



UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

Título
Métodos iterativos aplicados a la ecuación de Kepler
Autor/es
Manuel Aurelio Diloné Alvarado
Director/es
José Manuel Gutiérrez Jiménez
Facultad
Titulación
Departamento
Matemáticas y Computación
Curso Académico
2012-2013



Métodos iterativos aplicados a la ecuación de Kepler, tesis doctoral de Manuel Aurelio Diloné Alvarado, dirigida por José Manuel Gutiérrez Jiménez (publicada por la Universidad de La Rioja), se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported. Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.

- © El autor
- © Universidad de La Rioja, Servicio de Publicaciones, 2013
publicaciones.unirioja.es
E-mail: publicaciones@unirioja.es



FACULTAD DE CIENCIAS, ESTUDIOS AGROALIMENTARIOS E INFORMÁTICA
DEPARTAMENTO DE MATEMÁTICAS Y COMPUTACIÓN

TESIS DOCTORAL

Métodos iterativos aplicados a la ecuación de Kepler

18 de julio de 2013, Logroño, España

Dr. Manuel Aurelio Diloné Alvarado

Dirigida por:

Dr. José Manuel Gutiérrez Jiménez

TESIS DOCTORAL:
Métodos iterativos aplicados a la ecuación de Kepler

Tesis presentada por el Dr. Manuel Aurelio Diloné Alvarado
para la obtención del grado de Doctor en Matemáticas

Máxima puntuación: Sobresaliente

18 de julio de 2013

18:15-21:17

Composición del tribunal:

Presidente: Dr. Sergio Amat Plata

Secretario: Dr. Victor Lanchares Barrasa

Vocales: Dra. María Jesús Rubio, Dra. Alicia Cordero y Dr. Jesús Palacián

Suplentes: Dra. Sonia Busquier Sáez y Dr. Juan Ramón Torregrosa Sánchez

Agradecimientos

Han sido muchas las personas que de alguna manera hicieron posible la terminación de esta memoria. A todas les estaré profundamente agradecido y con ellas tendré un compromiso eterno de trabajo y cooperación. Mención especial merece la Dra. Ligia Amada Melo de Cardona, Ministra de Educación Superior, Ciencia y Tecnología, pues gracias a su persona fue posible mi ingreso y permanencia en la Universidad de La Rioja.

Gracias a mi director de Tesis, Dr. José Manuel Gutiérrez, su dedicación, paciencia y confianza han hecho posible culminar este pequeño aporte al mundo de las matemáticas. Dr. Gutiérrez, el puente entre las orillas del río Ebro y las cálidas aguas caribeñas por el que las Matemáticas puedan caminar a sus anchas, está tendido. Es un inmenso honor ser el primero en cruzarlo.

Gracias a mis hermanos: Nicanor, Fernando, Berna, Josecito e Ilander, por sus muestras de preocupación constante en todo el periodo de investigación.

Gracias a mis padres Don José Diloné y Doña Gladys Alvarado, ambos idos a destiempo, a quienes les hubiera encantado ver culminar esta memoria. Donde quiera que se encuentren, les digo: ¡¡¡misión cumplida !!!

Finalmente dedico mi Tesis a las personas más importantes que me rodean: a mi esposa Jocelyn D. Espinal y a mis hijos, Anneuris Manuel y Manuel Aurelio, por estar a mi lado a pesar de las ausencias, las distancias y el tiempo robado necesario para alcanzar esta meta.

Prólogo

Es bien conocido que la resolución de ecuaciones no lineales constituye un aspecto fundamental del estudio de cualquier área de las matemáticas. Esto así, porque el aporte de esta ciencia al desarrollo de la humanidad, es directamente proporcional a la capacidad que tienen los matemáticos de modelar la realidad, y este modelo, normalmente viene expresado a través de una ecuación o sistemas de ecuaciones. Esta relevancia cobra mayor magnitud si la ecuación en cuestión está identificada como una de las más importantes en el desarrollo de toda una rama del saber: la astronomía, pues con ella se puede determinar la posición de un planeta en un tiempo determinado, amén de tener el honor de ser la primera ecuación en ser tratada por un método algorítmico de solución. La ecuación a la que nos referimos, es la famosa ecuación de Kepler:

$$M = E - e \sin E.$$

El problema consiste en encontrar un valor de E una vez conocidos los valores de M y e . Tanto los parámetros M y e como la incógnita E tienen una clara interpretación astronómica, como veremos en el primer capítulo. En concreto, M es la anomalía media de una órbita elíptica, e es la excentricidad de dicha órbita y E es la anomalía excéntrica. Notemos que la ecuación de Kepler es de tipo trascendental.

Pero si importante es la solución de una ecuación, no menos importante lo es el método con el que se pretende dar solución a la misma. En ese sentido, y uniendo dos áreas apasionantes de las matemáticas como lo es el Análisis Numérico, constituido por la ecuación de Kepler, y las Matemáticas Aplicadas, representado por los métodos iterativos de solución de ecuaciones, se inició la tarea de investigar el comportamiento de la ecuación de Kepler bajo la acción de los métodos punto a punto (tales como el método de Newton, Halley, Chebyshev, super Halley y Danby) y de los métodos multipunto (como el método de la Secante, Bisección y Yun-Petkovic). El motivo principal de valorar a la ecuación de Kepler es que, tal y como lo establece Danby [51], la misma es una de las ecuaciones más estudiadas en la bibliografía matemática. Por otro lado, la elección de los métodos radicó en que los mismos son una muestra representativa del conjunto de métodos más aplicados para la resolución de ecuaciones no lineales.

Así, el Capítulo 1, La ecuación de Kepler, está conformado por una una breve descripción histórica del desarrollo y aparición de la misma. Presentamos además, las demostraciones de

las leyes fundamentales de la astronomía constituidas en las tres leyes de Kepler, la forma de obtener dicha ecuación y las soluciones geométricas y analíticas que se les ha dado a través del tiempo.

En el Capítulo 2, Métodos iterativos, presentamos los métodos iterativos del estudio: Newton, Halley, Chebyshev, super-Halley, Secante, Yun-Petković y Danby. Cada uno de los métodos iterativos está estructurado en tres secciones: notas históricas, descripción del método y convergencia. Dicho capítulo se inicia con el método de Newton. Debemos aclarar que la posición en el enfoque de tratar en primer lugar al método de Newton no es casual. Se trató en primer lugar al método de Newton, por ser, si no el más, al menos, uno de los métodos más conocidos y aplicados en la solución de una ecuación no lineal. Un segundo motivo es que sus teoremas de convergencia semilocal son básicos para entender la convergencia semilocal de los demás métodos. Los principales aportes de la tesis en este capítulo lo constituyen la demostración de un nuevo teorema de convergencia semilocal para el método de Chebyshev, publicado en *Bulletin of the Australian Mathematical Society* [61]. Además, los teoremas de convergencia local y global del método de Danby, en vía de publicación, la demostración en \mathbb{R} de los teoremas de convergencia semilocal del método de Newton (Teorema de Kantorovich y Smale), Halley (Teorema de Zhen-Robbie), super-Halley (Teorema de Gutiérrez), Secante (Teoremas de Yakoubsohn bajo condiciones de Kantorovich y Smale), así como el código para implementar el método de Yun-Petković en Mathematica.

Con las teorías obtenidas a partir del Capítulo 1 y el Capítulo 2, introducimos el Capítulo 3: Métodos iterativos y la ecuación de Kepler. Iniciamos con un estudio relacionado con los métodos punto a punto aplicados a la ecuación de Kepler. El objetivo de dicho estudio fue comparar la efectividad de las funciones de iteración de los métodos de Newton, Chebyshev, Halley y super-Halley, todos ellos obtenidos a partir de la familia uniparamétrica

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \alpha L_f(x_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}, \quad L_f(x_n) = \frac{f(x_n)f''(x_n)}{f'(x_n)^2},$$

y el método de Danby, cuando se aplican a la ecuación de Kepler. Para tal comparación, seleccionamos algunos valores de los parámetros e y M que han sido utilizados en otras investigaciones: [49], [51], [62], [147] y [198]. Además, complementamos el estudio con una muestra formada por valores reales asociados a los asteroides Amor, Apollo, Chiron, Damocles, Hidalgo, Hungaria, Pholus y al Planeta Mercurio [208].

En la sección 3.3, Métodos iterativos multipunto aplicados a la ecuación de Kepler, presentamos un estudio relacionado con la aplicación del método de la Secante, el método de Bisección y el método de Yun-Petković a la ecuación de Kepler. El objetivo de la sección consistió en comparar el comportamiento y efectividad de dichos métodos al ser aplicados a la ecuación de Kepler. Para tal acción utilizamos las 14 condiciones establecidas para los valores de los parámetros e y M , dados en el Cuadro 3.2 y en el Cuadro 3.3. La mayor contribución de este estudio fue comprobar el buen comportamiento que tienen los métodos

de la Secante, Yun-Petković y Bisección, comparados con los métodos de punto a punto. En efecto, el estudio demostró que el porcentaje de fracasos de los métodos de la Secante, Yun-Petković y Bisección, al ser aplicados a la ecuación de Kepler bajo las condiciones dadas en los cuadros 3.2 y 3.3, fue nula, es decir, para todos los casos establecidos, los métodos consiguieron converger.

Para el desarrollo de esta investigación, tomamos en cuenta que un aspecto clave a la hora de estudiar la convergencia de un proceso iterativo a la solución de una ecuación no lineal, lo constituye la elección del punto de partida. En ese sentido, en la Sección 3.4, se consideró a la anomalía media como punto de partida $E_0 = M$. Debemos destacar que dicho punto inicial también ha sido tomado en cuenta en las investigaciones realizadas por Diloné y Gutiérrez [62] y Zhen-Shi-Ming [198]. De forma análoga se analizaron las situaciones $E_0 = \pi$, propuesta por Charles y Tatum [35] y

$$\begin{cases} E_0 = M + (\sqrt[3]{6M} - M)e^2 & \text{si } M < 0.1, \\ E_0 = M + 0.85e & \text{si } M \geq 0.1, \end{cases}$$

dada por Danby-Burkardt [51] y Palacios [147].

Entre las conclusiones obtenidas, se destaca que los peores resultados de convergencia se presentaron cuando los métodos iterativos se aplicaron a partir de la condición inicial $E_0 = M$. Especialmente llamativo es el caso del método de Chebyshev, cuyo porcentaje de fracasos fue de un 97.5 %. Bajo las condiciones iniciales dadas en (3.7), los métodos iterativos, aplicados a la ecuación de Kepler, experimentaron su mejor comportamiento de convergencia. Por otro lado, para la condición inicial $E_0 = M$, se verificó que el método de Danby supera a los demás métodos, mientras que para las condiciones iniciales $E_0 = \pi$ y las dadas en (3.7), se comprueba que el método super-Halley es el más efectivo.

En el estudio comparativo de los métodos iterativos aplicados a la ecuación de Kepler se puso de manifiesto la aparición de ciclos o comportamiento caótico. Por ese motivo, en la sección 3.5, Ciclos atractores para el método de Newton aplicado a la ecuación de Kepler, nos centramos en el análisis del método de Newton. Para tal acción consideraremos la ecuación de Kepler como una función $f(x) = x - a\pi - b\text{sen } x$ definida en la recta real. Nuestro objetivo consistió en caracterizar los parámetros a y b para mostrar la existencia de ciclos superatractores de periodo 2. Los resultados originales de esta sección están contenidos en la demostración de los teoremas 3.2, 3.3, 3.4, 3.5, 3.5 y 3.7. Dicho análisis sirvió de base para la investigación titulada *On the existence of attractive cycles for Newton's method applied to Kepler's equation* ciclos, publicada por la Journal of Advanced Research in Applied Mathematics, [63].

En la sección 3.6 se utilizan algunas técnicas clásicas de la teoría del caos para analizar el comportamiento dinámico de diversos métodos iterativos aplicados a la ecuación de Kepler. Así, se extienden algunos resultados dados por Stumpf, [175], para el caso del método de Newton a otros métodos considerados en esta memoria. En concreto, se analiza el signo de

la derivada schwarziana de la función de iteración obtenida al aplicar dichos métodos a la ecuación de Kepler. Un análisis gráfico de la derivada schwarziana revela que no se puede demostrar que sea siempre negativa, luego, al igual que en el caso del método de Newton, no se pueden descartar comportamientos no convergentes.

La sección 3.7 complementa el estudio dinámico de la ecuación de Kepler iniciado en secciones anteriores. Siguiendo el trabajo iniciado en [126], se entiende la ecuación de Kepler como un sistema dinámico en el plano complejo, donde se admite que tanto la variable como los parámetros implicados pueden tomar valores complejos. En este caso, la principal novedad es que se muestra la aparición de unas interesantes estructuras de tipo Mandelbrot o de tipo Julia en el análisis gráfico de la ecuación de Kepler compleja.

En el Capítulo 4, Resultados de convergencia semilocal para el método de Newton aplicados a la ecuación de Kepler, se presenta un novedoso estudio de aplicación de la convergencia semilocal del método de Newton, a partir de los criterios establecidos por Kantorovich, Smale, Gutiérrez y Wang-Zhao. El mismo fue publicado en el libro *Una introducción al método de Newton para resolver ecuaciones no lineales*, escrito, en el 2009, por quien suscribe, para optar al Diploma de Estudios Avanzado de Doctorado y publicado posteriormente en *A comparative study of different semilocal convergence results applied to Kepler's equation*, que forma parte de la monografía científica que con el título *Computational Mathematics: Theory, methods and Applications* publicó la editorial Nova Science Publishers, Inc. en 2011 (véase [62]).

En este capítulo observaremos que los resultados de convergencia semilocal, además de garantizar la convergencia del método de Newton-Raphson a una solución de la ecuación considerada, proporcionan unos resultados sobre la existencia y la unicidad de solución. En ese sentido, procedimos a comparar los resultados sobre existencia de unicidad que proporcionan cuando se aplican a una misma ecuación: la ecuación de Kepler. Las conclusiones de este estudio comprobaron que bajo las condiciones de Cauchy-Kantorovich, la ecuación de Kepler posee el mayor rango de solución ($e \leq 0.4142$). La comparación también nos permitió verificar que bajo las condiciones fuertes de Kantorovich, la ecuación de Kepler posee un rango de solución, ($e < 0.3759$), mayor que bajo las condiciones de la α -teoría de Wang-Zhao, ($e < 0.3678$), y de la α -teoría de Smale ($e \leq 0.3594$).

Índice general

1. La ecuación de Kepler	1
1.1. Orígenes históricos de la ecuación de Kepler	1
1.2. Obtención de la ecuación de Kepler	5
1.3. Resolución de la ecuación de Kepler	7
2. Métodos iterativos	11
2.1. Introducción	11
2.2. El método de Newton	17
2.2.1. Notas históricas sobre el método de Newton	18
2.2.2. Descripción del método de Newton	20
2.2.3. Convergencia del método de Newton	23
2.3. El método de Halley	42
2.3.1. Notas históricas sobre el método de Halley	42
2.3.2. Descripción del método de Halley	44
2.3.3. Convergencia del método de Halley	50
2.4. El método de Chebyshev	62
2.4.1. Notas históricas sobre el método de Chebyshev	62
2.4.2. Descripción del método de Chebyshev	65
2.4.3. Convergencia del método de Chebyshev	72
2.5. El método super-Halley	80
2.5.1. Notas históricas sobre el método super-Halley	80
2.5.2. Descripción del método super-Halley	81
2.5.3. Convergencia del método super-Halley	85
2.6. El método de la Secante	90
2.6.1. Notas históricas sobre el método de la Secante	91
2.6.2. Descripción del método de la Secante	94
2.6.3. Convergencia del método de la Secante	96
2.7. Otros métodos iterativos	111
2.7.1. El método de Yun-Petković	111
2.7.2. El método de Danby	119

3. Métodos iterativos y la ecuación de Kepler	125
3.1. Introducción	125
3.2. Métodos iterativos punto a punto y la ecuación de Kepler	127
3.3. Métodos iterativos multipunto y la ecuación de Kepler	133
3.4. Análisis de los puntos de inicio	139
3.5. Ciclos atractores	143
3.6. La derivada schwarziana y la ecuación de Kepler	151
3.7. Estudio dinámico de la ecuación de Kepler	160
3.7.1. La función de Kepler en el plano complejo	166
4. Convergencia semilocal	169
4.1. Condiciones de Kantorovich	170
4.2. Condiciones fuertes de Kantorovich	172
4.3. α -teoría de Smale	176
4.4. α -teoría de Wang-Zhao	179
4.5. Existencia y unicidad de soluciones	182
4.6. Líneas futuras de investigación y Conclusiones	184
4.7. Anexo I: Cuadros con experimentos numéricos	186
4.8. Anexo II: Cuadros con experimentos numéricos	187

Índice de figuras

1.1. Claudio Ptolomeo (100AC – 170AC) y Aristóteles (384AC – 322AC).	3
1.2. Tycho Brahe (1546 – 1601) y Johannes Kepler (1571 – 1630).	4
1.3. Descripción geométrica de las anomalías verdadera θ , excéntrica E y media M de la órbita de un planeta.	6
2.1. Sir Isaac Newton (1643 – 1727) y Thomas Simpson (1710 – 1761).	18
2.2. Joseph Louis Lagrange (1736 – 1813) y Joseph Fourier (1768 – 1830).	20
2.3. Interpretación geométrica del método de Newton-Raphson	22
2.4. Augustin Louis Cauchy (1789 – 1857) y Leonid V. Kantorovich (1912 – 1986).	28
2.5. Gráfica del polinomio mayorizante $p(t) = \frac{1}{2}\gamma t^2 - t + \beta$	30
2.6. Stephen Smale (1930) y Mike Shub (1943).	33
2.7. Gráfica de la función mayorizante $\phi(t) = \beta - t + \frac{\gamma t^2}{1-\gamma t}$	34
2.8. Interpretación geométrica del teorema de Fourier.	39
2.9. Edmund Halley (1656 – 1742) y Ernst Schröder, (1841 – 1902).	42
2.10. A la izquierda, gráfica de $\varphi(t) = \frac{N}{6}t^3 + \frac{M}{2}t^2 - t + \eta$. A la derecha, gráfica de $\varphi'(t) = \frac{N}{2}t^2 + Mt - 1$	54
2.11. P. Chebyshev, (1821 – 1894) y L. Euler, (1707 – 1783).	63
2.12. Interpretación geométrica del método de Chebyshev.	67
2.13. N. Obreshkov, (1896 – 1963)	71
2.14. Francis Viète (1540 – 1603) y René Descartes (1596 – 1650).	91
2.15. Sir Isaac Newton (1643 – 1727) y José Mariano Vallejo (1779 – 1846).	92
2.16. Interpretación geométrica del método de la Secante	95
2.17. Miodrag S. Petković, (1948).	113
2.18. J. M. A. Danby (1929 – 2009).	120
3.1. Ciclo atractor del método obtenido al sustituir $\alpha = 2$ en (3.3), para $e = 0.996$, $M = 0.04\pi$ y $x_0 = M$	131
3.2. A la izquierda, gráfico de densidad del método de Newton aplicado a la ecuación de Kepler. A la derecha, gráfico de densidad ampliado del método de Newton aplicado a la ecuación de Kepler.	133

3.3. A la izquierda, gráfico de densidad ampliado del método de Halley aplicado a la ecuación de Kepler. A la derecha, gráfico de densidad ampliado del método super-Halley aplicado a la ecuación de Kepler. 134

3.4. Gráfico de densidad ampliado del método de Chebyshev aplicado a la ecuación de Kepler. 134

3.5. A la izquierda, gráfica de $f(E) = M - E + e \operatorname{sen}(E)$, para $e = 0.99$, $M = 0.2$ radianes y $E_0 = \pi$. A la derecha, gráfica de $f(E) = M - E + e \operatorname{sen}(E)$, para $e = 0.99$, $M = 0.2$ radianes y $E_0 = M$ 140

3.6. Gráfica de (3.16), para $M = 0$ y $e = j/10$ $j = 0, 1 \dots 10$ 141

3.7. Gráficas de $N_f(N_f(x))$ y $N_f(N_f(N_f(x)))$ para $f(x)$ definida en (3.15), con $M = 1$ y $e = 0.9$ 144

3.8. A la izquierda, un ciclo de periodo 2 de la forma $\{0, 8\pi\}$ para el método de Newton aplicado a $f(x) = x - 4\pi - 0.5 \operatorname{sen} x$. A la derecha, un ciclo de periodo 2 de la forma $\{\pi, 9\pi\}$ para el método de Newton aplicado a $f(x) = x - 5\pi + 0.5 \operatorname{sen} x$. 149

3.9. A la izquierda, un ciclo de periodo 2 en la forma $\{6\pi, 9\pi/2\}$, para el método de Newton aplicado a $f(x) = x - (\frac{9+18\pi}{2+3\pi})\pi - (\frac{3\pi}{2+3\pi}) \operatorname{sen} x$. A la derecha, un ciclo de periodo 2 de la forma $\{6\pi, 11\pi/2\}$, para el método de Newton aplicado a $f(x) = x - (\frac{6\pi-11}{\pi-2})\pi - (\frac{\pi}{\pi-2}) \operatorname{sen} x$ 150

3.10. A la izquierda, un ciclo de periodo 2 de la forma $\{-3\pi, \pi/2\}$, para el método de Newton aplicado a $f(x) = x - (\frac{1-21\pi}{2+7\pi})\pi + (\frac{7\pi}{2+7\pi}) \operatorname{sen} x$. A la derecha, un ciclo de periodo 2 de la forma $\{-5\pi, -5\pi/2\}$, para el método de Newton aplicado a $f(x) = x - (\frac{5-25\pi}{5\pi-2})\pi - (\frac{5\pi}{2-5\pi}) \operatorname{sen} x$ 150

3.11. Curvas correspondientes a la derivada schwarziana de los métodos de Newton (izquierda) y Halley (derecha), aplicados a la ecuación de Kepler. Los colores de las curvas representan distintos valores de e con $M = 0.2$ 153

3.12. Curvas correspondientes a la derivada schwarziana de los métodos de Chebyshev (izquierda) y super-Halley (derecha), aplicados a la ecuación de Kepler. Los colores de las curvas representan distintos valores de e con $M = 0.2$ 153

3.13. A la izquierda, derivada schwarziana del método de Danby. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$. Los colores de las curvas representan distintos valores de e con $M = 0.2$ 153

3.14. Curvas correspondientes a la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. Los colores de las curvas representan distintos valores de e con $M = 0.2$ 154

3.15. A la izquierda, la derivada schwarziana del método de Newton. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 156

3.16. A la izquierda, la derivada schwarziana del método de Halley. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 156

3.17. A la izquierda, la derivada schwarziana del método de Chebyshev. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 156

3.18. A la izquierda, la derivada schwarziana del método de super-Halley. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$. . . 157

3.19. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 157

3.20. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 157

3.21. A la izquierda, la derivada schwarziana del método de Danby. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$ 158

3.22. A la izquierda, la derivada schwarziana del método de Newton. A la derecha, la derivada schwarziana del método de Chebyshev. En ambos casos $M = 0.2$ y $e = 0.997$ 159

3.23. A la izquierda, derivada schwarziana del método de Danby. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. En ambos casos $M = 0.2$ y $e = 0.997$ 159

3.24. Derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$, con $M = 0.2$ y $e = 0.997$ 159

3.25. A la izquierda, la derivada schwarziana del método super-Halley. A la derecha, la derivada schwarziana del método de Halley. En ambos casos $M = 0.2$ y $e = 0.997$ 160

3.26. Diagrama de Feigenbaum para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para el intervalo $0.5 \leq r \leq 4$. A la derecha, para el intervalo $2 \leq r \leq 4$. Puede observarse la ruta del caos: sucesivos desdoblamiento de periodo que desembocan en un periodo infinito. 161

3.27. Diagrama de Telaraña para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para $r = 1$. A la derecha, para $r = 1.5$ 162

3.28. Diagrama de Telaraña para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para $r = 2$. A la derecha, para $r = 2.5$ 162

