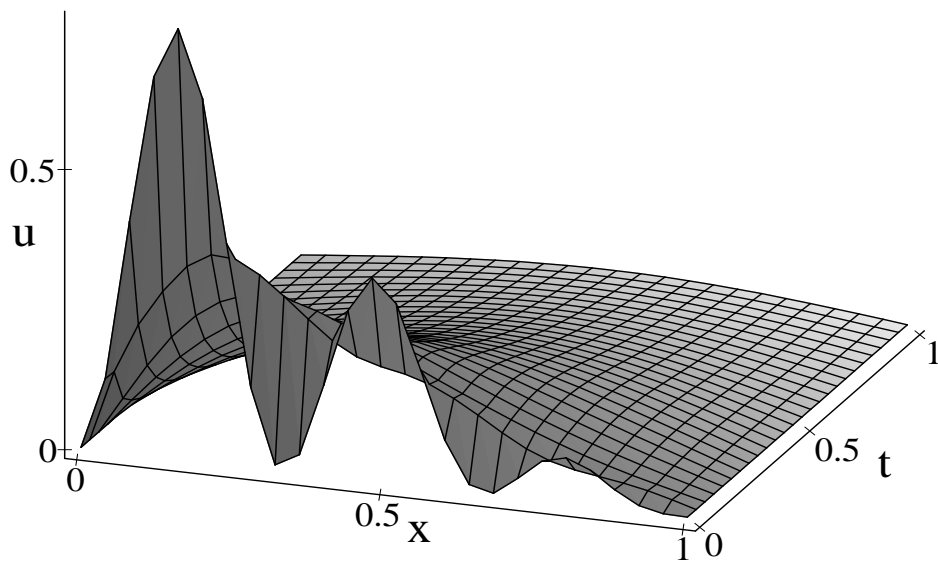
se obtiene el siguiente sistema de EDOs asociado a la ecuación de Burgers

$$\begin{aligned} u_i' &= -\frac{u_{i+1}^2 - u_{i-1}^2}{4\Delta x} + \nu \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2}, \quad i = 1, 2, \dots, N \\ u_i(0) &= (\sin(3\pi i\Delta x))^2 \cdot (1 - i\Delta x)^{3/2}, \quad i = 1, 2, \dots, N, \end{aligned} \quad (4.21)$$

que pertenece al tipo de sistemas al que es aplicable nuestro método y donde se tiene que  $u_i = u_i(t)$ . De las condiciones frontera se deduce que ha de verificarse que  $u_0(t) = u_{N+1}(t) = 0$ . Tras las oportunas exclusiones y sustituciones es fácil comprobar que el sistema anterior tiene asociada una matriz Jacobiana tridiagonal.

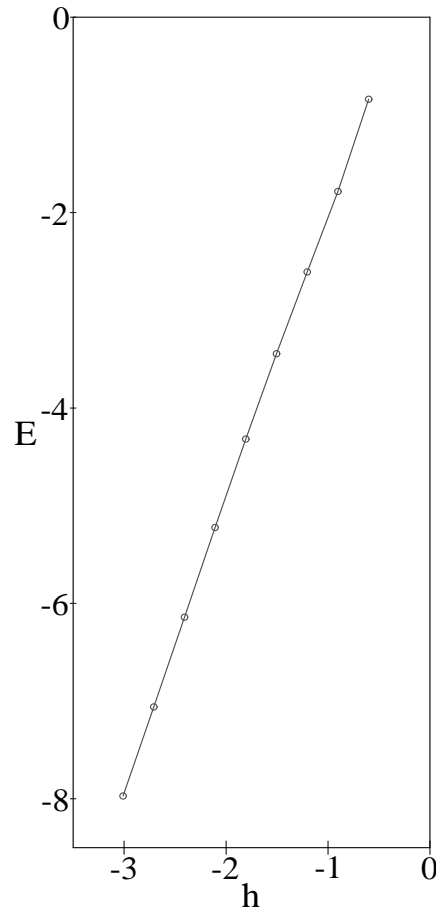
Ahora aplicaremos el método del apartado anterior al sistema de EDOs (4.21) que se obtiene tomando  $N = 24$  y  $\nu = 0.2$ . Para estos valores el sistema es de dimensión 24 y tiene estructura de banda. No es difícil comprobar que el sistema es moderadamente rígido. De hecho tenemos que los autovalores de la matriz Jacobiana asociada son todos reales y pertenecen al intervalo  $[-499, -1]$  (con autovalor dominante cercano a  $-498$ ) para el intervalo de integración considerado.

En primer lugar, integrando con paso fijo  $h = 0.04$  se obtiene la figura 4.1, que nos da una idea gráfica de la solución numérica que se obtiene. Por otra parte, con paso fijo  $h = 2^{-k}$  para los valores de  $k = 2, 3, \dots, 10$  obtenemos el valor de las soluciones numéricas en  $t = 1$ , esto es, la solución numérica obtenida tras  $2^k$  pasos utilizando el método propuesto. Calculamos asimismo una aproximación a la solución exacta (integrando con paso fijo  $h = 0.0001$ ) en el punto  $t = 1$ . El error  $E$  cometido en  $t = 1$  (medido utilizando la norma euclídea en  $\mathbb{R}^{24}$ ) para los distintos tamaños de paso  $h$  considerados, se muestra en la figura 4.2 en doble escala logarítmica (en base 10 para ambos ejes). En la escala logarítmica utilizada para esta figura, el error viene representado muy aproximadamente por una recta cuya pendiente coincide con el orden del método. La figura muestra claramente el orden tres (en línea con lo anunciado con anterioridad).



**Figura 4.1:** Solución numérica: aproximación por el método de líneas a la ecuación de Burgers tomando  $\nu = 0.2$ ,  $N = 24$  y  $h = 0.04$ .

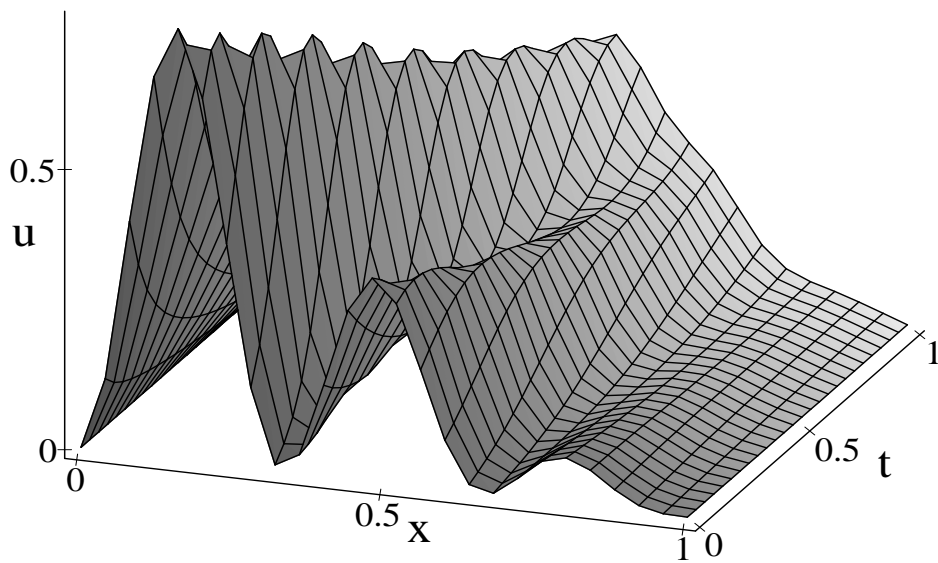
Repetimos nuestro experimento numérico tomando ahora  $N = 24$  y  $\nu = 0.004$  en (4.21). De nuevo el sistema de EDOs obtenido al aplicar el método de líneas a la ecuación de Burgers es de dimensión 24 y tiene estructura de banda, pero ahora los autovalores de la matriz



**Figura 4.2:** Error como función del paso  $h$  (doble escala logarítmica) para nuestro método.

Jacobiana (que es tridiagonal) son en su mayoría complejos (con las partes reales de dichos autovalores comprendidas entre  $-10$  y  $0$  y las partes imaginarias comprendidas entre  $-14$  y  $14$ ).

En la figura 4.3 mostramos la solución numérica obtenida al aplicar nuestro método a dicho problema. La solución posee "shock waves" que acaban uniéndose. Nótese en este punto que la ecuación de Burgers puede ser considerada como ejemplo de un problema hiperbólico con un término artificial de difusión para valores de  $\nu$  pequeños. Conviene observar también que las partes imaginarias de los autovalores de la matriz Jacobiana introducen oscilaciones en las soluciones numéricas (como se puede observar en la figura).



**Figura 4.3:** Solución numérica: aproximación por el método de líneas a la ecuación de Burgers tomando  $\nu = 0.004$ ,  $N = 24$  y  $h = 0.04$ .

Pasamos a continuación a ver algún experimento numérico más para completar esta primera aproximación a nuestros métodos para sistemas.

**4.1.5 Varios experimentos numéricos de interés.** Estudiaremos la siguiente familia de problemas tomada de [23], pp. 34 (véase también [37], pp. 233 y [21]).

$$\begin{aligned} y_1' &= -(b + an)y_1 + by_2^n & y_1(0) &= c^n \\ y_2' &= y_1 - ay_2 - y_2^n & y_2(0) &= c \end{aligned} \quad (4.22)$$

cuya solución viene dada por

$$y_1(x) = c^n e^{-anx}, \quad y_2(x) = ce^{-ax}. \quad (4.23)$$

Tomaremos los valores  $a = 0.1$ ,  $c = 1$ ,  $n = 4$  y  $b = 100^i$  ( $i = 0, 1, 2, 3$ ) en el intervalo  $0 \leq x \leq 10$ . Cuanto mayor es la constante  $b$  más rígido es el sistema (de hecho los autovalores a lo largo de la solución exacta verifican que  $\lambda_1 \approx -b$  and  $\lambda_2 \approx -a$ ) y cuanto mayor es  $n$  mayor es el carácter no lineal del problema (véase [23] para más detalles).

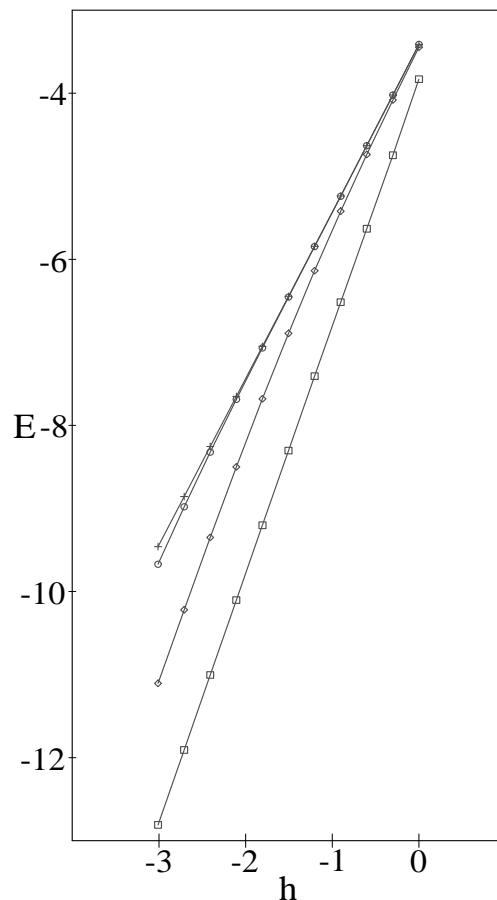
Con paso fijo  $h = 2^{-k}$  para los valores  $k = 0, 1, 2, \dots, 10$  integramos a lo largo de  $10 \cdot 2^k$  pasos y obtenemos las aproximaciones a la solución exacta  $y$  que nos proporciona nuestro método de dos etapas. El tamaño del error global  $E$  cometido (en norma  $L_2$ ), para los diferentes tamaños de paso  $h$  y para cada uno de los distintos problemas considerados, se muestra en la figura 4.4 en doble escala logarítmica. Se observa en dicha figura que el orden de convergencia para el problema (4.22) es menor que tres (de hecho dos) cuando tomamos  $b = 1000000$  (viene representada por cruces en la figura) y  $b = 10000$  (círculos) para el rango de tamaños de paso que estamos considerando. Cuando tomamos  $b = 100$  (diamantes) el orden se incrementa de dos a tres cuando reducimos el tamaño del paso considerado. El fenómeno de la reducción en el orden de convergencia está relacionado con el concepto de B-convergencia y muchos métodos implícitos muestran este comportamiento cuando son aplicados a algunas ecuaciones diferenciales rígidas (no lineales). Para el valor  $b = 1$  (cuadrados) el problema no es rígido y nuestro método muestra orden tres como era de esperar.

Consideraremos asimismo dos problemas rígidos no autónomos. El problema A está tomado de [36], pp. 27 y viene dado por

$$y' = -10^6 y + \cos x + 10^6 \sin x, \quad y(0) = 1, \quad (4.24)$$

con solución

$$y(x) = \sin x + e^{-10^6 x}. \quad (4.25)$$



**Figura 4.4:** Error como función del tamaño de paso en doble escala logarítmica para nuestro método (problema autónomo).

Nótese que este problema es un caso particular de la familia de ecuaciones escalares propuesta por Prothero y Robinson en [34].

El problema B viene dado por el siguiente sistema no autónomo

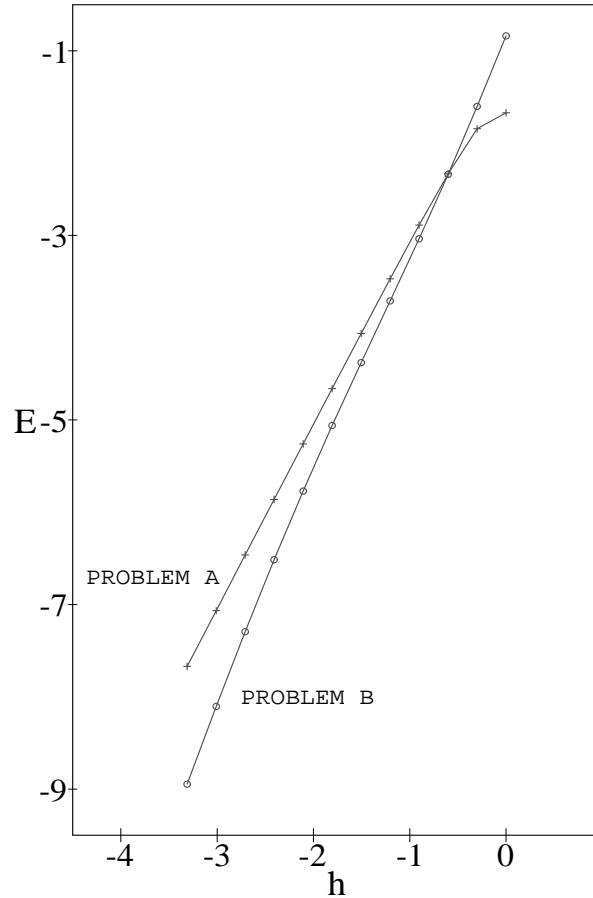
$$\begin{aligned} y_1' &= -2y_1 + y_2 + 2\sin x & y_1(0) &= 2 \\ y_2' &= 998y_1 - 999y_2 + 999(\cos x - \sin x) & y_2(0) &= 3 \end{aligned} \quad (4.26)$$

tomado de [27], pp. 213, para el cual la solución exacta viene dada por

$$y_1(x) = 2e^{-x} + \sin x, \quad y_2(x) = 2e^{-x} + \cos x. \quad (4.27)$$

Ambos problemas son rígidos (los autovalores son  $\lambda = -1000000$  para el problema A y  $\lambda_1 = -1$ ,  $\lambda_2 = -1000$  para el problema B).

Aplicamos el método de dos etapas a estos problemas con paso fijo  $h = 2^{-k}$  para los valores de la constante  $k = 0, 1, 2, \dots, 11$  sobre  $10 \cdot 2^k$  pasos. La figura 4.5 muestra el error global  $E$  (en norma  $L_2$ )



**Figura 4.5:** Error como función del tamaño de paso en doble escala logarítmica para nuestro método (problemas no autónomos).

frente al tamaño de paso en doble escala logarítmica. Como ocurría en un ejemplo anterior, podemos observar que el orden de convergencia que manifiesta nuestra fórmula al ser aplicada al problema A parece ser dos (para el rango de tamaños de paso  $h$  que hemos considerado). Tomando un tamaño de paso suficientemente pequeño el orden cambia a tres como era de esperar. Para el problema B el orden parece variar de dos a tres según vamos reduciendo el tamaño del paso, como puede



observarse en la figura. Como ya hemos comentado anteriormente, este fenómeno puede ser explicado en términos del concepto de B-convergencia.

## 4.2 Orden de consistencia de los métodos para sistemas.

### 4.2.1 Condiciones de orden para los métodos de dos etapas.

Como ya hemos mostrado en los experimentos numéricos anteriores, es posible obtener métodos de dos etapas para sistemas del tipo (4.1) con orden de consistencia tres. De hecho, en este caso, las condiciones de orden tres coinciden con las que se obtienen del caso escalar autónomo. Análogamente a como se vió en el caso escalar, basta deducir las condiciones de orden para los métodos de tipo polinomial, pues a partir de éstas se pueden deducir las condiciones de orden para los demás métodos.

Análogamente a como ocurría en el caso escalar, los métodos de dos etapas de tipo polinomial (para sistemas) vienen definidos por

$$y_{n+1} = y_n + hG_3(S_2)k_1, \quad (4.28)$$

donde las etapas son

$$k_1 = f(y_n), \quad k_2 = f(y_n + hc_2k_1), \quad (4.29)$$

$S_2$  es la matriz

$$\begin{pmatrix} \frac{f_{11}(y_{n(1)}+hc_2e_1k_1)-f_{11}(y_{n(1)})}{c_2e_1k_1} & \dots & \frac{f_{1m}(y_{n(m)}+hc_2e_mk_1)-f_{1m}(y_{n(m)})}{c_2e_mk_1} \\ \vdots & & \vdots \\ \frac{f_{m1}(y_{n(1)}+hc_2e_1k_1)-f_{m1}(y_{n(1)})}{c_2e_1k_1} & \dots & \frac{f_{mm}(y_{n(m)}+hc_2e_mk_1)-f_{mm}(y_{n(m)})}{c_2e_mk_1} \end{pmatrix} \quad (4.30)$$

y donde la función  $G_3$  en términos de  $S_2$  toma la forma

$$G_3(S_2) = c_3 \left( I + \sum_{i=1}^{r_3} a_i S_2^i \right), \quad r_3 \in \mathbb{N}, \quad (4.31)$$

y donde con  $I$  denotamos la matriz identidad de dimensión  $m$ . Las condiciones de orden (para orden tres) coinciden con las recogidas

en (3.57–3.60) para el caso escalar autónomo y por eso no las reproducimos aquí. Tras resolver las condiciones de orden tres, se deduce que  $c_2 = 2/3$  y que la función  $G_3$  asociada a cualquier método de orden tres de tipo polinomial admite una expresión del tipo

$$G_3(S_2) = I + \frac{1}{2}S_2 + \frac{1}{6}S_2^2 + \sum_{i=3}^{r_3} a_i S_2^i, \quad r_3 \in \mathbb{N}, \quad (4.32)$$

donde los parámetros libres  $a_i$  (con  $i \geq 3$ ) pueden tomarse arbitrariamente.

Las condiciones adicionales para orden cuatro difieren ahora de las del caso escalar. De hecho tenemos cuatro condiciones de orden adicionales (en vez de las tres del caso escalar) que vienen dadas por

$$c_3 c_2^2 a_1 = 1/4, \quad (4.33)$$

$$c_3 c_2 a_2 = 1/12, \quad (4.34)$$

$$c_3 c_2 a_2 = 1/4, \quad (4.35)$$

$$c_3 a_3 = 1/24, \quad (4.36)$$

dos de las cuales son claramente incompatibles ((4.34) y (4.35)). Es fácil comprobar que las tres primeras ecuaciones no se satisfacen para los valores de las constantes que ha de tomar un método de dos etapas y orden tres, por lo que no es posible obtener métodos de orden cuatro con sólo dos etapas (como ya habíamos deducido en el caso escalar). Sin embargo, es posible satisfacer la cuarta ecuación tomando  $a_3 = 1/24$  y minimizando de este modo la parte principal del error local de truncación. Además, para esta elección de la constante  $a_3$  (tomando además  $a_1 = 1/2$ ,  $a_2 = 1/6$ ,  $c_2 = 2/3$  y  $c_3 = 1$  para que el método sea de orden tres) las fórmulas resultantes exhiben orden cuatro cuando las aplicamos a problemas lineales a coeficientes constantes.