3.29. A la izquierda, diagrama de Telaraña, a la derecha, el Diagrama de Feigenbaum para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$, con $r = 3$ 163

3.30. A la izquierda, cascada de bifurcación de duplicación del periodo para la función $f(x) = e \text{ sen}(x)$. A la derecha, desplazamiento de las primeras bifurcaciones de duplicación del periodo cuando M aumenta. Para $M = 0$, la curva es azul, para $M = 0.25$, la curva es verde y para $M = 0.5$, la curva es roja 164

3.31. Iteración de de la función $E_{n+1} = M + e \operatorname{sen} E_n$ en el plano (M, e) para $E_0 = 0$.
 Se observan los contornos de las funciones $f^k(M, e; 0)$, donde $f(M, e; 0) = M + e \operatorname{sen} z$ 165

4.1. Gráficas de las funciones mayorizantes $\phi_W(t)$, $\phi_G(t)$, $\phi_S(t)$ y $\phi_K(t)$ para el caso del planeta Mercurio (izquierda) y el asteroide Hungaria (derecha). 184

4.2. Conjuntos del tipo $\mathcal{K}(M; z_0) = \{e \in \mathbb{C} | \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}$, para distintos valores de M y z_0 211

4.3. Conjuntos del tipo $\mathcal{J}(M, e) = \{z \in \mathbb{C} | \{f^n(M, e, z)\}_{n \in \mathbb{N}} \text{ acotada}\}$ para distintos valores de M y e 212

4.4. Conjuntos del tipo $\mathcal{I}(e, z_0) = \{M \in \mathbb{C} | \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}$ para distintos valores de e y z_0 213

Índice de cuadros

2.1. Iteraciones obtenidas al aplicar el método de Yun-Petković a (2.216).	115
2.2. Iteraciones obtenidas al aplicar el método de Yun-Petković a (2.217).	115
2.3. Iteraciones obtenidas al aplicar el método de Bisección a (2.216).	116
2.4. Iteraciones obtenidas al aplicar el método de Bisección a (2.217).	116
2.5. Iteraciones obtenidas al aplicar el método de la Secante a (2.216).	116
2.6. Iteraciones obtenidas al aplicar el método de la Secante a (2.216).	117
3.1. Función de iteración, orden de convergencia (p), constante asintótica del error (β) e índice de eficiencia (δ) de los distintos métodos iterativos que vamos a aplicar a la ecuación de Kepler, donde las constantes A_k están definidas en (3.6).127	
3.2. Valores de la anomalía media y de la excentricidad utilizados por M. Palacios [147].	128
3.3. Valores reales asociados a la anomalía media y a la excentricidad de diferentes asteroides y el planeta Mercurio [208].	128
3.4. Media m del número de iteraciones de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.	130
3.5. Media ponderada \bar{m} del número de iteraciones de de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.	130
3.6. Porcentaje de fracasos de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.	130
3.7. Función de iteración, orden de convergencia (p) e índice de eficiencia (δ) de los métodos iterativos de la Secante, Yun-Petković y Bisección.	135
3.8. Media m del número de iteraciones necesarias para la convergencia de los métodos de la Secante, Yun-Petković y Bisección aplicados a la ecuación de Kepler, tomando como puntos de partida los extremos del intervalo $[0, \pi]$	138
3.9. Porcentaje de fracasos de de los métodos de la Secante, Yun-Petković y Bisección aplicados a la ecuación de Kepler, tomando como puntos de partida los extremos del intervalo $[0, \pi]$	138

3.10. Número de iteraciones necesarias para alcanzar la solución de la ecuación Kepler, usando el método de la Secante, Yun-Petković y Bisección, y tomando como punto de partida los extremos del intervalo $[0, \pi]$ 138

3.11. Valores obtenidos a partir de (3.15), en donde $M = 0.2$, $e = 0.99$ y $E_0 = \pi$. . . 141

3.12. Valores obtenidos a partir de (3.15), en donde $M = 0.2$, $e = 0.99$ y $E_0 = M$. . . 142

3.13. Valores obtenidos a partir de (3.15), en donde $M = 1$, $e = 0.9$ y $E_0 = 5.32$. . . 143

3.14. Presencia de ciclos atractores para el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) para diferentes opciones de los parámetros a y b dados por los Teoremas 3.2–3.7. 149

4.1. Función mayorizante para el método de Newton aplicado a la ecuación de Kepler para las diferentes teorías consideradas, junto con la condición en términos de la excentricidad para que dichas teorías puedan ser aplicadas. 183

4.2. Valores de t^* y t^{**} para las 4 teorías consideradas. En todos los casos, hemos denotado $\delta = \frac{e}{1-e}$ 183

4.3. Existencia de solución para dos casos reales. 183

4.4. Unicidad de solución para dos casos reales. 184

4.5. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos $E_0 = M$ 187

4.6. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos $E_0 = \pi$ 188

4.7. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos las condiciones dadas en (3.7). 189

4.8. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Newton y tomando como punto de partida $E_0 = M$. 190

4.9. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Newton y tomando como punto de partida $E_0 = \pi$. 191

4.10. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Newton y tomando como punto de partida las condiciones dadas en (3.7). 192

4.11. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Chebyshev y tomando como punto de partida $E_0 = M$ 193

4.12. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Chebyshev y tomando como punto de partida $E_0 = \pi$ 194

4.13. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Chebyshev y tomando como punto de partida las condiciones dadas en (3.7). 195

4.14. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida $E_0 = M$. 196

4.15. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida $E_0 = \pi$. 197

4.16. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida las condiciones dadas en (3.7). 198

4.17. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida $E_0 = M$ 199

4.18. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida $E_0 = \pi$ 200

4.19. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida las condiciones dadas en (3.7). 201

4.20. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = \frac{3}{2}$ y tomando como punto de partida $E_0 = M$. 202

4.21. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = \frac{3}{2}$ y tomando como punto de partida $E_0 = \pi$. 203

4.22. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = \frac{3}{2}$ y tomando como punto de partida las condiciones dadas en (3.7). 204

4.23. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = 2$ y tomando como punto de partida $E_0 = M$. 205

4.24. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = 2$ y tomando como punto de partida $E_0 = \pi$. 206

4.25. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = 2$ y tomando como punto de partida las condiciones dadas en (3.7). 207

4.26. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida $E_0 = M$. 208

4.27. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida $E_0 = \pi$. 209

4.28. Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida las condiciones dadas en (3.7). 210

4.29. Código de Mathematica para generar fractales tipo Mandelbrot con la ecuación de Kepler. Está desarrollado para el caso particular en que $M = 0.01$, $e = 0.9$ y nivel de resolución de 0.1 214

4.30. Código de Mathematica para generar fractales tipo Julia con la ecuación de Kepler. Está desarrollado para el caso particular en que $M = 0.01$, $e = 0.9$ y nivel de resolución de 40 iteraciones o pasos 215

Capítulo 1

La ecuación de Kepler

1.1. Orígenes históricos de la ecuación de Kepler

Determinar la posición de los planetas en la bóveda celeste ha sido un problema que ha preocupado al hombre desde la antigüedad, tanto por cuestiones religiosas y místicas como también por cuestiones de índole práctica. Hoy sabemos, gracias al desarrollo de la Mecánica Celeste y del Cálculo Infinitesimal, que para determinar la posición de un planeta, o cualquier objeto que se encuentra sometido a la atracción gravitatoria de otro, es necesario resolver la denominada *ecuación de Kepler*

$$M = E - e \sin E. \quad (1.1)$$

Se trata de encontrar un valor de E una vez conocidos los valores de M y e . Los parámetros involucrados en esta ecuación tienen una clara interpretación física. En concreto, $M \in [0, 2\pi]$ es la anomalía media de una órbita elíptica, que es el ángulo girado por el planeta, a partir de una posición conocida, si se moviera con movimiento circular uniforme; $e \in [0, 1)$ es la excentricidad de la elipse que describe el planeta; E es la anomalía excéntrica, que es el ángulo que le corresponde a la posición del planeta en la parametrización natural de la elipse.

Debido a su relevancia y a su carácter no lineal, la ecuación de Kepler fue una de las primeras que se intentaron resolver mediante métodos numéricos. Desde los días de Newton este problema ha sido estudiado sin descanso por astrónomos y matemáticos y su solución se ha conseguido de muchas y diversas maneras, dependiendo de los factores del contexto en donde debe resolverse la ecuación. Esto demuestra, tal como dice Danby [51], que el tema sigue abierto y que todavía existen muchas posibilidades para buscar nuevas alternativas de solución. De hecho, existe un gran cúmulo de trabajos que se relaciona con su solución numérica o analítica. En efecto, Multon [134], presenta una lista incompleta de 123 artículos sobre la referida ecuación.

Para buscar el origen de la ecuación de Kepler hay que remontarse al momento en el que el hombre se dio cuenta de que hay astros que no ocupan una posición fija en el cielo, si se los compara con las estrellas fijas. A simple vista pueden distinguirse siete de estos astros, el

Sol, la Luna, Mercurio, Venus, Marte, Júpiter y Saturno. Mientras el Sol y la Luna parecen moverse dando vueltas alrededor de la Tierra en sentido de oeste a este, el resto de planetas presentan un movimiento peculiar. Su movimiento principal es como el del Sol y la Luna, dando vueltas alrededor de la Tierra de oeste a este, pero, de vez en cuando, detienen su movimiento para retroceder y moverse durante un tiempo en sentido este oeste, para retomar de nuevo su rumbo habitual.

Los primeros modelos sobre el movimiento de los planetas que han llegado hasta nuestros días corresponden a los griegos. Pensaban, como era tradición y la observación así lo evidenciaba, que la Tierra ocupaba un lugar central inmóvil y que el resto de astros giraba en torno a ella. Como el cielo era el lugar de lo perfecto, todo movimiento que tuviera lugar en él debería ser también perfecto, es decir circular. Eudoxo, un matemático griego del siglo IV a.C., propuso un modelo para el movimiento de los planetas basado en un complejo mecanismo de esferas engranadas que reproducía, en buena medida, lo que se podía observar a simple vista. Cada planeta estaba situado sobre una esfera que se unía a la siguiente mediante un eje alrededor del cual giraba. La primera esfera era la de la Luna, luego la de Mercurio, la de Venus, la del Sol, la de Marte, la de Júpiter, la de Saturno y finalmente la esfera de las estrellas fijas. Calipo, también en el siglo IV a.C., hizo mejoras sobre este sistema añadiendo esferas intermedias para refinar los movimientos. Básicamente este es el modelo que adopta Aristóteles, presentando el Universo como una cebolla compuesta por diferentes esferas encajadas que contienen a los planetas y en la última capa el límite del Universo formado por la esfera de las estrellas fijas.

El modelo de Aristóteles no permitía explicar algunos hechos que se podían percibir a simple vista, tales como la variación del brillo de los planetas, ya que siempre se encontraban a una misma distancia de la Tierra. Esto era así, porque el modelo de Aristóteles era más bien un modelo conceptual y no práctico, ya que además no podía predecir con precisión la posición de los planetas. Fue Apolonio (siglo III a.C.) el que inventó un sistema de epiciclos para explicar el movimiento esquivo de los planetas. Supuso que los planetas giraban en círculos alrededor de un punto que daba vueltas alrededor de la Tierra. Con esto se explicaba por qué a veces parecía que el planeta se movía hacia atrás y también su cambio de brillo, ya que ahora su distancia a la Tierra era variable.

Hiparco, en el siglo II a.C., y posteriormente Ptolomeo en el II de nuestra era mejoraron el sistema de epiciclos hasta crear una maquinaria casi perfecta, ya que se podía predecir la posición de los planetas casi con total exactitud. Todo queda compendiado en una gran obra, el *Almagesto*, una de las más influyentes de todos los tiempos y que sintetiza todo el saber astronómico de la época.

La hegemonía de los modelos planetarios de Aristóteles y Ptolomeo se mantuvo durante toda la edad media, hasta que en el Siglo XVI, el astrónomo polaco Nicolás Copérnico (1473-1543) propuso un nuevo modelo, conocido como Sistema Heliocéntrico, en el cual el Sol, y no la Tierra, era el centro del Sistema Solar. En este nuevo modelo, la explicación de la retrogra-



Figura 1.1: Claudio Ptolomeo (100AC – 170AC) y Aristóteles (384AC – 322AC).

dación de los planetas y el movimiento diario de las estrellas se explica con mayor sencillez. Sin embargo, Copérnico seguía pensando que los planetas efectuaban un movimiento circular uniforme, por lo que su sistema seguía manteniendo alguno de los artificios de Ptolomeo. Siendo más preciso que el modelo de Ptolomeo, la gran dificultad que encontró el Sistema Heliocéntrico fue de corte conceptual, al desplazar a la Tierra del centro del Universo.

Tratando de conciliar el Sistema Heliocéntrico con una posición hegemónica de la Tierra, Tycho Brahe, propuso un sistema híbrido, donde la Tierra ocupaba el centro del Universo con el Sol y la Luna dando vueltas alrededor de la Tierra y el resto de planetas orbitando en torno al Sol. Sin embargo, esta no fue la mayor contribución de Tycho Brahe, sino su colección de observaciones de los planetas con la mayor precisión conocida hasta la fecha. Gracias a ellas, se sabía que tanto el modelo de Ptolomeo, como el de Copérnico no eran capaces de aproximar bien la posición de algunos planetas, como era el caso de Marte, al que Brahe observó sin descanso.

Brahe se trasladó a Praga y ahí tomó a Johannes Kepler (1571-1630) como su asistente. Brahe y Kepler tenían algunas diferencias desde el punto de vista científico. Una de ellas era que Brahe creía que la Tierra era el centro del Sistema Solar, mientras que Kepler creía en el modelo de Copérnico, con el Sol en el centro. A la muerte de Brahe, y en poder de las tablas de observaciones de Marte, Kepler se lanza al reto de domar la órbita de este planeta, el más díscolo de todos, ya que era el que más se desviaba de su posición predicha. Tras muchos intentos, Kepler se da cuenta de que hay que rechazar la teoría de los movimientos circulares uniformes, después de probar con una órbita circular donde el Sol se encontraba en el punto ecuante.¹ Sin embargo, mientras la mayoría de posiciones de Marte encajaban perfectamente en la nueva órbita así definida, dos observaciones se separaban unos 8 grados de la posición teórica. Esto obligó a Kepler a renunciar a la hipótesis de que la órbita tenía que ser circular y

¹El ecuante es un punto cerca del centro de la órbita de un planeta respecto al cual el movimiento del planeta es uniforme. Este punto fue inventado por Ptolomeo en el siglo II de nuestra era.

tras probar con varias curvas, concluyó que los planetas se movían en elipses. De esta forma, estableció tres leyes empíricas, que se conocen como las tres *leyes de Kepler* y que son las siguientes:

1. Un planeta gira alrededor del Sol describiendo una órbita elíptica con el Sol ubicado en uno de los focos.
2. La recta imaginaria que une al Sol con el planeta barre áreas iguales en tiempos iguales.
3. El cuadrado del periodo orbital del planeta es proporcional al cubo del semieje mayor de la órbita.



Figura 1.2: Tycho Brahe (1546 – 1601) y Johannes Kepler (1571 – 1630).

La consecuencia inmediata de estas tres leyes, en concreto de la segunda, es que el planeta no se mueve con movimiento uniforme, desplazándose más rápido cuando se encuentra cerca del Sol que cuando se encuentra más alejado. Esto dificulta el cálculo de la posición del planeta en su órbita y para ello Kepler tuvo que deducir lo que se denomina *ley horaria* del movimiento planetario, que trata de dar la posición del planeta en función del tiempo. Y es aquí donde tiene su origen la ecuación de Kepler, ya que es preciso resolverla para poder determinar la posición de un planeta para un instante de tiempo dado.

Por finalizar este contexto histórico en el que se enmarca la ecuación de Kepler, debemos citar a Newton, quien dedujo las leyes de Kepler a partir de la ley de gravitación universal y dio el espaldarazo definitivo al Sistema Heliocéntrico. Pero Newton no solo contribuyó a eso, sino que también sentó las bases del Cálculo Infinitesimal y de la Mecánica Celeste, gracias a la cual se ha desarrollado la exploración espacial, de la que nos sentimos ahora tan dependientes. Los satélites que están en órbita siguen las leyes de Kepler y, por tanto, resolver la ecuación de Kepler tiene hoy en día más sentido, si cabe, que en la época de Kepler o Newton. Existen excelentes textos clásicos que estudian en profundidad los orígenes de la

ecuación de Kepler y sus aplicaciones en Astrodinámica y Mecánica Celeste. Estos textos también muestran las diferentes técnicas, tanto analíticas, como geométricas o numéricas que se han desarrollado para su resolución. El siguiente listado es sólo una pequeña muestra de los libros que tratan la ecuación de Kepler y que han sido empleados para la redacción de esta memoria: Abad [1], Abad, Docobo y Elipe [2], Colwell [49], Danby [53] y Or's, Català y Núñez de Murga [56].

1.2. Obtención de la ecuación de Kepler

Existen dos formas de obtener la ecuación de Kepler, la geométrica y la analítica, como puede verse en la literatura existente sobre el tema ([1], [49], [56], [171]). Nosotros presentamos aquí la deducción geométrica, basada en una transformación de coordenadas polares, referida al foco de la elipse, a otra referida al centro de la misma.

Supongamos que un planeta se mueve alrededor del Sol siguiendo una órbita elíptica de excentricidad e , $0 < e < 1$. Sin pérdida de generalidad, y para simplificar los cálculos, supongamos que el Sol se encuentra en un foco de la elipse de coordenadas $S = (0, 0)$. Llamemos a al semieje mayor de la elipse, b al menor, C al centro de la elipse, A al perihelio, o punto más cercano de la elipse al foco S , y T al periodo o tiempo que tarda el planeta en completar una vuelta, véase la Figura 1.3. Sea $P = P(t)$ la posición que ocupa el planeta en el instante t y sean (x, y) sus coordenadas cartesianas y (r, θ) sus coordenadas polares, es decir

$$P = (x, y) = (r \cos \theta, r \sin \theta).$$

El ángulo θ recibe el nombre de anomalía verdadera y corresponde al ángulo girado por el planeta desde el perihelio. El objeto que persigue la ecuación de Kepler es relacionar este ángulo con el tiempo t , transcurrido desde que el planeta pasó por el perihelio. Es por ello que se introduce un ángulo auxiliar M , denominado anomalía media, que mide el desplazamiento angular del planeta desde el perihelio, suponiendo que se desplaza con velocidad uniforme. Así, si denotamos por τ al instante en el que el planeta pasa por el punto A , resulta

$$M = \frac{2\pi}{T}(t - \tau). \quad (1.2)$$

Por otra parte, podemos considerar el ángulo E , o anomalía excéntrica, que resulta de la parametrización natural de la elipse referida a su centro. Este ángulo puede visualizarse trazando una circunferencia de radio a y centro C . En efecto, si trazamos una perpendicular al eje mayor que pasa por P y denotamos por Q a la intersección de dicha perpendicular con la circunferencia, la anomalía excéntrica es el ángulo $\angle QCA$.

Mientras que las anomalías verdadera y excéntrica tienen una relación geométrica clara con la elipse, la anomalía media no la tiene. Sin embargo, podemos introducir un punto Y sobre la circunferencia de centro C y radio a , de manera que el ángulo $\angle YCA$ corresponda

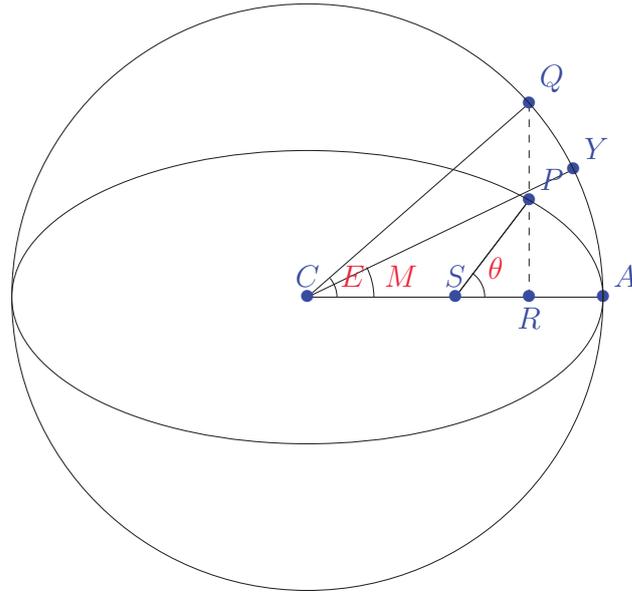


Figura 1.3: Descripción geométrica de las anomalías verdadera θ , excéntrica E y media M de la órbita de un planeta.

con M . Gracias a la segunda ley de Kepler, o ley de las áreas, si llamamos A_1 al área del sector elíptico PSA , resulta

$$\frac{A_1}{\pi ab} = \frac{t - \tau}{T} \Rightarrow A_1 = \frac{1}{2} abM. \quad (1.3)$$

Puesto que el área del sector circular YCA es $\frac{1}{2} a^2 M$, resulta que el área del sector circular YCA es a/b veces el área del sector elíptico PSA .

De la geometría de la Figura 1.3 se deduce que $\cos E = (x + CS)/a$ y como, por definición, $CS/CA = e$, se tiene que

$$x = a(-e + \cos E). \quad (1.4)$$

Por otra parte, la elipse está centrada en $C = (-CS, 0)$, por lo tanto su ecuación en coordenadas cartesianas es

$$\frac{(x + CS)^2}{a^2} + \frac{y^2}{b^2} = \frac{(a \cos E)^2}{a^2} + \frac{y^2}{b^2} = 1,$$

y de aquí, $y = b \sin E$ o, teniendo en cuenta que $e^2 = 1 - b^2/a^2$,

$$y = a\sqrt{1 - e^2} \sin E. \quad (1.5)$$

De aquí resulta

$$r = a(1 - e \cos E), \quad (1.6)$$

donde $r = x^2 + y^2$ es la distancia del punto P al foco S .

Tengamos ahora en cuenta que la ecuación de una elipse en coordenadas polares es

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta},$$

de donde se deduce que

$$\cos \theta = \frac{1}{e} \left(\frac{1 - e^2}{1 + e \cos E} - 1 \right) = \frac{\cos E - e}{1 - e \cos E}.$$

De aquí podemos concluir las siguientes igualdades

$$1 - \cos \theta = \frac{(1 + e)(1 - \cos E)}{1 - e \cos E},$$

$$1 + \cos \theta = \frac{(1 - e)(1 - \cos E)}{1 + e \cos E},$$

y, a partir de fórmulas trigonométricas conocidas:

$$\tan \left(\frac{\theta}{2} \right) = \sqrt{\frac{1 - \cos \theta}{1 + \cos \theta}} = \sqrt{\frac{1 + e}{1 - e}} \tan \left(\frac{E}{2} \right). \quad (1.7)$$

La relación (1.7) nos da la relación entre la anomalí excéntrica y la verdadera, faltando la relación entre una de ellas y la anomalía media M . Para llegar a esta relación, consideremos el punto R , sobre el lado SA , obtenido como intersección con la recta perpendicular a SA que pasa por P (véase la Figura 1.3). Denotemos por A_1 , A_2 y A_3 a las áreas de las regiones PSA , PSR y PRA respectivamente. El área de A_1 viene dada por (1.3), mientras que, de (1.4) y (1.5), se sigue que

$$A_2 = \frac{xy}{2} = \frac{1}{2} ab \sin E (\cos E - e). \quad (1.8)$$

Denotamos ahora por A_4 , A_5 y A_6 a las áreas de las regiones QRA , QCA y QCR respectivamente. Sin más que calcular las áreas comprendidas bajo las curvas $x^2 + y^2 = a^2$ (circunferencia) y $b^2x^2 + a^2y^2 = a^2b^2$ (elipse), tenemos que

$$A_3 = \frac{b}{a} A_4 = \frac{b}{a} (A_5 - A_6) = \frac{ba}{2} (E - \sin E \cos E). \quad (1.9)$$

Como $A_1 = A_2 + A_3$, a partir de (1.3), (1.8) y (1.9), se tiene

$$M = E - e \sin E,$$

que no es otra cosa que la famosa ecuación de Kepler.

1.3. Resolución de la ecuación de Kepler

La solución a la ecuación de Kepler (1.1) ha sido un tema ampliamente abordado desde su formulación en 1650. Desde entonces, los matemáticos han aplicado básicamente tres maneras de resolverla: la geométrica, la analítica y la numérica. En el libro de Colwell [49] se hace un estudio detallado de las distintas formas de resolver la ecuación de Kepler en sus casi cuatro

siglos de historia, junto con un listado exhaustivo de autores que han trabajado en este tema y las correspondientes referencias.

Veamos algunos momentos importantes en la solución de (1.1). En cuanto las demostraciones geométricas podemos citar la del astrónomo inglés Jeremiah Horrocks quien, en 1638, describe un método basado en diferencias aproximadas para resolver la ecuación de Kepler. Posteriormente, el danés Peder Horrebow obtuvo una solución de la ecuación de Kepler que está estrechamente relacionada a la solución de Horrocks. Los detalles de la solución están contenidos en unos artículos publicados en 1717 y 1718, los cuales fueron compilados en 1742 en un manuscrito titulado *Adytum Astronomiae*. El astrónomo francés Jacques Cassini publicó en 1719 un artículo en el cual emplea un método geométrico para dar solución a la ecuación de Kepler.

Cristopher Wren, en 1658, ofrece una solución de la ecuación de Kepler basada en las cicloides. La misma aparece publicada en la sección *De problemate Kepleriano per Cycloidem Solvendo* del tratado *De cycloide Tractatus* publicado por John Wallis en 1695. El propio Isaac Newton utiliza la construcción de Wren en sus *Principia* (1686) para situar la posición de un planeta en una órbita elíptica.

John Couch Adams, en [3], presenta un método muy sencillo para resolver la ecuación de Kepler a partir de la curva de la función seno. El propio Adams se sorprendió de su sencillez y él mismo afirma que es probable que este método hubiera sido propuesto con anterioridad.

En cuanto a las soluciones analíticas de la ecuación de Kepler, sin duda, las soluciones que usan series de potencias son las más elegantes. De nuevo, siguiendo a Colwell [49] podemos destacar cuatro de estas soluciones. Así, por ejemplo, Joseph-Louis Lagrange obtuvo en 1771, a partir de un resultado general sobre desarrollos en serie de Cauchy, la siguiente fórmula para la anomalía excéntrica E como serie de potencias de la excentricidad e :

$$E = M + \sum_{n=1}^{\infty} a_n(M)e^n, \quad \text{con} \quad a_n(M) = \frac{1}{n!} \frac{d^{n-1}}{dM^{n-1}}(\sin^n(M)).$$

Un siglo después, los trabajos de Gustav Lejeune Dirichlet y Leopold Fejér, entre otros, permitieron dar una solución de la ecuación de Kepler como serie de Fourier de senos y en cuyos coeficientes aparecen las funciones de Bessel. En concreto,

$$E = M + \sum_{n=1}^{\infty} \frac{2}{n} J_n(ne) \sin(nM),$$

donde

$$J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin \theta - n\theta) d\theta.$$

El matemático italiano Tullio Levi-Civita presentó en 1904 una novedosa solución de la ecuación de Kepler, basada en series infinitas y usando técnicas de variable compleja. Levi-Civita demostró que si se denota

$$z = \frac{e \exp(\sqrt{1-e^2})}{1 + \sqrt{1-e^2}}, \quad 0 \leq e < 1,$$

entonces la ecuación de Kepler tiene como solución

$$E = \sum_{n=0}^{\infty} L_n(M)z^n,$$

que es convergente para $0 \leq z < 1$. El cálculo de los coeficientes $L_n(M)$ en función de M se puede hacer de forma recursiva, tal como se muestra en [49, Apéndice E], aunque el elevado coste operacional empleado para su cálculo hacen desaconsejable su uso práctico.

Por último, en 1968 K. Stumpff [174] aplica una herramienta bastante sofisticada, como son las series de Lie a diversos problemas de Mecánica Celeste. En particular, Stumpff consigue un desarrollo de E como serie de potencias de M donde los coeficientes son funciones racionales de e . Por ejemplo, los primeros términos de este desarrollo son

$$E \approx \frac{M}{1-e} - \frac{M^3 e}{3!(1-e)^4} + \frac{M^5 e(9e+1)}{5!(1-e)^7}.$$

En cuanto a las soluciones de la ecuación de Kepler (1.1) mediante técnicas iterativas, los primeros resultados se atribuyen a Leonhard Euler, quién en 1740 ya utilizaba una iteración de punto fijo para aproximar las soluciones de (1.1). En concreto, Euler buscaba un punto fijo de la ecuación

$$f(E) = M + e \sin E$$

a partir de la sucesión

$$E_0 = M, \quad E_{n+1} = M + e \sin E_n, \quad n \geq 0.$$

Euler era consciente de que para valores de $e > 1/2$ hacían falta muchos términos para alcanzar una solución con una buena precisión. Sin embargo, para excentricidades e pequeñas, como es el caso de la órbita de la mayoría de los planetas, unas pocas iteraciones son suficientes.

Otra aportación interesante fue la que dio C. F. W. Peters en 1891, quien propuso como valor para E :

$$E \approx E_2 - \frac{(E_2 - E_1)^2}{E_2 - 2E_1 + E_0},$$

donde $E_1 = M + e \sin E_0$, $E_2 = M + e \sin E_1$. Notemos que la propuesta de Peters no es otra cosa que la aceleración de Aitken aplicada a la sucesión $E_{n+1} = M + e \sin E_n$.

Sin entrar en detalles sobre las aportaciones de muchos otros científicos que buscaban métodos rápidos y seguros para que los astrónomos pudieran hacer sus cálculos en la determinación de las órbitas de los planetas, podríamos afirmar que el método de Newton es uno de los procesos iterativos más usados en la práctica para resolver la ecuación de Kepler (1.1). En concreto, para unos valores de M y e dados y a partir de una aproximación inicial E_0 , el método de Newton aplicado a (1.1) define una sucesión

$$E_{n+1} = E_n - \frac{E_n - e \sin E_n - M}{1 - e \cos E_n}, \quad (1.10)$$

que, en condiciones adecuadas, converge a una solución de (1.1). El estudio de la convergencia de la sucesión (1.10), la elección de puntos de partida E_0 adecuados, el análisis de la sensibilidad de la sucesión ante pequeños cambios en las condiciones iniciales o tratamientos estadísticos para determinar la media del número de iteraciones necesario para que el método de Newton alcance la solución con una precisión prefijada, han sido objeto de estudio por parte de diversos autores, como se pone de manifiesto en los dos últimos capítulos de esta memoria.

Capítulo 2

Métodos iterativos para resolver ecuaciones no lineales

2.1. Introducción

Muchos problemas relacionados con las Matemáticas se reducen a resolver una ecuación

$$f(x) = 0, \tag{2.1}$$

donde, en principio, supondremos que f es una función real de variable real. Pese a su sencillez de planteamiento, éste ha sido un problema complicado que ha sido tratado por numerosos matemáticos célebres: Cardano, Newton, Ruffini, Galois, etc. Incluso para funciones «sencillas» como son los polinomios, el cálculo de sus raíces es un tema complicado. En un principio, se intentó encontrar dichas raíces en función de los coeficientes del polinomio, como ocurre con la ecuación de segundo grado:

$$ax^2 + bx + c = 0 \iff x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

También se conocen fórmulas para polinomios de grados 3 y 4 aunque, debido a su complejidad, no se suelen usar en la práctica. A principios del s. XIX, Galois probó que no existen este tipo de fórmulas para polinomios de grado mayor o igual que 5.

El problema de no poder encontrar de forma exacta las soluciones de una ecuación surgía también al trabajar con ecuaciones trascendentes, como la ecuación de Kepler, relacionada con el cálculo de órbitas planetarias:

$$f(E) = E - e \operatorname{sen} E - M$$

para distintos valores de E y de M .

Estas limitaciones obligan a buscar métodos para encontrar las soluciones de una ecuación de forma aproximada. Básicamente, estos métodos generan una sucesión $\{x_1, x_2, \dots, x_n, \dots\}$

que, en condiciones adecuadas, converge a una solución de la ecuación (2.4). En muchas ocasiones, esta sucesión se genera de forma recursiva

$$x_{n+1} = F(x_n, \dots, x_{n-p+1}),$$

a partir de una función de iteración, F , que puede depender de uno o varios argumentos.

Los intentos por resolver estos tipos de ecuaciones se vienen reflejando desde la antigüedad. Leyendo a Knill [124], nos damos cuenta que desde las civilizaciones griegas, aproximadamente dos mil años antes de Cristo, ya se conocían algoritmos iterativos, tal es el caso de la fórmula de Herón para el cálculo de raíces cuadradas, donde partiendo de una aproximación inicial s de \sqrt{S} se propone como nueva aproximación $(s + S/s)/2$. Con esta fórmula, Herón podía aproximar el valor de la raíz cuadrada que aparece en la fórmula del área de un triángulo:

$$S = \sqrt{p(p-a)(p-b)(p-c)},$$

donde $p = (a + b + c)/2$ es el semiperímetro del triángulo de lados a , b y c .

La utilización de un algoritmo iterativo conlleva el planteamiento de varios problemas teóricos ya que es básico examinar su velocidad de convergencia, determinar el número de iteraciones necesarias para obtener la precisión acordada, estudiar los efectos de la aritmética del computador sobre la sucesión, así como resolver el problema de buscar una buena aproximación inicial.

El problema de resolver una ecuación presenta aspectos muy diferentes. Por ejemplo, con respecto a la función que se presente, la misma puede ser polinomial, trascendental o no disponer de una expresión explícita de la misma, tal y como ocurre cuando se trata de una ecuación diferencial. Por otra parte, está el problema que presentan las soluciones, pues es sabido que una ecuación puede no tener solución, tener una única solución, poseer soluciones múltiples o tener infinitas soluciones.

Definición 2.1 (Error Absoluto). *Dada una sucesión $\{x_n\}_{n=0}^{\infty}$ convergente a un límite x^* , definimos el error absoluto como la diferencia entre el valor exacto x_n de la sucesión y x^* .*

$$a_n = x_n - x^*.$$

En ocasiones se usa la expresión

$$|a_n| = |x_n - x^*|,$$

para referirse al módulo del error absoluto.

Definición 2.2 (Error Relativo). *El error relativo del límite x^* de una sucesión $\{x_n\}_{n=0}^{\infty}$ con respecto al valor exacto x_n de la sucesión, es el cociente entre el módulo del error absoluto y el valor absoluto del valor exacto.*

$$r_n = \frac{|x_n - x^*|}{|x_n|}.$$

El error relativo y el módulo del error absoluto tienen un uso muy especial en la determinación del criterio de parada y el orden de convergencia de un algoritmo iterativo.

Definición 2.3 (Criterios de parada). *Los criterios de parada son los que definen el momento en que se detendrá el algoritmo que se está ejecutando, pues establecen cuando se está lo suficientemente cerca del valor buscado.*

El proceso habitual para fijar un criterio de parada consiste en fijar un parámetro pequeño llamado tolerancia, denotado usualmente por TOL , deteniéndose el proceso cuando se cumple una de las siguientes condiciones

1. $|x_n - x_{n-1}| < TOL$.
2. $\frac{|x_n - x_{n-1}|}{|x_n|} < TOL$.

Estos criterios están justificados en el hecho de que para muchas sucesiones, las expresiones 1 y 2 intervienen en las cotas del error, es decir

$$|x_n - x^*| \leq K_n |x_n - x_{n-1}|,$$

de esta manera se consigue que si $|x_n - x_{n-1}|$ es pequeño, entonces el error cometido, cuando K_n sea de un tamaño moderado, también será pequeño. Un error muy a menudo cometido cuando se resuelve una ecuación no lineal (2.4) es el hecho de usar, como único criterio de parada, la expresión

$$|f(x_n)| < TOL,$$

pues suponer siempre que x_n es una buena aproximación de la raíz si se cumple que $f(x_n)$ es pequeño, puede resultar completamente incorrecto para algunas funciones. Este tipo de caso luego se verá en los ejemplos que presentaremos.

Definición 2.4 (Orden de convergencia). *El orden de convergencia nos indica la velocidad con la cual una sucesión converge a su límite. En términos matemáticos asume la siguiente expresión:*

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^p} = C, \quad \text{con } C \in \mathbb{R}_+ \text{ y } p \geq 1,$$

donde $\{x_n\}$ es la sucesión, x^* el número al que converge la sucesión, p es el orden de convergencia y C es la constante asintótica de error. Bajo esta definición, si $p = 1$, se debe cumplir que $C < 1$.

Si $p = 1, 2, 3, 4, 5$, se dice que es de orden lineal, cuadrático, cúbico, cuártico y quíntico respectivamente.

Se debe recalcar la importancia práctica que presenta el orden de convergencia. En efecto, cuando x_{i+1} está lo suficientemente cerca de x^* , se cumple

$$|x_{i+1} - x^*| \approx \mu |x_i - x^*|^p,$$

y haciendo

$$D_n = -\log |x_i - x^*|,$$

se obtiene

$$D_{n+1} = pD_n + r, \quad \text{donde } r = -\log \mu.$$

Esto significa que en cada iteración se multiplica aproximadamente por p , el número de cifras exactas, de forma que, en general, un algoritmo con orden de convergencia alto converge más rápidamente que un algoritmo de orden menor.

Cuando se verifica que

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|} = 0 \quad \text{ó} \quad \lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^2} = 0,$$

el orden de convergencia de la sucesión $\{x_n\}_{n=1}^{\infty}$ a x^* se llama superlineal o supercuadrático, respectivamente.

Una caracterización del orden de convergencia la encontramos en el teorema de Schröder, cuya demostración damos a continuación. Parte de la demostración la pueden encontrar en [117] y [180].

Teorema 2.1 (de Schröder). *Consideremos un método iterativo*

$$x_{n+1} = g(x_n), \tag{2.2}$$

en el que g es continua y q veces diferenciable en un entorno de x^* , un punto fijo de g , $x^* = g(x^*)$. Entonces la sucesión $\{x_n\}$ converge a x^* con orden de convergencia q si

$$\frac{d^j g(x^*)}{dx^j} = 0, \quad j = 1, 2, \dots, q-1, \quad \frac{d^q g(x^*)}{dx^q} \neq 0.$$

Demostración. Desarrollando (2.2) en serie de Taylor se obtiene

$$x_{n+1} = g(x_n) = x^* + g'(x^*)(x_n - x^*) + \dots + \frac{1}{q!} \frac{d^q g(x)}{dx^q} (x_n - x^*)^q, \quad x \in (x_n, x^*).$$

Luego, usando la condición del enunciado, se sigue

$$x_{n+1} - x^* = \frac{1}{q!} \frac{d^q g(x^*)}{dx^q} (x_n - x^*)^q,$$

con lo que

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^q} = \frac{1}{q!} \frac{d^q g(x^*)}{dx^q} \neq 0.$$

Por lo tanto el método tiene orden de convergencia q . □

Otro importante aspecto que tiene que ver con la convergencia es el que se refiere a la velocidad de convergencia de dos sucesiones, ya que ésta nos permitirá comparar la velocidad de los diferentes algoritmos iterativos.

Definición 2.5 (Velocidad de convergencia de dos sucesiones). *Dada una sucesión $\{x_n\}$ que converge a x y una sucesión $\{y_n\}$ que converge a y , se dice que $\{y_n\}$ converge más rápidamente que $\{x_n\}$ si*

$$\lim_{n \rightarrow \infty} \frac{y - y_n}{x - x_n} = 0.$$

Otro elemento que también se emplea para comparar la velocidad de convergencia de dos sucesiones es la constante asintótica del error. En efecto, dadas dos sucesiones con el mismo orden de convergencia, la que tenga la constante asintótica del error más pequeña convergerá más rápidamente que la otra. El recíproco de esta proposición no es cierto, pues dos sucesiones del mismo orden y con la misma constante asintótica de error no tienen por qué converger con la misma velocidad. En [145], se muestran un ejemplo en el que verifican la no reciprocidad de la proposición. En el mismo encontramos que dadas $y_n = a^n/n^2$ y $x_n = a^n/n$ con $0 < a \leq 1$, el orden de convergencia es 1 y a es, para ambos casos, la constante asintótica del error, pero $\{y_n\}$ converge más rápidamente a cero. Para verificar su demostración, nótese que al sustituir cualquier valor de $0 < a < 1$ y $n \geq 1$ se obtiene que

$$0 < y_n < x_n, \quad \forall n \geq 1.$$

Para una mayor profundidad de los aspectos concernientes a orden y velocidad de convergencia, ver a [10], [47], [145] y [163] entre otros.

En la determinación de las raíces de una ecuación no lineal, el aumento del orden de convergencia permite reducir el número de iteraciones. Pero el aumento en el orden requiere una mayor cantidad de evaluaciones de funciones y sus derivadas, aumentando el tiempo de cálculo en cada iteración. Esto podría significar que el aumento del orden de convergencia del método lleva implícito que el mismo sea menos eficiente, pero no siempre es así ya que en realidad lo que se establece es un equilibrio entre el orden de convergencia y el coste operacional del algoritmo.

Desde el punto de vista práctico, lo realmente importante no es ni el orden de convergencia de una sucesión ni el número de iteraciones necesarias para llegar a la precisión deseada, sino el tiempo que emplea el computador en alcanzar tal precisión, por lo tanto el coste del cómputo de una iteración es fundamental. En otras palabras, el índice de eficiencia del método iterativo es básico. Veamos su definición:

Definición 2.6 (Índice de eficiencia computacional de un método). *De acuerdo con [93], existen dos tipos de índice de eficiencia:*

- i) *El de aritmética con precisión fija y pequeña. El índice de eficiencia viene definido por la expresión*

$$IEC(p, d) = p^{1/d}, \quad (2.3)$$

donde $p > 1$ es el orden de convergencia del método y d es el número de evaluaciones de la función f o de sus derivadas.

ii) El de aritmética con múltiple precisión. En este caso el índice de eficiencia del método se denota por la fórmula

$$IEC(d, p, \alpha) = p^{\frac{1}{\mathcal{C}(d, p, \alpha)}} = p^{\frac{1-p^{-\alpha}}{d}},$$

Donde p y d tienen el mismo significado que en la ecuación (2.3) y

$$\begin{aligned} \mathcal{C}(d, p, \alpha) &= \frac{d}{1-p^{-\alpha}} \quad \text{es el coste total o tiempo computacional de evaluar a } f, \text{ en donde.} \\ \alpha &= \log_{k+1}(2k+1), \quad \text{para } k \geq 1. \end{aligned}$$

y k es una constante que indica el número de veces que se debe evaluar la función f .

Computacionalmente hablando, cuanto mayor sea el índice de eficiencia mejor es el método iterativo. Es importante señalar que una mejora en el orden de convergencia de un método iterativo, no siempre implica una mejora en la eficiencia computacional del mismo. Generalmente estas mejoras traen como consecuencia el incremento del coste de evaluar la función, en perjuicio de su eficiencia.

Existen diferentes formas de clasificar los procesos iterativos, dependiendo de los aspectos que se vayan a tener en cuenta. Por ejemplo, en función del número de elementos de la sucesión $\{x_n\}$ que necesitamos para calcular x_{n+1} , los podemos catalogar en dos grandes bloques

(i) **Métodos de un punto.** Aquellos en los que la sucesión de iteración es de la forma

$$x_{n+1} = g(x_n).$$

Entre estos se encuentran el método de: Newton-Raphson, Halley, Super-Halley, Chebyshev, Danby, etc.

(ii) **Métodos multipunto.** En estos la función de iteración es de la forma

$$x_{n+1} = g(x_n, x_{n-1}, \dots, x_{n-p+1}), \quad \text{con } p \geq 1.$$

Es decir, son métodos en los que el n -ésimo término viene dado en función de los p términos anteriores. A este grupo pertenecen los métodos de: la Secante, Bisección, Yun-Petković, Falsa Posición o Regula Falsi, etc.

Pero ésta no es una clasificación única. Por ejemplo, Aubanell y Delshams [17] presentan una clasificación de los métodos iterativos atendiendo a dos rasgos que los caracterizan:

1. La herramienta que se emplea en el desarrollo de la función. De acuerdo a la herramienta que se emplea en el desarrollo de la función, los métodos iterativos se clasifican en:

- **Métodos de Taylor.** Son aquéllos que utilizan el desarrollo de Taylor, es decir, calculan a x_{n+1} a partir del conocimiento de la función y de sus derivadas en un punto x_n . Ejemplos: método de Euler, método de Punto medio.
- **Métodos de Interpolación.** En éstos se calcula x_{n+1} a partir del conocimiento de la función en x_n, x_{n-1}, \dots . Ejemplos: método de Lagrange, método de Hermite.

2. La función que se desarrolla o se interpola. Por la función que se desarrolla o se interpola, es decir, si se emplea $g = f^{-1}$, que es equivalente a buscar a x^* tal que $f(x^*) = 0$ a buscar x^* tal que $g(0) = x^*$, es posible dar dos enfoques a los métodos que se aplicaran en la solución del problema según se emplee el desarrollo de Taylor o la interpolación de la función f ó g :

- **Métodos Directos.** Si usamos la función f . Ejemplos: método de Newton-Raphson, Método de Halley.
- **Métodos Inversos.** Si empleamos la función $g = f^{-1}$. Ejemplo: método de Chebyshev.

Además, teniendo en cuenta el comportamiento especial para raíces múltiples, consideraremos un tercer bloque

(iii) **Métodos para raíces múltiples.**

Es conocido que algunos métodos, i.e el método de Newton-Raphson y el método de la Secante, disminuyen su orden de convergencia cuando la ecuación que intentan resolver tienen raíces múltiples. Para evitar esta consecuencia se han creado una serie de métodos que introducen en el método original una constante de multiplicidad que permite aumentar nuevamente el orden de convergencia.

Entre los métodos para encontrar raíces múltiples tenemos el de Newton-Raphson modificado, Newton Relajado, King-Van de Vel, etc.

2.2. El método de Newton

El método de Newton, también conocido como método de Newton-Raphson, es un proceso iterativo que, en condiciones adecuadas, permite aproximar una solución x^* de una ecuación no lineal

$$f(x) = 0. \tag{2.4}$$

Partiendo de una aproximación inicial x_0 de la raíz x^* , el método de Newton-Raphson define una sucesión de la forma

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, 3, \dots \quad (2.5)$$

En esta sección, analizaremos el origen y desarrollo histórico de este método, distintas técnicas que dan lugar a la fórmula (2.5) y los resultados que garantizan la convergencia del método a la raíz buscada.

2.2.1. Notas históricas sobre el método de Newton

La idea general de resolver una ecuación mediante un algoritmo iterativo o estimando la solución mediante la adición de un término corrector ha sido empleada por muchas culturas milenarias. Por ejemplo, en la antigua Grecia y Babilonia se emplearon métodos parecidos a los utilizados por Newton, para aproximar la solución de radicales. Métodos algebraicamente equivalente al de Newton también fueron aplicados en el siglo *XII*, por el algebrista Sharaf al-Din al-Tusi. En el siglo *XV*, el matemático árabe Al-Kashi empleó un algoritmo numérico para encontrar las raíces p -ésimas de N en la ecuación $x^p = N$. También Henry Briggs, en su obra *Trigonometria Britannica* publicada en 1633, usó un método parecido al de Newton.

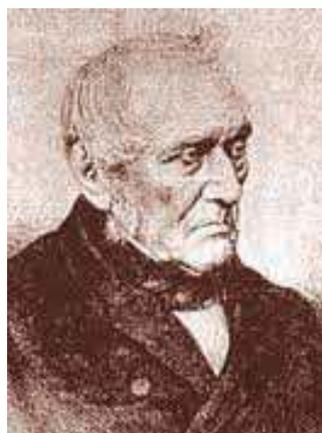


Figura 2.1: Sir Isaac Newton (1643 – 1727) y Thomas Simpson (1710 – 1761).