Ahora es fácil deducir las condiciones de orden para los métodos para sistemas de dos etapas definidos en términos de funciones racionales a partir de las obtenidas para las fórmulas de tipo polinomial, siguiendo un procedimiento parecido al que vimos al tratar el caso escalar. Un método de dos etapas de tipo racional para sistemas viene dado (generalizando la situación ya estudiada en el caso escalar) por las fórmulas (4.12)–(4.14), donde la función  $G_3$  en términos de  $S_2$

admite la siguiente expresión

$$G_3(S_2) = c_3 \left( I + \sum_{i=1}^{d_3^*} d_i S_2^i \right)^{-1} \left( I + \sum_{i=1}^{n_3^*} n_i S_2^i \right), \quad n_3^*, d_3^* \in \mathbb{N}, \quad (4.37)$$

Análogamente a como se hizo en el caso escalar, es suficiente tomar  $G_2 = c_2 I$  con  $c_2 = 2/3$  e imponer que la función  $G_3$  dada por (4.37) verifique la condición

$$G_3(S_2) = I + \frac{1}{2} S_2 + \frac{1}{6} S_2^2 + O(S_2^3). \quad (4.38)$$

Bastará pues con considerar los valores  $n_3^* = d_3^* = 2$  en (4.37) a la hora de obtener las condiciones de orden tres para el método de tipo racional a partir de las correspondientes al caso polinomial. Se obtienen de este modo sin dificultad las cuatro condiciones recogidas en (3.97)–(3.100).

Si además se desea minimizar la parte principal del error local de truncación, bastará tomar  $n_3^* = d_3^* = 3$  en (4.37) y desarrollar el segundo miembro en (4.38) un orden más, obteniendo así la condición adicional recogida en (3.101).

A partir de las consideraciones anteriores se deduce que la expresión general de un método de dos etapas y orden tres de tipo racional para sistemas (del tipo considerado en este capítulo) es

$$y_{n+1} = y_n + hG_3(S_2) k_1, \quad (4.39)$$

donde las etapas son

$$k_1 = f(y_n), \quad k_2 = f\left(y_n + \frac{2}{3} h k_1\right), \quad (4.40)$$

la matriz  $S_2$  es

$$\begin{pmatrix} \frac{f_{11}(y_{n(1)} + \frac{2}{3} h e_1 k_1) - f_{11}(y_{n(1)})}{\frac{2}{3} e_1 k_1} & \dots & \frac{f_{1m}(y_{n(m)} + \frac{2}{3} h e_m k_1) - f_{1m}(y_{n(m)})}{\frac{2}{3} e_m k_1} \\ \vdots & & \vdots \\ \frac{f_{m1}(y_{n(1)} + \frac{2}{3} h e_1 k_1) - f_{m1}(y_{n(1)})}{\frac{2}{3} e_1 k_1} & \dots & \frac{f_{mm}(y_{n(m)} + \frac{2}{3} h e_m k_1) - f_{mm}(y_{n(m)})}{\frac{2}{3} e_m k_1} \end{pmatrix} \quad (4.41)$$

y la función  $G_3$  es

$$G_3(S_2) = \left( I + d_1 S_2 + d_2 S_2^2 + \sum_{i=3}^{d_3^*} d_i S_2^i \right)^{-1} \cdot \left( I + \frac{1+2d_1}{2} S_2 + \frac{1+3d_1+6d_2}{6} S_2^2 + \sum_{i=3}^{n_3^*} n_i S_2^i \right). \quad (4.42)$$

La expresión general de un método de dos etapas y orden tres de tipo racional para sistemas, con minimización de la parte principal del error local de truncación, vendrá dada por las fórmulas (4.39–4.41), donde ahora la función  $G_3$  toma la forma

$$G_3(S_2) = \left( I + d_1 S_2 + d_2 S_2^2 + d_3 S_2^3 + \sum_{i=4}^{d_3^*} d_i S_2^i \right)^{-1} \cdot \left( I + \frac{1+2d_1}{2} S_2 + \frac{1+3d_1+6d_2}{6} S_2^2 + \frac{1+4d_1+12d_2+24d_3}{24} S_2^3 + \sum_{i=4}^{n_3^*} n_i S_2^i \right). \quad (4.43)$$

#### 4.2.2 Condiciones de orden para los métodos de tres etapas.

Ahora la situación cambia respecto del caso de dos etapas. Por una parte la definición de la matriz  $S_3$  que introduciremos, muestra alguna diferencia con respecto al caso escalar como ya hemos adelantado en el apartado 4.1.2. Por otra parte, al intervenir dos matrices  $S_2$  y  $S_3$ , la posible no conmutatividad de los productos de ambas habrá de ser tenida en cuenta e introducirá mayores diferencias con respecto al caso escalar estudiado en capítulos precedentes que las que se daban en el caso de dos etapas.

Empecemos considerando la expresión general de los métodos GRK de tres etapas (para sistemas del tipo descrito en (4.1.1)) que se deduce de (4.1.2) particularizando para  $p = 3$ . Se tiene

$$y_{n+1} = y_n + hG_4(S_2, S_3) k_1, \quad (4.44)$$

donde  $y_n$ ,  $y_{n+1}$  y  $k_1$  son vectores de  $\mathbb{R}^m$  y los  $S_i$  son matrices cuadradas de dimensión  $m$ .

Las etapas  $k_i$  (con  $1 \leq i \leq 3$ ) vienen dadas en la forma

$$\begin{aligned} k_1 &= f(y_n) \\ k_2 &= f(y_n + hG_2 k_1) \\ k_3 &= f(y_n + hG_3(S_2) k_1), \end{aligned} \quad (4.45)$$

(con  $G_2 = c_2 I$ ) en términos de las matrices cuadradas  $S_2$  definida como en (4.14) y  $S_3$  dada por

$$\begin{pmatrix} \frac{f_{11}(y_{n(1)} + \alpha_1) - f_{11}(y_{n(1)})}{e_1 G_3(S_2) k_1} & \dots & \frac{f_{1m}(y_{n(m)} + \alpha_m) - f_{1m}(y_{n(m)})}{e_m G_3(S_2) k_1} \\ \vdots & & \vdots \\ \frac{f_{m1}(y_{n(1)} + \alpha_1) - f_{m1}(y_{n(1)})}{e_1 G_3(S_2) k_1} & \dots & \frac{f_{mm}(y_{n(m)} + \alpha_m) - f_{mm}(y_{n(m)})}{e_m G_3(S_2) k_1} \end{pmatrix}, \quad (4.46)$$

donde  $\alpha_j = h e_j G_3(S_2) k_1$  (con  $1 \leq j \leq m$ ) y con  $e_j$  denotamos el vector de  $\mathbb{R}^m$  que tiene todas sus componentes nulas a excepción de la que ocupa la componente  $j$ -ésima que vale 1. Nótese que el elemento que ocupa la fila  $p$  y columna  $q$  (con  $1 \leq p, q \leq m$ ) en la matriz  $S_3$  viene dado por

$$\frac{f_{pq}(y_{n(q)} + h e_q G_3(S_2) k_1) - f_{pq}(y_{n(q)})}{e_q G_3(S_2) k_1}. \quad (4.47)$$

Evidentemente,  $G_i k_1$  denota el producto de la matriz  $G_i$  por el vector  $k_1$ .

En primer lugar nos centraremos en obtener las condiciones de orden para los métodos de tres etapas de tipo polinomial, esto es, aquellos para los cuales se tiene que las funciones  $G_2$ ,  $G_3$  y  $G_4$  que los caracterizan son polinomiales. Más concretamente,  $G_2 = c_2 I$  (con  $c_2$  constante), y las otras dos funciones pueden expresarse como

$$\begin{aligned} G_3(S_2) &= c_3 \left( I + \sum_{i=1}^{r_3} a_{3,22\dots 2} S_2^i \right), \quad r_3 \in \mathbb{N}, \\ G_4(S_2, S_3) &= c_4 \left( I + \sum_{i=1}^{r_4^*} a_{4,\sigma_1 \sigma_2 \dots \sigma_i}^* S_{\sigma_1} S_{\sigma_2} \dots S_{\sigma_i} \right), \quad r_4^* \in \mathbb{N}, \end{aligned} \quad (4.48)$$

donde los coeficientes  $a_{3,22\dots 2}$  que aparecen en el primer sumatorio de (4.48) sustituyen a los coeficientes  $a_i$  que utilizábamos en el caso

escalar y en los métodos de dos etapas para sistemas separados. Cada uno de los subíndices  $\sigma_k$  que aparece en el segundo sumatorio de (4.48) pueden tomar los valores 2 o 3. Nos hemos visto obligados a introducir estos subíndices  $\sigma_k$ , por la posible no conmutatividad de los productos de las matrices  $S_2$  y  $S_3$ , que obliga a distinguir los coeficientes de los productos de dichas matrices cuando, pese a estar formados por los mismos factores, éstos aparecen multiplicados en diferente orden.

Como ya comentamos en el caso escalar, sólo algunos de los parámetros libres de los que disponemos aparecerán en las condiciones de orden (pues las matrices  $S_2$  y  $S_3$  son  $O(h)$ ), pero aún así las condiciones de orden resultantes son relativamente complicadas. Por ello definiremos una nueva matriz  $\tilde{S}_3$  que sustituirá a la matriz  $S_3$  y simplificará el estudio que sigue. Tomaremos  $\tilde{S}_3 = S_3 - S_2$  con lo cual se verificará que  $\tilde{S}_3 = O(h^2)$  y de este modo el número de parámetros involucrados en las condiciones de orden será inferior, simplificando éstas. Con la introducción de  $\tilde{S}_3$  los métodos de tres etapas de tipo polinomial pueden describirse ahora así

$$y_{n+1} = y_n + h\tilde{G}_4(S_2, \tilde{S}_3) k_1, \quad (4.49)$$

donde  $\tilde{S}_3 = S_3 - S_2$ , con  $S_2$  y  $S_3$  definidos como hemos visto anteriormente en términos de las etapas dadas por (4.45) y donde ahora

$$\tilde{G}_4(S_2, \tilde{S}_3) = c_4 \left( I + \sum_{i=1}^{r_4} a_{4, \sigma_1 \sigma_2 \dots \sigma_i} S_{\sigma_1} S_{\sigma_2} \dots S_{\sigma_i} \right), \quad r_4 \in \mathbb{N}, \quad (4.50)$$

(las funciones  $G_2$  y  $G_3$  vienen dadas como hemos visto anteriormente). Los subíndices  $\sigma_k$ , que aparecen en el sumatorio de (4.50), pueden tomar los valores 2 o 3 y pondremos  $S_{\sigma_k} = S_2$  cuando  $\sigma_k = 2$  y  $S_{\sigma_k} = \tilde{S}_3$  cuando sea  $\sigma_k = 3$ . Obsérvese que en el sumatorio de (4.50), al ser  $S_2 = O(h)$  y  $\tilde{S}_3 = O(h^2)$ , bastará considerar únicamente aquellos sumandos para los cuales se verifique que  $n_2 + 2n_3 \leq r_4$ , donde con  $n_2$  denotamos el número de subíndices  $\sigma_k$  (del coeficiente  $a_{4, \sigma_1 \sigma_2 \dots \sigma_i}$ ) que toman el valor 2, y con  $n_3$  el número de subíndices  $\sigma_k$  que toman el valor 3. De este modo, sólo aparecerán en el sumatorio aquellos términos (y sus coeficientes asociados) que intervienen en las condiciones de orden menor o igual que  $r_4 + 1$  y prescindimos de aquellos otros términos superfluos que intervienen únicamente en condiciones de orden superiores a  $r_4 + 1$ .

Veremos que ahora no es posible obtener métodos de orden cinco con sólo tres etapas, a diferencia de lo que ocurría en el caso escalar. Sin embargo, es posible obtener orden cuatro y retener el resto de propiedades de estabilidad lineal observadas en el caso escalar. Más aún, como veremos, es posible satisfacer casi todas las condiciones de orden cinco con lo que las fórmulas así obtenidas muestran orden cinco cuando son aplicadas a algunos problemas concretos.

Para orden cinco es suficiente tomar  $r_3 = 2$  en (4.48) y  $r_4 = 4$  en (4.50), por lo que las condiciones de orden cinco quedan expresadas y completamente determinadas en función de los parámetros  $c_i$  ( $2 \leq i \leq 4$ ),  $a_{3,2}$ ,  $a_{3,22}$  y  $a_{4,\sigma_1\sigma_2\dots\sigma_i}$  (con  $n_2 + 2n_3 \leq 4$ ).

Las condiciones de orden que ha de satisfacer una fórmula de tres etapas de tipo polinomial para tener orden cinco, vienen dadas por

$$c_4 = 1, \quad (4.51)$$

$$c_4 a_{4,2} = 1/2, \quad (4.52)$$

$$c_4 (c_2(a_{4,2} - a_{4,3}) + c_3 a_{4,3}) = 1/3, \quad (4.53)$$

$$c_4 a_{4,22} = 1/6, \quad (4.54)$$

$$c_4 (c_2^2(a_{4,2} - a_{4,3}) + c_3^2 a_{4,3}) = 1/4, \quad (4.55)$$

$$c_4 (c_2(a_{4,22} - a_{4,23}) + c_3 a_{4,23}) = 1/12, \quad (4.56)$$

$$c_4 (c_2(a_{4,22} - a_{4,32}) + c_3(a_{4,32} + a_{3,2}a_{4,3})) = 1/4, \quad (4.57)$$

$$c_4 a_{4,222} = 1/24, \quad (4.58)$$

$$c_4 (c_2^3(a_{4,2} - a_{4,3}) + c_3^3 a_{4,3}) = 1/5, \quad (4.59)$$

$$c_4 (c_2^2(a_{4,22} - a_{4,32}) + c_3^2(a_{4,32} + 2a_{3,2}a_{4,3})) = 3/10, \quad (4.60)$$

$$c_4 (c_2^2(a_{4,22} - a_{4,23}) + c_3^2 a_{4,23}) = 1/20, \quad (4.61)$$

$$c_4 (c_2^2(a_{4,22} - a_{4,23} - a_{4,32} + a_{4,33}) + c_2 c_3 (a_{4,23} + a_{4,32} - 2a_{4,33} + a_{3,2}a_{4,3}) + c_3^2 a_{4,33}) = 2/15, \quad (4.62)$$

$$c_4 (c_2(a_{4,222} - a_{4,322}) + c_3(a_{4,322} + a_{3,22}a_{4,3})) = 1/15, \quad (4.63)$$

$$c_4 c_3 a_{3,2} a_{4,32} = 1/20, \quad (4.64)$$

$$c_4 (c_2(a_{4,222} - a_{4,232} + c_3(a_{4,232} + a_{3,2}a_{4,23})) = 1/20, \quad (4.65)$$

$$c_4 (c_2(a_{4,222} - a_{4,223} + c_3 a_{4,223}) = 1/60, \quad (4.66)$$

$$c_4 a_{4,2222} = 1/120. \quad (4.67)$$

Es fácil comprobar que no es posible satisfacer todas las condiciones

anteriores simultáneamente, si bien, como ya habíamos adelantado, es posible satisfacer la mayoría de ellas. Por tanto, no es posible obtener ninguna fórmula de tres etapas y orden cinco para sistemas separados. Sin embargo, se obtiene una familia infinita de métodos de tres etapas y orden cuatro resolviendo las condiciones de orden (4.51–4.58). Para ello bastará seguir los pasos siguientes:

*Paso 1.* De (4.51) deducimos que ha de ser  $c_4 = 1$  (que es la condición de consistencia). Tras sustituir este valor en el resto de ecuaciones, se obtienen directamente los coeficientes  $a_{4,2}$ ,  $a_{4,22}$  y  $a_{4,222}$  de las ecuaciones (4.52), (4.54) y (4.58).

*Paso 2.* Sustituyendo los valores obtenidos en las restantes ecuaciones y resolviendo el sistema no lineal dado por las ecuaciones (4.53) y (4.55) obtenemos los valores de  $c_2$  y  $a_{4,3}$  (en función del parámetro  $c_3$ ).

*Paso 3.* Finalmente, se obtienen los coeficientes  $a_{4,23}$  y  $a_{3,2}$  a partir de las ecuaciones (4.56) y (4.57) (en términos de los parámetros  $c_3$  y  $a_{4,32}$ ).

Las soluciones obtenidas de este modo vienen dadas por

$$\begin{aligned} c_2 &= \frac{3 - 4c_3}{2(2 - 3c_3)}, \quad c_4 = 1, \quad a_{4,2} = \frac{1}{2}, \quad a_{4,22} = \frac{1}{6}, \quad a_{4,222} = \frac{1}{24}, \\ a_{3,2} &= \frac{(9 - 39c_3 + 58c_3^2 - 30c_3^3) + (54 - 288c_3 + 600c_3^2 - 576c_3^3 + 216c_3^4)a_{4,32}}{2c_3(2 - 3c_3)}, \\ a_{4,3} &= \frac{1}{6(3 - 8c_3 + 6c_3^2)}, \quad a_{4,23} = \frac{1 - c_3}{6(3 - 8c_3 + 6c_3^2)}. \end{aligned} \quad (4.68)$$

El resto de parámetros se pueden tomar de manera arbitraria.

Las condiciones adicionales para orden cinco, que vienen dadas por las ecuaciones (4.59–4.67), pueden ser satisfechas en su mayoría. Concretamente, es posible satisfacer todas las condiciones a excepción de una (las ecuaciones (4.60) y (4.64) no pueden ser satisfechas simultáneamente). Para ello basta proceder como sigue:

*Paso 1.* Sustituimos los valores (4.68), obtenidos tras resolver las condiciones de orden cuatro, en las nueve condiciones adicionales para orden cinco. Tras esta sustitución, las condiciones (4.59) y (4.61) dan lugar a la misma ecuación. Obtenemos de este modo un sistema no lineal de ocho ecuaciones con ocho incógnitas (los parámetros que restan por determinar) que no tiene solución.



*Paso 2.* Se obtienen directamente dos valores para el coeficiente  $c_3$  que satisfacen la ecuación (4.59) (que coincide con (4.61)), y un valor para  $a_{4,222}$  a partir de la ecuación (4.67).

*Paso 3.* Tras las oportunas sustituciones de los valores calculados en el paso anterior, es posible ahora obtener directamente (sin más que despejar) los coeficientes  $a_{4,33}$  y  $a_{4,223}$  de las ecuaciones (4.62) y (4.66) respectivamente. También se observa que las ecuaciones (4.60) y (4.64) no pueden satisfacerse simultáneamente por depender ambas únicamente del parámetro  $a_{4,32}$  y ser incompatibles. Nosotros optamos por satisfacer la primera de estas dos ecuaciones (también se podría optar por satisfacer la segunda obteniéndose entonces un par de valores distintos para el coeficiente  $a_{4,32}$ ).

*Paso 4.* Finalmente, tras las oportunas sustituciones de valores ya calculados, se obtienen los restantes coeficientes  $a_{3,22}$  y  $a_{4,232}$  de las ecuaciones (4.63) y (4.65) ( $a_{3,22}$  viene dado en términos del parámetro libre  $a_{4,322}$ ).

Los valores de los coeficientes que se obtienen siguiendo el procedimiento anterior son

$$\begin{aligned}
 c_2 &= \frac{6 \mp \sqrt{6}}{10}, & c_3 &= \frac{6 \pm \sqrt{6}}{10}, & c_4 &= 1, & a_{3,2} &= \frac{-3 \pm 2\sqrt{6}}{5}, \\
 a_{3,22} &= \frac{(17 + 144 a_{4,322}) \mp (3 + 96 a_{4,322})\sqrt{6}}{50}, & a_{4,2} &= \frac{1}{2}, \\
 a_{4,3} &= \frac{9 \pm \sqrt{6}}{36}, & a_{4,22} &= \frac{1}{6}, & a_{4,23} &= \frac{6 \mp \sqrt{6}}{72}, & a_{4,32} &= \frac{-1 \pm \sqrt{6}}{8}, \\
 a_{4,33} &= \frac{1 \pm 4\sqrt{6}}{72}, & a_{4,222} &= \frac{1}{24}, & a_{4,232} &= \frac{-3 \pm 2\sqrt{6}}{48}, \\
 a_{4,223} &= \frac{3 \mp \sqrt{6}}{144}, & a_{4,2222} &= \frac{1}{120}.
 \end{aligned} \tag{4.69}$$

El parámetro  $a_{4,322}$  y el resto de parámetros (que no aparecen en las condiciones de orden cinco) pueden tomar valores arbitrarios.