Leyendo a Tjalling J. Ypma [195] encontramos que la primera formulación del método de Newton-Raphson apareció en un tratado de Newton titulado *De analysi per aequationes numero terminorum infinitas*, el cual fue, probablemente, presentado a mediados del año 1669 y publicado en 1711 por William Jones. Esta primera formulación es considerablemente diferente a la que conocemos hoy en día y estaba definida para resolver sólo ecuaciones polinómicas. Además la forma de realizar los cálculos son mucho más tediosos que en la formulación actual. En 1671 Newton dio a conocer un tratado, con una versión ligeramente más

comprensible, titulado *De methodis fluxionum et serierum infinitarum*, la cual fue publicada por John Colson en 1736.

El proceso descrito por Newton requiere un cálculo explícito de una sucesión polinómica de la forma g_1, g_2, \dots , que lo convierte en un método muy laborioso. También se observa que la estimación final x^* es calculada al final del proceso como $x^* = x_0 + c_0 + c_1 + \dots$ en vez de hacer las estimaciones sucesivas de x_n , lo cual evidencia que Newton no utilizó una técnica recursiva.

La primera vez que Newton publica su método aplicado a ecuaciones no polinómicas fue en la segunda y tercera edición de su obra *Philosophiae Naturalis Principia Mathematica*, la cual fue presentada en 1687 en Londres. En ellas describe las técnicas para resolver la ecuación de Kepler

$$f(E) = E - e \operatorname{sen} E - M, \quad (2.6)$$

donde M es la anomalía media, E es la anomalía excéntrica y e es la excentricidad de la elipse.

La segunda formulación del método de Newton se debe a Joseph Raphson, quien en 1690 publicó un tratado titulado *Analysis aequationum universalis* en el cual se presentó un método nuevo para resolver ecuaciones polinómicas. Una segunda edición de dicho tratado se publicó como libro en 1697, con un apéndice que hace referencia a Newton. Se comprueba que los dos métodos son matemáticamente equivalentes, la diferencia es que Raphson no necesitó generar una sucesión de polinomios, como lo hacía Newton.

Raphson consideró una ecuación de la forma $a^3 - ba - c = 0$, donde a es una variable. Prosigue indicando que si g es una estimación de la solución x^* , entonces, una mejor aproximación se consigue haciendo a $g + x$, donde

$$x = \frac{c + bg - g^3}{3g^2 - b}.$$

Formalmente esto es de la forma

$$g + x = g - \frac{f(g)}{f'(g)}, \quad \text{con } f(a) = a^3 - ba - c,$$

pero como en el caso de Newton, Raphson también procedió de manera puramente algebraica, aunque prefirió las reglas del cálculo al utilizar la forma de derivada, pues en cada ejemplo escribió expresiones, tales como $f(x)$ y $f'(x)$, para referirse a polinomios completos.

Los contactos entre Newton y Raphson parece que fueron muy limitados, aunque parece que Newton aprovechó las circunstancias de la muerte de Raphson para adjuntar un apéndice al último libro de Raphson, titulado *Historia fluxionum*. En el prefacio del tratado de 1690, Raphson se refiere a los trabajos de Newton, indicando que su método no es único.

La tercera versión del método de Newton se debe a Thomas Simpson y la misma está descrita en un libro titulado *Essays on Several Curious and Useful Subjects in Speculative*

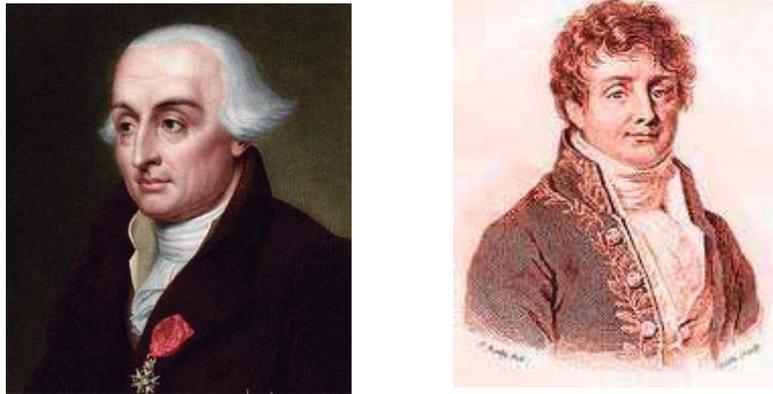


Figura 2.2: Joseph Louis Lagrange (1736 – 1813) y Joseph Fourier (1768 – 1830).

and Mixed Mathematicks, publicado en 1740 en Londres. En el ensayo número seis, Simpson describe un nuevo método para la solución de ecuaciones numéricas, sin hacer referencia a los trabajos de ningún predecesor. En el prefacio de dicho ensayo, contrasta las técnicas de su método, las cuales estaban basadas en el análisis matemático y cálculo numérico, con los métodos algebraicos de entonces. En éste se dan ejemplos de solución de ecuaciones polinómicas y no polinómicas, con detalles para resolver ecuaciones en una o dos variables mediante instrucciones que involucran el signo \dot{x} , el cual es equivalente a dx/dt . Además aparece la primera formulación de un método iterativo para resolver, en general, ecuaciones no lineales basados en el cálculo numérico. Finalmente, fue Lagrange, en 1798, quien introdujo el método de Newton-Raphson con la simbología que se emplea actualmente. La obra en que aparece esta notación moderna se titula *Sur la méthode d'Approximation donnée par Newton* y en la misma no hace referencia a los trabajos de Simpson, aunque si menciona a Newton y Raphson.

Para un análisis más profundo del desarrollo histórico del método de Newton-Raphson, véase a [195].

2.2.2. Descripción del método de Newton

Existen varias formas de deducir la expresión (2.5). A continuación presentamos, de forma resumida, algunas de ellas.

A partir del desarrollo de Taylor

La técnica que explicaremos a continuación aparece en varios libros clásicos de Análisis Numérico, entre ellos [47], [75], [117] y [129].

Sea x_0 una aproximación de x^* , la raíz de $f(x) = 0$. Buscamos un término corrector h de

forma que $x + h$ sea la raíz buscada, es decir,

$$f(x_0 + h) = f(x^*) = 0.$$

Si desarrollamos por la serie de Taylor $f(x_0 + h)$ se obtiene:

$$0 = f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!}f''(x_0) + \dots \quad (2.7)$$

La solución exacta de esta ecuación nos daría el término corrector h tal que $x_0 + h = x^*$.

Si en lugar de considerar la ecuación (2.7), tomamos su linealizada

$$f(x_0) + hf'(x_0) = 0,$$

obtenemos de forma sencilla su solución

$$h = -\frac{f(x_0)}{f'(x_0)}.$$

Si bien este valor no es la solución exacta de (2.7), en condiciones adecuadas podría ser una buena aproximación. De esta forma

$$x_0 + h = x_0 - \frac{f(x_0)}{f'(x_0)},$$

sería una mejor aproximación a x^* que la dada inicialmente por x_0 .

Si denotamos por

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

y reiteramos el proceso obtenemos el método de Newton-Raphson: (2.5).

Construcción geométrica: método de la tangente

Esta técnica aparece en una inmensa bibliografía de Métodos Matemáticos y Análisis Numérico, entre los que podemos citar los siguientes: [31], [34], [46], [133], [146], [152], [153] y muchísimos más.

La técnica de iteración de este método consiste en tomar un punto inicial x_0 , véase la Figura (2.3), para luego trazar por el punto $(x_0, f(x_0))$ la tangente geométrica a la gráfica de la función $f(x)$. Luego, la intersección de dicha tangente con el eje X nos da un nuevo punto x_1 , por el cual trazamos una nueva recta tangente a la gráfica de la función, pero esta vez por el punto $(x_1, f(x_1))$. En condiciones adecuadas la sucesión así obtenida

$$x_0, x_1, x_2, \dots, x_{n-1}, x_n, \dots$$

tiende a la raíz exacta x^* .

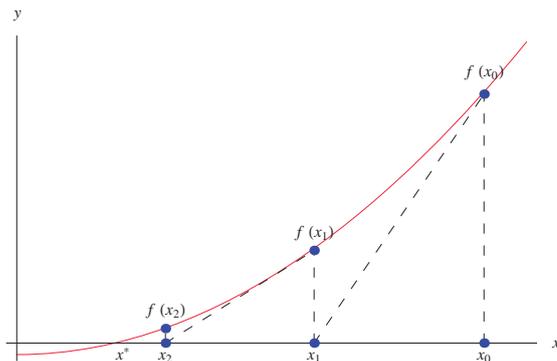


Figura 2.3: Interpretación geométrica del método de Newton-Raphson

Por la ecuación de la recta tangente obtenemos

$$y - f(x_0) = f'(x_0)(x_1 - x_0).$$

Luego, en la intersección con el eje X , ($y = 0$), se sigue

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Reiterando el proceso se llega a la función iterativa (2.5) para el punto x_n .

Iteración del punto fijo

El método de Newton-Raphson puede ser obtenido a partir de del método de iteración del punto fijo. Para ello escribiremos la ecuación $f(x) = 0$ como

$$x = x - f(x), \quad (2.8)$$

luego introducimos en (2.8) un parámetro λ , con lo cual

$$x + \lambda x = \lambda x + x - f(x),$$

de donde

$$x = \frac{\lambda}{\lambda + 1}x + \frac{1}{1 + \lambda}(x - f(x)) = G(x). \quad (2.9)$$

Si x^* es una raíz de $f(x)$, x^* es un punto fijo de $G(x)$.

Si imponemos que $G'(x^*) = 0$ es un punto fijo superatractor, entonces el parámetro λ debe cumplir

$$\frac{\lambda}{1 + \lambda} + \frac{1}{1 + \lambda}(1 - f'(x^*)) = 0,$$

de donde se obtiene

$$\lambda = f'(x^*) - 1. \quad (2.10)$$

Sustituyendo (2.10) en (2.9) y utilizando la iteración del punto fijo, sigue

$$\begin{aligned} x_{n+1} &= G(x_n) \\ &= \frac{f'(x^*) - 1}{f'(x^*)}x_n + \frac{1}{f'(x^*)}(x_n - f(x_n)) \\ &= x_n - \frac{f(x_n)}{f'(x^*)}. \end{aligned}$$

Pero como x^* es, en principio desconocido, podemos usar la aproximación definida por $x^* \sim x_n$ para llegar al siguiente método iterativo.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

que es, de nuevo, el método de Newton-Raphson.

2.2.3. Convergencia del método de Newton

En los apartados anteriores, hemos visto como definir una sucesión

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

con el objetivo de aproximar la solución de x^* de la ecuación $f(x) = 0$.

Tenemos que ser conscientes de que la convergencia de $\{x_n\}$ hacia x^* no siempre va a suceder. Para que esto ocurra, se tienen que dar una serie de condiciones sobre la función f , el punto de partida x_0 o la raíz x^* . En concreto, distinguiremos 3 tipos de resultados de convergencia:

1. Local: Se dan condiciones sobre la raíz x^* .
2. Semilocal: Se dan condiciones sobre el punto de partida x_0 .
3. Global: Se dan condiciones sobre un intervalo.

Existen una gran cantidad de publicaciones con diversos resultados de convergencia para el método de Newton-Raphson (véanse [10], [17], [77], [133], [137],[141], [152] y [180]). No es nuestro objetivo dar un listado exhaustivo de todos ellos. No obstante, a modo de ejemplos, expondremos uno o dos resultados de cada tipo de convergencia.

Convergencia local del método de Newton

Existen diferentes técnicas que permiten estudiar el orden de convergencia de un proceso iterativo. Así por ejemplo, en [60] se muestra un teorema que permite garantizar la convergencia del método de Newton en términos del denominado grado de convexidad logarítmico de una función:

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}. \quad (2.11)$$

Leyendo a [98] encontramos que el grado de convexidad logarítmico de una función es una medida puntual de la convexidad, dada por la resistencia de una función a ser «concavizada» por un operador logarítmico. En otras palabras, es el número de veces que hay que aplicar un operador logarítmico a una función convexa para obtener como resultado una función cóncava.

Además, el grado de convexidad logarítmico está relacionado con la velocidad de convergencia del método de Newton-Raphson. No en vano se tiene que

$$N'_f(x) = L_f(x),$$

es decir, el grado de convexidad logarítmico es la derivada de la función de iteración del método de Newton $N_f(x)$ definida por (2.5). En consecuencia, la expresión (2.11) es muy utilizada en las demostraciones de la convergencia de muchos métodos iterativos, tales como el método de Halley y el método de Chebyshev, así como en libros y artículos de Análisis Numérico, entre los que se pueden citar a [5], [7], [9], [67], [68], [70], [83], [98], [99] y [154].

Una técnica diferente consiste en utilizar el teorema de Schröder, introducido en el Teorema 2.1. Dicho teorema puede ser considerado ocasionalmente como una definición del orden de convergencia de un proceso iterativo para encontrar las raíces de una ecuación, en lugar de la definición . Esto es sugerido por algunos autores como [105] o [123]. El propio Traub comenta esta posibilidad (véase [180, p. 10]).

En adelante, tendremos en cuenta la siguiente notación introducida por Traub [180] para funciones suficientemente derivables:

$$A_j = \frac{f^{(j)}(x^*)}{j!f'(x^*)}. \quad (2.12)$$

Además denotamos

$$e_n = x_n - x^* \quad (2.13)$$

a los términos del error y

$$O_n = O(e_n^p). \quad (2.14)$$

Con estas notaciones, si un método iterativo tiene orden p se satisface la siguiente ecuación del error:

$$e_{n+1} = Ce_n^p + O_{n+1}. \quad (2.15)$$

Utilizando el Teorema de Schröder (véase el teorema 2.1), podemos probar que el método de Newton tiene convergencia cuadrática en un entorno de una raíz simple x^* de $f(x) = 0$. Usando las notaciones introducidas en (2.12), (2.13), (2.14), (2.15), podemos formular el siguiente resultado.

Teorema 2.2. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz simple de $f(x) = 0$, ($f(x^*) = 0$ y $f'(x^*) \neq 0$). Supongamos que f es tres veces*

derivable en I . Entonces el método de Newton definido por $x_{n+1} = N_f(x_n)$,

$$N_f(x) = x - \frac{f(x)}{f'(x)},$$

tiene al menos convergencia cuadrática localmente. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = A_2 e_n^2 + O_3.$$

donde la constante A_2 , definida en (2.12), es la constante de error asintótico.

Demostración. (Usando el Teorema de Schröder) Notemos que se necesita que f''' sea continua para poder satisfacer la hipótesis del Teorema de Schröder de que N_f'' sea continua. Un simple cálculo nos permite comprobar que

$$N_f(x^*) = x^*, \quad N_f'(x^*) = 0, \quad N_f''(x^*) = \frac{f''(x^*)}{f'(x^*)} \neq 0.$$

Además tenemos que

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^2} = \frac{f''(x^*)}{2f'(x^*)} = A_2,$$

□

Demostración. (Usando desarrollos de Taylor) Podemos obtener las conclusiones del teorema siguiendo una técnica diferente. En concreto, desarrollamos $f(x_n)$ y $f'(x_n)$ en serie de Taylor y truncamos en la potencia N -ésima. Obtenemos así:

$$f(x_n) = f(x_n - x^* + x^*) = f(x^* + e_n) = f'(x^*) \left(e_n + \sum_{i=2}^N A_i e_n^i \right) + O_{N+1}$$

$$f'(x_n) = f'(x^*) \left(1 + \sum_{i=2}^N i A_i e_n^{i-1} \right) + O_N,$$

donde A_i , e_n y O_N se han introducido en (2.12), (2.13) y (2.14) respectivamente.

Además, se tiene que

$$\frac{f(x_n)}{f'(x_n)} = e_n - A_2 e_n^2 + O_3.$$

Por lo tanto,

$$e_{n+1} = x_{n+1} - x^* = x_n - x^* - \frac{f(x_n)}{f'(x_n)} = e_n - (e_n - A_2 e_n^2 + O_3) = A_2 e_n^2 + O_3,$$

con lo que hemos demostrado el resultado. □

Las dos técnicas de demostración expuestas anteriormente han sido empleadas por diversos autores en el estudio de la convergencia local de procesos iterativos. Por ejemplo podemos citar los siguientes: [12], [20], [39], [59], [88], [103], [104], [108], [114] y [188],

Es bien conocido [180] que cuando se trata de aproximar una raíz múltiple, el orden de convergencia del método de Newton es solamente lineal. Diremos que x^* es una raíz múltiple o un cero de multiplicidad m de f si

$$f(x^*) = f'(x^*) = \cdots = f^{(m-1)}(x^*) = 0 \quad \text{y} \quad f^{(m)}(x^*) \neq 0. \quad (2.16)$$

En este caso, se puede escribir

$$\begin{aligned} f(x) &= (x - x^*)^m g(x), \quad \text{con } g(x^*) \neq 0, \\ f'(x) &= (x - x^*)^{m-1} (mg(x) + (x - x^*)g'(x)), \\ \frac{f(x)}{f'(x)} &= (x - x^*) \frac{g(x)}{mg(x) + (x - x^*)g'(x)}. \end{aligned}$$

Teniendo en cuenta estas consideraciones, podemos formular el siguiente resultado.

Teorema 2.3. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz de multiplicidad $m \geq 2$ de $f(x) = 0$. Supongamos que f es $m + 1$ veces derivable en I . Para $1 \leq q \leq m$ consideramos los procesos iterativos*

$$x_{n+1} = x_n - q \frac{f(x_n)}{f'(x_n)}. \quad (2.17)$$

Entonces:

1. Si $q < m$, el proceso anterior tiene convergencia lineal a la raíz. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = C e_n + O_2, \quad C = 1 - \frac{q}{m}.$$

2. Si $q = m$, el proceso anterior tiene convergencia cuadrática a la raíz. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = C e_n + O_2, \quad C = \frac{A_{m+1}}{m! A_m}.$$

Demostración. Para valores de x cercanos a x^* podemos escribir

$$\frac{f(x)}{f'(x)} = (x - x^*) \frac{1 + \varepsilon(x)}{m}, \quad \lim_{x \rightarrow x^*} \varepsilon(x) = 0.$$

Para este método definido en (2.17) se tiene que

$$x_{n+1} - x^* = (x_n - x^*) \left(1 - \frac{q}{m}\right) - \frac{q}{m} (x_n - x^*) \varepsilon(x_n). \quad (2.18)$$

Por lo tanto, si $q < m$,

$$\lim_{n \rightarrow \infty} \frac{(x_{n+1} - x^*)}{(x_n - x^*)} = 1 - \frac{q}{m} \neq 0.$$

Por consiguiente, se ha perdido la convergencia cuadrática del proceso, pudiéndose garantizar únicamente la convergencia lineal.

Sin embargo, si se toma $q = m$, el valor de la multiplicidad de la raíz y se supone que $f^{(m+1)}(\alpha)$ está definido, entonces el método

$$x_{n+1} = x_n - m \frac{f(x_n)}{f'(x_n)}. \quad (2.19)$$

recupera la convergencia cuadrática. Es más, en este caso,

$$\lim_{n \rightarrow \infty} \frac{(x_{n+1} - \alpha)}{(x_n - \alpha)^2} = \frac{f^{(m+1)}(\alpha)}{(m+1)!f^{(m)}(\alpha)}.$$

Llegados a este punto, conviene destacar que si no se puede asegurar la existencia de $f^{(m+1)}(\alpha)$, únicamente se podría garantizar la convergencia superlineal del método correspondiente. En efecto, de la fórmula (2.18) se deduce

$$\lim_{n \rightarrow \infty} \frac{(x_{n+1} - \alpha)}{(x_n - \alpha)} = - \lim_{n \rightarrow \infty} \frac{q}{m} (x_n - \alpha) \varepsilon(x_n) = 0.$$

□

Corolario 2.4. *El método de Newton para aproximar una raíz de multiplicidad $m > 1$ tiene convergencia lineal.*

Demostración. Basta tener en cuenta que el método de Newton se obtiene para $q = 1$ en (2.17) y aplicar el teorema anterior. □

Convergencia semilocal del método de Newton

En este trabajo demostraremos el teorema de Kantorovich y la α -teoría de Smale como una versión en \mathbb{R} que podrá ser extendida al caso general de sistemas de ecuaciones o ecuaciones definidas en espacios de Banach.

Históricamente hablando, en 1829, Cauchy presentó el primer resultado de convergencia semilocal para el método de Newton-Raphson. Los detalles de dicho teorema se encuentran en [32], mientras que en [192] aparece un resumen parecido al que presentamos a continuación.

Teorema 2.5 (Cauchy). *Sea $X = \mathbb{R}$, $F = f \in C^2$, $x_0 \in X$, $f'(x_0) \neq 0$, $\sigma_0 = -\frac{f(x_0)}{f'(x_0)}$, $\eta = |\sigma_0|$,*

$$I = \langle x_0, x_0 + 2\sigma_0 \rangle \equiv \begin{cases} [x_0, x_0 + 2\sigma_0], & \text{si } \sigma_0 \geq 0, \\ [x_0 + 2\sigma_0, x_0], & \text{si } \sigma_0 < 0. \end{cases}$$

y $|f''(x_0)| \leq K \in I$. Entonces se cumplen los siguientes resultados:

i) Si

$$2K\eta < |f'(x_0)|,$$

entonces $F(x)$ tiene a x^* como solución única en I .

ii) Si $m \in I$ y $|f'(x_0)| \geq m$,

$$2K\eta < m,$$

entonces, la sucesión del método de Newton definida por (2.5), iniciando en x_0 , cumple que

$$|x_{k+1} - x_k| \leq \frac{K}{2m} |x_k - x_{k-1}|^2, \quad k \geq 1$$

y

$$x^* \in \langle x_k, x_k + 2\sigma_k \rangle,$$

donde

$$\sigma_k = -\frac{f(x_k)}{f'(x_k)} = x_{k+1} - x_k,$$

así que

$$\begin{aligned} |x^* - x_k| &\leq 2|x_{k+1} - x_k|, \quad k \geq 0, \\ &\leq \frac{K}{m} |x_k - x_{k-1}|^2, \quad k \geq 1, \\ &\leq 2\eta \left(\frac{K\eta}{2m}\right)^{2^k - 1}, \quad k \geq 0. \end{aligned}$$



Figura 2.4: Augustin Louis Cauchy (1789 – 1857) y Leonid V. Kantorovich (1912 – 1986).

Este resultado, inicialmente establecido para funciones reales de variable real, fue generalizado años más tarde por Kantorovich en su célebre teorema enunciado para operadores definidos en espacios de Banach.

Leyendo a [83], encontramos que fue a finales del 1940 cuando L. V. Kantorovich y L. B. Rall introdujeron el método de Newton-Raphson en estos tipos de espacios. Su planteamiento establece que si F es un operador suficientemente diferenciable definido entre dos espacios de Banach X e Y , a partir de un punto $x_0 \in X$ se construye la sucesión de Newton-Raphson

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad (2.20)$$

donde $F'(x_n)^{-1}$ es el inverso del operador lineal $F'(x_n)$.

A partir de ciertas condiciones, Kantorovich demostró que la sucesión $\{x_n\}$, generada a partir de (2.20) está bien definida y es convergente a una solución de $F(x) = 0$. Para una visión mas profunda del teorema de Kantorovich se puede consultar a [109], [110], [111], [141] y [143].

Teorema 2.6 (Teorema de Kantorovich). *Sea $f : I \rightarrow \mathbb{R}$, en donde I es un intervalo abierto en \mathbb{R} , una función diferenciable en I . Asumamos, sin pérdida de generalidad, que $f(x_0) \neq 0$ y que*

i) $x_0 \in I$.

ii) $f'(x_0) \neq 0$.

iii) $\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta$.

iv) $\left| \frac{f'(x) - f'(y)}{f'(x_0)} \right| \leq \gamma|x - y|, \quad \forall x, y \in I$.

v) $h = \gamma\beta \leq 1/2$.

vi) $t^* = \frac{1 - \sqrt{1 - 2h}}{2\beta}$ es tal que $(x_0 - t^*, x_0 + t^*) \subset I$.

Entonces el método de Newton-Raphson

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0, \quad (2.21)$$

está bien definido y $\{x_n\}_{n=0}^{\infty}$ converge a $x^* \in I$, donde $f(x^*) = 0$.

Demostración. Para encontrar una sucesión que mayorice a $\{x_n\}_{n=0}^{\infty}$ vamos a considerar el polinomio

$$p(t) = \frac{1}{2}\gamma t^2 - t + \beta, \quad (2.22)$$

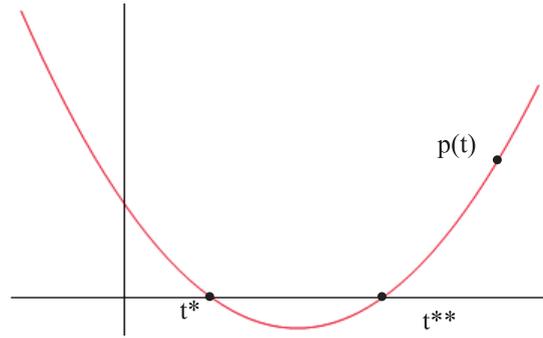


Figura 2.5: Gráfica del polinomio mayorizante $p(t) = \frac{1}{2}\gamma t^2 - t + \beta$

cuyas raíces son

$$t^*, t^{**} = \frac{1 \pm \sqrt{1 - 2\gamma\beta}}{\gamma},$$

tal y como se muestra en la gráfica 2.5

Notemos que (2.22) tendrá solución en \mathbb{R} cuando $1 - 2\gamma\beta \geq 0$, de donde $\gamma\beta \leq \frac{1}{2}$. Esto explica la condición v).

Aplicando el método de Newton-Raphson a (2.22) se sigue

$$\begin{cases} t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)} = t_n - \left[\frac{-\beta + t_n - \frac{1}{2}\gamma t_n^2}{1 - \gamma t_n} \right] = \frac{-\frac{1}{2}\gamma t_n^2 + \beta}{1 - \gamma t_n}. \\ t_0 = 0. \end{cases} \quad (2.23)$$

Notemos que el polinomio $p(t)$ cumple las condiciones de Fourier (Teorema 2.8). Así, tomando a $t_0 = 0$ se genera una sucesión monótona creciente cuyo límite es t^* . Ahora, por inducción probaremos que

$$|x_{n+1} - x_n| \leq t_{n+1} - t_n, \quad \forall n \geq 0.$$

Sustituyendo a $n = 0$ en (2.23) se sigue

$$t_1 - t_0 = -\frac{p(t_0)}{p'(t_0)} = -\frac{p(0)}{p'(0)} = \beta, \quad t_1 = \beta.$$

Sustituyendo $n = 0$ en (2.21) y aplicando *iii*)

$$|x_1 - x_0| \leq \beta = t_1 - t_0.$$

En consecuencia $x_1 \in I$ ya que $|x_1 - x_0| < t_1 < t^*$.

Supongamos que $x_0, x_1, \dots, x_m \in I$ y además que $|x_{i+1} - x_i| \leq t_{i+1} - t_i$ para $i = 0, 1, \dots, m-1$, demostraremos que $|x_{m+1} - x_m| \leq t_{m+1} - t_m$.

En efecto, como

$$\begin{aligned} |x_m - x_0| &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \dots + |x_1 - x_0|. \\ &\leq (t_m - t_{m-1}) + (t_{m-1} - t_{m-2}) + \dots + (t_1 - t_0). \\ &= t_m - t_0 = t_m < t^*, \end{aligned}$$

se tiene que $x_m \in I$.

Por otra parte, notemos que

$$1 - \gamma|x_m - x_0| \geq 1 - \gamma t_m \Rightarrow \frac{1}{1 - \gamma|x_m - x_0|} \leq \frac{1}{1 - \gamma t_m}.$$

Además

$$|x_{m+1} - x_m| = \left| \frac{f(x_m)}{f'(x_m)} \right| = \left| \frac{f(x_m)}{f'(x_0)} \right| \left| \frac{f'(x_0)}{f'(x_m)} \right|. \quad (2.24)$$

Aplicando la condición *iv*) se sigue

$$\left| \frac{f'(x_m) - f'(x_0)}{f'(x_0)} \right| \leq \gamma|x_m - x_0| \Rightarrow \left| 1 - \frac{f'(x_m)}{f'(x_0)} \right| \leq \gamma|x_m - x_0| \leq \gamma(t_m - t_0) < \gamma t^* < 1.$$

Luego

$$-\gamma|x_m - x_0| \leq 1 - \frac{f'(x_m)}{f'(x_0)} \leq \gamma|x_m - x_0|.$$

Entonces

$$-\gamma|x_m - x_0| - 1 \leq -\frac{f'(x_m)}{f'(x_0)} \leq \gamma|x_m - x_0| - 1.$$

Luego

$$-1 < 1 - \frac{f'(x_m)}{f'(x_0)} \Rightarrow -\frac{f'(x_m)}{f'(x_0)} < 0.$$

Así

$$\left| \frac{f'(x_m)}{f'(x_0)} \right| = \frac{f'(x_m)}{f'(x_0)} \geq 1 - \gamma|x_m - x_0|.$$

De donde

$$\left| \frac{f'(x_0)}{f'(x_m)} \right| \leq \frac{1}{1 - \gamma|x_m - x_0|}. \quad (2.25)$$

Ahora bien, de (2.5) se deduce

$$f(x_m) = f(x_m) - f(x_{m-1}) - f'(x_{m-1})(x_m - x_{m-1}).$$

Luego

$$\left| \frac{f(x_m)}{f'(x_0)} \right| = \left| \frac{f(x_m) - f(x_{m-1}) - f'(x_{m-1})(x_m - x_{m-1})}{f'(x_0)} \right| = \left| \int_{x_{m-1}}^{x_m} \left[\frac{f'(t) - f'(x_{m-1})}{f'(x_0)} \right] dt \right|.$$

Haciendo el cambio de variable $t = x_{m-1} + s(x_m - x_{m-1})$ se sigue

$$\left| \frac{f(x_m)}{f'(x_0)} \right| \leq \int_0^1 \left| \frac{f'(x_{m-1} + s(x_m - x_{m-1})) - f'(x_{m-1})}{f'(x_0)} \right| |x_m - x_{m-1}| ds. \quad (2.26)$$

Aplicando la condición *iv*) a (2.26) se sigue

$$\left| \frac{f(x_m)}{f'(x_0)} \right| \leq \int_0^1 \gamma s |x_m - x_{m-1}|^2 ds = \gamma |x_m - x_{m-1}|^2 \int_0^1 s ds \leq \frac{1}{2} \gamma |x_m - x_{m-1}|^2. \quad (2.27)$$

Sustituyendo (2.25) y (2.27) en (2.24) se sigue

$$|x_{m+1} - x_m| \leq \frac{\frac{1}{2}\gamma(t_m - t_{m-1})^2}{1 - \gamma t_m}. \quad (2.28)$$

Por otra parte

$$t_{m+1} - t_m = -\frac{p(t_m)}{p'(t_m)}. \quad (2.29)$$

De (2.22) se obtiene

$$p'(t) = \gamma t - 1. \quad (2.30)$$

$$p''(t) = \gamma. \quad (2.31)$$

$$p^{(i)}(t) = 0, \quad i \geq 3.$$

Aplicando el desarrollo de Taylor a (2.22) se sigue

$$p(t_m) = p(t_{m-1}) + p'(t_{m-1})(t_m - t_{m-1}) + \frac{1}{2}p''(t_{m-1})(t_m - t_{m-1})^2 + 0. \quad (2.32)$$

Pero

$$t_m - t_{m-1} = -\frac{p(t_{m-1})}{p'(t_{m-1})}. \quad (2.33)$$

Luego, sustituyendo (2.31) y (2.33) en (2.32) se sigue

$$p(t_m) = \frac{1}{2}\gamma(t_m - t_{m-1})^2. \quad (2.34)$$

Sustituyendo (2.34) y (2.30) en (2.29)

$$t_{m+1} - t_m = \frac{\frac{1}{2}\gamma(t_m - t_{m-1})^2}{1 - \gamma t_m}. \quad (2.35)$$

Finalmente, sustituyendo (2.35) en (2.28)

$$|x_{m+1} - x_m| \leq t_{m+1} - t_m.$$

El proceso inductivo queda probado. Como $\{t_n\}_{n=0}^{\infty}$ es convergente, se sigue que $\{x_n\}_{n=0}^{\infty}$ también es convergente.

De (2.27) se deduce que cuando $m \rightarrow \infty$, entonces $f(x_m) \rightarrow 0$. Por tanto $\{x_m\}$ converge a x^* , la solución de $f(x^*) = 0$.

□

A finales de los años 80, surgió una nueva teoría sobre la convergencia semilocal de procesos iterativos. Esta teoría, conocida como α -teoría fue introducida por M. Kim y S. Smale [55]. Kim la introdujo en su tesis doctoral titulada *Computational Complexity of the Euler Type Algorithms for the Roots of Polynomials*, publicada en febrero del 1986 y en un artículo publicado en 1988 bajo el título *On approximate zeroes and root finding algorithms for a*

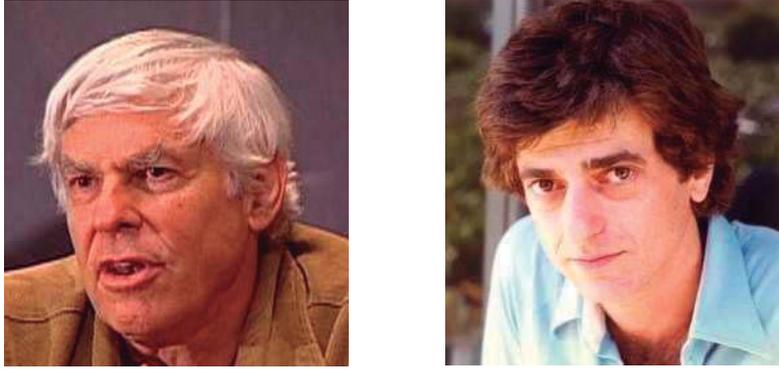


Figura 2.6: Stephen Smale (1930) y Mike Shub (1943).

complex polynomial. En ambos trabajos Kim aplicó la α -teoría a ecuaciones polinómicas de una variable. Por su parte, Smale en un libro titulado *New Directions in Pure, Applied and Computational Mathematics*, publicado en 1986, introdujo la α -teoría para sistemas de ecuaciones, los cuales tienen solución en base a tres invariantes de la forma

$$\begin{aligned}\beta(F, x) &= \|\mathbf{D}F(x)^{-1}F(x)\|, \\ \gamma(F, x) &= \sup \|\mathbf{D}F(x)^{-1} \frac{\mathbf{D}^k F(x)}{k!}\|^{1/(k-1)}, \quad k \geq 2, \\ \alpha(F, x) &= \beta(F, x)\gamma(F, x).\end{aligned}$$

donde F es un operador definido entre dos espacios de Banach. Para mayores detalles de este teorema en espacios de Banach, véase a [115] [116] y [170].

Teorema 2.7 (α -teoría de Smale). *Sea $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función continua y diferenciable en un intervalo D en el que se cumplen las siguientes condiciones:*

- i) $\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta$. $f'(x_0) \neq 0$.
- ii) $\frac{1}{k!} \left| \frac{f^{(k)}(x_0)}{f'(x_0)} \right| \leq \gamma^{k-1}$, para $k \geq 2$.
- iii) $\alpha = \beta\gamma \leq 3 - 2\sqrt{2}$.

Entonces el método de Newton-Raphson definido por (2.5) converge a x^* solución de $f(x) = 0$.

Demostración. Dado $\{x_n\}$, se demostrará que $|x_{n+1} - x_n| \leq t_{n+1} - t_n$, donde $\{t_n\}$ es una sucesión real obtenida al aplicar el método de Newton-Raphson a la función mayorante $\phi(t)$.

Consideremos la función

$$\phi(t) = \beta - t + \sum_{k \geq 2} \gamma^{k-1} t^k = \beta - t + \frac{\gamma t^2}{1 - \gamma t}, \quad \text{para } 0 \leq t < \frac{1}{\gamma}, \quad (2.36)$$

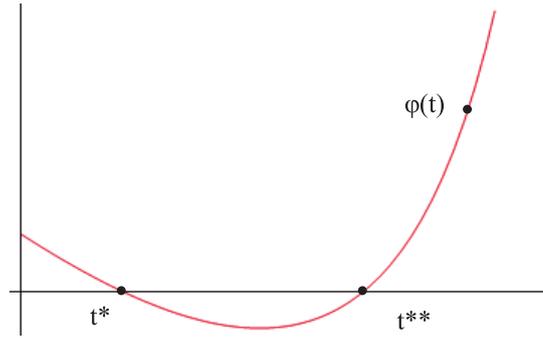


Figura 2.7: Gráfica de la función mayorizante $\phi(t) = \beta - t + \frac{\gamma t^2}{1-\gamma t}$

cuya gráfica es

Notemos que para $\phi(t) = 0$ se debe cumplir

$$2\gamma t^2 - t(\alpha + 1) + \beta = 0,$$

de donde

$$t^*, t^{**} = \frac{(\alpha + 1) \pm \sqrt{(\alpha + 1)^2 - 8\alpha}}{4\gamma}. \quad (2.37)$$

Luego, para que (2.37) tenga raíz en \mathbb{R} se debe cumplir

$$(\alpha + 1)^2 - 8\alpha \geq 0, \quad \text{de donde } \alpha^2 - 6\alpha + 1 \geq 0, \text{ por tanto } \alpha \leq 3 - 2\sqrt{2}.$$

Como vemos, este resultado se deduce directamente de la condición *iii*).

Ahora, sea $\{t_n\}$ la sucesión obtenida al aplicar el método de Newton-Raphson a (3.1), es decir:

$$\begin{cases} t_{n+1} = t_n - \frac{\phi(t_n)}{\phi'(t_n)} = \frac{\beta(1 - \gamma t_n)^2 - t_n(1 - \gamma t_n)^2 + \gamma t_n^2(1 - \gamma t_n)}{-1 + 2\gamma t_n(2 - \gamma t_n)}. \\ t_0 = 0. \end{cases} \quad (2.38)$$

Observemos que para $n = 0$, de (2.38) se obtiene

$$\begin{aligned} t_1 - t_0 &= \beta, \\ \left| \frac{\phi(t_0)}{\phi'(t_0)} \right| &\leq \beta, \\ \frac{\phi(t_0)}{\phi'(t_0)} &= \beta, \\ -\frac{\phi(0)}{\phi'(0)} &= \beta. \end{aligned}$$

Asimismo, si sustituimos $n = 0$ en (2.5) y aplicamos *i*) se sigue

$$|x_1 - x_0| \leq \left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta = t_1 - t_0. \quad (2.39)$$

$$\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta = -\frac{\phi(t_0)}{\phi'(t_0)}. \quad (2.40)$$

Por otra parte, notemos que

$$\phi'(t) = -1 + \sum_{k \geq 2} k \gamma^{k-1} t^{k-1} \Rightarrow \phi'(t_0) = \phi'(0) = -1. \quad (2.41)$$

Además

$$\begin{aligned} \phi''(t) &= \sum_{k \geq 2} k(k-1) \gamma^{k-1} t^{k-2}. \\ \phi'''(t) &= \sum_{k \geq 3} k(k-1)(k-2) \gamma^{k-1} t^{k-3}. \\ &\vdots \\ \phi^{(j)}(t) &= \sum_{k \geq j} k(k-1) \dots (k-j+1) \gamma^{k-1} t^{k-j}. \end{aligned}$$

Entonces

$$\phi^{(j)}(t_0) = \phi^{(j)}(0) = k! \gamma^{k-1}. \quad (2.42)$$

Luego, de (2.42) y (2.41) se obtiene

$$\gamma^{k-1} = \frac{\phi^{(j)}(t_0)}{k!} = -\frac{1}{k!} \frac{\phi^{(j)}(t_0)}{\phi'(t_0)}. \quad (2.43)$$

Aplicando *ii*) en (2.43) se sigue

$$\frac{1}{k!} \left| \frac{f^{(k)}(x_0)}{f'(x_0)} \right| \leq \gamma^{k-1} = -\frac{1}{k!} \frac{\phi^{(k)}(t_0)}{\phi'(t_0)}. \quad (2.44)$$

Probaremos que las desigualdades (2.39), (2.40) y (2.44) se cumplen siempre. En otras palabras, demostraremos que si (2.39), (2.40) y (2.44), se cumplen, entonces

$$\begin{aligned} \left| \frac{f(x_j)}{f'(x_j)} \right| &\leq -\frac{\phi(t_j)}{\phi'(t_j)}, \quad \text{para } j = 0, 1, 2, \dots \\ \frac{1}{k!} \left| \frac{f^{(k)}(x_j)}{f'(x_j)} \right| &\leq -\frac{1}{k!} \frac{\phi^{(k)}(t_j)}{\phi'(t_j)}, \quad \text{para } k \geq 2 \text{ y } j = 0, 1, 2, \dots \\ |x_{j+2} - x_{j+1}| &\leq t_{j+2} - t_{j+1}. \end{aligned} \quad (2.45)$$

La demostración será por inducción.

Para $j = 0$; (2.39), (2.40) y (2.44) son ciertas por hipótesis. Comprobaremos que si son ciertas para j , también lo serán para $j + 1$.

Aplicando el desarrollo de Taylor a $f(x_{j+1})$ se sigue

$$f(x_{j+1}) = f(x_j) + f'(x_j)(x_{j+1} - x_j) + \sum_{k \geq 2} \frac{1}{k!} f^{(k)}(x_j)(x_{j+1} - x_j)^k. \quad (2.46)$$

Pero

$$(x_{j+1} - x_j) = -\frac{f(x_j)}{f'(x_j)}. \quad (2.47)$$

Sustituyendo (2.47) en (2.46) se sigue

$$f(x_{j+1}) = f(x_j) - f'(x_j) \frac{f(x_j)}{f'(x_j)} + \sum_{k \geq 2} \frac{1}{k!} f^{(k)}(x_j) (x_{j+1} - x_j)^k.$$

Así

$$f(x_{j+1}) = \sum_{k \geq 2} \frac{1}{k!} f^{(k)}(x_j) (x_{j+1} - x_j)^k.$$

Luego

$$\left| \frac{f(x_{j+1})}{f'(x_j)} \right| \leq \sum_{k \geq 2} \frac{1}{k!} \left| \frac{f^{(k)}(x_j)}{f'(x_j)} \right| \left| \frac{f(x_j)}{f'(x_j)} \right|^k \leq \sum_{k \geq 2} \left[-\frac{1}{k!} \frac{\phi^{(k)}(t_j)}{\phi'(t_j)} \right] \left[-\frac{\phi(t_j)}{\phi'(t_j)} \right]^k. \quad (2.48)$$

Si ahora aplicamos el desarrollo de Taylor a $\phi(t_{j+1})$ se sigue

$$\phi(t_{j+1}) = \phi(t_j) + \phi'(t_j)(t_{j+1} - t_j) + \sum_{k \geq 2} \frac{1}{k!} \phi^{(k)}(t_j) (t_{j+1} - t_j)^k. \quad (2.49)$$

Pero

$$(t_{j+1} - t_j) = -\frac{\phi(t_j)}{\phi'(t_j)}. \quad (2.50)$$

Sustituyendo (2.50) en (2.49) sigue

$$\phi(t_{j+1}) = \phi(t_j) - \phi'(t_j) \frac{\phi(t_j)}{\phi'(t_j)} + \sum_{k \geq 2} \frac{1}{k!} \phi^{(k)}(t_j) (t_{j+1} - t_j)^k.$$

Así

$$\phi(t_{j+1}) = \sum_{k \geq 2} \frac{1}{k!} \phi^{(k)}(t_j) (t_{j+1} - t_j)^k.$$

Luego

$$-\frac{\phi(t_{j+1})}{\phi'(t_j)} = \sum_{k \geq 2} -\frac{1}{k!} \frac{\phi^{(k)}(t_j)}{\phi'(t_j)} (t_{j+1} - t_j)^k = \sum_{k \geq 2} \left[-\frac{1}{k!} \frac{\phi^{(k)}(t_j)}{\phi'(t_j)} \right] \left[-\frac{\phi(t_j)}{\phi'(t_j)} \right]^k.$$

Entonces, (2.48) se transforma en

$$\left| \frac{f(x_{j+1})}{f'(x_j)} \right| \leq -\frac{\phi(t_{j+1})}{\phi'(t_j)}. \quad (2.51)$$

Por otra parte, aplicando el desarrollo de Taylor a la derivada de (2.46) se sigue

$$f'(x_{j+1}) = \sum_{k \geq 0} \frac{1}{k!} f^{(k+1)}(x_j) (x_{j+1} - x_j)^k.$$

Luego

$$\left| 1 - \frac{f'(x_{j+1})}{f'(x_j)} \right| \leq \sum_{k \geq 1} \frac{1}{k} \left| \frac{f^{(k+1)}(x_j)}{f'(x_j)} \right| |x_{j+1} - x_j|^k \leq \sum_{k \geq 1} -(k+1) \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} \left[\frac{\phi(t_j)}{\phi'(t_j)} \right]^k. \quad (2.52)$$

Por otro lado

$$\phi'(t_{j+1}) = \phi'(t_j) + \sum_{k \geq 1} (k+1) \phi^{(k+1)}(t_j) \left[\frac{\phi(t_j)}{\phi'(t_j)} \right]^k.$$

De donde

$$\begin{aligned} \phi'(t_j) - \phi'(t_{j+1}) &= \sum_{k \geq 1} -(k+1) \phi^{(k+1)}(t_j) \left[\frac{\phi(t_j)}{\phi'(t_j)} \right]^k. \\ \frac{\phi'(t_j) - \phi'(t_{j+1})}{\phi'(t_j)} &= \sum_{k \geq 1} -(k+1) \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} \left[\frac{\phi(t_j)}{\phi'(t_j)} \right]^k. \\ 1 - \frac{\phi'(t_{j+1})}{\phi'(t_j)} &= \sum_{k \geq 1} -(k+1) \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} \left[\frac{\phi(t_j)}{\phi'(t_j)} \right]^k. \end{aligned} \quad (2.53)$$

Sustituyendo (2.53) en (2.52) se sigue

$$\left| 1 - \frac{f'(x_{j+1})}{f'(x_j)} \right| \leq 1 - \frac{\phi'(t_{j+1})}{\phi'(t_j)} < 1,$$

pues ϕ es de signo constante en $(0, t^*)$.