Si en el paso 3 hubiésemos optado por satisfacer la ecuación (4.64) (en vez de la ecuación (4.60)) la solución descrita en (4.69) se vería modificada únicamente en lo que respecta a los coeficientes  $a_{3,2}$ ,  $a_{4,32}$  y  $a_{4,232}$ , que pasarían a venir dados por

$$a_{3,2} = \frac{\pm\sqrt{6}}{5}, \quad a_{4,32} = \frac{-1 \pm \sqrt{6}}{12}, \quad a_{4,232} = \frac{-1 \pm \sqrt{6}}{48}, \tag{4.70}$$

o bien por

$$a_{3,2} = \frac{6 \mp 2\sqrt{6}}{5}, \quad a_{4,32} = \frac{4 \pm \sqrt{6}}{24}, \quad a_{4,232} = \frac{5 \mp \sqrt{6}}{48}, \quad (4.71)$$

pues en este caso se obtienen dos valores para el coeficiente  $a_{4,32}$  (para cada una de las soluciones ya obtenidas).

Seguidamente estudiaremos el caso racional. Un método de tres etapas para sistemas, definido en términos de funciones racionales, viene dado, con las modificaciones vistas en el caso polinomial, por una expresión como la considerada en (4.49), esto es

$$y_{n+1} = y_n + h\tilde{G}_4(S_2, \tilde{S}_3) k_1. \quad (4.72)$$

Sin embargo, aunque las etapas  $k_1$ ,  $k_2$  y  $k_3$  vienen dadas formalmente como en el caso polinomial (véase (4.45)) y las matrices  $S_2$  y  $\tilde{S}_3$  también (véanse (4.14) y (4.46)), ahora las funciones  $G_3$  y  $\tilde{G}_4$  toman la forma

$$\begin{aligned} G_3(S_2) &= c_3 \left( I + \sum_{i=1}^{d_3^*} d_{3,22\dots 2} S_2^i \right)^{-1} \left( I + \sum_{i=1}^{n_3^*} n_{3,22\dots 2} S_2^i \right), \\ \tilde{G}_4(S_2, \tilde{S}_3) &= c_4 \left( I + \sum_{i=1}^{d_4^*} d_{4,\sigma_1\sigma_2\dots\sigma_i} S_{\sigma_1} S_{\sigma_2} \cdots S_{\sigma_i} \right)^{-1} \\ &\quad \cdot \left( I + \sum_{i=1}^{n_4^*} n_{4,\sigma_1\sigma_2\dots\sigma_i} S_{\sigma_1} S_{\sigma_2} \cdots S_{\sigma_i} \right). \end{aligned} \quad (4.73)$$

La función  $G_2$  viene dada por  $G_2 = c_2 I$  (con  $c_2$  constante) como viene siendo habitual.

Recordemos que los subíndices  $\sigma_k$  de (4.73) pueden tomar los valores 2 o 3 y pondremos  $S_{\sigma_k} = S_2$  cuando  $\sigma_k = 2$  y  $S_{\sigma_k} = \tilde{S}_3$  cuando sea  $\sigma_k = 3$ . Además, en los sumatorios de (4.73) que definen la función  $\tilde{G}_4$ , consideraremos únicamente aquellos sumandos para los cuales se verifique que  $n_2 + 2n_3$  sea menor o igual que  $d_4^*$  (en el primer sumatorio) o  $n_4^*$  (en el segundo), donde con  $n_2$  denotamos el número de subíndices  $\sigma_k$  del coeficiente correspondiente que toman el valor 2, y con  $n_3$  el número de subíndices  $\sigma_k$  que toman el valor 3.

Pasamos a ver a continuación las condiciones de orden correspondientes a los métodos de tres etapas de tipo racional. Como viene

siendo habitual, deduciremos dichas condiciones a partir de las ya obtenidas para el caso polinomial. Procederemos análogamente a como hicimos en el caso escalar y para los métodos de dos etapas, si bien ahora hay que tener en cuenta la posible no conmutatividad de los productos a la que ya nos hemos referido con anterioridad.

Empecemos considerando las condiciones de orden cuatro que, para los métodos de tipo polinomial, venían dadas por el sistema de ecuaciones (4.51–4.58), cuya solución recogíamos en (4.68). Una vez determinados los coeficientes correspondientes a los métodos de tres etapas y orden cuatro de tipo polinomial, bastará imponer que las funciones  $G_2$ ,  $G_3$  y  $\tilde{G}_4$  asociadas a las fórmulas de tipo racional coincidan con las de tipo polinomial hasta el orden adecuado. Para ello bastará tomar  $n_3^* = d_3^* = 1$  y  $n_4^* = d_4^* = 3$  en (4.73).

Igualando se obtiene que  $G_2$  viene dado como en el caso polinomial, esto es,  $G_2 = c_2 I$  con la constante  $c_2$  dada en términos del parámetro libre  $c_3$  como veíamos en (4.68). Al comparar las funciones  $G_3$  y  $\tilde{G}_4$  definidas por (4.73) con las correspondientes al caso polinomial, se obtiene que el parámetro  $c_3$  se puede tomar arbitrariamente y que ha de ser  $c_4 = 1$ . Los restantes coeficientes que definen la función  $G_3$  vienen dados en términos de los calculados en el caso polinomial (véase (4.68)) a través de las relaciones

$$n_{3,2} = d_{3,2} + a_{3,2}, \quad (4.74)$$

$$n_{4,2} = d_{4,2} + a_{4,2}, \quad (4.75)$$

$$n_{4,3} = d_{4,3} + a_{4,3}, \quad (4.76)$$

$$n_{4,22} = d_{4,22} + a_{4,2} d_{4,2} + a_{4,22}, \quad (4.77)$$

$$n_{4,23} = d_{4,23} + a_{4,3} d_{4,2} + a_{4,23}, \quad (4.78)$$

$$n_{4,32} = d_{4,32} + a_{4,2} d_{4,3} + a_{4,32}, \quad (4.79)$$

$$n_{4,222} = d_{4,222} + a_{4,2} d_{4,22} + a_{4,22} d_{4,2} + a_{4,222}. \quad (4.80)$$

Nótese que hay que sustituir los coeficientes (4.68) en dichas fórmulas (no lo hacemos para no complicar las relaciones innecesariamente).

Si además tratamos de minimizar en lo posible el error local de truncación de modo que se satisfagan casi todas las condiciones de orden cinco (todas menos una como ya vimos en el caso polinomial), bastará tomar  $n_3^* = d_3^* = 2$  y  $n_4^* = d_4^* = 4$  en (4.73). Se obtiene así que los coeficientes  $c_i$  (con  $2 \leq i \leq 4$ ) vienen dados como en (4.69) y

se han de verificar las relaciones

$$n_{3,2} = d_{3,2} + a_{3,2}, \quad (4.81)$$

$$n_{3,22} = d_{3,22} + a_{3,2} d_{3,2} + a_{3,22}, \quad (4.82)$$

$$n_{4,2} = d_{4,2} + a_{4,2}, \quad (4.83)$$

$$n_{4,3} = d_{4,3} + a_{4,3}, \quad (4.84)$$

$$n_{4,22} = d_{4,22} + a_{4,2} d_{4,2} + a_{4,22}, \quad (4.85)$$

$$n_{4,23} = d_{4,23} + a_{4,3} d_{4,2} + a_{4,23}, \quad (4.86)$$

$$n_{4,32} = d_{4,32} + a_{4,2} d_{4,3} + a_{4,32}, \quad (4.87)$$

$$n_{4,33} = d_{4,33} + a_{4,3} d_{4,3} + a_{4,33}, \quad (4.88)$$

$$n_{4,222} = d_{4,222} + a_{4,2} d_{4,22} + a_{4,22} d_{4,2} + a_{4,222}, \quad (4.89)$$

$$n_{4,223} = d_{4,223} + a_{4,3} d_{4,22} + a_{4,23} d_{4,2} + a_{4,223}, \quad (4.90)$$

$$n_{4,232} = d_{4,232} + a_{4,2} d_{4,23} + a_{4,32} d_{4,2} + a_{4,232}, \quad (4.91)$$

$$n_{4,322} = d_{4,322} + a_{4,2} d_{4,32} + a_{4,22} d_{4,3} + a_{4,322}, \quad (4.92)$$

$$n_{4,2222} = d_{4,2222} + a_{4,2} d_{4,222} + a_{4,22} d_{4,22} + a_{4,222} d_{4,2} + a_{4,2222}. \quad (4.93)$$

Recordemos que en las relaciones anteriores hay que sustituir los coeficientes (4.69) (modificados y completados si se desea con (4.70) y (4.71)).

### 4.3 Propiedades de estabilidad lineal de los métodos para sistemas.

**4.3.1 Función de estabilidad lineal de los métodos para sistemas.** El estudio de las propiedades de estabilidad lineal se complica ahora respecto al que hicimos en el caso escalar autónomo. Esto es así porque ahora (al estudiar los métodos para sistemas separados) hemos de considerar en vez de la ecuación test  $y' = \lambda y$  con  $\lambda \in \mathbb{C}$ , el sistema test  $y' = Ay$  donde  $A$  es una matriz cuadrada de dimensión  $m$  con  $m$  autovalores distintos  $\lambda_i \in \mathbb{C}$  ( $1 \leq i \leq m$ ) cuyas partes reales son negativas. Las condiciones anteriores permiten asegurar que las soluciones (exactas) del sistema considerado tienden a cero cuando  $x$  tiende a infinito y que existe una matriz no singular  $Q$  tal que se verifica  $Q^{-1}AQ = \Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_m]$ . La transformación  $y = Qz$  nos permitirá simplificar el estudio de la estabilidad lineal de los métodos. De hecho dicho estudio se reduce a considerar la ecuación escalar

test que hemos mencionado anteriormente, pues tanto el sistema test como las fórmulas en diferencias que definen los métodos quedan desacopladas tras aplicar la transformación. Para ver que esto es así, basta aplicar un método cualquiera de  $p$  etapas para sistemas (definido por (4.4–4.6)) al sistema test  $y' = Ay$  (verificando las condiciones que hemos mencionado) tras lo cual se obtiene

$$y_{n+1} = y_n + hG_{p+1}(S_2, S_3, \dots, S_p) k_1, \quad (4.94)$$

donde ahora las etapas  $k_i$  vienen dadas recursivamente por

$$\begin{aligned} k_1 &= Ay_n \\ k_2 &= A(I + hG_2 A) y_n \\ k_3 &= A(I + hG_3(hA) A) y_n \\ &\vdots \\ k_p &= A(I + hG_p(hA, hA, \dots, hA) A) y_n, \end{aligned} \quad (4.95)$$

pues, aplicadas a este sistema, las matrices cuadradas  $S_i$  verifican que  $S_i = hA$  (para  $i = 2, 3, \dots, p$ ).

Definimos ahora  $y_n = Q w_n$  y  $k_i = Q \alpha_i$  (para  $i = 1, 2, \dots, p$ ), esto es, realizamos la transformación para desacoplar el sistema. Tras multiplicar por  $Q^{-1}$  por la izquierda las relaciones (4.94) y (4.95), y realizar la transformación (recordando que  $Q^{-1} A Q = \Lambda$ ) obtenemos

$$w_{n+1} = w_n + hG_{p+1}(h\Lambda, h\Lambda, \dots, h\Lambda) \alpha_1, \quad (4.96)$$

con las etapas dadas recursivamente por

$$\begin{aligned} \alpha_1 &= \Lambda w_n \\ \alpha_2 &= \Lambda(I + hG_2 \Lambda) w_n \\ \alpha_3 &= \Lambda(I + hG_3(h\Lambda) \Lambda) w_n \\ &\vdots \\ \alpha_p &= \Lambda(I + hG_p(h\Lambda, h\Lambda, \dots, h\Lambda) \Lambda) w_n, \end{aligned} \quad (4.97)$$

como se deduce sin dificultad observando que se verifica la relación  $Q^{-1} G_i(hA, hA, \dots, hA) A Q = G_i(h\Lambda, h\Lambda, \dots, h\Lambda) \Lambda$ .

Hemos visto de este modo que los sistemas en diferencias que definen nuestros métodos (cuando son aplicados al sistema test) quedan desacoplados y es evidente que tras realizar la transformación

$y = Qw$ , el sistema test también queda desacoplado (pasa a tener la forma  $w' = \Lambda w$ ). Por esta razón, en lo que sigue bastará considerar la ecuación escalar test dada por  $y' = \lambda y$  con  $\lambda \in \mathbb{C}$  y  $\Re(\lambda) < 0$ , a la hora de realizar el estudio de las propiedades de estabilidad lineal de los métodos aplicables a sistemas separados.

Al aplicar un método de  $p$  etapas para sistemas definido por (4.4–4.6) a la función escalar test obtenemos

$$y_{n+1} = R(z) y_n, \quad (4.98)$$

donde  $R(z)$  es la función de estabilidad lineal asociada con  $z = h\lambda$ . A partir de (4.4–4.6) se obtienen recursivamente

$$\begin{aligned} k_1 &= \lambda y_n \\ S_2 &= z \\ S_3 &= z \\ &\vdots \\ S_p &= z, \end{aligned} \quad (4.99)$$

y por tanto, la función de estabilidad lineal toma la forma

$$R(z) = 1 + z G_{p+1}(z, z, \dots, z). \quad (4.100)$$

Obsérvese que con la modificación en la definición de los términos (matriciales)  $S_i$  (mediante la inclusión de las funciones  $G_i$ ) a la que nos hemos visto obligados en el caso de sistemas separados, la función de estabilidad lineal viene dada de forma más sencilla.

Conviene comentar en este punto que, con la modificación introducida cuando hemos estudiado los métodos de tres etapas para sistemas mediante la introducción del término matricial  $\tilde{S}_3 = S_3 - S_2$  en el apartado 4.2.2, se verificará trivialmente a partir de (4.99) que  $\tilde{S}_3 = 0$  (cuando aplicamos el método a la ecuación escalar test). Por tanto, la función de estabilidad lineal vendrá dada en este caso por  $R(z) = 1 + z \tilde{G}_4(z, 0)$ , como es fácil deducir de (4.49).

**4.3.2 Un primer ejemplo de método L-estable de tres etapas y orden cuatro.** En el apartado 4.1.3 vimos un primer ejemplo de método L-estable de dos etapas y orden tres para sistemas separados, así como varios experimentos numéricos en los dos apartados

siguientes que nos sirvieron para ilustrar el buen comportamiento de dicha fórmula al ser aplicada a diversos problemas rígidos.

En (4.2.2) hemos obtenido la expresión general de los métodos de tres etapas y orden cuatro para sistemas separados. Evidentemente la mayor parte de dichos métodos no serán de interés práctico, pues el coste computacional que requieren para su implementación puede ser demasiado alto, hasta el punto de anular la principal ventaja de los mismos (disponer de aproximaciones a la matriz Jacobiana sin evaluaciones adicionales). Para reducir dicho coste computacional, como ya hemos comentado con anterioridad, nos centraremos en aquellas fórmulas que requieran de una única factorización LU por cada paso de integración. Además, buscaremos fórmulas con buenas propiedades de estabilidad lineal (tales como la A-estabilidad y la L-estabilidad) para que sean aplicables a problemas rígidos.

A continuación pasamos a ver un primer ejemplo de método L-estable de tres etapas y orden cuatro para sistemas del tipo considerado. Análogamente a como ocurre al estudiar fórmulas diagonalmente implícitas de tipo Runge-Kutta (SDIRK), hemos de elegir una función de estabilidad lineal de tipo racional con un único polo real (de este modo bastará realizar una única factorización LU por cada paso de integración). Se puede comprobar que para que el método sea de orden cuatro y L-estable, es necesario que la función de estabilidad lineal tenga el único polo (real) de multiplicidad al menos cuatro. Además, existe un único valor (del polo) que permite obtener las propiedades anteriores con multiplicidad exactamente cuatro, y viene dado por la raíz del polinomio

$$24x^4 - 96x^3 + 72x^2 - 16x + 1, \quad (4.101)$$

que toma aproximadamente el valor  $a \approx 0.57281606$  (véase [20], pp. 96–98 para más detalles). La correspondiente función de estabilidad lineal es una función racional  $R(z)$  cuyo numerador es de grado tres y cuyo denominador es de grado cuatro, y que viene dada en términos del valor  $a$  al que hemos hecho referencia anteriormente por

$$R(z) = \frac{6 + 6(1 - 4a)z + 3(1 - 8a + 12a^2)z^2 + (1 - 12a + 36a^2 - 24a^3)z^3}{6(1 - az)^4}. \quad (4.102)$$

Una forma sencilla de obtener un método verificando todo lo anterior, es proceder como sigue:

*Paso 1.* Damos a los parámetros libres  $a_{4,32}$  y  $c_3$  en (4.68) los valores

$$a_{4,32} = 0, \quad c_3 = \frac{6 + \sqrt{6}}{10}. \quad (4.103)$$

Nótese que el valor que tomamos para  $c_3$  es el que nos permitía obtener orden cinco con tres etapas en el caso escalar.

*Paso 2.* Sustituimos estos valores en (4.68), y a su vez la solución así obtenida (que nos proporcionaba orden cuatro para los métodos de tres etapas de tipo polinomial) en las relaciones (4.74–4.80). Obtenemos así una familia de métodos de tres etapas de tipo racional y orden cuatro. Nótese que una fórmula de dicha familia queda completamente determinada a partir de los valores de los coeficientes  $d_*$ .

*Paso 3.* Finalmente, de entre las muchas opciones que tenemos a la hora de obtener un método cuya función de estabilidad lineal asociada sea la que deseamos, optamos por anular los siguientes coeficientes

$$d_{4,3} = d_{4,23} = d_{4,32} = 0, \quad (4.104)$$

y por tomar

$$d_{3,2} = -a, \quad (4.105)$$

obteniendo de este modo en términos del valor  $a$  que mencionamos anteriormente la solución

$$\begin{aligned} c_2 &= \frac{6 - \sqrt{6}}{10}, \quad c_3 = \frac{6 + \sqrt{6}}{10}, \quad c_4 = 1, \quad d_{3,2} = -a, \\ d_{4,2} &= -4a, \quad d_{4,3} = 0, \quad d_{4,22} = 6a^2, \quad d_{4,23} = 0, \quad d_{4,32} = 0, \\ d_{4,222} &= -4a^3, \quad d_{4,2222} = a^4, \quad n_{3,2} = \frac{(6 - 5a) - \sqrt{6}}{5}, \\ n_{4,2} &= \frac{1 - 8a}{2}, \quad n_{4,3} = \frac{9 + \sqrt{6}}{36}, \quad n_{4,22} = \frac{36a^2 - 12a + 1}{6}, \\ n_{4,23} &= \frac{6(1 - 12a) - (1 + 8a)\sqrt{6}}{72}, \quad n_{4,32} = 0, \\ n_{4,222} &= \frac{-96a^3 + 72a^2 - 16a + 1}{24}. \end{aligned} \quad (4.106)$$

El resto de parámetros los consideramos nulos.

Nótese que en la fórmula anterior se ha introducido el coeficiente  $d_{4,2222}$  pese a que, como ya sabemos, no interviene en las condiciones



de orden cuatro. La razón que nos mueve a introducir dicho coeficiente (y el término  $S_2^4$  que lo acompaña) no es otra que la de lograr que el método tenga asociada la función de estabilidad lineal (4.102) (cuyo denominador es de grado 4) y sea por tanto L-estable.

Pasamos a resumir las razones que nos han movido a anular los coeficientes en (4.104). Hemos anulado en (4.104) los coeficientes  $d_*$  correspondientes a aquellos términos que aparecerían multiplicados por algún factor  $\tilde{S}_3$  (esto es, los que tienen algún tres en el subíndice), de modo que a la hora de hacer la factorización LU sólo intervenga la matriz  $S_2$ .

Por otra parte, hemos tomado el parámetro  $d_{3,2} = -a$  en (4.105) para que de este modo no sean necesarias factorizaciones LU adicionales a la hora de calcular las etapas (concretamente la etapa  $k_3$ ). Podría parecer mejor reducir el coste computacional asociado al método, anulando el coeficiente  $d_{3,2}$  en (4.105), pero diversos experimentos numéricos muestran que esta elección del parámetro conduce en ocasiones a errores muy grandes al integrar algunos problemas rígidos con tamaños del paso  $h$  moderados.

Conviene comentar en este punto que la fórmula L-estable de tres etapas y orden cuatro que acabamos de construir (y otras que podríamos haber construido análogamente) mejora, en lo que respecta al número de evaluaciones de la función  $f$  que se requieren en cada paso de la integración, los resultados que se obtienen con las fórmulas diagonalmente implícitas de tipo Runge-Kutta (SDIRK). En [20], pp. 98, se muestra que son necesarias al menos cuatro etapas si se desea obtener una fórmula SDIRK de orden cuatro y L-estable.

A continuación ilustraremos mediante varios experimentos numéricos el comportamiento de la fórmula que acabamos de describir.

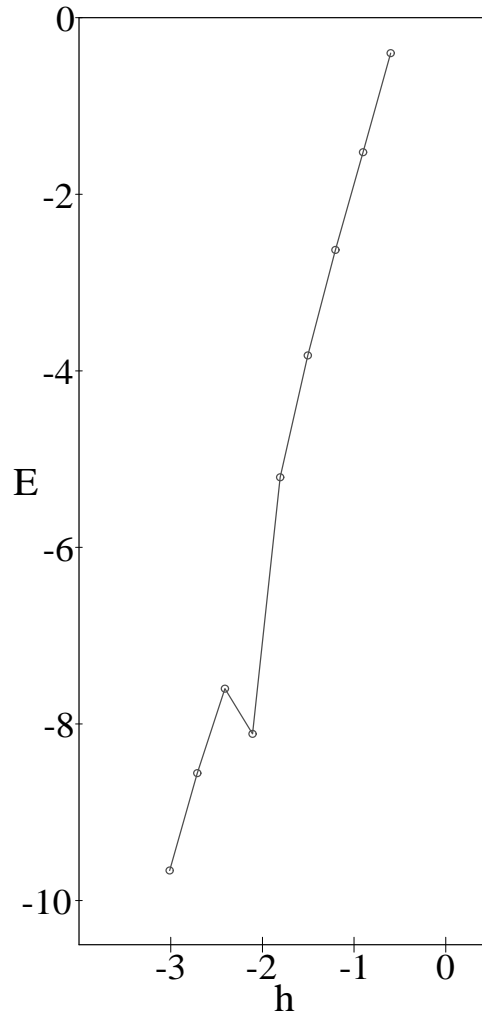
### 4.3.3 Varios experimentos numéricos con problemas rígidos.

Repetimos a continuación los experimentos numéricos que realizamos en su momento en los apartados 4.1.4 y 4.1.5, pero utilizando ahora el método L-estable de tres etapas y orden cuatro que acabamos de describir.

Al aplicar nuestro método al sistema de EDOs asociado a la ecuación de Burgers (obtenido por el método de líneas) que describimos en (4.21), tomando los valores  $N = 24$  y  $\nu = 0.2$  e integrando con paso fijo  $h = 0.04$ , se obtiene una representación gráfica de la solu-

ción numérica que no mostramos por ser indistinguible de la recogida en la figura 4.1 del apartado 4.1.4.

En la figura 4.6 se muestra en doble escala logarítmica el error  $E$



**Figura 4.6:** Error como función del paso  $h$  (doble escala logarítmica) para nuestro método.

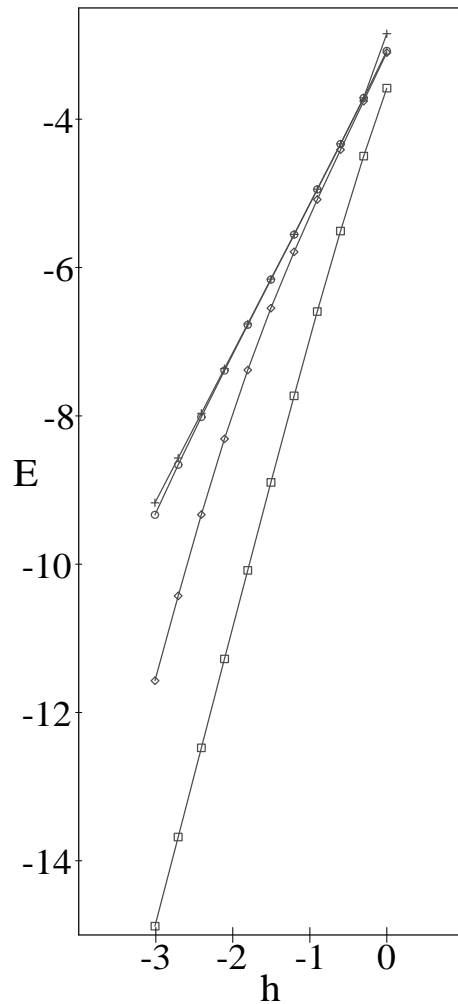
(utilizando la norma euclídea en  $\mathbb{R}^{24}$ ) en el punto  $t = 1$  que se produce al integrar con paso fijo  $h = 2^{-k}$  para los valores  $k = 2, 3, \dots, 10$  el problema anterior. Al comparar dicha gráfica con la figura 4.2 que obteníamos en el apartado 4.1.4 se observa que para tamaños del paso

$h$  grandes se comporta algo mejor el método de orden tres, pero al ir reduciendo el tamaño del paso  $h$  acaba siendo más preciso el de orden cuatro como cabía esperar.

Tomando los valores  $N = 24$  y  $\nu = 0.004$  en (4.21) e integrando con paso fijo  $h = 0.04$  se obtiene una representación gráfica de la solución numérica que tampoco reproducimos aquí por ser indistinguible de la mostrada en la figura 4.3 (obtenida con el método de orden tres).

En la gráfica 4.7 representamos en doble escala logarítmica el error  $E$  en la norma  $L_2$  que se comete al integrar a lo largo de  $10 \cdot 2^k$  pasos el problema (4.22), utilizando los tamaños de paso  $h = 2^{-k}$  (con  $k = 0, 1, \dots, 10$ ). Los valores de los parámetros en dicho problema se han tomado como en el apartado 4.1.5 para facilitar la comparación de los resultados obtenidos con este método de tres etapas con aquellos que obtuvimos para la fórmula de dos etapas (véase la figura 4.4). Se observa (comparando dichas figuras) que para los valores  $b = 1000000$  (cruces en la figura) y  $b = 10000$  (círculos) se obtienen resultados muy parecidos con ambos métodos. En ambos casos, para el rango de tamaños de paso considerados, se observa un orden de convergencia próximo a dos (claramente inferior al que cabría esperar). Al tomar  $b = 100$  (diamantes) podemos ver en la figura 4.7 que el orden se incrementa de dos a cuatro cuando reducimos el tamaño del paso considerado y, para pasos pequeños, el error es bastante menor que el obtenido con el método de dos etapas. Como ya hemos comentado anteriormente, el fenómeno de la reducción en el orden de convergencia está relacionado con el concepto de B-convergencia y muchos métodos implícitos muestran este comportamiento cuando son aplicados a algunas ecuaciones diferenciales rígidas (no lineales). Finalmente, para el valor  $b = 1$  (cuadrados) el problema no es rígido y nuestro método muestra orden cuatro (como era de esperar) y proporciona una aproximación a la solución exacta mucho mejor que la que obteníamos con el método de dos etapas.

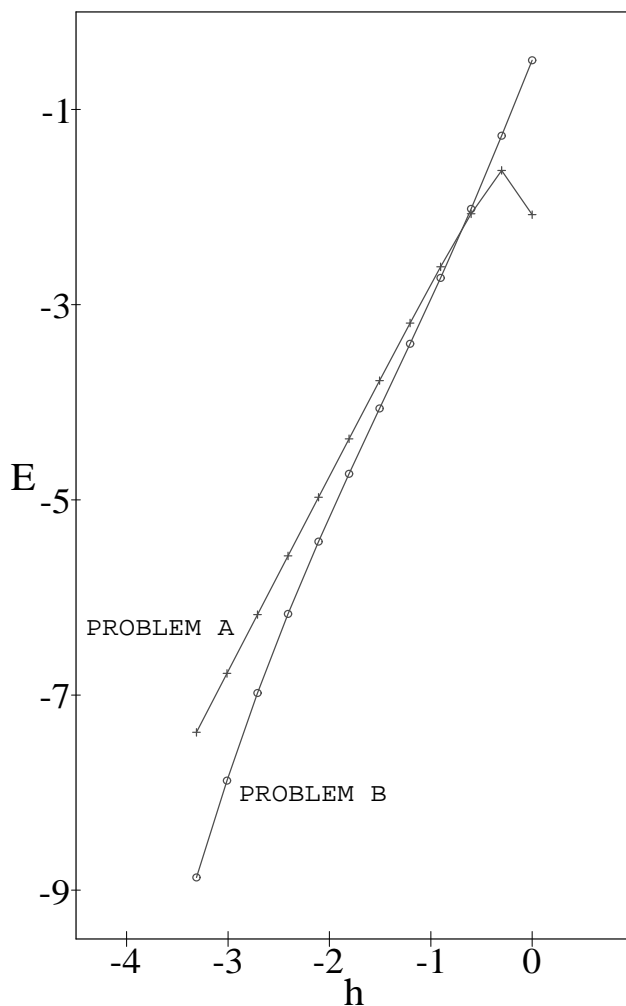
Finalmente, hemos repetido los experimentos numéricos que realizamos en el apartado 4.1.5 para los problemas A y B (véanse (4.24) y (4.26) resp.) y los resultados obtenidos se muestran en la figura 4.8. Comparándola con la gráfica 4.5 que obteníamos al aplicar el método de orden tres se observa que, en estos dos problemas, las diferencias entre los resultados obtenidos aplicando ambos métodos son poco apreciables para el rango de tamaños de paso que hemos considerado.



**Figura 4.7:** Error como función del tamaño de paso en doble escala logarítmica para nuestro método (problema autónomo).

**4.3.4 Otros ejemplos de métodos de dos y tres etapas A-estables y L-estables.** En apartados anteriores hemos obtenido un par de métodos L-estables para sistemas separados y nos proponemos ahora describir algún método más para completar nuestro estudio.

Partiendo de la expresión general de los métodos de dos etapas y orden tres de tipo racional para sistemas separados (véase (4.39–4.42)) construiremos ahora un método A-estable. Para ello partiremos de



**Figura 4.8:** Error como función del tamaño de paso en doble escala logarítmica para nuestro método (problemas no autónomos).

una función  $G_3(S_2)$  de la forma  $(I - a S_2)^{-2} N(S_2)$  para una cierta constante  $a \in \mathbb{R}$  y una cierta función polinomial  $N$  de grado dos, de modo que sólo sea necesaria una factorización LU por paso de integración. Como deseamos que el método resultante sea de orden tres, lo anterior equivale a tomar en (4.42) los valores  $d_1 = -2a$  y  $d_2 = a^2$  (el resto de parámetros  $n_i = d_i = 0$  para  $i \geq 3$ ). Ahora sólo falta

determinar el valor de la constante  $a$  de modo que la fórmula resulte ser A-estable. Recordando que la función de estabilidad lineal asociada  $R$  viene dada en la forma  $R(z) = 1 + z G_3(z)$  (véase el apartado 4.3.1) se tiene en nuestro caso que

$$R(z) = \frac{6 + 6(1 - 2a)z + 3(1 - 4a + 2a^2)z^2 + (1 - 6a + 6a^2)z^3}{6(1 - az)^2}, \quad (4.107)$$

y obtenemos el valor de  $a$  al anular el coeficiente del término en  $z^3$  del numerador de (4.107), esto es resolviendo  $1 - 6a + 6a^2 = 0$ . De los dos posibles valores que se obtienen, el que proporciona la A-estabilidad viene dado por

$$a = \frac{3 + \sqrt{3}}{6}. \quad (4.108)$$

Para más detalles relativos a la A-estabilidad que proporciona la anterior función  $R(z)$ , [20], pp. 97 es de nuevo una buena referencia. Resumiendo, los parámetros no nulos que determinan el método A-estable de dos etapas y orden tres para sistemas separados vendrán dados por

$$c_2 = \frac{2}{3}, \quad c_3 = 1, \quad d_1 = -\frac{3 + \sqrt{3}}{3}, \quad d_2 = \frac{2 + \sqrt{3}}{6}, \quad n_1 = -\frac{3 + 2\sqrt{3}}{6}. \quad (4.109)$$

Veamos ahora cómo obtener un método de dos etapas y orden tres de tipo racional para sistemas, con minimización de la parte principal del error local de truncación y L-estable. Para ello partimos de la expresión general que vendrá dada por (4.39–4.41), pero donde ahora la función  $G_3$  toma la forma (4.43) y nos garantiza todas las propiedades que buscamos a excepción de la L-estabilidad. Razones análogas a las ya comentadas al construir otros métodos nos mueven a tomar  $G_3(S_2)$  de la forma  $(I - aS_2)^{-4}N(S_2)$  para una cierta constante  $a \in \mathbb{R}$  y una cierta función polinomial  $N$  de grado tres. Esto equivale evidentemente a tomar en (4.43) los valores  $d_1 = -4a$ ,  $d_2 = 6a^2$ ,  $d_3 = -4a^3$  y  $d_4 = a^4$  (el resto de parámetros  $d_i$  nulos y los  $n_i = 0$  para  $i \geq 4$ ). El valor de la constante  $a$  lo determinaremos de modo que la fórmula resulte ser L-estable. Tras anular el coeficiente del término en  $z^4$  del numerador de la función de estabilidad lineal asociada  $R(z)$ , es fácil

comprobar que se obtiene la siguiente función de estabilidad lineal

$$R(z) = \frac{6+6(1-4a)z+3(1-8a+12a^2)z^2+(1-12a+36a^2-24a^3)z^3}{6(1-az)^4}, \quad (4.110)$$

para el método L-estable resultante, en función del valor de  $a$  que viene dado por la raíz del polinomio  $1-16a+72a^2-96a^3+24a^4=0$  que toma aproximadamente el valor

$$a \approx 0.57281606. \quad (4.111)$$

En la tabla de [20], pp. 98 se recoge este valor que proporciona la L-estabilidad. Obsérvese que el valor de  $a$  en (4.111) es el mismo que obteníamos en el apartado 4.3.2 para el método L-estable de tres etapas y orden cuatro. Más aún, la función de estabilidad lineal asociada a dicho método (véase (4.102)) coincide con la asociada al método que acabamos de construir.

La fórmula L-estable de dos etapas y orden tres con minimización de la parte principal del error local de truncación queda completamente determinada (en función del valor de  $a$  recogido en (4.111)) por los valores

$$\begin{aligned} c_2 &= \frac{2}{3}, & c_3 &= 1, & d_1 &= -4a, & d_2 &= 6a^2, & d_3 &= -4a^3, \\ d_4 &= a^4, & n_1 &= \frac{1-8a}{2}, & n_2 &= \frac{1-12a+36a^2}{6}, \\ n_3 &= \frac{1-16a+72a^2-96a^3}{24}, \end{aligned} \quad (4.112)$$

y tomando el resto de parámetros nulos.

A continuación describiremos un método de tres etapas y orden cuatro de tipo racional para sistemas separados que es A-estable. Partiendo de la expresión general obtenida en el apartado 4.2.2 y en la línea de lo que venimos comentando al construir los métodos, describimos los pasos seguidos para obtener la fórmula:

*Paso 1.* Damos a los parámetros libres  $a_{4,32}$  y  $c_3$  en (4.68) los valores

$$a_{4,32} = 0, \quad c_3 = \frac{6 + \sqrt{6}}{10}. \quad (4.113)$$

El valor que tomamos para  $c_3$  es el que nos permitía obtener orden cinco con tres etapas en el caso escalar.

*Paso 2.* Tras sustituir estos valores en (4.68) trasladamos la solución así obtenida (que nos proporcionaba orden cuatro para los métodos de tres etapas de tipo polinomial) a las relaciones (4.74–4.80). Obtenemos así una familia de métodos de tres etapas de tipo racional y orden cuatro que queda completamente determinada a partir de los valores de los coeficientes  $d_*$ .

*Paso 3.* De entre las muchas opciones que tenemos a la hora de obtener un método cuya función de estabilidad lineal asociada sea la que deseamos, optamos por anular los siguientes coeficientes

$$d_{4,3} = d_{4,23} = d_{4,32} = 0, \quad (4.114)$$

y por tomar el resto de parámetros libres como sigue

$$d_{3,2} = -a, \quad d_{4,2} = -3a, \quad d_{4,22} = 3a^2, \quad d_{4,222} = -a^3, \quad (4.115)$$

para que sólo sea necesaria una factorización LU por cada paso de integración.

*Paso 4.* Finalmente determinamos el valor de  $a$  de modo que la fórmula resulte ser A-estable. Para ello obtenemos  $a$  como una raíz del polinomio  $-1 + 12a - 36a^2 + 24a^3 = 0$  que toma aproximadamente el valor

$$a \approx 1.06857902, \quad (4.116)$$

y la función de estabilidad lineal del método vendrá dada en términos de  $a$  por

$$R(z) = \frac{6 + 6(1 - 3a)z + 3(1 - 6a + 6a^2)z^2 + (1 - 9a + 18a^2 - 6a^3)z^3}{6(1 - az)^3}. \quad (4.117)$$

La tabla de [20], pp. 97 recoge este valor de  $a$  que proporciona la A-estabilidad.

La fórmula A-estable de tres etapas y orden cuatro queda completamente determinada (en función del valor de  $a$  recogido en (4.116)) por los valores

$$\begin{aligned} c_2 &= \frac{6 - \sqrt{6}}{10}, & c_3 &= \frac{6 + \sqrt{6}}{10}, & c_4 &= 1, & d_{3,2} &= -a, \\ d_{4,2} &= -3a, & d_{4,3} &= 0, & d_{4,22} &= 3a^2, & d_{4,23} &= 0, \\ d_{4,32} &= 0, & d_{4,222} &= -a^3, & n_{3,2} &= \frac{(6 - 5a) - \sqrt{6}}{5}, \end{aligned}$$



$$\begin{aligned}
n_{4,2} &= \frac{1 - 6a}{2}, & n_{4,3} &= \frac{9 + \sqrt{6}}{36}, & n_{4,22} &= \frac{18a^2 - 9a + 1}{6}, \\
n_{4,23} &= \frac{6(1 - 9a) - (1 + 6a)\sqrt{6}}{72}, & n_{4,32} &= 0, \\
n_{4,222} &= \frac{-24a^3 + 36a^2 - 12a + 1}{24}, & & & & (4.118)
\end{aligned}$$

y tomando el resto de parámetros nulos.

Por último describiremos un método de tres etapas y orden cuatro de tipo racional para sistemas, con minimización de la parte principal del error local de truncación y L-estable. En el apartado 4.2.2 obtuvimos la expresión general de las fórmulas de tres etapas y orden cuatro de tipo racional para sistemas, con minimización de la parte principal del error local de truncación. Concretamente (aunque hay otras posibilidades) y en línea con lo que venimos comentando al construir los otros métodos, describimos los pasos seguidos para obtener la fórmula:

*Paso 1.* Anulamos el parámetro libre  $a_{4,322}$  en (4.69) y sustituimos los valores así obtenidos en las relaciones (4.81–4.93). Obtenemos así los coeficientes de una familia de métodos de tres etapas de tipo racional y orden cuatro con minimización de la parte principal del error local de truncación. Un método de dicha familia queda completamente determinado al fijar los valores de los coeficientes  $d_*$ .

*Paso 2.* De entre las distintas posibilidades que tenemos a la hora de obtener un método cuya función de estabilidad lineal asociada sea la que deseamos, optamos por anular todos los coeficientes  $d_*$  que aparecen multiplicados por términos que contienen a  $S_3$ . El resto de parámetros libres los tomaremos como sigue

$$\begin{aligned}
d_{3,2} &= -2a, & d_{3,22} &= a^2, & d_{4,2} &= -5a, & d_{4,22} &= 10a^2, \\
d_{4,222} &= -10a^3, & d_{4,2222} &= 5a^4, & d_{4,22222} &= -a^5. & & (4.119)
\end{aligned}$$

En la fórmula anterior hemos introducido el coeficiente  $d_{4,22222}$  pese a que, como ya sabemos, no interviene en las condiciones de orden cuatro ni en las de orden cinco (que utilizamos para minimizar la parte principal del error local de truncación). Hacemos esto para lograr que el método tenga asociada una función de estabilidad lineal que veremos a continuación (cuyo denominador es de grado cinco) y sea por tanto L-estable.