Así

$$\left| \frac{f'(x_j)}{f'(x_{j+1})} \right| \leq \frac{1}{1 - \left[1 - \frac{\phi'(t_{j+1})}{\phi'(t_j)} \right]} = \frac{\phi'(t_j)}{\phi'(t_{j+1})}. \quad (2.54)$$

Aplicando el desarrollo de Taylor a $f^{(k)}(x_{j+1})$ se sigue

$$f^{(k)}(x_{j+1}) = \sum_{i \geq 0} \frac{1}{i!} f^{(k+1)}(x_j) (x_{j+1} - x_j)^i.$$

Luego

$$\begin{aligned} \frac{1}{k!} \left| \frac{f^{(k)}(x_{j+1})}{f'(x_j)} \right| &\leq \sum_{i \geq 0} \frac{(k+1)!}{k! i!} \frac{1}{(k+1)!} \left| \frac{f^{(k+1)}(x_j)}{f'(x_j)} \right| |x_{j+1} - x_j|^i \leq \\ &\leq \sum_{i \geq 0} \binom{k+i}{i} \left(\frac{-1}{(k+1)!} \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} \right) (t_{j+1} - t_j)^i = \\ &= -\frac{1}{k!} \sum_{i \geq 0} \frac{1}{i!} \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} (t_{j+1} - t_j)^i = -\frac{1}{k!} \frac{\phi^{(k)}(t_{j+1})}{\phi'(t_j)}. \end{aligned}$$

Además

$$\frac{1}{k!} \frac{\phi^{(k)}(t_{j+1})}{\phi'(t_j)} = \frac{1}{k!} \sum_{i \geq 0} \frac{1}{i!} \frac{\phi^{(k+1)}(t_j)}{\phi'(t_j)} (t_{j+1} - t_j)^i.$$

Así que

$$\frac{1}{k!} \left| \frac{f^{(k)}(x_{j+1})}{f'(x_j)} \right| \leq -\frac{1}{k!} \frac{\phi^{(k)}(t_{j+1})}{\phi'(t_j)}. \quad (2.55)$$

Entonces

$$\left| \frac{f(x_{j+1})}{f'(x_{j+1})} \right| \leq \left| \frac{f'(x_j)}{f'(x_{j+1})} \right| \left| \frac{f(x_{j+1})}{f'(x_j)} \right|. \quad (2.56)$$

Sustituyendo (2.54) y (2.51) en (2.56) se sigue

$$\left| \frac{f(x_{j+1})}{f'(x_{j+1})} \right| \leq -\frac{\phi'(t_j)}{\phi'(t_{j+1})} \frac{\phi(t_{j+1})}{\phi'(t_j)} = -\frac{\phi(t_{j+1})}{\phi'(t_{j+1})}.$$

Asimismo

$$\frac{1}{k!} \left| \frac{f^{(k)}(x_{j+1})}{f'(x_{j+1})} \right| \leq \frac{1}{k!} \left| \frac{f^{(k)}(x_{j+1})}{f'(x_j)} \right| \left| \frac{f'(x_j)}{f'(x_{j+1})} \right|. \quad (2.57)$$

Sustituyendo (2.54) y (2.55) en (2.57) sigue

$$\frac{1}{k!} \left| \frac{f^{(k)}(x_{j+1})}{f'(x_{j+1})} \right| \leq -\frac{1}{k!} \frac{\phi^{(k)}(t_{j+1})}{\phi'(t_j)} \frac{\phi'(t_j)}{\phi'(t_{j+1})} = -\frac{1}{k!} \frac{\phi^{(k)}(t_{j+1})}{\phi'(t_{j+1})}.$$

Por último

$$|x_{j+2} - x_{j+1}| \leq \left| \frac{f(x_{j+1})}{f'(x_{j+1})} \right| \leq -\frac{\phi(t_{j+1})}{\phi'(t_{j+1})} = t_{j+2} - t_{j+1}.$$

Luego, hemos probado que

$$|x_{k+1} - x_k| \leq t_{k+1} - t_k.$$

Además $\{t_k\}$ es convergente a t^* , por tanto $\{t_k\}$ es de Cauchy, lo cual implica que $\{x_k\}$ también es de Cauchy y que $\{x_k\}$ converge a x^* .

Por otra parte, notemos que de (2.45) se deduce que

$$\lim_{j \rightarrow \infty} \left| \frac{f(x_j)}{f'(x_j)} \right| \leq \lim_{j \rightarrow \infty} -\frac{\phi(x_j)}{\phi'(x_j)} = -\frac{\phi(t^*)}{\phi'(t^*)} = 0.$$

Entonces $f(x^*) = 0$, luego x^* es la solución de $f(x)$. □

Convergencia global del método de Newton

En el Análisis Numérico principalmente se emplean dos teoremas que permiten conocer el comportamiento global de los métodos iterativos de un punto. Éstos son

1. El teorema del Punto Fijo.
2. El teorema de Fourier.

El teorema del Punto Fijo nos da las condiciones suficientes para la convergencia de un algoritmo dado por una función contractiva a partir de un valor inicial tomado en un intervalo apropiado. Una demostración detallada de la convergencia del método de Newton usando este tipo de técnicas puede verse en [60]

En esta memoria nos centraremos en una serie de condiciones suficientes para la convergencia del método de Newton, que tienen una clara interpretación geométrica y que muchos autores denominan como *condiciones de Fourier*. Para un estudio más detallado, véase a [96], [132], [142] y [180]

Teorema 2.8 (Condiciones de convergencia de Fourier). Sea $f(x) : [a, b] \rightarrow \mathbb{R}$ una función $f \in \mathcal{C}^2[a, b]$ que cumple las siguientes condiciones:

- i) $f(a)f(b) < 0$.
- ii) $f'(x) \neq 0, \forall x \in [a, b]$.
- iii) $f''(x)$ no cambia de signo en $[a, b]$.
- iv) $\max\{|\frac{f(a)}{f'(a)}|, |\frac{f(b)}{f'(b)}|\} \leq b - a$.

Entonces existe una única raíz x^* de (2.4) en $[a, b]$ y la sucesión $\{x_n\}_{n=0}^{\infty}$, definida por (2.5) converge hacia x^* para cualquier valor inicial $x_0 \in [a, b]$.

Demostración. De acuerdo a las condiciones que se establecen en las hipótesis, pueden darse varias posibilidades. Para fijar ideas asumiremos que $f'(x) < 0$ y que $f''(x) \geq 0$ en $[a, b]$, con lo cual se asegura que ii) y iii) se cumplan.

Notemos que en estas condiciones x^* es la única raíz de $f(x) = 0$ en $[a, b]$. Además, al ser $f'(x) < 0$ y verificar i), se sigue que $f(a) > 0$ y $f(b) < 0$, tal y como se muestra en la gráfica (2.8).

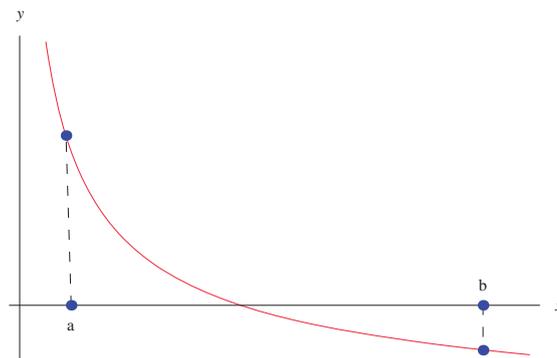


Figura 2.8: Interpretación geométrica del teorema de Fourier.

Visto esto, y asumiendo que $x_0 \in [a, x^*]$ demostraremos que:

- $\{x_n\}_{n=0}^{\infty}$ es una sucesión creciente.
- El límite de $\{x_n\}_{n=0}^{\infty}$ es la raíz x^*

Veamos la primera parte. Si $x_0 \in [a, x^*]$ entonces $x_n < x_{n+1} < x^*$. La demostración será por inducción.

Para $n = 0$ verificaremos que $x_0 < x_1 < x^*$.

Por (2.5) sabemos que

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Pero $f'(x_0) < 0$ y $f(x_0) > 0$, luego

$$x_1 - x_0 = -\frac{f(x_0)}{f'(x_0)} > 0,$$

por tanto $x_0 < x_1$.

Ahora demostraremos que $x_1 < x^*$, para esto comprobaremos que $x^* - x_1 > 0$. Aplicando el teorema del valor medio, se sigue

$$x^* - x_1 = N_f(x^*) - N_f(x_0) = N'_f(\gamma)(x^* - x_0),$$

donde

$$N'_f(\gamma) = \frac{f(\gamma)f''(\gamma)}{f'(\gamma)^2} > 0, \quad \text{para } \gamma \in (x_0, x^*).$$

Luego

$$x^* - x_1 > 0.$$

Asumiendo que $x_k < x_{k+1} < x^*$ demostraremos que $x_{k+1} < x_{k+2} < x^*$. Lo haremos aplicando un procedimiento análogo al anterior. Primero demostraremos que $x_{k+1} < x_{k+2}$.

Nuevamente, por (2.5) se obtiene

$$x_{k+2} = x_{k+1} - \frac{f(x_{k+1})}{f'(x_{k+1})}.$$

Pero acá también se cumple que $f'(x_{k+1}) < 0$ y $f(x_{k+1}) > 0$, por tanto

$$x_{k+2} - x_{k+1} = -\frac{f(x_{k+1})}{f'(x_{k+1})} > 0,$$

luego

$$x_{k+1} < x_{k+2}.$$

Para demostrar que $x_{k+2} < x^*$ comprobaremos que $x^* - x_{k+2} > 0$. Aplicando de nuevo el teorema del valor medio, se sigue

$$x^* - x_{k+2} = N_f(x^*) - N_f(x_{k+1}) = N'_f(\theta)(x^* - x_{k+1}),$$

donde

$$N'_f(\theta) = L_f(\theta) = \frac{f(\theta)f''(\theta)}{f'(\theta)^2} > 0, \quad \text{para } \theta \in (x_{k+1}, x^*).$$

Por lo tanto $x^* - x_{k+2} > 0$, $x_{k+2} < x^*$. Así queda demostrado que $\{x_n\}_{n=0}^{\infty}$ es una sucesión creciente acotada superiormente por x^* y, en consecuencia, $\{x_n\}$ es una sucesión convergente.

Veamos ahora que el límite de la sucesión $\{x_n\}_{n=0}^{\infty}$ es la raíz x^* .

Aplicando límites a (2.5) se sigue

$$\lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \left(x_n - \frac{f(x_n)}{f'(x_n)} \right).$$

Pero

$$\lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} x_n = L.$$

Además, como f y f' son funciones continuas en $[a, b]$,

$$L = L - \frac{f(L)}{f'(L)} \implies \frac{f(L)}{f'(L)} = 0.$$

Como $f'(L) < 0$ se sigue que $f(L) = 0$ y, por tanto, $L = x^*$.

□

Quedan por comprobar los casos siguientes

- 1) $f'(x) < 0$, $f''(x) \leq 0$, $x \in [a, b]$.
- 2) $f'(x) > 0$, $f''(x) \geq 0$, $x \in [a, b]$.
- 3) $f'(x) > 0$, $f''(x) \leq 0$, $x \in [a, b]$.

Sin embargo, el caso 3) se reduce al caso estudiado cambiando a f por $-f$. A su vez, el caso 1) se reduce al 2) de la misma forma. En ese sentido hagamos el siguiente análisis.

Haciendo un cambio de variable $z = -x$ y llamando q a la función resultante se sigue

$$f(x) = f(-z) = q(z),$$

esta función está definida en $[-b, -a]$ y se comprueba que

$$f'(x) = \frac{dq(z)}{dz} \frac{dz}{dx} = -q'(z).$$

Análogamente se comprueba que $f''(x) = q''(z)$. Esto significa que la raíz x^* de (2.4) en $[a, b]$ se habrá transformado en la raíz $-x^*$ de $q(z) = 0$ en $[-b, -a]$.

Entonces, el caso 2) se reduce también al que hemos estudiado, porque si $f'(x) > 0$ y $f''(x) \geq 0$ en $[a, b]$, $q'(z) < 0$ y $q''(z) \geq 0$ en $[-b, -a]$, luego el teorema será válido para q . Deshaciendo el cambio, también será cierto para el caso 2) y para el caso 1).

De acuerdo a [98], las condiciones de Fourier son insuficientes para asegurar la convergencia del método de Newton-Raphson, pues tomando en cuenta la interpretación geométrica del método de Newton-Raphson, nos damos cuenta que su convergencia está asegurada siempre y cuando la segunda derivada de la función $f''(x)$ no cambie de signo en el intervalo $[a, b]$ en que aparece la solución x^* . Es decir, cuando la función tenga un punto de inflexión en su dominio. En ese sentido, [98] realiza un estudio del método de Newton-Raphson en función del grado de convexidad logarítmico (2.11). Para mayores detalles de este estudio refiérase nuevamente a [98].

2.3. El método de Halley

Otro de los métodos iterativos más utilizados para resolver ecuaciones no lineales (2.4) es el método de Halley, cuya expresión viene dada por

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{2[f'(x_n)]^2 - f(x_n)f''(x_n)}, \quad n = 0, 1, 2, \dots \quad (2.58)$$

De forma equivalente, podemos expresar este método en términos del grado de convexidad logarítmico definido en (2.11):

$$x_{n+1} = x_n - \left(\frac{1}{1 - 1/2L_f(x_n)} \right) \frac{f(x_n)}{f'(x_n)} \quad n = 0, 1, 2, \dots \quad (2.59)$$

Siguiendo el esquema de la sección anterior, analizaremos la evolución histórica del método de Halley, las técnicas para deducirlo y algunos resultados que garantizan su convergencia.

2.3.1. Notas históricas sobre el método de Halley

Edmond Halley, nació el 8 de noviembre de 1656 en Haggerston, localidad ubicada en los alrededores de Londres, y murió el 14 de enero de 1742 en Greenwich, Inglaterra.

Siguiendo a [160] encontramos que Halley fue impresionado por los trabajos de un ingenioso profesor de matemática llamado Thomas Fautet de Lagny, quien, en un libro publicado en 1621 en París, presentó algunas fórmulas que servían para extraer la raíz de potencias puras, especialmente, la raíz cúbica.



Figura 2.9: Edmund Halley (1656 – 1742) y Ernst Schröder, (1841 – 1902).

El resultado de Lagny que impresionó a Halley es que la raíz de $\sqrt[3]{a^3 + b}$ con $a, b \in \mathbb{Z}$ está entre

$$a + \frac{ab}{3a^3 + b} \quad \text{y} \quad \frac{a}{2} + \sqrt{\frac{a^2}{4} + \frac{b}{3a}} \quad \text{para} \quad a^3 > b > 0.$$

Halley llamó a esto la forma racional e irracional respectivamente, prefiriendo emplear en sus cálculos la forma irracional pues la racional poseía un gran divisor que dificultaría su aplicación.

Leyendo a [67] nos enteramos que Halley inició la demostración de la fórmula de Lagny buscando un termino corrector ε de forma que $(a + \varepsilon)^3 = a^3 + b$, de donde se obtiene que $b = 3a^2\varepsilon + 3a\varepsilon^2 + \varepsilon^3$. Si despreciamos el término ε^3 obtenemos que $b \approx 3a^2\varepsilon + 3a\varepsilon^2$ de donde

$$\frac{b}{3a^2} \approx \varepsilon + \frac{\varepsilon^2}{a} \approx \varepsilon \quad \text{y} \quad (2.60)$$

$$\varepsilon \approx \frac{b}{3a^2 + 3a\varepsilon}. \quad (2.61)$$

Luego, sustituyendo (2.60) en (2.61) se comprueba que

$$\varepsilon \approx \frac{b}{3a^2 + 3a\frac{b}{3a^2}} = \frac{ab}{3a^3 + b},$$

luego, la raíz aproximada de la ecuación de Lagny sería

$$\sqrt[3]{a^3 + b} \approx \frac{ab}{3a^3 + b}. \quad (2.62)$$

Si hacemos una traducción de este resultado empleando la notación actual y llamando $f(x) = x^3 - a^3 - b$ con $x_0 = a$, entonces (2.62) se correspondería con el primer paso del método de Halley, el cual viene definido por la expresión

$$x_{n+1} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}, \quad (2.63)$$

donde

$$L_f(x_n) = \frac{f(x_n)f''(x_n)}{f'(x_n)^2}. \quad (2.64)$$

De acuerdo con [67], Halley también admiró los trabajos de F. Viète, cuyo método es un algoritmo linealmente convergente parecido al método de la Bisección y que normalmente se utiliza para obtener los valores iniciales del método de Newton-Raphson y de otros algoritmos iterativos de orden superior. Esto, al parecer, significa que el método de Viète fué un importante precursor del método de Halley y de Newton-Raphson. Pero Halley al escribir sus notas, aparentemente, no se dio cuenta de esto ni de que su método suponía el uso de derivadas; lo cual es algo lógico que así fuera, pues fue Simpson, en 1740, quien por primera vez estableciera la conexión entre las derivadas y el método de Newton-Raphson. Lo verdaderamente sorprendente es que Taylor, en 1712 reconociera de manera tan fácil las derivadas en el método de Halley.

A pesar del éxito de Taylor, él no pudo dar una fórmula general para el método de Halley. Esto lo consiguió Schröder, más de un siglo después, al derivar la función de iteración del método de Halley tal y como la conocemos hoy. En su trabajo, [162], Schröder hace referencia

a la fórmula de Halley de manera muy pasajera. Luego Kobal, [119], en un breve trabajo publicado en 1891, derivó la fórmula de Halley, pero desafortunadamente la demostración es poco clara y confusa. Después Frame, [74], logró, por primera vez, derivar la función de iteración empleando el polinomio de Taylor de segundo grado.

2.3.2. Descripción del método de Halley

Al igual que otros métodos para encontrar raíces, el método de Halley es un algoritmo numérico que se emplea para resolver una ecuación no lineal. Existen una gran cantidad de libros y artículos en los cuales aparece la descripción analítica del método de Halley, véase [81], [83], [160], [188], [205],[206] y [207]. En lo que sigue daremos un listado de diversas formas de obtener el método de Halley, tal y como se hace por ejemplo en [67]. Además, añadimos a dicho listado una nueva construcción geométrica que usa interpolación mediante funciones exponenciales.

A partir del desarrollo de Taylor

Leyendo a [67] encontramos que la idea de construir la función de iteración del método de Halley mediante la aplicación del teorema de Taylor, es probablemente la forma más usada de redescubrir dicho método. Tanto en [160] como en [205] podemos encontrar bonitas demostraciones, veamos la que presenta [160].

Dado el siguiente polinomio de Taylor de segundo grado

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2!}(x - x_n)^2,$$

donde x_n es otra vez una raíz aproximada de (2.4). Tal como en el método de Newton-Raphson, el objetivo aquí es determinar un punto x_{n+1} donde la gráfica de $f(x)$ corta al eje X , esto es , resolver la ecuación

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{f''(x_n)}{2!}(x_{n+1} - x_n)^2, \quad (2.65)$$

para x_{n+1} .

Siguiendo a [74], factorizamos el término $(x_{n+1} - x_n)$ de (2.65) y obtenemos

$$0 = f(x_n) + (x_{n+1} - x_n) \left(f'(x_n) + \frac{1}{2} f''(x_n)(x_{n+1} - x_n) \right),$$

de donde

$$x_{n+1} - x_n = \frac{f(x_n)}{f'(x_n) + \frac{1}{2} f''(x_n)(x_{n+1} - x_n)}.$$

Aproximando la diferencia $x_{n+1} - x_n$ de la ecuación anterior por la corrección de Newton $\frac{-f(x_n)}{f'(x_n)}$ se obtiene

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}.$$

La cual representa la función de iteración del método de Halley.

Construcción geométrica: método de las hipérbolas tangentes

Como vimos anteriormente, el método de Newton-Raphson tiene una fácil interpretación geométrica, a tal punto que ha sido el motivo por el cual es también llamado el método de la tangente.

Leyendo a [160] encontramos que, aparentemente, fue Salehov el primero en sugerir que la función de iteración del método de Halley se podía obtener a partir de una función oscultriz racional de la forma

$$y(x) = \frac{x + c}{ax + b}. \quad (2.66)$$

Recordemos que una curva oscultriz de una función f en x_n es aquella que satisface la ecuación

$$y^{(k)}(x_n) = f^{(k)}(x_n), k = 0, 1, 2. \quad (2.67)$$

Siguiendo a [160] encontramos la siguiente construcción de la fórmula de iteración del método de Halley a partir de la interpretación geométrica.

Sea la ecuación

$$y(x) = \frac{(x - x_n) + c}{a(x - x_n) + b}, \quad (2.68)$$

la cual define una hipérbola equivalente a (2.66).

Combinando a (2.67) y (2.68) obtenemos el siguiente sistema de ecuaciones

$$\begin{aligned} \frac{c}{b} &= f(x_n) \\ \frac{b - ac}{b^2} &= f'(x_n) \\ \frac{2a(ac - b)}{b^3} &= f''(x_n). \end{aligned}$$

Cuya solución es

$$a = \frac{-f''(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)}, \quad (2.69)$$

$$b = \frac{2f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)},$$

$$c = \frac{2f(x_n)f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)}. \quad (2.70)$$

De (2.68) sigue que si tomamos como nueva aproximación su intersección con el eje y , es decir si $y(x_{n+1}) = 0$, entonces se tiene que

$$x_{n+1} = x_n - c. \quad (2.71)$$

Sustituyendo (2.70) en (2.71) se obtiene (2.59).

Otra interpretación geométrica: interpolación exponencial

En lo que sigue daremos una nueva descripción de obtener el método de Halley, basada en una interpolación mediante funciones exponenciales y polinomiales. Esta técnica está basada en un desarrollo dado por Kanwar et al. en [23] y [112].

Teniendo en cuenta la aproximación de una función $f(x)$ por una función osculatriz $y(x)$ de forma que $y^j(x_n) = f^j(x_n)$, para $j = 0, 1, 2$ vista en [5], consideraremos ahora una curva de la forma

$$y(x) = e^{a(x-x_n)}[b(x-x_n) + c], \quad (2.72)$$

donde a, b y c son tres parámetros a determinar.

Por otro lado, si denotamos por x_{n+1} a la intersección de la curva (2.72) con el eje de las abscisas, tenemos que $y(x_{n+1}) = 0$, por tanto

$$x_{n+1} = x_n - \frac{c}{b}.$$

Podemos obtener así diversos procesos iterativos que nos permitirán resolver la ecuación no lineal $f(x) = 0$. Notemos que si buscamos que el método deseado tenga convergencia cuadrática, se debe cumplir que

$$y(x_n) = f(x_n), \quad y'(x_n) = f'(x_n),$$

lo que nos conduce a las igualdades $c = f(x_n)$, $b = f'(x_n) - a$.

En consecuencia, se obtiene la familia uniparamétrica

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n) - af(x_n)}, \quad a \in \mathbb{R}. \quad (2.73)$$

Para un estudio más profundo de (2.73,) refiérase a [112].

Si ahora deseamos que el método dado en (2.73) tengan convergencia cúbica, debemos exigir que

$$y''(x_n) = f''(x_n),$$

o lo que es lo mismo

$$a^2 f(x_n) - 2af'(x_n) + f''(x_n) = 0. \quad (2.74)$$

Para valores de $a \neq 0$, muy próximos a cero, tenemos que el término a^2 dado en (2.74) puede ser despreciado respecto al resto. Por tanto, podemos obtener la solución aproximada

$$a = \frac{f''(x_n)}{2f'(x_n)}. \quad (2.75)$$

Sustituyendo (2.75) en (2.73) obtenemos el método de Halley:

$$x_{n+1} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}.$$

Notemos que aunque en teoría esta construcción es válida sólo para valores del parámetro a dado por (2.69) muy próximos a 0, se tiene que el proceso obtenido tiene convergencia cúbica para funciones $f(x)$ en general, con la única restricción de que se aproximen raíces simples.

Construcción a partir de la interpolación racional inversa

Para la obtención del método de Halley a partir de la interpolación racional inversa se considera resolver una ecuación $f(x) = 0$, considerando que a cada x se le hace corresponder la función $y = f(x)$, denotando a $x = \phi(y)$ como la función su inversa. Luego, la solución aproximada de $f(x) = 0$ se obtiene encontrando la aproximación de $x = \phi(y)$. Para tal acción aproximaremos a $x = \phi(y)$ aplicando el desarrollo de Taylor. Es importante destacar que esta idea se atribuye al matemático Chebyshev. Para una visión más profunda de esta construcción, refiérase a [67].