*Paso 3.* Finalmente determinamos el valor de  $a$  de modo que la fórmula resulte ser L-estable. Para ello debemos tomar  $a$  como la raíz del polinomio  $-1 + 25a - 200a^2 + 600a^3 - 600a^4 + 120a^5 = 0$  que toma aproximadamente el valor

$$a \approx 0.27805384, \quad (4.120)$$

y la función de estabilidad lineal  $R(z)$  del método vendrá dada en términos de  $a$  por

$$R(z) = \frac{24 + 24(1-5a)z + 12(1-10a+20a^2)z^2}{24(1-az)^5} + \frac{4(1-15a+60a^2-60a^3)z^3 + (1-20a+120a^2-240a^3+120a^4)z^4}{24(1-az)^5}. \quad (4.121)$$

En [20], pp. 98 se recoge este valor de  $a$  que proporciona la L-estabilidad.

Resumiendo, los coeficientes del método obtenido vienen dados por

$$\begin{aligned} c_2 &= \frac{6 - \sqrt{6}}{10}, & c_3 &= \frac{6 + \sqrt{6}}{10}, & c_4 &= 1, & d_{3,2} &= -2a, \\ d_{3,22} &= a^2, & d_{4,2} &= -5a, & d_{4,22} &= 10a^2, & d_{4,222} &= -10a^3, \\ d_{4,2222} &= 5a^4, & d_{4,22222} &= -a^5, & n_{3,2} &= \frac{-(3 + 10a) + 2\sqrt{6}}{5}, \\ n_{3,22} &= \frac{(17 + 60a + 50a^2) - (3 + 40a)\sqrt{6}}{50}, & n_{4,2} &= \frac{1 - 10a}{2}, \\ n_{4,3} &= \frac{9 + \sqrt{6}}{36}, & n_{4,22} &= \frac{60a^2 - 15a + 1}{6}, \\ n_{4,23} &= \frac{6(1 - 15a) - (1 + 10a)\sqrt{6}}{72}, & n_{4,32} &= \frac{-1 + \sqrt{6}}{8}, \\ n_{4,33} &= \frac{1 + 4\sqrt{6}}{72}, & n_{4,222} &= \frac{-240a^3 + 120a^2 - 20a + 1}{24}, \\ n_{4,223} &= \frac{3(1 - 20a + 120a^2) + (-1 + 10a + 40a^2)\sqrt{6}}{144}, \\ n_{4,232} &= \frac{3(-1 + 10a) + 2(1 - 15a)\sqrt{6}}{48}, \\ n_{4,2222} &= \frac{600a^4 - 600a^3 + 200a^2 - 25a + 1}{120}, \end{aligned} \quad (4.122)$$

en función del valor de  $a$  recogido en (4.120) y tomando los restantes parámetros nulos.

#### 4.3.5 Experimentos numéricos con los métodos de dos y tres etapas A-estables y L-estables: La ecuación de Van der Pol.

Con el fin de ilustrar el comportamiento numérico de los métodos GRK que hemos descrito, realizaremos a continuación varios experimentos numéricos. Para ello aplicaremos nuestros métodos a varios problemas y compararemos los resultados así obtenidos con aquellos que se obtienen al utilizar algunas fórmulas clásicas.

El primer problema que consideraremos surge de la conocida ecuación de Van der Pol, que viene dada por

$$\varepsilon z'' + (z^2 - 1)z' + z = 0, \quad z(0) = a, \quad z'(0) = b. \quad (4.123)$$

Nótese que nuestros métodos no son aplicables al sistema de EDOs de primer orden dado por

$$\begin{aligned} z' &= y & z(0) &= a \\ y' &= \frac{1}{\varepsilon}((1 - z^2)y - z) & y(0) &= b \end{aligned} \quad (4.124)$$

que se obtiene procediendo como es habitual a la hora de transformar este tipo de ecuaciones de segundo orden, por no ser dicho sistema separado. Sin embargo, utilizando coordenadas de Liénard (véase [20], pp. 372 para más detalles) el problema (4.123) toma la forma

$$\begin{aligned} y' &= -z & y(0) &= \varepsilon b - a + \frac{a^3}{3} \\ z' &= \frac{1}{\varepsilon} \left( y + z - \frac{z^3}{3} \right) & z(0) &= a \end{aligned} \quad (4.125)$$

y nuestros métodos si son aplicables.

Integramos el problema anterior en el intervalo  $[0, 0.5]$ , tomando  $\varepsilon = 10^{-5}$ . Consideramos las condiciones iniciales siguientes

$$y(0) = 0.66666000001234554549467, \quad z(0) = 2., \quad (4.126)$$

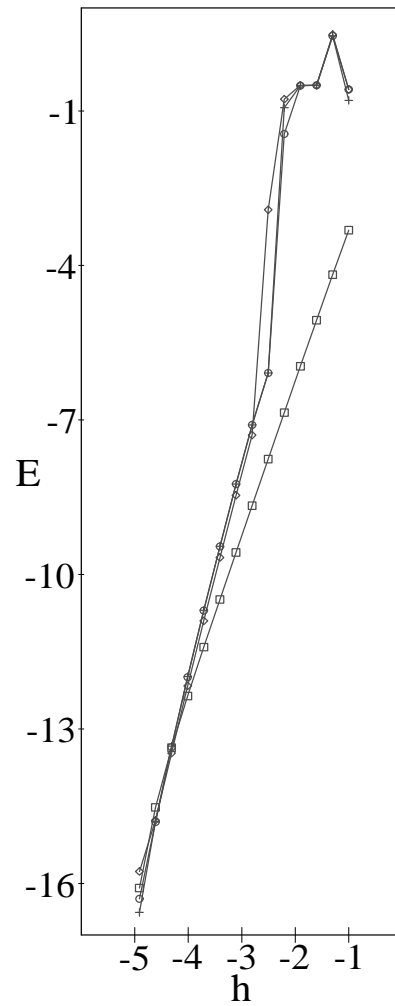
obtenidas a partir de las consideradas en [20], pp. 403, tras pasar a coordenadas de Liénard. Con los seis métodos numéricos obtenidos

en este capítulo, hemos integrado el problema anterior con paso fijo  $h$  y hemos medido el error relativo cometido en ambas componentes en el punto 0.5. Los errores se han medido respecto de una solución de referencia obtenida con un método de paso variable (concretamente el "gear" implementado en MAPLEV, trabajando con 30 dígitos y una tolerancia de  $10^{-20}$ ) suficientemente precisa para nuestros propósitos.

A efectos comparativos hemos realizado el mismo experimento numérico utilizando el método de tipo Rosenbrock de dos etapas y orden tres recogido en [12], pp. 334 y conocido como Calahan.

En las figuras 4.9 y 4.10 mostramos los errores relativos correspondientes a la primera y a la segunda componente respectivamente como función del tamaño de paso  $h$  (tomando  $h = 0.1 \cdot 2^{-k}$  para  $k = 0, 1, 2, \dots, 12$ ) en doble escala logarítmica para los distintos métodos de orden tres. El método de Calahan viene representado en dichas figuras por cuadrados. Con diamantes representamos el método de orden tres y A-estable, con círculos la primera fórmula de orden tres y L-estable que obtuvimos y con cruces el último método de orden tres y L-estable que hemos descrito.

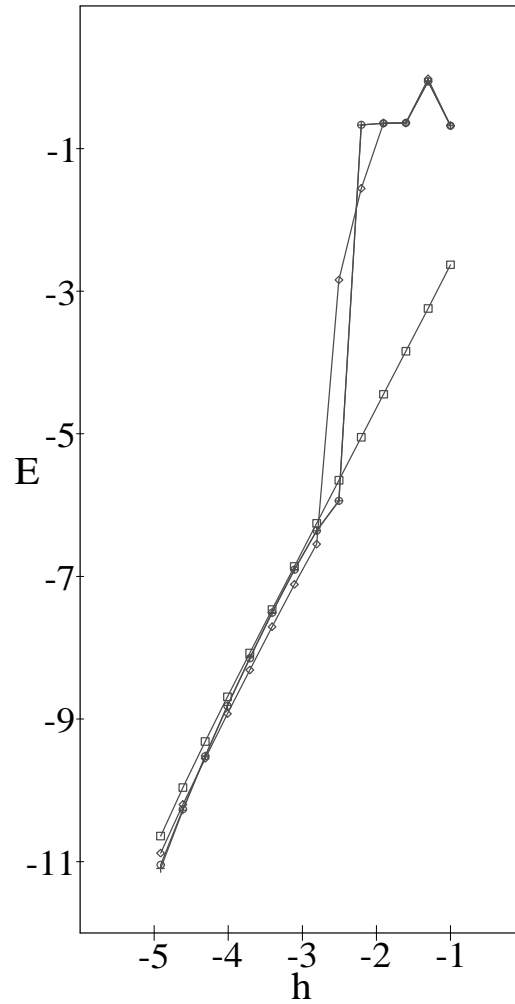
Se observa en dichas figuras que el método de Calahan al que hemos hecho referencia anteriormente se comporta mejor que los tres propuestos para tamaños de paso "grandes". Cuando vamos disminuyendo el tamaño del paso, los métodos se comportan de manera muy parecida y alguno de ellos mejoran los resultados proporcionados por la fórmula de Calahan. Conviene señalar en este punto que el problema considerado es muy rígido, con uno de los autovalores de la matriz Jacobiana variando en  $[-300000, -150000]$  y el otro pequeño y negativo a lo largo del intervalo de integración. Recordemos que el método de Calahan hace uso en cada paso de la integración de una evaluación de la matriz Jacobiana asociada al problema, en tanto que nuestros métodos, como ya hemos comentado con anterioridad, utilizan la aproximación a dicha matriz Jacobiana que proporciona la matriz  $S_2/h$ . Esto explica el comportamiento que se observa para tamaños del paso  $h$  grandes, pues entonces no cabe esperar que la aproximación proporcionada por la matriz  $S_2$  sea demasiado precisa (lo que en algunos problemas deteriora la aproximación obtenida). Cabe esperar que al utilizar dichas fórmulas con paso variable la situación mejore considerablemente, pero en esta memoria no abordaremos la construcción de métodos de dicho tipo y dejamos esta cuestión para



**Figura 4.9:** Error relativo (primera componente) como función del tamaño de paso en doble escala logarítmica para los métodos de orden tres.

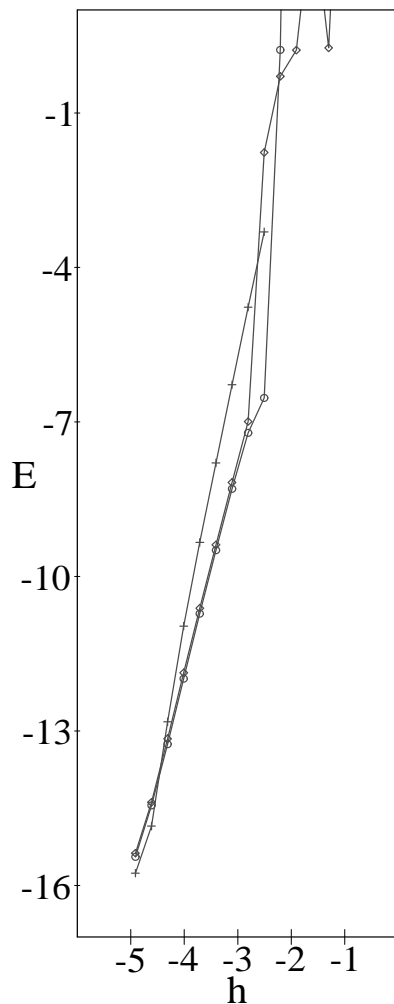
posteriores investigaciones.

Hemos repetido el experimento anterior con los tres métodos de orden cuatro. Los resultados se recogen en las figuras 4.11 y 4.12. En dichas figuras representamos con diamantes el método A-estable, con círculos la primera fórmula de orden cuatro y L-estable que obtuvimos y con cruces el último método de orden cuatro y L-



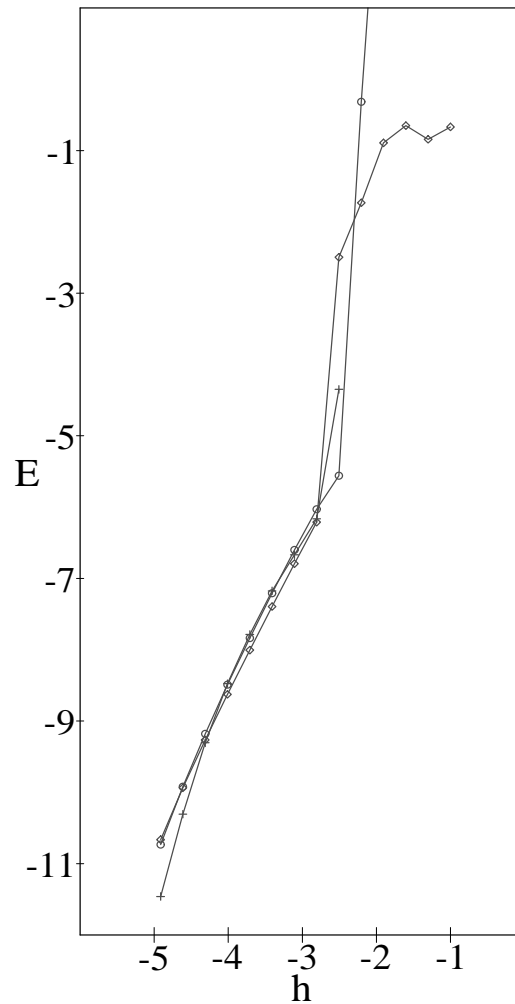
**Figura 4.10:** Error relativo (segunda componente) como función del tamaño de paso en doble escala logarítmica para los métodos de orden tres.

estable que hemos propuesto. En la fórmula descrita con cruces en dichas figuras hemos suprimido los resultados que se obtienen para tamaños del paso  $h$  grandes. Ello se debe a que en alguna de las integraciones la matriz a invertir resulta ser singular en alguno de los pasos. Observamos además que para tamaños de paso grandes los resultados numéricos no son en modo alguno satisfactorios, siendo



**Figura 4.11:** Error relativo (primera componente) como función del tamaño de paso en doble escala logarítmica para los métodos de orden cuatro.

peores que los que obteníamos con las fórmulas de orden tres. Las razones son las mismas que las comentadas anteriormente para los métodos de dos etapas, pero la situación se ve incluso agravada por ser la aproximación  $S_3/h$  a la matriz Jacobiana (para tamaños grandes del paso) aún peor que  $S_2/h$ . Se observa sin embargo que al reducir el tamaño del paso la situación mejora considerablemente.

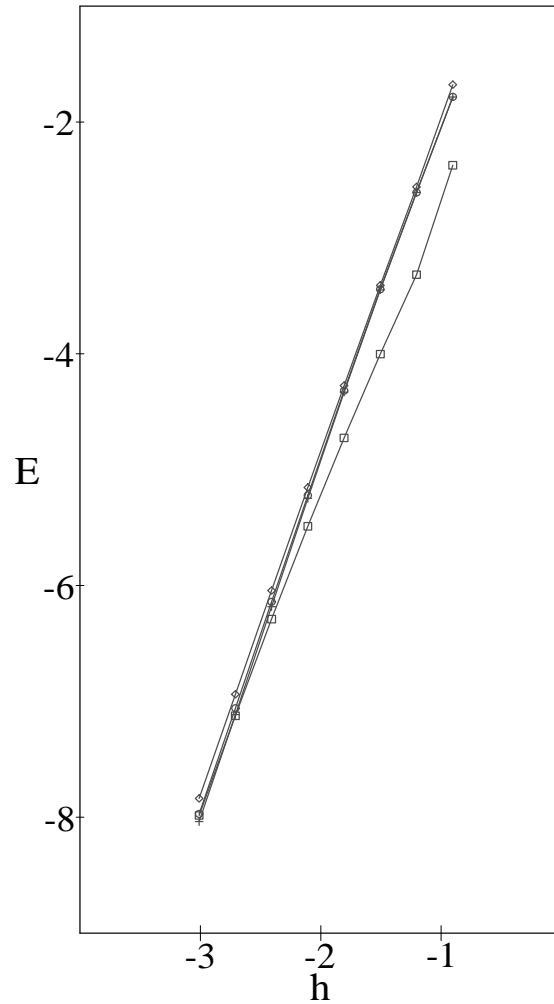


**Figura 4.12:** Error relativo (segunda componente) como función del tamaño de paso en doble escala logarítmica para los métodos de orden cuatro.

En cualquier caso, los métodos de dos y tres etapas que hemos introducido proporcionan buenas aproximaciones a la solución exacta del problema cuando consideramos tamaños de paso  $h$  moderados. De hecho, se obtienen buenas aproximaciones a la solución integrando con tamaños de paso mucho mayores que los que requeriría cualquier método explícito clásico. Las reducciones en el orden efectivo de los mé-



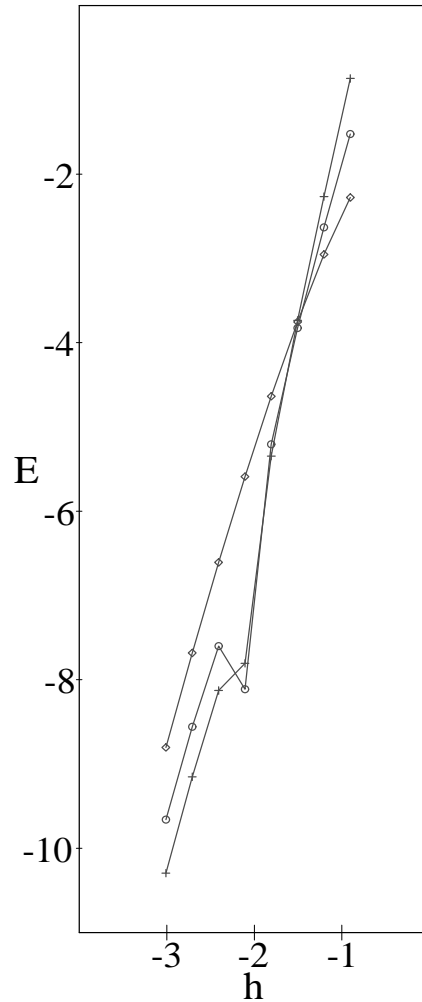
todos que se observan para este problema ocurren también al utilizar otros métodos A-estables y L-estables de tipo Runge-Kutta (véase [20], pp. 403–404).



**Figura 4.13:** Error como función del tamaño de paso en doble escala logarítmica para los métodos de orden tres.

**4.3.6 Experimentos numéricos con los métodos de dos y tres etapas A-estables y L-estables: La ecuación de Burgers.** Pasemos a considerar otro problema. Concretamente repetiremos el experi-

mento numérico realizado en el apartado 4.1.4 y consistente en aplicar los métodos obtenidos al sistema de EDOs que describimos en (4.21). En dicho problema, obtenido por el método de líneas a partir de la ecuación de Burgers, tomaremos los valores  $N = 24$  y  $\nu = 0.2$ .



**Figura 4.14:** Error como función del tamaño de paso en doble escala logarítmica para los métodos de orden cuatro.

En la figura 4.13 se muestra en doble escala logarítmica el error  $E$  (utilizando la norma euclídea en  $\mathbb{R}^{24}$ ) en el punto  $t = 1$  que se produce al integrar con paso fijo  $h = 2^{-k}$  para los valores  $k = 3, 4, \dots, 10$  el

problema anterior con los distintos métodos de orden tres. Hemos representado los métodos con los mismos símbolos que utilizamos en las figuras 4.9 y 4.10. Se observa en la gráfica que el método de Calahan se comporta ligeramente mejor para los tamaños más grandes de paso  $h$  considerados, pero según vamos reduciendo el tamaño de paso las diferencias se reducen y en algunos casos acaban siendo menores para las fórmulas que hemos introducido. Además, para este problema, los tres métodos introducidos proporcionan aproximaciones muy parecidas y son difíciles de distinguir en la figura. Es fácil comprobar que las pendientes de las curvas representadas en la gráfica son muy próximas a tres (como cabía esperar).

Repitiendo el experimento anterior con los tres métodos de orden cuatro se obtiene la figura 4.14. Hemos utilizado los mismos símbolos que en las figuras 4.11 y 4.12 para representar los métodos. Es fácil observar que para tamaños grandes del paso  $h$  las aproximaciones numéricas que proporcionan los métodos de orden cuatro son en algunos casos peores que las que obteníamos con las fórmulas de orden tres. Sin embargo, cuando vamos considerando tamaños de paso más pequeños las fórmulas de orden cuatro acaban siendo más precisas que las de orden tres. Observamos asimismo que el método que proporciona las mejores aproximaciones para tamaños de paso pequeños es el representado en la figura por cruces, aunque también es el que lleva asociado un mayor coste computacional. Las pendientes de las curvas representadas en la figura se aproximan a cuatro a medida que vamos reduciendo el tamaño de paso considerado.



## Capítulo 5

# Métodos para ecuaciones especiales de segundo orden.

En muchas aplicaciones prácticas, aparecen ecuaciones diferenciales de segundo orden del tipo  $y'' = f(x, y)$ . Es una práctica común en estos casos transformar este sistema en otro equivalente de primer orden y aplicar al sistema resultante uno de los muchos y bien conocidos métodos numéricos existentes. Sin embargo, existen métodos específicos (para ecuaciones de segundo orden) que obtienen importantes ventajas de la aplicación directa a este problema. Entre otros podemos citar los métodos multipaso de Störmer [42] y los métodos Runge-Kutta-Nyström [33] (para más detalles véase por ejemplo [18]).

Nos proponemos estudiar a modo de ejemplo una modificación de los métodos propuestos en capítulos precedentes, de modo que sean aplicables a algunas EDOs de segundo orden del tipo anterior. Veremos que esto es posible y que, análogamente a como ocurría para ecuaciones de primer orden, se obtienen importantes ventajas respecto de los métodos clásicos (incluyendo los específicamente diseñados para este tipo de problemas) que se aplican a este tipo de ecuaciones.

En este capítulo, tras motivar la introducción de los nuevos métodos y dar un primer ejemplo de fórmula de dos etapas y orden cuatro, pasaremos a dar la expresión general de los métodos de dos etapas que consideraremos. Posteriormente estudiaremos las condiciones de orden de los métodos y obtendremos fórmulas de dos etapas y orden cuatro con buenas propiedades de estabilidad cuando son aplicadas a problemas oscilatorios. Finalmente realizaremos algunos experimentos numéricos con problemas oscilatorios.

## 5.1 Métodos de dos etapas para ecuaciones especiales de segundo orden.

**5.1.1 Expresión general de la ecuación especial de segundo orden.** Los métodos que proponemos en este capítulo son aplicables a problemas escalares autónomos del tipo

$$y''(x) = f(y(x)), \quad y(x_0) = y_0, \quad y'(x_0) = z_0, \quad f : \mathbb{R} \rightarrow \mathbb{R}, \quad (5.1)$$

esto es, ecuaciones escalares autónomas de segundo orden que no dependen de la primera derivada  $y'(x)$ , con condiciones iniciales de tipo Dirichlet.

Es bien conocido que para este tipo de problemas existen métodos clásicos que sacan partido de la no dependencia de la función  $f$  respecto de  $y'$  a la hora de conseguir una reducción en el número de evaluaciones por paso necesarias para obtener un determinado orden de consistencia (entre otras ventajas). Por ejemplo, entre la familia de métodos de tipo multipaso, encontramos los llamados métodos de Störmer, diseñados específicamente para reducir el número de evaluaciones necesarias para alcanzar un determinado orden en problemas generales de este tipo. En [18], pp. 461–474, se estudian estos métodos con cierto detalle. Los métodos de tipo Runge-Kutta diseñados para reducir el número de etapas necesario a la hora de obtener un determinado orden, se conocen como métodos Runge-Kutta-Nyström (abreviadamente RKN), o simplemente métodos de Nyström. Para más detalles véase por ejemplo [18], pp. 283–301.

Los métodos que propondremos a continuación están más relacionados con los RKN pues, como éstos, vienen dados en términos de ciertas etapas. Las fórmulas que veremos a continuación son explícitas y permitirán obtener, para el tipo de problemas considerado, órdenes de consistencia mayores que en el caso de los métodos RKN del mismo número de etapas.

En esta memoria no abordaremos el estudio de las posibles generalizaciones del tipo de métodos que introduciremos a continuación a otros problemas más generales de segundo orden. Tampoco pretendemos estudiar el caso general de  $p$  etapas, si bien es evidente el procedimiento a seguir a partir de lo visto en anteriores capítulos para EDOs de primer orden. El propósito de este capítulo es mostrar, a

modo de ejemplo, cómo se pueden aplicar las ideas vistas en capítulos precedentes a otro tipo de problemas.

En algunos momentos, será conveniente reescribir el problema (5.1) en forma de sistema. Para ello denotaremos por  $z(x) = y'(x)$  tras lo cual el problema toma la forma

$$\begin{aligned} y'(x) &= z(x), & y(x_0) &= y_0, \\ z'(x) &= f(y(x)), & z(x_0) &= z_0. \end{aligned} \quad (5.2)$$

Pasamos a ver ya un primer ejemplo.

**5.1.2 Un primer ejemplo de método para la ecuación especial de segundo orden.** Para empezar, describiremos un método explícito de orden cuatro y dos etapas para el problema visto en el apartado anterior, que resulta ser exacto (salvo errores de redondeo) para  $y'' = -\alpha y + \beta$  con  $\alpha > 0, \beta \in \mathbb{R}$ . El método viene descrito por

$$\begin{aligned} y_{n+1} &= y_n + h \left( \left( \frac{\cos \sqrt{-s} - 1}{s} \right) (k_1 - c_1 s y'_n) + \left( \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}} \right) y'_n \right), \\ y'_{n+1} &= y'_n + \left( (\cos \sqrt{-s} - 1) y'_n + \left( \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}} \right) (k_1 - c_1 s y'_n) \right), \end{aligned} \quad (5.3)$$

donde

$$\begin{aligned} k_1 &= hf(y_n + hc_1 y'_n), \\ k_2 &= hf(y_n + h(c_2 y'_n + d_2 k_1)), \\ s &= \frac{k_2 - k_1}{(c_2 - c_1) y'_n + d_2 k_1}, \end{aligned} \quad (5.4)$$

y los valores de los parámetros vienen dados por

$$c_1 = \frac{3 - \sqrt{3}}{6}, \quad c_2 = \frac{3 + \sqrt{3}}{6}, \quad d_2 = \frac{\sqrt{3}}{6}. \quad (5.5)$$

El método está especialmente indicado en el caso de problemas altamente oscilatorios. Es fácil ver que el término  $s$  proporciona una aproximación a  $h^2 f_y$ , y que de hecho se verifica que  $s = h^2 f_y(y_n) + O(h^3)$ . En apartados posteriores realizaremos un estudio detallado de las condiciones de orden y propiedades de los métodos de dos etapas.

Obsérvese que el método (5.3) aparentemente no está definido para valores de  $s$  positivos (pues entonces tendríamos una raíz de un número negativo) o nulos. Sin embargo, veremos más adelante que en realidad sí que está bien definido también en estos casos.

De momento, ilustramos lo anterior mediante el siguiente experimento numérico: para el problema (oscilador perturbado)

$$y'' = -\alpha y + \varepsilon y^3, \quad y(0) = 1, \quad y'(0) = 0, \quad (5.6)$$

tomando distintos valores de los parámetros  $\alpha$  y  $\varepsilon$ , y distintos tamaños de paso (con paso fijo), hemos integrado el problema anterior con el método anteriormente descrito y con el método de Nyström de orden cuatro y tres evaluaciones por paso (véase [18], pág. 285). En las siguientes tablas se recogen los errores obtenidos. Como el problema anterior no admite solución en términos de funciones elementales, hemos optado por medir el error en la integral primera  $(\alpha y^2 + y'^2)/2 - \varepsilon y^4/4$ . En todos los casos se ha medido el error en  $x = 1000$  y tomando  $\varepsilon = 10^{-3}$ . Obsérvese que nuestro método es siempre superior al

**Tabla 5.1:** Errores para nuestro método.

error	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$\alpha = 100$	$5.912D - 03$	$2.103D - 04$	$6.751D - 06$	$2.122D - 07$
$\alpha = 10$	$6.912D - 06$	$2.192D - 07$	$6.930D - 09$	$2.203D - 10$
$\alpha = 1$	$6.621D - 09$	$1.970D - 10$	$5.520D - 12$	$1.296D - 13$

**Tabla 5.2:** Errores para el método de Nyström de orden 4.

error	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$\alpha = 100$	$5.000D + 01$	$3.309D + 01$	$1.667D + 00$	$5.294D - 02$
$\alpha = 10$	$1.701D - 01$	$5.389D - 03$	$1.675D - 04$	$5.172D - 06$
$\alpha = 1$	$1.715D - 05$	$5.307D - 07$	$1.626D - 08$	$4.882D - 10$

RKN, si bien la diferencia es más acusada cuando aumenta el carácter oscilatorio del problema (esto es, cuando  $\alpha \gg 1$ ).



**5.1.3 Una primera aproximación a los métodos de dos etapas.** Análogamente a como hicimos en el capítulo introductorio de esta memoria, a la hora de motivar la introducción de los nuevos métodos para EDOs autónomas de primer orden, nos proponemos estudiar una propiedad que nos lleva a formular los nuevos métodos para ecuaciones de segundo orden.

Comenzaremos suponiendo que la función  $f$  en (5.1) (y por tanto también en (5.2)) es lo suficientemente regular como para garantizar que los desarrollos que siguen tengan sentido. Es fácil deducir de la ecuación (5.1), derivando sucesivamente respecto de la variable  $x$  y utilizando la relación  $y'' = f(y)$ , que se verifican las relaciones

$$\begin{aligned} y''(x) &= f(y(x)), \\ y^3(x) &= f_y(y(x)) y'(x), \\ y^4(x) &= f_{yy}(y(x)) (y'(x))^2 + f_y(y(x)) y''(x) \\ &= f_{yy}(y(x)) (y'(x))^2 + f_y(y(x)) f(y(x)), \\ y^5(x) &= f_{yyy}(y(x)) (y'(x))^3 + 3f_{yy}(y(x)) y''(x) y'(x) + f_y(y(x)) y^3(x) \\ &= f_{yyy}(y(x)) (y'(x))^3 + 3f_{yy}(y(x)) f(y(x)) y'(x) + f_y^2(y(x)) y'(x). \end{aligned} \quad (5.7)$$

Consideremos ahora los desarrollos de Taylor de la solución y de su derivada en un entorno del punto  $x$  (esto es, centrado en  $h = 0$ )

$$\begin{aligned} y(x+h) &= y(x) + y'(x)h + \frac{1}{2}y''(x)h^2 + \frac{1}{6}y^3(x)h^3 + \frac{1}{24}y^4(x)h^4 \\ &\quad + O(h^5), \\ y'(x+h) &= y'(x) + y''(x)h + \frac{1}{2}y^3(x)h^2 + \frac{1}{6}y^4(x)h^3 + \frac{1}{24}y^5(x)h^4 \\ &\quad + O(h^5), \end{aligned} \quad (5.8)$$

o equivalentemente, tras sustituir las derivadas por sus correspondientes expresiones en términos de la función  $f$  y de la derivada  $y'(x)$  vistas en (5.7) (recordando las notaciones  $z(x) = y'(x)$ ), tendremos

$$\begin{aligned} y(x+h) &= y(x) + zh + \frac{1}{2}fh^2 + \frac{1}{6}f_y z h^3 \\ &\quad + \frac{1}{24}(f_{yy}z^2 + f_y f)h^4 + O(h^5), \\ z(x+h) &= z(x) + fh + \frac{1}{2}f_y z h^2 + \frac{1}{6}(f_{yy}z^2 + f_y f)h^3 \end{aligned}$$

$$+ \frac{1}{24} (f_{yyy} z^3 + 3 f_{yy} f z + f_y^2 z) h^4 + O(h^5), \quad (5.9)$$

donde hemos obviado el punto de evaluación que para los  $z$  es  $x$  (luego con  $z$  denotamos  $z(x) = y'(x)$ ) y para las funciones  $f$  y sus derivadas parciales viene dado por  $y(x)$ .

Para los métodos de dos evaluaciones por paso que pasamos a describir, las dos etapas vendrán dadas por

$$\begin{aligned} k_1 &= hf(y + hc_1z), \\ k_2 &= hf(y + h(c_2z + d_2k_1)), \end{aligned} \quad (5.10)$$

esto es, coinciden con las que tendríamos al considerar un método Nyström de dos etapas. Esto no debe resultar sorprendente, puesto que para los métodos de dos etapas para EDOs autónomas que hemos considerado en capítulos precedentes las etapas coincidían con las correspondientes a las fórmulas explícitas de Runge-Kutta del mismo número de etapas. Si hacemos los desarrollos de Taylor de dichas etapas (como funciones del paso  $h$  en un entorno de  $h = 0$ ) obtenemos

$$\begin{aligned} k_1(h) &= fh + c_1f_yz h^2 + \frac{1}{2} c_1^2 f_{yy} z^2 h^3 + \frac{1}{6} c_1^3 f_{yyy} z^3 h^4 + O(h^5), \\ k_2(h) &= fh + c_2f_yz h^2 + \frac{1}{2} (c_2^2 f_{yy} z^2 + 2 d_2 f_y f) h^3 \\ &\quad + \frac{1}{6} (c_2^3 f_{yyy} z^3 + 6 c_2 d_2 f_{yy} f z + 6 c_1 d_2 f_y^2 z) h^4 + O(h^5), \end{aligned} \quad (5.11)$$

donde, como antes, estamos obviando los puntos de evaluación. Definimos ahora un nuevo término  $s$  por

$$s = \frac{k_2 - k_1}{(c_2 - c_1)z + d_2 k_1}. \quad (5.12)$$

Es evidente que puede ser considerado la modificación del término  $s$  definido al tratar los métodos de dos etapas para EDOs autónomas de primer orden, de modo que sea aplicable a la nueva situación. Las razones que nos han movido a tomar de este modo el denominador en la definición del término  $s$  quedan más claras si observamos lo siguiente: haciendo el desarrollo de Taylor de  $f(y + a)$  en un entorno del punto  $a = 0$  hasta orden tres se tiene

$$f(y + a) = f + f_y a + \frac{1}{2} f_{yy} a^2 + \frac{1}{6} f_{yyy} a^3 + O(a^4). \quad (5.13)$$

Si ahora tomamos  $a = hc_1z$  y sustituimos en (5.13) se obtiene, tras multiplicar el resultado por  $h$ , el desarrollo de Taylor en potencias de  $h$  de  $k_1$  que vimos en (5.11). Análogamente, si tomamos  $a = h(c_2z + d_2k_1)$  y sustituimos en (5.13) obtenemos, tras multiplicar por  $h$ , el desarrollo en potencias de  $h$  para  $k_2$

$$\begin{aligned} k_2(h) &= fh + f_y(c_2z + d_2k_1)h^2 + \frac{1}{2}f_{yy}(c_2z + d_2k_1)^2h^3 \\ &+ \frac{1}{6}f_{yyy}(c_2z + d_2k_1)^3h^4 + O(h^5), \end{aligned} \quad (5.14)$$

que coincide, tras las oportunas simplificaciones (teniendo en cuenta cómo viene dado  $k_1$ ), con el resultado obtenido en (5.11). Para nuestros propósitos (ilustrar por qué tomamos de este modo el denominador en la definición de  $s$ ) es conveniente dejarlo así. Nótese que las composiciones en los desarrollos que estamos haciendo están justificadas puesto que en ambos casos se tiene que si  $h = 0$  entonces  $a = 0$ . Ahora se deduce fácilmente de lo anterior que

$$\begin{aligned} k_2 - k_1 &= f_y((c_2 - c_1)z + d_2k_1)h^2 + \frac{1}{2}f_{yy}((c_2z + d_2k_1)^2 - (c_1z)^2)h^3 \\ &+ \frac{1}{6}f_{yyy}((c_2z + d_2k_1)^3 - (c_1z)^3)h^4 + O(h^5) \\ &= ((c_2 - c_1)z + d_2k_1) \left( f_y h^2 + \frac{1}{2}f_{yy}((c_1 + c_2)z + d_2k_1)h^3 \right. \\ &+ \left. \frac{1}{6}f_{yyy}((c_1z)^2 + c_1z(c_2z + d_2k_1) + (c_2z + d_2k_1)^2)h^4 \right) \\ &+ O(h^5), \end{aligned} \quad (5.15)$$

lo que justifica el tomar el denominador en  $s$  del modo que lo hemos hecho. Además si dividimos el desarrollo anterior por  $(c_2 - c_1)z + d_2k_1$  se obtiene el desarrollo de  $s$ . Tras las oportunas sustituciones (de  $k_1$  por su desarrollo visto anteriormente) y simplificaciones, se comprueba que el desarrollo de Taylor de  $s$  como función de  $h$  en un entorno del punto  $h = 0$  viene dado en la forma

$$\begin{aligned} s(h) &= f_y h^2 + \frac{1}{2}(c_1 + c_2)f_{yy}z h^3 \\ &+ \frac{1}{6}((c_1^2 + c_1c_2 + c_2^2)f_{yyy}z^2 + 3d_2f_{yy}f)h^4 + O(h^5) \end{aligned} \quad (5.16)$$

Ahora es fácil comprobar que, al igual que ocurría con las fórmulas para ecuaciones de primer orden, es posible obtener muchos parámetros adicionales que permiten aspirar a obtener órdenes de consistencia más altos que los que se obtienen con los métodos Runge-Kutta-Nyström del mismo número de etapas. Para ello, observemos que además de las propias etapas  $k_1$  y  $k_2$ , podemos incorporar a las fórmulas que definen los métodos los productos de  $h$  por expresiones de la forma

$$z \sum_{i=1}^r \alpha_i s^i, \quad \text{y} \quad k_1 \sum_{i=1}^{r^*} \beta_i s^i, \quad (5.17)$$

para las aproximaciones a la solución y a la derivada, obteniendo de este modo muchos parámetros libres adicionales.

Es sencillo comprobar que los desarrollos en serie de Taylor de dichos factores (convenientemente escalados con  $h$  cuando sea necesario) son idénticos a los de la solución exacta (y su derivada) en lo que respecta a las diferenciales elementales que intervienen en los mismos (aunque los coeficientes pueden variar). Más aún, se pueden incorporar a las fórmulas que definen los métodos, productos de la forma  $zG(s)$  y  $k_1H(s)$  (para ciertas funciones  $G$  y  $H$ ).

Estamos ya en condiciones de dar la forma general de los métodos de dos etapas que consideraremos.

**5.1.4 Expresión general de los métodos de dos etapas.** En lo que sigue, consideraremos métodos de dos etapas para el problema

$$y''(x) = f(y(x)), \quad y(x_0) = y_0, \quad y'(x_0) = z_0, \quad f : \mathbb{R} \rightarrow \mathbb{R}, \quad (5.18)$$

o equivalentemente

$$\begin{aligned} y'(x) &= z(x), & y(x_0) &= y_0, \\ z'(x) &= f(y(x)), & z(x_0) &= z_0, \end{aligned} \quad (5.19)$$

en forma vectorial. Estos métodos de dos etapas vienen dados por

$$\begin{aligned} y_{n+1} &= y_n + h (P_1(s) z_n + P_2(s) k_1), \\ z_{n+1} &= Q_1(s) z_n + Q_2(s) k_1. \end{aligned} \quad (5.20)$$

Las funciones  $P_1$ ,  $P_2$ ,  $Q_1$  y  $Q_2$  se pueden tomar en principio de forma arbitraria. Las etapas y el término  $s$  vendrán dados por

$$\begin{aligned} k_1 &= hf(y_n + hc_1z_n), \\ k_2 &= hf(y_n + h(c_2z_n + d_2k_1)), \\ s &= \frac{k_2 - k_1}{(c_2 - c_1)z_n + d_2k_1}. \end{aligned} \quad (5.21)$$

Por ejemplo, el método que hemos utilizado en el apartado 5.1.2 y que venía definido por las fórmulas (5.3–5.4), se obtiene tomando en las expresiones anteriores los valores de los parámetros vistos en (5.5) y las funciones definidas por

$$\begin{aligned} P_1(s) &= \frac{\text{sen } \sqrt{-s}}{\sqrt{-s}} + \frac{3 - \sqrt{3}}{6} (1 - \cos \sqrt{-s}), \\ P_2(s) &= \frac{\cos \sqrt{-s} - 1}{s}, \\ Q_1(s) &= \cos \sqrt{-s} + \frac{3 - \sqrt{3}}{6} \sqrt{-s} \text{sen } \sqrt{-s}, \\ Q_2(s) &= \frac{\text{sen } \sqrt{-s}}{\sqrt{-s}}. \end{aligned} \quad (5.22)$$

Pasamos ahora al estudio de las condiciones de orden correspondientes a los métodos de dos etapas.