Asumamos que la aproximación de $\phi(y)$ viene dada por una función racional de la forma

$$\phi(y) \simeq \frac{(y - y_0) + a}{b(y - y_0) + c} = R(y),$$

además que $\phi^{(k)}(y_0) = R^{(k)}(y_0)$, para $k = 0, 1, 2$, luego

$$R(y) = \frac{(2f'(x_0) + x_0 f''(x_0))(y - y_0) + 2x_0 f'(x_0)^2}{f''(x_0)(y - y_0) + 2f'(x_0)^2}.$$

Tomando a $x_1 = R(0)$ como una nueva aproximación, sigue

$$x_1 = \frac{-f(x_0)(2f'(x_0) + x_0 f''(x_0)) + 2x_0 f'(x_0)^2}{-f(x_0) f''(x_0) + 2f'(x_0)^2},$$

de donde, después de operar adecuadamente se obtiene el método de Halley:

$$x_1 = x_0 - \frac{2}{2 - L_f(x_0)} \frac{f(x_0)}{f'(x_0)}.$$

Construcción utilizando fracciones continuas

La idea de obtener el método de Halley a partir de fracciones continuas fue aportada por Frame, la misma aparece en [74]. La idea de Frame descansa en la obtención del término corrector ε y su posterior desarrollo en serie de potencias.

Dado que este procedimiento de construcción es una variante de la interpolación inversa, mantendremos la misma notación, tanto de $x = \phi(y)$ como para $y = f(x)$.

Dado el término corrector ε en la forma

$$\begin{aligned}\varepsilon &= r - x_0 = \phi(0) - \phi(y_0) = \sum_{k \geq 1} c_k \nu^k \\ &= \nu - \frac{f''(x_0)}{2f'(x_0)} \nu^2 + \frac{3f''(x_0) - f'''(x_0)f'(x_0)}{6f'(x_0)} \nu^3 + \dots,\end{aligned}$$

donde r representa la solución de $f(x) = 0$, $\nu = f(x_0)/f'(x_0)$ y $c_1 = 1$, $c_2 = -f''(x_0)/2f'(x_0)$, etc.

Desarrollando en fracciones continuas el desarrollo en serie de potencias del término corrector anterior, se obtiene

$$\varepsilon = \sum_{k \geq 1} c_k \nu^k = b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \dots \frac{a_j}{b_j + \rho_{j+1}}}}}, \quad (2.76)$$

donde

$$\rho_j = \frac{a_j}{b_j + \rho_{j+1}}.$$

Dependiendo del lugar en que se trunque a (2.76), se obtendrán diversos procesos iterativos. Para nuestro caso, debemos truncharla en el segundo término y calcular a a_2 . Notemos que a_2 es el primer término del desarrollo en serie de potencias de ρ_2 . Como

$$\rho_2 = \frac{a_1 - \rho_1}{\rho_1} = -c_2 \nu + (c_2^2 - c_3) \nu^3 + \dots$$

Luego $a_2 = -c_2 \nu = -L_f(x_0)/2$. Por tanto

$$\varepsilon \simeq \frac{a_1}{1 + a_2} = -\frac{2}{2 - L_f(x_0)} \frac{f(x_0)}{f'(x_0)},$$

obteniéndose así el término corrector del método de Halley.

Construcción acelerando el método de Newton

En el proceso de solución de una ecuación $f(x) = 0$, existen casos en los que es preferible resolver una ecuación modificada de la forma $f(x)g(x) = 0$, donde $g(x)$ es una función convenientemente elegida, pues deberá cumplir que no se anule en un entorno de la raíz buscada, esto es, si r es tal que $f(r) = 0$, entonces $g(x) \neq 0$ en un entorno de r .

Para nuestro caso, si aproximamos la solución de una ecuación aplicando el método de Newton, lo preferible es que la función sea una recta, pues en este caso la solución exacta se obtendrá en la primera iteración. A pesar de que en la práctica no tiene sentido utilizar el método de Newton para resolver una ecuación lineal, la idea de utilizar a $g(x)$ para «linealizar» lo más posible la ecuación $f(x)g(x) = 0$ fue sugerida por [26], el cual comprobó que un aumento de la «linealidad» de la ecuación en un entorno de r , tiene incidencia en la velocidad de convergencia del método de Newton. Retomando esta idea, Gerlach [81] demuestra que cuando se aplica el método de Newton a una ecuación $\Psi(x) = 0$, donde la solución r cumple que $\Psi'(r) \neq 0$, $\Psi''(r) = 0$ y $\Psi'''(r) \neq 0$, entonces el método de Newton alcanza la convergencia de orden 3. Bajo las condiciones impuestas por Gerlach, ¿que forma tiene el método de Newton?

Si $\Psi(x) = f(x)g(x)$, buscamos a $g(x)$ de manera que $\Psi''(r) = 0$. Así

$$f''(r)g(r) + 2f'(r)g'(r) = 0.$$

Ecuación diferencial que al ser resuelta, salvo constantes multiplicativas, se obtiene que

$$g(x) = \frac{1}{\sqrt{f'(x)}}.$$

Luego, la idea de Brown y Gerlach consiste en aplicar el método de Newton a la función $\Psi(x) = \frac{f(x)}{\sqrt{f'(x)}}$. De donde se obtiene

$$x_{n+1} = x_n - \frac{\Psi(x)}{\Psi'(x)} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}.$$

Luego, bajo las condiciones de aceleración impuestas por Brown y Gerlach, el método de Newton se convierte en (2.59).

Construcción a partir de iteraciones de Laguerre

La construcción de la iteración de Laguerre considera la derivación logarítmica de un polinomio $p(x)$ de grado n , con raíces reales $x_1 \leq x_2 \leq \dots \leq x_n$ de la forma:

$$p(x) = \prod_{i=1}^n (x - x_i).$$

Después de varios cálculos Laguerre obtuvo la ecuación

$$nu^2 - 2Au - (n-1)B + A^2 = 0, \quad (2.77)$$

en donde

$$A = p'(x)/p(x) = (\log p(x))' = \sum_{i=1}^n \frac{1}{(x - x_i)} = \sum_{i=1}^n a_i$$

y

$$B = -dA/dx = -\frac{d(p'(x)/p(x))}{dx} = \frac{p'(x)^2 - p(x)p''(x)}{p(x)^2} = \sum_{i=1}^n \frac{1}{(x-x_i)^2} = \sum_{i=1}^n a_i^2.$$

En [67] y [142] encontramos una demostración detallada de dicho método.

Cuatro años después de la publicación de [142], Hansen y Patrick [95] extendieron la iteración de Laguerre a funciones cualesquiera. Su idea consistió en sustituir el grado del polinomio por un parámetro λ en (2.77) manteniendo la notación

$$A = \frac{f'(x)}{f(x)}, \quad B = -A' = \frac{f'(x)^2 - f(x)f''(x)}{f(x)^2}.$$

La iteración que aparece en el trabajo de [95] es

$$x_+ = x - \frac{\lambda}{1 \pm \sqrt{(1-\lambda)(1-\lambda + \lambda L_f(x))}} \frac{f(x)}{f'(x)}. \quad (2.78)$$

Al racionalizar a (2.78) nos queda

$$x_+ = x - \frac{1 \pm \sqrt{(1-\lambda)(1-\lambda + \lambda L_f(x))}}{2 - \lambda - (1-\lambda)L_f(x)} \frac{f(x)}{f'(x)},$$

luego, haciendo $\lambda = 0$ y eligiendo el signo $+$ en el numerador, se llega a (2.59).

2.3.3. Convergencia del método de Halley

Presentamos a continuación una selección de resultados de convergencia para el método de Halley. Ajustándonos al mismo esquema seguido para el método de Newton en la sección anterior, los clasificamos en tres tipos: local, semilocal y global.

Convergencia local del método de Halley

Siguiendo en la línea de resultados locales obtenidos para el método de Newton, presentamos a continuación un teorema de convergencia local para el método de Halley. Podemos encontrar resultados similares en los trabajos [160] y [205].

Teorema 2.9. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz simple de $f(x) = 0$, ($f(x^*) = 0$ y $f'(x^*) \neq 0$). Supongamos que f es tres veces derivable con continuidad en I . Entonces el método de Halley (2.59) definido por $x_{n+1} = H_f(x_n)$,*

$$H_f(x) = x - \frac{2}{2 - L_f(x)} \frac{f(x)}{f'(x)},$$

tiene al menos convergencia cúbica localmente. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = (A_2^2 - A_3)e_n^3 + O_4.$$

donde A_i , e_n y O_N se han introducido en (2.12), (2.13) y (2.14) respectivamente.

Demostración. (Usando el Teorema de Schröder) El cálculo de las derivadas de H_f nos permite deducir que

$$H_f(x^*) = x^*, \quad H'_f(x^*) = H''_f(x^*) = 0, \quad H'''_f(x^*) = \frac{3f''(x^*)^2 - 2f'(x^*)f'''(x^*)}{2f'(x^*)^2} \neq 0.$$

Por lo tanto,

$$e_{n+1} = x_{n+1} - x^* = H_f(x_n) - x^* = \frac{H'''_f(x^*)}{3!}(x_n - x^*)^3 + O_4,$$

de donde se deduce que

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^3} = \frac{H'''_f(x^*)}{3!} = A_2^2 - A_3.$$

Por lo tanto, el método de Halley es un método con orden de convergencia local igual a 3 y constante de error asintótico $A_2^2 - A_3$. \square

Demostración. (Usando desarrollos de Taylor)

En esta ocasión aplicamos el teorema de Taylor, pero seguimos unos pasos diferentes a los dados en el Teorema 2.2 sobre la convergencia del método de Newton. Se obtiene en este caso:

$$0 = f(x^*) = f(x_n) + f'(x_n)(x^* - x_n) + \frac{f''(x_n)}{2!}(x^* - x_n)^2 + \frac{f'''(\alpha)}{3!}(x^* - x_n)^3, \quad (2.79)$$

y también

$$0 = f(x^*) = f(x_n) + f'(x_n)(x^* - x_n) + \frac{f''(\beta)}{2!}(x^* - x_n)^2, \quad (2.80)$$

donde α y β son números que están entre x^* y x_n .

Si ahora multiplicamos a (2.79) por $2f'(x_n)$ y le restamos a (2.80) multiplicada por $f''(x_n)(x^* - x_n)$, se obtiene

$$\begin{aligned} 0 &= 2f(x_n)f'(x_n) + 2[f'(x_n)]^2(x^* - x_n) + f'(x_n)f''(x_n)(x^* - x_n)^2 + \frac{f'(x_n)f'''(\alpha)}{3}(x^* - x_n)^3 \\ &\quad - f(x_n)f''(x_n)(x^* - x_n) - f'(x_n)f''(x_n)(x^* - x_n)^2 - \frac{f''(x_n)f''(\beta)}{2}(x^* - x_n)^3. \end{aligned} \quad (2.81)$$

Luego, si cancelamos los términos iguales con signos opuestos y reorganizamos adecuadamente, nos queda

$$\begin{aligned} 0 &= 2f(x_n)f'(x_n) + (2[f'(x_n)]^2 - f(x_n)f''(x_n))(x^* - x_n) + \\ &\quad + \left(\frac{f'(x_n)f'''(\alpha)}{3} - \frac{f''(x_n)f''(\beta)}{2} \right) (x^* - x_n)^3. \end{aligned} \quad (2.82)$$

Luego, si en (2.82) despejamos a $(x^* - x_n)$ y dividimos por $2[f'(x_n)]^2 - f(x_n)f''(x_n)$, obtenemos

$$x^* - x_n = \frac{-2f(x_n)f'(x_n)}{2[f'(x_n)]^2 - f(x_n)f''(x_n)} - \frac{2f'(x_n)f'''(\alpha) - 3f''(x_n)f''(\beta)}{6(2[f'(x_n)]^2 - f(x_n)f''(x_n))}(x^* - x_n)^3.$$

Así

$$x^* - x_{n+1} = -\frac{2f'(x_n)f'''(\alpha) - 3f''(x_n)f''(\beta)}{12[f'(x_n)]^2 - 6f(x_n)f''(x_n)}(x^* - x_n)^3.$$

Por lo tanto,

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^3} = -\frac{2f'(x^*)f'''(x^*) - 3f''(x^*)^2}{12[f'(x^*)]^2} = \left(\frac{f''(x^*)}{2f'(x^*)}\right)^2 - \frac{f'''(x^*)}{6f'(x^*)} = A_2^2 - A_3.$$

□

Convergencia semilocal del método de Halley

Para demostrar la convergencia semilocal de un método, normalmente se aplican los criterios establecidos por Kantorovich, [110] o los de la alpha-teoría de Smale, [170], las cuales han sido presentadas en la demostración de la convergencia semilocal del método de Newton.

En lo que respecta al método de Halley, existe varios autores, [40], [159], [199], que asumiendo condiciones tipo Kantorovich, han escrito sobre la convergencia semilocal, en espacios de Banach, de dicho método. A modo de ilustración, presentamos algunos de estos teoremas, pero en \mathbb{R} .

Teorema 2.10 (Teorema de Salehov-Mertvetsova). *Sea $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ una función 2 veces diferenciable e I un intervalo abierto en \mathbb{R} . Supongamos que para $x_0 \in I$, se cumple que $|f''(x)| \leq M_1$ y $|f'''(x)| \leq N_1$. Además que se satisfacen las siguientes condiciones:*

$$\begin{aligned} \left| \frac{1}{f'(x_0)} \right| &\leq B_0, \quad |x_1 - x_0| \leq \eta'_0. \\ h'_0 &= M_1 B_0 \eta'_0 \leq \frac{1}{2}. \\ S &= [x_0, 2\eta'_0] \subset I. \\ G_0 &= \left[\frac{N_1}{M_1^2 B_0} (2 + h'_0) + 3 \right] (1 + h'_0) \leq 9. \end{aligned}$$

Entonces, en S existe una solución x^* de $f(x) = 0$ y el método de Halley converge a x^* . Además se cumple que

$$|x^* - x_n| \leq (2h'_0)^{3^n - 1} \frac{\eta'_0}{2^{n-1}}, \quad n = 1, 2, \dots$$

Teorema 2.11 (Teorema de Zheng). *Sea $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ una función 2 veces diferenciable e I un intervalo abierto en \mathbb{R} . Supongamos que para $x_0 \in I$, se cumple que $|f''(x)| \leq M_1$ y $|f'''(x)| \leq N_1$. Además que se satisfacen las siguientes condiciones:*

$$\begin{aligned} \left| \frac{1}{f'(x_0)} \right| &\leq B_0, & \left| \frac{f(x_0)}{f'(x_0)} \right| &\leq \eta'_0. \\ h_0 &= KB_0\eta_0 \leq \frac{1}{2}, \text{ donde } K = \sqrt{M_1^2 + \frac{2N_1}{3B_0(1 - \frac{1}{2}M_1B_0\eta_0)}}. \\ S &= (x_0, (1 + \theta_0)\eta_0) \subset I, \text{ donde } \theta_0 = \frac{1 - \sqrt{1 - 2h_0}}{1 + \sqrt{1 + 2h_0}} \end{aligned} \tag{2.83}$$

Entonces, se cumple que

$$(i) \ x_n \in S, n = 0, 1, \dots$$

$$(ii) \ \lim_{n \rightarrow \infty} x_n = x^* \in [x_0, (1 + \theta_0)\eta_0], f(x^*) = 0.$$

$$(iii) \ |x^* - x_n| \leq (1 + \theta_0)\eta_0\theta_0^{3^n - 1} / \sum_{i=0}^{3^n - 1} \theta_0^i \quad n = 1, 2, \dots$$

En 1995, S. Zhen y D. Robbie, [200], aplicando condiciones tipo Kantorovich, publicaron un interesante teorema de convergencia local para el método de Halley en espacios de Banach. Parte del mismo se presenta en esta tesis, pero adaptado al espacio \mathbb{R} , en donde las funciones son reales de variable real.

Teorema 2.12 (Teorema de Zhen-Robbie). *Sea $f : I \rightarrow \mathbb{R}$, en donde I es un intervalo abierto en \mathbb{R} , una función dos veces diferenciable en I . Asumamos, sin pérdida de generalidad que $f(x_0) \neq 0$, $f'(x_0) \neq 0$ y que $x_0 \in I$. Además*

$$\left| \frac{f''(x) - f''(y)}{f'(x_0)} \right| \leq N|x - y|, \forall x, y \in I. \tag{2.84}$$

$$\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \eta, \quad \left| \frac{f''(x_0)}{f'(x_0)} \right| \leq M. \tag{2.85}$$

$$9N^2\eta^2 + 18NM\eta + 6M^3\eta \leq 3M^2 + 8N. \tag{2.86}$$

Sea

$$S = \{x : |x - x_0| \leq t_1^*\} \subset I,$$

donde t_1^* es la raíz positiva más pequeña de la función real

$$\varphi(t) = \frac{N}{6}t^3 + \frac{M}{2}t^2 - t + \eta, \tag{2.87}$$

entonces la sucesión $\{x_n\}_{n=0}^{\infty}$ generada por la función de iteración del método de Halley está bien definida y converge a $x^* \in I$, donde $f(x^*) = 0$.

Previo a la demostración del Teorema 2.12, Zhen y Robbie demuestran los siguientes lemas.

Lema 2.13. *Bajo la condición dada en (2.85), la función $\varphi(t)$ definida por (2.87) posee tres raíces reales.*

Demostración. Derivando a (2.87) se obtiene

$$\varphi'(t) = \frac{N}{2}t^2 + Mt - 1, \quad (2.88)$$

resolviendo (2.88) se obtienen las raíces reales

$$t_+ = \frac{-M + \sqrt{M^2 + 2N}}{N} \geq 0. \quad (2.89)$$

$$t_- = \frac{-M - \sqrt{M^2 + 2N}}{N} \leq 0. \quad (2.90)$$

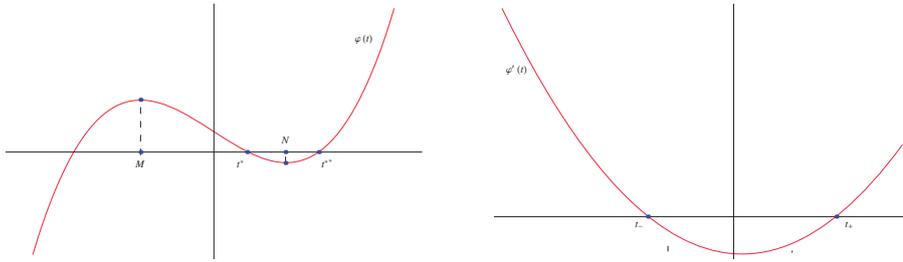


Figura 2.10: A la izquierda, gráfica de $\varphi(t) = \frac{N}{6}t^3 + \frac{M}{2}t^2 - t + \eta$. A la derecha, gráfica de $\varphi'(t) = \frac{N}{2}t^2 + Mt - 1$.

Si $\varphi'(t_+) = 0$, entonces $t_+^2 = \frac{2(1-Mt_+)}{N}$. Además

$$\begin{aligned} \varphi(t) &= \frac{t}{3} \left[\frac{N}{2}t^2 + \frac{3}{2}Mt - 3 \right] + \eta. \\ &= \frac{t}{3} \left[\varphi'(t) + \frac{M}{2}t - 2 \right] + \eta, \end{aligned}$$

luego

$$\begin{aligned} \varphi(t_+) &= \frac{M}{6}t_+^2 - \frac{2}{3}t_+ + \eta = \frac{M}{3N}(1 - Mt_+) - \frac{2}{3}t_+ + \eta. \\ &= \frac{M}{3N} - \frac{2N + M^2}{3N}t_+ + \eta. \\ &= \frac{1}{3N^2} [3NM + M^3 + 3N^2\eta - (2N + M^2)^{3/2}]. \end{aligned} \quad (2.91)$$

Tomando en cuenta las condiciones impuestas en (2.85), se comprueba que

$$(3NM + M^3 + 3N^2\eta)^2 \leq (2N + M^2)^3,$$

luego $\varphi(t_+) \leq 0$. Además, es evidente que $\varphi(-\infty) = -\infty$, $\varphi(\infty) = \infty$ y $\varphi(0) > 0$ □

Siguiendo a Davies y Dawson, [54], podemos enunciar el siguiente resultado. La demostración que mostramos a continuación utiliza notaciones y técnicas que pueden verse en el libro de Hernández y Salanova [102].

Lema 2.14. *Bajo la condición dada en (2.85), la sucesión $\{t_n\}$, obtenida al aplicar la función de iteración del método de Halley a (2.87), definida por*

$$\begin{cases} t_0 &= 0, \\ t_{n+1} &= t_n - \left(\frac{2}{2 - L_\varphi(t_n)} \right) \frac{\varphi(t_n)}{\varphi'(t_n)}, \end{cases}$$

es monótonamente creciente y converge a la raíz t_1^ de (2.87).*

Demostración. De la geometría de la función $\varphi(t)$ (véase la gráfica 2.10) y del Lema de Altman 2.28 se deduce que $t_{n+1} > t_n$, ya que $\varphi(t) > 0$, $\varphi'(t) < 0$ y $L_\varphi(t) < 1/2$. Por otra parte, la sucesión $\{t_n\}$ está acotada superiormente por t_1^* . En efecto, es evidente que $t_1^* - t_0 > 0$ y, por inducción,

$$t_1^* - t_{n+1} = H_\varphi(t_1^*) - H_\varphi(t_n) = H'_\varphi(\zeta_n)(t_1^* - t_n),$$

donde $H_\varphi(t)$ es la función de iteración del método de Halley aplicado al polinomio $\varphi(t)$ y ζ_n es un punto comprendido entre t_n y t_1^* .

Como muestran Hernández y Salanova [102], la derivada de $H_\varphi(t)$ es

$$H'_\varphi(t) = \frac{L_\varphi(t)^2}{(2 - L_\varphi(t))^2} (3 - 2L_{\varphi'}(t)).$$

En nuestro caso (véase la gráfica de $\varphi'(t)$ en la Figura 2.10) tenemos que $\varphi'(t) < 0$ para todo $t \in (0, t_1^*)$ y $\varphi'''(t) = M > 0$, por lo tanto

$$L_{\varphi'}(t) = \frac{\varphi'(t)\varphi'''(t)}{\varphi''(t)^2} < 0.$$

En consecuencia, $H'_\varphi(\zeta_n) > 0$ y $t_1^* - t_{n+1} > 0$. □

Lema 2.15. *Bajo las condiciones del Teorema 2.12, si $|x - x_0| < t_+$, donde $t_+ > 0$ es la raíz positiva de $\varphi'(t)$, entonces $1/f'(x)$ existe y*

$$\left| \frac{f'(x_0)}{f'(x)} \right| \leq \frac{-1}{\varphi'(|x - x_0|)} \quad (2.92)$$

$$\left| \frac{f''(x)}{f'(x_0)} \right| \leq \varphi''(|x - x_0|). \quad (2.93)$$

Demostración. Del Lema 2.13, vemos que cuando $|x - x_0| < t_+$, entonces

$$\varphi'(|x - x_0|) = \frac{N}{2}|x - x_0|^2 + M|x - x_0| - 1 < 0.$$

Así, bajo las condiciones del Teorema 2.12, tenemos que

$$\left| \frac{f''(x_0)(x-x_0)}{f'(x_0)} + \int_0^1 \frac{[f''(x_0+t(x-x_0)) - f''(x_0)]dt(x-x_0)}{f'(x_0)} \right| \leq M|x-x_0| + \frac{N}{2}|x-x_0|^2 < 1.$$

Del Teorema del valor medio, obtenemos

$$f'(x) = f'(x_0) + \int_0^1 f''(x_0+t(x-x_0))dt(x-x_0),$$

luego

$$\frac{f'(x)}{f'(x_0)} = 1 + \frac{f''(x_0)(x-x_0)}{f'(x_0)} + \int_0^1 \frac{[f''(x_0+t(x-x_0)) - f''(x_0)]dt(x-x_0)}{f'(x_0)}.$$

Por el Lema de Banach podemos afirmar que la inversa $\frac{f'(x)}{f'(x_0)}$ existe y que

$$\left| \frac{f'(x_0)}{f'(x)} \right| \leq \frac{1}{1 - M|x-x_0| - N/2|x-x_0|^2} = \frac{1}{\varphi'(|x-x_0|)}.$$

Luego, es evidente que

$$\left| \frac{f''(x)}{f'(x_0)} \right| \leq \left| \frac{f''(x_0)}{f'(x_0)} \right| + \left| \frac{[f''(x_0) - f''(x)]}{f'(x_0)} \right| \leq M + N|x-x_0| = \varphi''(|x-x_0|).$$

Luego el lema está demostrado. \square

Veamos ahora la demostración del Teorema 2.12.