### 5.1.5 Condiciones de orden para los métodos de dos etapas.

Empezaremos considerando métodos de dos etapas de tipo polinomial, esto es, fórmulas definidas por las expresiones (5.20–5.21) en las cuales las funciones  $P_1$ ,  $P_2$ ,  $Q_1$  y  $Q_2$  que los definen son polinomios. Describiremos dichas funciones polinómicas por

$$\begin{aligned} P_1(s) &= \sum_{i=0}^{p_1} p_{1i} s^i, & P_2(s) &= \sum_{i=0}^{p_2} p_{2i} s^i, \\ Q_1(s) &= \sum_{i=0}^{q_1} q_{1i} s^i, & Q_2(s) &= \sum_{i=0}^{q_2} q_{2i} s^i, \end{aligned} \quad (5.23)$$

análogamente a como hemos hecho en capítulos precedentes.

Pasamos a definir la convergencia de orden  $q$  de los métodos que hemos propuesto, imponiendo las mismas condiciones que han de satisfacerse en el caso de las fórmulas de Runge-Kutta-Nyström (véase por ejemplo [18], pág. 284).

**Definición 5.1** *Diremos que el método (5.20) es de orden  $q$  para problemas del tipo (5.18) suficientemente regulares, si se verifican las condiciones*

$$y(x_0 + h) - y_1 = O(h^{q+1}), \quad y'(x_0 + h) - y'_1 = O(h^{q+1}), \quad (5.24)$$

Es fácil comprobar que, para las condiciones de orden cuatro, basta tomar en (5.23) los valores  $p_1 = p_2 = q_2 = 1$  y  $q_1 = 2$ , pues los restantes parámetros sólo intervienen en las condiciones de órdenes superiores. Esto se deduce de que, como hemos visto anteriormente, se tiene que  $k_1 = O(h)$  y  $s = O(h^2)$ .

Las condiciones de orden cuatro para estos métodos vienen dadas por

$$p_{10} = 1, \quad (5.25)$$

$$p_{20} = \frac{1}{2}, \quad (5.26)$$

$$p_{11} + c_1 p_{20} = \frac{1}{6}, \quad (5.27)$$

$$(c_1 + c_2) p_{11} + c_1^2 p_{20} = \frac{1}{12}, \quad (5.28)$$

$$p_{21} = \frac{1}{24}, \quad (5.29)$$

$$q_{10} = 1, \quad (5.30)$$

$$q_{20} = 1, \quad (5.31)$$

$$q_{11} + c_1 q_{20} = \frac{1}{2}, \quad (5.32)$$

$$(c_1 + c_2) q_{11} + c_1^2 q_{20} = \frac{1}{3}, \quad (5.33)$$

$$q_{21} = \frac{1}{6}, \quad (5.34)$$

$$(c_1^2 + c_1 c_2 + c_2^2) q_{11} + c_1^3 q_{20} = \frac{1}{4}, \quad (5.35)$$

$$(c_1 + c_2) q_{21} + d_2 q_{11} = \frac{1}{4}, \quad (5.36)$$

$$q_{12} + c_1 q_{21} = \frac{1}{24}, \quad (5.37)$$

donde las cinco primeras corresponden a la aproximación a la solución  $y$  y las restantes a la aproximación a la derivada de dicha solución  $z = y'$ . La solución a dicho sistema de ecuaciones es sencilla de obtener y viene dada por

$$\begin{aligned} c_1 &= \frac{3 \mp \sqrt{3}}{6}, & c_2 &= \frac{3 \pm \sqrt{3}}{6}, & d_2 &= \frac{\pm \sqrt{3}}{6}, & (5.38) \\ p_{10} &= 1, & p_{20} &= \frac{1}{2}, & p_{11} &= \frac{-1 \pm \sqrt{3}}{12}, & p_{21} &= \frac{1}{24}, \\ q_{10} &= 1, & q_{20} &= 1, & q_{11} &= \frac{\pm \sqrt{3}}{6}, & q_{21} &= \frac{1}{6}, & q_{12} &= \frac{-3 \pm 2\sqrt{3}}{72}. \end{aligned}$$

Es posible también en este caso satisfacer algunas condiciones de orden cinco (aunque no todas). De hecho, para obtener las condiciones correspondientes a orden cinco, bastará tomar en (5.23) los valores  $p_2 = 1$  y  $p_1 = q_1 = q_2 = 2$ , y añadir a las ecuaciones correspondientes a orden cuatro (dadas por (5.25–5.37)) las asociadas a orden cinco, esto es

$$(c_1^2 + c_1 c_2 + c_2^2) p_{11} + c_1^3 p_{20} = \frac{1}{20}, \quad (5.39)$$

$$(c_1 + c_2) p_{21} + d_2 p_{11} = \frac{1}{20}, \quad (5.40)$$

$$p_{12} + c_1 p_{21} = \frac{1}{120}, \quad (5.41)$$

$$(c_1^3 + c_1^2 c_2 + c_1 c_2^2 + c_2^3) q_{11} + c_1^4 q_{20} = \frac{1}{5}, \quad (5.42)$$

$$(c_1^2 + c_1 c_2 + c_2^2) q_{21} + (c_1 + 2c_2) d_2 q_{11} = \frac{3}{10}, \quad (5.43)$$

$$2(c_1 + c_2) q_{12} + c_1(2c_1 + c_2) q_{21} + c_1 d_2 q_{11} = \frac{1}{12}, \quad (5.44)$$

$$d_2 q_{21} = \frac{1}{20}, \quad (5.45)$$

$$q_{22} = \frac{1}{120}, \quad (5.46)$$

de las cuales las tres primeras corresponden a la aproximación a  $y$  y las cinco restantes a la aproximación a la derivada  $z = y'$ .

Tras sustituir los valores obtenidos en (5.38) para orden cuatro en las ecuaciones (5.39–5.46) correspondientes a orden cinco, se comprueba que seis de ellas no pueden ser satisfechas (por no depender de los parámetros libres), por lo que no es posible obtener fórmulas de orden cinco del tipo considerado con sólo dos etapas. Sin embargo, sí es posible minimizar la parte principal del error local de truncación satisfaciendo las otras dos condiciones de orden cinco, concretamente las ecuaciones (5.41) y (5.46). Para ello basta tomar los valores

$$p_{12} = \frac{-9 \pm 5\sqrt{3}}{720}, \quad q_{22} = \frac{1}{120}, \quad (5.47)$$

y añadirlos a los obtenidos en (5.38) para orden cuatro.

En lo anterior hemos deducido las condiciones de orden para los métodos de dos etapas de tipo polinomial (dados por las expresiones (5.20), (5.21) y (5.23)) y las hemos resuelto obteniendo de este modo todas las fórmulas de orden cuatro. Es fácil obtener, análogamente a como ya hicimos en capítulos anteriores, las condiciones de orden correspondientes a métodos de dos etapas más generales. Para ello, basta considerar los desarrollos en potencias de  $s$  de las funciones  $P_1$ ,  $P_2$ ,  $Q_1$  y  $Q_2$  del método general de dos etapas, definido por (5.20) y (5.21), y comparar dichos desarrollos con los correspondientes a los métodos de dos etapas de tipo polinomial, con las etapas  $k_1$ ,  $k_2$  y el término  $s$  definidos del mismo modo. Ilustraremos lo anterior deduciendo que el método visto en el apartado 5.1.2, que evidentemente no es de tipo polinomial, es de orden cuatro.

Observemos en primer lugar que el método no polinomial al que aludimos era

$$\begin{aligned} y_{n+1} &= y_n + h \left( \left( \frac{\cos \sqrt{-s} - 1}{s} \right) (k_1 - c_1 s y'_n) + \left( \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}} \right) y'_n \right), \\ y'_{n+1} &= y'_n + \left( (\cos \sqrt{-s} - 1) y'_n + \left( \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}} \right) (k_1 - c_1 s y'_n) \right) \end{aligned} \quad (5.48)$$

donde

$$\begin{aligned} k_1 &= hf(y_n + hc_1 y'_n), \\ k_2 &= hf(y_n + h(c_2 y'_n + d_2 k_1)), \\ s &= \frac{k_2 - k_1}{(c_2 - c_1) y'_n + d_2 k_1}, \end{aligned} \quad (5.49)$$



y los valores de los parámetros vienen dados por

$$c_1 = \frac{3 - \sqrt{3}}{6}, \quad c_2 = \frac{3 + \sqrt{3}}{6}, \quad d_2 = \frac{\sqrt{3}}{6}; \quad (5.50)$$

o sea, es

$$\begin{aligned} y_{n+1} &= y_n + h(P_1(s)z_n + P_2(s)k_1), \\ z_{n+1} &= Q_1(s)z_n + Q_2(s)k_1 \end{aligned} \quad (5.51)$$

para las funciones

$$\begin{aligned} P_1(s) &= \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}} + \frac{3 - \sqrt{3}}{6} (1 - \cos \sqrt{-s}), \\ P_2(s) &= \frac{\cos \sqrt{-s} - 1}{s}, \\ Q_1(s) &= \cos \sqrt{-s} + \frac{3 - \sqrt{3}}{6} \sqrt{-s} \operatorname{sen} \sqrt{-s}, \\ Q_2(s) &= \frac{\operatorname{sen} \sqrt{-s}}{\sqrt{-s}}. \end{aligned} \quad (5.52)$$

Las etapas  $k_1$ ,  $k_2$  y el término  $s$  vienen dadas exactamente igual que en los métodos de tipo polinomial de orden cuatro, por venir definidas del mismo modo y coincidir las constantes  $c_1$ ,  $c_2$  y  $d_2$  con los valores de (5.38) tomando el signo superior. Los desarrollos en potencias de  $s$  de las funciones  $P_1$ ,  $P_2$ ,  $Q_1$  y  $Q_2$  (convenientemente redefinidas por continuidad en  $s = 0$ ) que definen dicho método, vienen dados por

$$\begin{aligned} P_1(s) &= 1 + \frac{-1 + \sqrt{3}}{12} s + \frac{-9 + 5\sqrt{3}}{720} s^2 + O(s^3), \\ P_2(s) &= \frac{1}{2} + \frac{1}{24} s + \frac{1}{720} s^2 + O(s^3), \\ Q_1(s) &= 1 + \frac{\sqrt{3}}{6} s + \frac{-3 + 2\sqrt{3}}{72} s^2 + O(s^3), \\ Q_2(s) &= 1 + \frac{1}{6} s + \frac{1}{120} s^2 + O(s^3), \end{aligned} \quad (5.53)$$

y permiten deducir trivialmente que el método es de orden cuatro con minimización de la parte principal del error local de truncación (compárense los coeficientes de (5.53) con los obtenidos en (5.38) y (5.47) para el caso polinomial.

**5.1.6 Construcción de métodos de dos etapas para problemas oscilatorios.** En los últimos años han surgido distintos métodos diseñados para la integración de problemas oscilatorios. Entre otros cabe señalar a modo de ejemplo los discutidos en [7], [30], [31], [46], [47], [48], [39] y [40].

Nos proponemos construir fórmulas de dos etapas del tipo descrito en (5.20–5.21) específicamente diseñadas para integrar eficientemente problemas altamente oscilatorios. Un primer ejemplo de fórmula de este tipo lo hemos utilizado para introducir el capítulo y viene descrito al final del apartado precedente. A continuación veremos como obtener dicha fórmula.

Nuestro punto de partida es la ecuación test

$$y'' = -\alpha y, \quad y(x_0) = y_0, \quad y'(x_0) = z_0, \quad (5.54)$$

con  $\alpha > 0$ . El problema anterior tiene una solución bien conocida que viene dada por

$$\begin{pmatrix} y(x) \\ z(x) \end{pmatrix} = \begin{pmatrix} \cos(\sqrt{\alpha}(x-x_0)) & \frac{\text{sen}(\sqrt{\alpha}(x-x_0))}{\sqrt{\alpha}} \\ -\sqrt{\alpha} \text{sen}(\sqrt{\alpha}(x-x_0)) & \cos(\sqrt{\alpha}(x-x_0)) \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \quad (5.55)$$

donde, como viene siendo habitual, con  $z$  denotamos  $y'$ . En adelante tomaremos, sin pérdida de generalidad,  $x_0 = 0$ . Nótese que para valores  $\alpha$  positivos y muy grandes la solución presenta rápidas oscilaciones que dificultan su integración numérica con pasos de tamaño razonable.

Al aplicar un método de dos etapas de la familia descrita por (5.20–5.21) a la ecuación test (5.54) obtenemos para las etapas

$$\begin{aligned} k_1 &= -h\alpha(y_n + h c_1 z_n), \\ k_2 &= -h\alpha((1 - h^2\alpha d_2)y_n + h(c_2 - h^2\alpha d_2 c_1)z_n), \end{aligned} \quad (5.56)$$

y el término  $s$  toma la forma

$$s = -h^2\alpha. \quad (5.57)$$

Por tanto, aplicado a la ecuación test, cualquier método descrito por (5.20) tomará la forma

$$\begin{pmatrix} y_{n+1} \\ z_{n+1} \end{pmatrix} = \begin{pmatrix} 1 - h^2\alpha P_2(-h^2\alpha) & h(P_1(-h^2\alpha) - h^2\alpha c_1 P_2(-h^2\alpha)) \\ -h\alpha Q_2(-h^2\alpha) & Q_1(-h^2\alpha) - h^2\alpha c_1 Q_2(-h^2\alpha) \end{pmatrix} \begin{pmatrix} y_n \\ z_n \end{pmatrix} \quad (5.58)$$

De la expresión (5.55) (con  $x_0 = 0$ ) que nos proporciona la solución exacta del problema (5.54) es fácil deducir que se verifica

$$\begin{aligned} \begin{pmatrix} y(x+h) \\ z(x+h) \end{pmatrix} &= \begin{pmatrix} \cos(\sqrt{\alpha}(x+h)) & \frac{\operatorname{sen}(\sqrt{\alpha}(x+h))}{\sqrt{\alpha}} \\ -\sqrt{\alpha} \operatorname{sen}(\sqrt{\alpha}(x+h)) & \cos(\sqrt{\alpha}(x+h)) \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\sqrt{\alpha}h) & \frac{\operatorname{sen}(\sqrt{\alpha}h)}{\sqrt{\alpha}} \\ -\sqrt{\alpha} \operatorname{sen}(\sqrt{\alpha}h) & \cos(\sqrt{\alpha}h) \end{pmatrix} \begin{pmatrix} y(x) \\ z(x) \end{pmatrix}. \end{aligned} \quad (5.59)$$

Cuando integramos con paso fijo  $h$  el problema (5.54), con cualquiera de los métodos que estamos considerando, pretendemos que  $y_n$  y  $z_n$  en (5.58) proporcionen aproximaciones a las cantidades  $y(nh)$  y  $z(nh)$  de (5.59) respectivamente.

Podemos ir más lejos e imponer que, salvo errores de redondeo, la solución numérica coincida con la solución exacta del problema considerado. Para ello basta imponer que la matriz considerada en (5.58) coincida con la segunda matriz de (5.59). Tras igualar coeficiente a coeficiente y despejar se obtiene que las funciones  $P_i$  y  $Q_i$  que definen los métodos han de verificar

$$\begin{aligned} P_1(-h^2\alpha) &= \frac{\operatorname{sen}(\sqrt{\alpha}h)}{\sqrt{\alpha}h} + c_1(1 - \cos(\sqrt{\alpha}h)), \\ P_2(-h^2\alpha) &= \frac{1 - \cos(\sqrt{\alpha}h)}{h^2\alpha}, \\ Q_1(-h^2\alpha) &= \cos(\sqrt{\alpha}h) + c_1\sqrt{\alpha}h \operatorname{sen}(\sqrt{\alpha}h), \\ Q_2(-h^2\alpha) &= \frac{\operatorname{sen}(\sqrt{\alpha}h)}{\sqrt{\alpha}h}. \end{aligned} \quad (5.60)$$

Finalmente, de la relación  $s = -h^2\alpha$  se deduce sin dificultad que el primer método visto en este capítulo integra exactamente problemas oscilatorios del tipo  $y'' = -\alpha y + \beta$  con  $\alpha > 0$  y  $\beta \in \mathbb{R}$ . Además, como ya vimos, la elección de los parámetros  $c_1$ ,  $c_2$  y  $d_2$  que hemos hecho en (5.5) nos permite asegurar que el método es de orden cuatro para problemas escalares autónomos del tipo (5.1) y minimiza la parte principal del error local de truncación.

Conviene observar en este punto que el estudio que acabamos de realizar es análogo al que en su momento realizamos al estudiar las

propiedades de estabilidad lineal de los métodos para problemas de primer orden. De hecho la segunda matriz en (5.59) juega un papel análogo al de la función de estabilidad lineal  $R$  que obteníamos a partir de la ecuación test  $y' = \lambda y$ . Más aún, vimos en la sección (2.3.4) que es posible obtener métodos de dos etapas a partir de funciones de estabilidad prefijadas, y siguiendo un procedimiento similar es también posible construir métodos de dos etapas para el problema (5.1) a partir de matrices prefijadas. De hecho el estudio previo es un ejemplo de como hacer ésto, pero no entraremos en más detalles a este respecto.

Aprovechamos para hacer un comentario relativo a la aplicabilidad del método obtenido a problemas para los cuales el término  $s$  resulte ser positivo o nulo, caso en el aparentemente la fórmula resultante no está definida (en el primer caso  $\sqrt{-s}$  no está definido en  $\mathbb{R}$  y en el segundo aparecen indeterminaciones). Esta situación, que no se da en las fórmulas de tipo polinomial por razones obvias, se soluciona sin más que observar que cuando  $s$  es positivo se puede poner  $\sqrt{-s} = i\sqrt{s}$  y utilizar las conocidas relaciones

$$\operatorname{ch}(z) = \cos(iz), \quad \operatorname{sh}(z) = -i \operatorname{sen}(iz), \quad z \in \mathbb{C}, \quad (5.61)$$

tras lo cual bastará reescribir como sigue las funciones  $P_i$  y  $Q_i$  que caracterizan nuestro método

$$\begin{aligned} P_1(s) &= \frac{\operatorname{sh} \sqrt{s}}{\sqrt{s}} + \frac{3 - \sqrt{3}}{6} (1 - \operatorname{ch} \sqrt{s}), \\ P_2(s) &= \frac{\operatorname{ch} \sqrt{s} - 1}{s}, \\ Q_1(s) &= \operatorname{ch} \sqrt{s} - \frac{3 - \sqrt{3}}{6} \sqrt{s} \operatorname{sh} \sqrt{s}, \\ Q_2(s) &= \frac{\operatorname{sh} \sqrt{s}}{\sqrt{s}}, \end{aligned} \quad (5.62)$$

cuando  $s$  resulte ser positivo. Cuando  $s = 0$  es fácil deducir (tras resolver las indeterminaciones que surgen) que bastará tomar

$$P_1(s) = 1, \quad P_2(s) = \frac{1}{2}, \quad Q_1(s) = 1, \quad Q_2(s) = 1. \quad (5.63)$$

De las observaciones anteriores podemos concluir que el primer método propuesto en este capítulo integrará exactamente problemas del tipo  $y'' = -\alpha y + \beta$  con  $\alpha, \beta \in \mathbb{R}$  (nótese que hemos suprimido la condición  $\alpha > 0$ ).

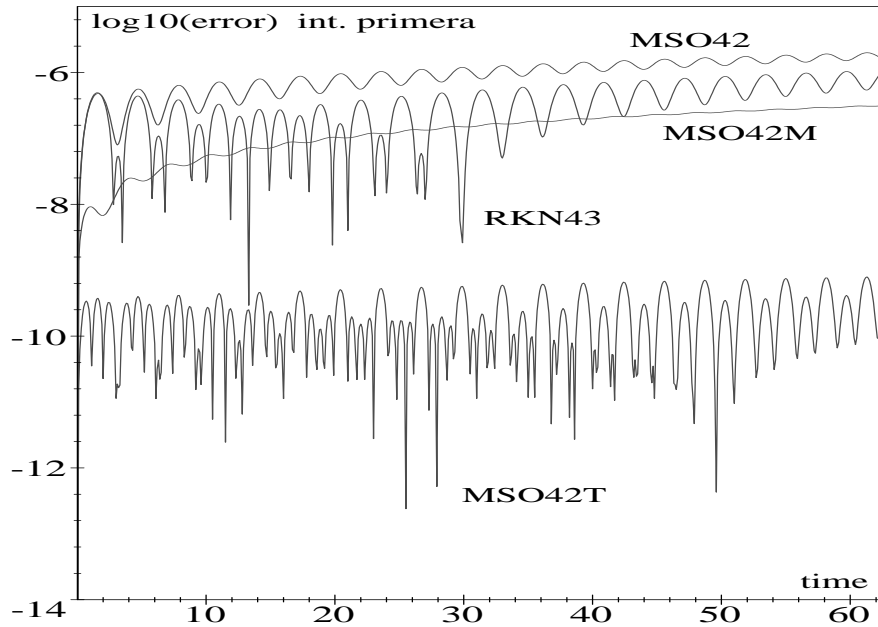
**5.1.7 Experimentos numéricos con los métodos de dos etapas.** Pasamos a ilustrar el comportamiento numérico de los métodos de dos etapas y orden cuatro que acabamos de obtener. Hemos visto un primer ejemplo en el apartado 5.1.2 para motivar el interés de las fórmulas que hemos introducido y estudiado en los apartados siguientes. Completamos a continuación los experimentos numéricos iniciados en dicho apartado.

Empezaremos integrando numéricamente problemas del tipo descrito en (5.6), tomando distintos valores de los parámetros  $\alpha$  y  $\varepsilon$ , y utilizando distintos tamaños de paso (siempre con paso fijo). Los métodos que utilizaremos para ello son:

1. El método Runge-Kutta-Nyström de orden cuatro y tres evaluaciones por paso descrito en [18], pág. 285 y que denotaremos por RKN43.
2. El método de orden cuatro y dos evaluaciones por paso de tipo polinomial descrito por las expresiones (5.20), (5.21) y (5.23), que se obtiene tomando el signo superior en (5.38) y el resto de parámetros nulos. A dicho método lo denotaremos por MSO42.
3. El método de orden cuatro y dos evaluaciones por paso de tipo polinomial descrito por las expresiones (5.20), (5.21) y (5.23), que se obtiene tomando el signo superior en (5.38) y (5.47). A dicho método (cuya parte principal del error local de truncación queda minimizada) lo denotaremos por MSO42M.
4. El método de orden cuatro y dos evaluaciones por paso, que es exacto para problemas del tipo  $y'' = -\alpha y + \beta$  con  $\alpha > 0, \beta \in \mathbb{R}$  y que describimos en los apartados 5.1.2 y 5.1.5. Lo denotaremos por MSO42T.

Como el problema (5.6) no admite una solución en términos de funciones elementales, hemos optado por medir el error cometido en la integral primera  $(\alpha y^2 + y'^2)/2 - \varepsilon y^4/4$  en los experimentos numéricos que realizamos. También aportaremos en algunos casos los errores cometidos para la solución y su derivada, medidos respecto de una solución de referencia suficientemente precisa.

La figura 5.1 muestra los resultados numéricos obtenidos al aplicar



**Figura 5.1:**  $y'' = -y + 10^{-3}y^3$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $h = 0.1$ ,  $t \in [0, 20\pi]$ .

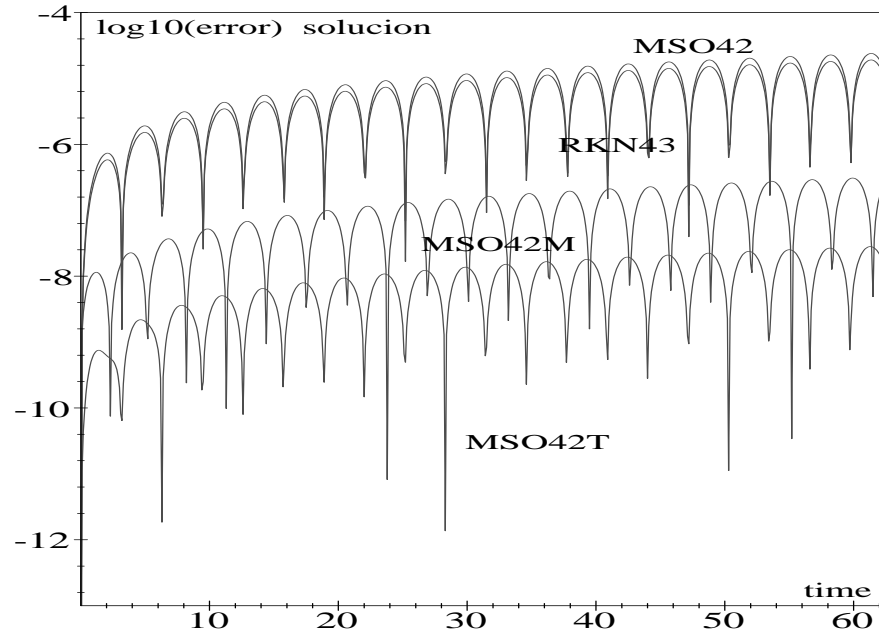
los métodos anteriores al problema

$$y'' = -\alpha y + \varepsilon y^3, \quad y(0) = 1, \quad y'(0) = 0, \quad (5.64)$$

tomando los valores  $\alpha = 1$  y  $\varepsilon = 10^{-3}$  e integrando con paso fijo  $h = 0.1$  en el intervalo  $[0, 20\pi]$  (esto es 10 revoluciones). Es evidente, a la vista de la gráfica para el error (medido respecto de la integral primera), que el método que muestra un mejor comportamiento es aquél que denotamos con MSO42T. Los métodos MSO42 y MSO42M muestran un comportamiento parecido al del método de referencia RKN43, si bien no olvidemos que ambos utilizan una evaluación de función menos por paso.

En las figuras 5.2 y 5.3 se muestran los errores medidos respecto de una solución de referencia cuidadosamente calculada, para la solución y la derivada de la solución respectivamente (tomando los mismos valores que en la figura 5.1).

Repetimos a continuación el mismo experimento numérico tomando ahora  $\alpha = 100$  (el resto de valores vienen dados como en el expe-



**Figura 5.2:**  $y'' = -y + 10^{-3}y^3$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $h = 0.1$ ,  $t \in [0, 20\pi]$ .

rimento que generó la gráfica 5.1) de modo que el problema muestre un mayor carácter oscilatorio. Los resultados obtenidos para el error medido respecto de la integral primera se muestran en la figura 5.4. Se observa en dicha gráfica que el único método que proporciona una aproximación satisfactoria es el de tipo trigonométrico que denotábamos por MSO42T. Esto es así por el tamaño de paso considerado y por ser éste el único de entre los considerados diseñado específicamente para problemas de carácter oscilatorio.

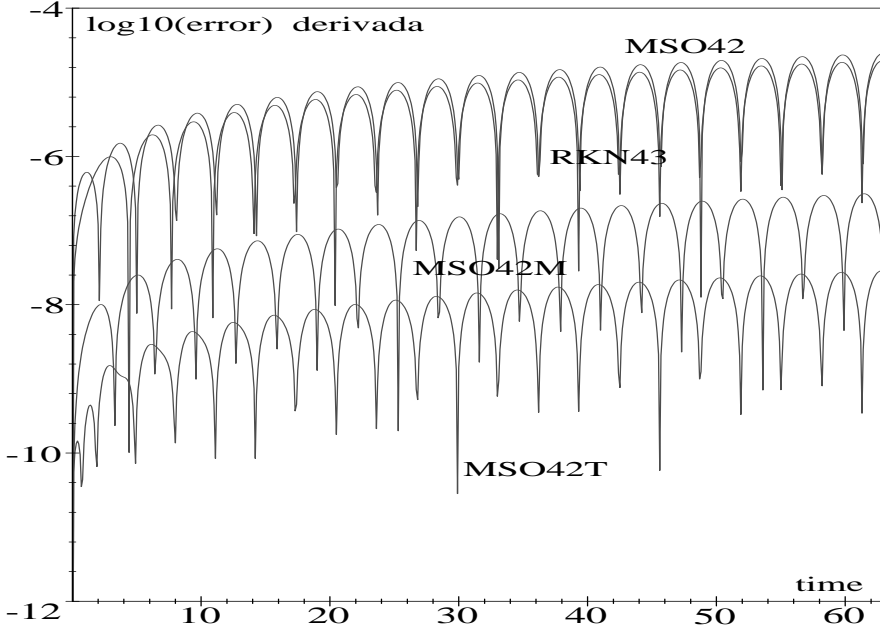


Figura 5.3:  $y'' = -y + 10^{-3}y^3$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $h = 0.1$ ,  $t \in [0, 20\pi]$ .



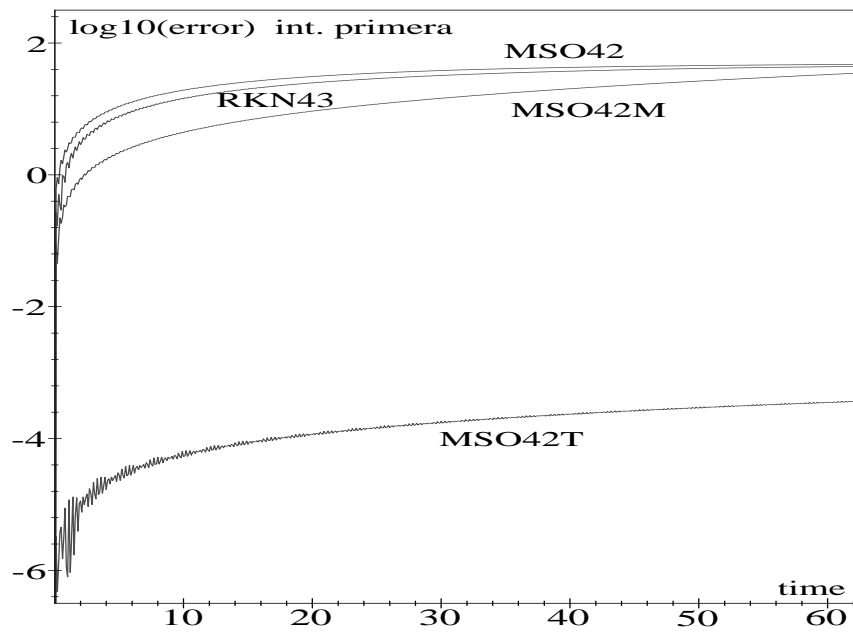


Figura 5.4:  $y'' = -100y + 10^{-3}y^3$ ,  $y(0) = 1$ ,  $y'(0) = 0$ ,  $h = 0.1$ ,  $t \in [0, 20\pi]$ .



## Bibliografía

- [1] R. Alexander, Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s, *SIAM J. Numer. Anal.* **14** (1977) 1006–1021.
- [2] J. Álvarez, Obtaining New Explicit Two-Stage Methods for the Scalar Autonomous IVP with Prefixed Stability Functions, *Intl. Journal of Applied Sc. & Computations* **6** (1999) 39–44.
- [3] J. Álvarez y J. Rojo, New A-stable explicit two-stage methods of order three for the scalar autonomous IVP, in: P. de Oliveira, F. Oliveira, F. Patrício, J.A. Ferreira, A. Araújo, eds., *Proc. of the 2nd. Meeting on Numerical Methods for Differential Equations, NMDE'98* (Coimbra, Portugal, 1998) 57–66.
- [4] J. Álvarez y J. Rojo, A New Family of Explicit Two-Stage Methods of order Three for the Scalar Autonomous IVP, *Intl. Journal of Applied Sc. & Computations* **5** (1999) 246–251.
- [5] J. Álvarez y J. Rojo, Special Methods for the Numerical Integration of some ODEs Systems, *Nonlinear Analysis: Theory, Methods and Applications* **47** (2001) 5703–5708.
- [6] J. Álvarez y J. Rojo, An improved class of generalized Runge-Kutta methods for stiff problems, Part I: scalar case, *Applied Mathematics and Computations* (aceptado, en prensa).
- [7] D.G. Bettis, Stabilization of Finite Difference Methods of Numerical Integration, *Celest. Mech.* **2** (1970) 282–295.
- [8] K. Burrage, The dichotomy of stiffness: Pragmatism versus theory, *Appl. Math. Comput.* **31** Spec. Issue (1989) 92–111.

- [9] J.C. Butcher, Implicit Runge-Kutta processes, *Math. Comp.* **18** (1964) 50–64.
- [10] J.C. Butcher, On the convergence of numerical solutions to ordinary differential equations, *Math. Comp.* **20** (1966) 1–10.
- [11] J.C. Butcher, An algebraic theory of integration methods, *Math. Comput.* **26** (1972) 79–106.
- [12] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods* (Wiley, Chichester, 1987).
- [13] R. Caira, C. Costabile and F. Costabile, A Class of pseudo Runge-Kutta methods, *BIT* **30** (1990) 642–649.
- [14] M. Calvo and M. Mar Quemada, On the stability of rational Runge-Kutta methods, *J. Comp. Appl. Math.* **8** (1982) 289–292.
- [15] G. Dahlquist, A special stability problem for linear multistep methods, *BIT* **3** (1963) 27–43.
- [16] J.D. Day and D.N.P. Murthy, Two Classes of Internally  $S$ -Stable Generalized Runge-Kutta Processes Which Remain Consistent With an Inaccurate Jacobian, *Math. Comp.* **39** (1982) 491–509.
- [17] E. Hairer, Unconditionally Stable Explicit Methods for Parabolic Equations, *Numer. Math.* **35** (1980) 57–68.
- [18] E. Hairer, S.P. Nørset and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems* (Springer-Verlag, Berlin, 1993).
- [19] E. Hairer and G. Wanner, On the Butcher group and general multi-value methods, *Computing* **13** (1974) 1–15.
- [20] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems* (Springer-Verlag, Berlin, 1996).

- [21] P. Kaps, Rosenbrock-type methods, in: G. Dahlquist and R. Jeltsch, eds., *Numerical methods for solving stiff initial value problems, Proceedings, Oberwolfach 28/6-4/7 1981*, Bericht Nr. 9, Inst. für Geometrie und Praktische Mathematik der RWTH Aachen (Aachen, Germany, 1981) 5 pp.
- [22] P. Kaps and A. Ostermann, Rosenbrock Methods using few LU-Decompositions, *IMA J. Numer. Anal.* **9** (1989) 15–27.
- [23] P. Kaps, S. W. H. Poon and T. D. Bui, Rosenbrock Methods for Stiff ODEs: A Comparison of Richardson Extrapolation and Embedding Technique, *Computing* **34** (1985) 17–40.
- [24] P. Kaps and P. Rentrop, Generalized Runge-Kutta Methods of Order Four with Step-size Control for Stiff Ordinary Differential Equations, *Numer. Math.* **33** (1979) 55–68.
- [25] P. Kaps and G. Wanner, A Study of Rosenbrock-Type Methods of High Order, *Numer. Math.* **38** (1981) 279–298.
- [26] J.D. Lambert, *Computational Methods in Ordinary Differential Equations*. (Wiley, Chichester, 1990).
- [27] J.D. Lambert, *Numerical Methods for Ordinary Differential Systems. The Initial Value Problem* (Wiley, Chichester, 1991).
- [28] J.D. Lambert and B. Shaw, On the numerical solution of  $y' = f(x, y)$  by a class of formulae based on rational approximation, *Math. Comp.* **19** (1965) 456–462.
- [29] M.Z. Liu, A Note on the Stability of Rational Runge-Kutta Methods, *Computing* **42** (1989) 329–339.
- [30] P. Martín and J.M. Ferrándiz, Behaviour of the SMF method for the numerical integration of satellite orbits, *Celest. Mech. Dyn. Astron.* **63** (1995) 29–40.
- [31] P. Martín and J.M. Ferrándiz, Multistep Numerical Methods Based on the Scheifele  $G$ -Functions with Application to Satellite Dynamics, *SIAM J. Numer. Anal.* **34** (1997) 359–375.

- [32] S.P. Nørsett and A. Wolfbrandt, Order Conditions for Rosenbrock Type Methods, *Numer. Math.* **32** (1979) 1–15.
- [33] E.J. Nyström, Ueber die numerische Integration von Differentialgleichungen, *Acta Soc. Sci. Fenn.* **50** (1925) 1–54.,
- [34] A. Prothero and A. Robinson, On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Math. Comput.* **28** (1974) 145–162.
- [35] H.H. Rosenbrock, Some general implicit processes for the numerical solution of differential equations, *Comput. J.* **5** (1963) 329–330.
- [36] J.M. Sanz-Serna and M.P. Calvo, *Numerical Hamiltonian Problems* (Chapman-Hall, London, 1993).
- [37] S. Scholz, Order Barriers for the B-Convergence of ROW Methods, *Computing* **41** (1989) 219–235.
- [38] B. Shaw, The use of non-polynomial interpolants in the numerical solution of ordinary differential equations, *Ph. D. thesis, University of St. Andrews* (1965).
- [39] T.E. Simos, Some modified Runge-Kutta methods for the numerical solution of initial-value problems with oscillating solutions, *J. Sci. Comput.* **13** (1998) 51–63.
- [40] T.E. Simos, E. Dimas and A.B. Sideridis, A Runge-Kutta-Nystroem method for the numerical integration of special second-order periodic initial-value problems, *J. Comput. Appl. Math.* **51** (1994) 317–326.
- [41] T. Steihaug and A. Wolfbrandt, An Attempt to Avoid Exact Jacobian and Nonlinear Equations in the Numerical Solution of Stiff Differential Equations, *Math. Comp.* **33** (1979) 521–534.
- [42] C. Störmer, Méthodes d'intégration numérique des équations différentielles ordinaires, *C. R. Congr. Intern. Math., Strasbourg* (1921) 243–257.

- [43] A.M. Urbani, Metodi Row modificati per problemi stiff, *Calcolo* **27** (1990) 89–102.
- [44] J. G. Verwer,  $S$ -Stability Properties for Generalized Runge-Kutta Methods, *Numer. Math.* **27** (1977) 359–370.
- [45] J. G. Verwer and S. Scholz, Rosenbrock methods and time-lagged Jacobian matrices, *Beitr. Numer. Math.* **11** (1983) 173–183.
- [46] J. Vigo-Aguiar and J.M. Ferrándiz, Higer-order variable-step algorithms adapted to the accurate numerical integration of perturbed oscillators, *Computers in Physics* **12** (1998) 467-470.
- [47] J. Vigo-Aguiar and J.M. Ferrándiz, A general procedure for the adaptation of multistep algorithms to the integration of oscillatory problems, *SIAM J. Numer. Anal.* **35** (1998) 1684-1708.
- [48] J. Vigo-Aguiar and J.M. Ferrándiz, VSVO multistep formulae adapted to perturbed second-order differential equations, *Appl. Math. Lett.* **11/3** (1998) 83-87.
- [49] A. Wambecq, Nonlinear methods in solving ordinary differential equations, *J. Comp. Appl. Math.* **2** (1976) 27-33.
- [50] A. Wambecq, Rational Runge-Kutta Methods for Solving Systems of Ordinary Differential Equations, *Computing* **20** (1978) 333-342.
- [51] A. Wambecq, Some properties of rational methods for solving ordinary differential equations, *Pade approximation and its applications, Proc. Conf. Antwerp 1979, Lect. Notes. Math.* **765** (1979) 352-365.
- [52] A. Wambecq, Solution of the equations associated with rational Runge-Kutta methods of orders up to four, *J. Comp. Appl. Math.* **6** (1980) 275-281.
- [53] H. Zedan, Avoiding the exactness of the Jacobian matrix in Rosenbrock formulae, *Comput. Math. Appl.* **19** (1990) 83–89.