Demostración del teorema de Zhen-Robbie. Probaremos que para todo entero positivo n , se cumple que $x_n \in S = \{x : |x-x_0| \leq t_1^*\}$, $1/f'(x_n)$ existe y que

$$\left| \frac{f(x_n)}{f'(x_0)} \right| \leq \varphi(t_n), \quad (2.94)$$

$$\left| \frac{f'(x_0)}{f'(x_n)} \right| \leq \frac{-1}{\varphi'(t_n)}, \quad (2.95)$$

$$\left| \frac{f''(x_n)}{f'(x_0)} \right| \leq \varphi''(t_n), \quad (2.96)$$

$$|x_{n+1} - x_n| \leq t_{n+1} - t_n. \quad (2.97)$$

Primero mostraremos que (2.97) es cierto cuando (2.94), (2.95) y (2.96) se cumplen. De hecho, de (2.94), (2.95) y (2.96) obtenemos

$$\left| \frac{f(x_n)}{f'(x_n)} \right| \leq \left| \frac{f'(x_0)}{f'(x_n)} \right| \left| \frac{f(x_n)}{f'(x_0)} \right| \leq -\frac{\varphi(t_n)}{\varphi'(t_n)}, \quad (2.98)$$

y

$$\left| \frac{f''(x_n)}{f'(x_n)} \right| \leq \left| \frac{f'(x_0)}{f'(x_n)} \right| \left| \frac{f''(x_n)}{f'(x_0)} \right| \leq -\frac{\varphi''(t_n)}{\varphi'(t_n)}. \quad (2.99)$$

Luego, aplicando la monotonicidad de la sucesión $\{t_n\}$ demostrada en el Lema 2.14, sigue

$$\frac{1}{2} \left| \frac{f''(x_n) f(x_n)}{f'(x_n)^2} \right| \leq \frac{1}{2} \frac{\varphi(t_n) \varphi''(t_n)}{\varphi'(t_n)^2} < 1.$$

Por el Lema de Banach, la inversa de

$$H_n = 1 - \frac{1}{2} \frac{f''(x_n) f(x_n)}{f'(x_n)^2} = 1 - \frac{1}{2} L_f(x_n) \quad (2.100)$$

existe y

$$\frac{1}{H_n} = \frac{1}{1 - \frac{1}{2} \frac{f''(x_n) f(x_n)}{f'(x_n)^2}} \leq \frac{1}{1 - \frac{1}{2} \frac{\varphi(t_n) \varphi''(t_n)}{\varphi'(t_n)^2}}.$$

Es fácil ver que de (2.59) y (2.98) se obtiene (2.97).

En lo que sigue probaremos por inducción las condiciones (2.94), (2.95) y (2.96). Es evidente que se cumplen para $n = 0$. Supongamos que (2.94), (2.95) y (2.96) se cumplen para $n \leq k$, entonces, de (2.97) se sigue que

$$|x_{k+1} - x_0| \leq t_{k+1} - t_0 = t_{k+1} < t_1^* \leq t_+,$$

luego $x_{k+1} \in S$. Notemos que por el Lema 2.15, existe $1/f'(x_{k+1})$ y que (2.95) y (2.96) se cumplen para $n = k + 1$.

Aplicando a Taylor a (2.59) obtenemos que

$$\begin{aligned} 0 &= f(x_k) + f'(x_k)(x_{k+1} - x_k) - \frac{1}{2} \frac{f''(x_k) f(x_k) (x_{k+1} - x_k)}{f'(x_k)} \\ &= f(x_k) + f'(x_k)(x_{k+1} - x_k) + \frac{1}{2} f''(x_k) (x_{k+1} - x_k)^2 - \frac{1}{4} \frac{f''(x_k)^2 f(x_k) (x_{k+1} - x_k)^2}{f'(x_k)^2}. \end{aligned}$$

Por tanto

$$f(x_{k+1}) = A_k + B_k,$$

donde

$$\begin{aligned} A_k &= f(x_{k+1}) - f(x_k) - f'(x_k)(x_{k+1} - x_k) - \frac{1}{2} f''(x_k) (x_{k+1} - x_k)^2, \\ &= \int_0^1 (1-u) [f''(x_k + u(x_{k+1} - x_k)) - f''(x_k)] du (x_{k+1} - x_k)^2. \\ B_k &= \frac{1}{4} \frac{f''(x_k)^2 f(x_k) (x_{k+1} - x_k)^2}{f'(x_k)^2}. \end{aligned}$$

De forma similar se comprueba que

$$\varphi(t_{k+1}) = \frac{N}{6} (t_{k+1} - t_k)^3 - \frac{1}{4} \frac{\varphi''(t_k)^2 \varphi(t_k)}{\varphi'(t_k)^3} (t_{k+1} - t_k).$$

Cuando $u \in [0, 1]$, se cumple

$$|x_k + u(x_{k+1} - x_k) - x_0| \leq t_k + u(t_{k+1} - t_k) \leq t_{k+1} < t_1^* \leq t_+$$

Del Lema 2.15 obtenemos que

$$\begin{aligned} \left| \frac{A_k}{f'(x_0)} \right| &\leq N \int_0^1 (1-u) u du (t_{k+1} - t_k)^3 = \frac{N}{6} (t_{k+1} - t_k)^3, \\ \left| \frac{B_k}{f'(x_0)} \right| &\leq -\frac{1}{4} \frac{\varphi''(t_k)^2 \varphi(t_k)}{\varphi'(t_k)^3} (t_{k+1} - t_k). \end{aligned}$$

Por tanto

$$\left| \frac{f(x_{k+1})}{f'(x_0)} \right| \leq \varphi(t_{k+1}),$$

luego (2.94) cumple para $n = k+1$. Así (2.94), (2.95) y (2.96) cumplen para todo $n = 0, 1, \dots$, por lo tanto (2.97) también se cumple. Por lo tanto

$$\lim_{n \rightarrow \infty} x_n = x^*$$

existe. Además, en (2.94), cuando $n \rightarrow \infty$, se obtiene que $f(x^*) = 0$. \square

A modo de ilustración, presentamos a continuación un teorema de convergencia local bajo condiciones tipo Smale en \mathbb{R} . La versión original del mismo aparece en [198] enunciado para espacios de Banach.

Teorema 2.16 (Teorema de Zhen. Condiciones tipo Smale). *Supongamos que $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función analítica. Sea*

$$\alpha = \alpha(x, f) = \beta\gamma, \quad \text{donde } x \in \mathbb{R},$$

$$\beta = \beta(x, f) = \left| \frac{f(x)}{f'(x)} \right|$$

y

$$\gamma = \gamma(x, f) = \sup_{n \geq 2} \left| \frac{f^{(n)}(x)}{n! f'(x)} \right|^{\frac{1}{n-1}}.$$

Si $\alpha = \alpha(x, f) \leq 3 - 2\sqrt{2}$, entonces la función de iteración del método de Halley, cuando $x_0 = x$, está bien definida, además existe $\lim_{n \rightarrow \infty} x_n = x^*$ tal que $f(x^*) = 0$.

Convergencia global del método de Halley

Por considerarlo de interés para esta este acápite, damos la siguiente definición.

Definición 2.7 (Función signo de una función $f(x)$). *Es una función de valor real cuya regla de correspondencia viene dada por:*

$$\text{sgn}(f(x)) = \begin{cases} 1, & \text{si } x > 0, \\ 0, & \text{si } x = 0, \\ -1, & \text{si } x < 0, \end{cases}$$

En [132] encontramos que el método de Halley, al igual que el método de Chebyshev, está basado en una aproximación de tercer orden, que si no degenera en una línea, es una función racional de la forma

$$a + \frac{b}{x+c} \quad (2.101)$$

a la cual se imponen las siguiente 3 condición de aproximación en un punto x_0 :

$$a + \frac{b}{x_0+c} = f(x_0), \quad (2.102)$$

$$-\frac{b}{(x_0+c)^2} = f'(x_0), \quad (2.103)$$

$$\frac{2b}{(x_0+c)^3} = f''(x_0). \quad (2.104)$$

Tomando en cuenta a (2.102), (2.103) y (2.104), Melman [132] demuestra el siguiente teorema de convergencia global.

Teorema 2.17. *Sea $f(x)$ una función tal que en un intervalo I que contiene a la raíz x^* de $f(x)$ se cumple que $f'''(x)$ es continua, $f'(x) = 0$ y $\left(1/\sqrt{n f'(x)}\right)'' \geq 0$, siendo $n = \text{sgn}(f'(x))$. Entonces (2.63), desde cualquier punto del intervalo I , converge monótonamente a la solución x^* .*

Demostración. Nos limitaremos al caso en que $f'(x) > 0$ en I , pues la demostración para $f'(x) < 0$ es análoga.

Sea $x \in I$, asumiremos que $f''(x) \neq 0$, pues así las aproximaciones de (2.63) no degenerarán en una línea. Tomando en cuenta las condiciones de aproximación dadas en (2.102), (2.103) y (2.104), se asegura que para algún punto $x \in I$, se cumple que $b < 0$, pues $f'(x) > 0$.

A partir de las condiciones (2.103) y (2.104) podemos interpretar que el término corrector $\left(\frac{-\varepsilon}{\sqrt{b}}\right)(x+c)$, con $-\varepsilon = \text{sgn}(f''(x))$ aproxima a $\frac{1}{\sqrt{f'(x)}}$ hasta el segundo grado.

Puesto que $f'(x) > 0$ y teniendo en cuenta las condiciones impuestas en el teorema, podemos decir que en I se cumple que $\left(\frac{1}{\sqrt{f'(x)}}\right)'' > 0$, luego $\left(\frac{1}{\sqrt{f'(x)}}\right)$ es convexa, por lo tanto

$$\left(\frac{-\varepsilon}{\sqrt{-b}}\right)(x+c) \leq \left(\frac{1}{\sqrt{f'(x)}}\right).$$

Puesto que sólo estamos interesados en los puntos donde $\left(\frac{-\varepsilon}{\sqrt{-b}}\right)(x+c)$ está por encima del eje x , nos limitaremos a los puntos en los cuales se cumple que

$$0 < \left(\frac{-\varepsilon}{\sqrt{-b}}\right)(x+c) \leq \left(\frac{1}{\sqrt{f'(x)}}\right) \Leftrightarrow \frac{(x+c)^2}{-b} \leq \frac{1}{f'(x)} \Leftrightarrow -\frac{b}{(x+c)^2} \geq f'(x).$$

Luego, para $x \geq x_0$ sigue que

$$\int_{x_0}^x \frac{-b}{(t+c)^2} dt \geq \int_{x_0}^x f'(t) dt \Leftrightarrow a + \frac{b}{x+c} - a - \frac{b}{x_0+c} \geq f(x) - f(x_0).$$

Por otro lado, tomando en cuenta a (2.103) sigue

$$a + \frac{b}{(x+c)} \geq f(x). \quad (2.105)$$

La desigualdad inversa de (2.105) también es válida para $x \leq x_0$.

Cuando $f''(x_0) = 0$ las aproximaciones de (2.63) degeneran en una línea, es decir, la función de aproximación toma la forma $ax + b$ con $a > 0$. Esto significa que la función constante $\frac{1}{\sqrt{a}}$ se aproxima a $\frac{1}{\sqrt{f'(x)}}$ hasta el segundo orden. Esto mismo se deriva de la convexidad de $\frac{1}{\sqrt{f'(x)}}$.

Por tanto, si iniciamos en un punto que está a la izquierda de x^* , la aproximación generada por (2.63) convergerá monótonamente hacia x^* . Si comenzamos en un punto que está a la derecha de la raíz x^* , (2.105) cumple con la desigualdad inversa y la convergencia será monótona desde la derecha. □

En el libro de M. Hernández y A. Salanova [98], aparece un interesante teorema que enfoca la convergencia global de (2.63) a partir del grado de convexidad logarítmico de las funciones $f(x)$ y $f'(x)$. Parte del mismo es presentado a continuación.

Teorema 2.18. *Supongamos que $L_{f'}(x) \leq 3/2$ en un intervalo cerrado $[a, b]$. Si $f(x_0) > 0$ y $L_f(x) < 2$ en $[a, b]$, entonces la sucesión $\{x_n\}$ dada por (2.63) es decreciente y convergente a x^* . Si $f(x_0) < 0$, entonces $\{x_n\}$ es creciente y convergente a x^* .*

Demostración. Supongamos que $x_0 \in [a, b]$ tal que $f(x_0) > 0$. Demostraremos que $x_n \geq x^*$ y que $x_{n-1} \geq x_n$ para todo $n \in \mathbb{N}$.

Notemos que $x_0 - x^* \geq 0$, pues $f(x_0) > 0$. Además que

$$x_1 - x^* = F(x_0) - F(x^*) = F'(\xi_0)(x_0 - x^*),$$

con $F(x) = x_{n+1}$ definida por (2.63).

Teniendo en cuenta que

$$F'(x) = \frac{L_f(x)^2}{(2 - L_f)^2} (3 - 2L_{f'}(x)),$$

y $L_{f'}(x) \leq 3/2$, entonces $F'(x) \geq 0$ en $[a, b]$. Por tanto $x_1 \geq x^*$.

Por otra parte, como $L_f(x) < 2$, sigue

$$x_1 - x_0 = \frac{-f(x_0)}{f'(x_0)} \frac{2}{2 - L_f(x_0)} \leq 0 \Leftrightarrow x_0 \geq x_1.$$

Si continuamos aplicando hipótesis de inducción sobre n , sigue

$$x_n \geq x^*$$

y que

$$x_{n-1} \geq x_n,$$

para todo $n \in \mathbb{N}$. Entonces la sucesión $\{x_n\}$ es decreciente y acotada, por tanto es convergente. Pasando al límite de $\{x_n\}$, se obtiene que $\lim_{x \rightarrow \infty} \{x_n\} = x^*$.

□

En [88], aparece la demostración de un teorema de convergencia global de una familia de métodos iterativos de la forma

$$x_{n+1} = x_n - \left[1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \alpha L_f(x_n)} \right] \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}. \quad (2.106)$$

Al sustituir el parámetro α por la constante $1/2$, $\alpha = 1/2$, obtenemos la función de iteración del método de Halley. El teorema que sigue, es una adaptación del teorema anteriormente mencionado.

Teorema 2.19. *Supongamos que $f(x)$ es una función decreciente y convexa en un intervalo $[0, x^*]$, con $f'''(x) > 0$ en $[0, x^*]$. En estas condiciones, la sucesión dada por (2.59), empezando en $x_0 = 0$ es creciente y converge a x^* , solución de $f(x) = 0$.*

Demostración. Llamamos

$$H(x) = x - \left[1 + \frac{1}{2} \frac{L_f(x)}{1 - \frac{1}{2}L_f(x)} \right] \frac{f(x)}{f'(x)}.$$

Se comprueba que H es creciente en $[0, x^*]$. En efecto,

$$\begin{aligned} H'(x) &= L_f(x) - \frac{1}{2} \frac{L_f'(x)}{\left(1 - \frac{1}{2}L_f(x)\right)^2} \frac{f(x)}{f'(x)} - \frac{1}{2} \frac{L_f(x)}{1 - \frac{1}{2}L_f(x)} (1 - L_f(x)) \\ &= L_f(x) - \frac{1}{2} \frac{1}{\left(1 - \frac{1}{2}L_f(x)\right)^2} \left(L_f(x) + \frac{f^2(x)f'''(x)}{(f'(x))^3} - 2L_f^2(x) \right) - \frac{1}{2} \frac{L_f(x)}{1 - \frac{1}{2}L_f(x)} + \frac{1}{2} \frac{L_f^2(x)}{1 - \frac{1}{2}L_f(x)} \\ &= -\frac{1}{2} \frac{1}{\left(1 - \frac{1}{2}L_f(x)\right)^2} \frac{f^2(x)f'''(x)}{(f'(x))^3} + \frac{1}{2} \frac{L_f^2(x)}{\left(1 - \frac{1}{2}L_f(x)\right)^2} \left(\frac{1}{2}L_f(x) + \frac{3}{2} - \frac{1}{2}L_f^2(x) \right). \end{aligned}$$

El primer sumando es positivo ya que $f'''(x) \geq 0$ y $f'(x) < 0$ para $x \in [0, x^*]$. El segundo sumando es positivo si lo es

$$\frac{1}{2}L_f(x) + \frac{3}{2} - \frac{1}{2}L_f^2(x) = \frac{3}{2} + \frac{1}{2}L_f(x)(1 - L_f(x)).$$

De acuerdo al Lema 2.28 de Altman [4], si f es decreciente y convexa, y además verifica $f'''(x) \geq 0$ para $x \in [0, x^*]$, tenemos que $0 \leq L_f(x) < 1/2$ si $x \in [0, x^*]$.

En consecuencia $H'(t) > 0$ y H es una función creciente en el intervalo $[0, x^*]$. De aquí se sigue que la sucesión $\{x_n\}$ está acotada por x^* . En efecto, $x_1 = H(x_0) = H(0) < H(x^*) = x^*$, y por inducción se sigue que $x_n < x^*, \forall n \in \mathbb{N}$. Como además

$$x_{n+1} - x_n = - \left[1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \frac{1}{2} L_f(x_n)} \right] \frac{f(x_n)}{f'(x_n)} \geq 0$$

se tiene que $\{x_n\}$ también es creciente, y por tanto convergente. □

2.4. El método de Chebyshev

Otro de los métodos para aproximar la solución de ecuaciones no lineales que vamos a considerar en este trabajo es el método de Chebyshev definido, a partir de un valor inicial x_0 , por la recurrencia

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} L_f(x_n) \right) \frac{f(x_n)}{f'(x_n)}. \quad n \geq 0, \quad (2.107)$$

Al igual que en la sección dedicada al método de Halley, hemos expresado el método de Chebyshev en términos del grado de convexidad logarítmico

$$L_f(x_n) = \frac{f(x_n)f''(x_n)}{f'(x_n)^2}.$$

2.4.1. Notas históricas sobre el método de Chebyshev

Históricamente hablando, el método de Chebyshev (2.107), está contenido en una familia de procesos iterativos de la forma $x_{n+1} = E_m(x_n)$, donde

$$E_m(x) = x + \sum_{n=1}^{m-1} (-1)^n \frac{f(x)^n}{n!} (f^{-1})^n(f(x)), \quad (2.108)$$

La familia de métodos definida por (2.108) se conoce como métodos de Schröder de primer tipo. Recordar que el orden de convergencia de cada método definido por (2.108) es m . Notemos que para $m = 2$ se obtiene el método de Newton-Raphson, mientras que para $m = 3$, se obtiene el método de Chebyshev.

Hemos de hacer notar que el nombre de *método de Chebyshev* no se ha utilizado de forma uniforme en la literatura matemática. De hecho, en varias referencias bibliográficas la expresión (2.107) recibe varios nombres, como puede verse en el listado siguiente.

1. Método de Euler: [25], [118], [180].
2. Método de Euler–Chebyshev: [118], [180].

3. Método Súper-Newton: [123].
4. Método de Schröder: [82].
5. Método de las parábolas tangentes: [10].
6. Método de la interpolación cuadrática inversa: [14], [76], [98]

La controversia existente acerca de la autoría de (2.107) (o de (2.108)) no es nueva. En efecto, en el conocido libro de J. F. Traub [180, p. 81] sobre métodos iterativos para resolver ecuaciones algebraicas se afirma que la familia E_m se atribuye tanto a Euler como a Chebyshev. En consecuencia, Traub usa ambos nombres de forma conjunta para referirse a la familia E_m .

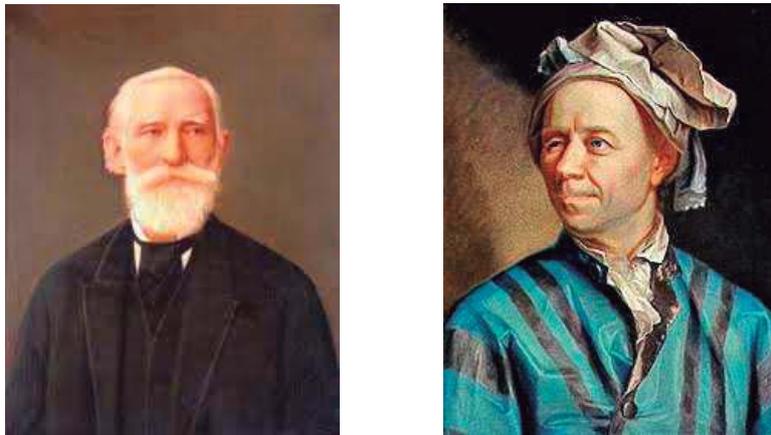


Figura 2.11: P. Chebyshev, (1821 – 1894) y L. Euler, (1707 – 1783).

La duda sobre a quién atribuir el método (2.107) es compartida por otros autores, como por ejemplo M. S. Petković, L. D. Petković y D. Herceg [148, p. 1759], quienes advierten que en la literatura rusa la secuencia E_m definida en (2.108) se atribuye a un trabajo de Chebyshev escrito en 1837 o 1838 titulado *Cálculo de las raíces de una ecuación*, y hacen referencia a la cita anterior de Traub sobre este aspecto. Sin embargo, también indican que otros autores como E. Bodewig [25] o S. Smale y M. Shub [167] adjudican E_m a Euler y, en concreto, a su *Opera Omnia*, Ser. I, Vol. X.

En el trabajo de M. García y J.M. Gutiérrez, [78], se han pretendido resolver la disyuntiva surgida en cuanto a la autoría del método de Chebyshev, analizando los textos históricos en los que el método pudo ser desarrollado por primera vez. Según [78], la única edición existente del *Cálculo de las raíces de una ecuación* es la que aparece en ruso en el Capítulo 5 de las obras completas de Chebyshev [38]. En la Sección II de dicho trabajo, Chebyshev establece un método general que permita encontrar las raíces de una ecuación de la forma

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0. \quad (2.109)$$

Para lograr dicho objetivo, se basa en la aplicación de series de potencia, en concreto, dado un desarrollo de la forma

$$Y = \sum_{n=1}^{\infty} c_n X^n,$$

indica un procedimiento para calcular el desarrollo en serie de

$$X = \sum_{n=1}^{\infty} d_n Y^n,$$

expresando los coeficientes de d_n en función de c_n como

$$\begin{aligned} d_1 &= \frac{1}{c_1}. \\ d_2 &= -\frac{c_2}{c_1^3}. \\ d_3 &= \frac{2c_2^2 - c_1 c_3}{c_1^5}. \\ d_4 &= \frac{-5c_2^3 + 5c_1 c_2 c_3 - c_1^2 c_4}{c_1^7}. \end{aligned}$$

Para obtener la función inversa, Chebyshev aplicó la misma idea. Así, sea x_0 una aproximación inicial de la raíz x^* , luego

$$y = f(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n, \quad (2.110)$$

donde

$$c_0 = \beta = f(x_0), c_n = \frac{f^{(n)}(x_0)}{n!} (n \geq 1).$$

Definido β , Chebyshev aplica el desarrollo en serie de la función inversa en la forma

$$x = f^{-1}(y) = \sum_{n=0}^{\infty} d_n (y - \beta)^n, \quad (2.111)$$

luego, en vez de hacer $y = 0$ en (2.110) y resolver la ecuación obtenida con respecto a x , Chebyshev sustituye $y = 0$ en (2.111), obteniendo

$$x = x^* + \sum_{n=1}^{\infty} (-1)^n d_n \beta^n.$$

A partir de esta expresión, se obtiene el desarrollo

$$x^* = x_0 - \frac{f(x_0)}{f'(x_0)} - \left(\frac{f(x_0)}{f'(x_0)}\right)^2 \frac{f''(x_0)}{2f'(x_0)} - \left(\frac{f(x_0)}{f'(x_0)}\right)^3 \left(\frac{f''(x_0)^2}{2f'(x_0)^2} - \frac{f'''(x_0)}{6f'(x_0)}\right) - \dots \quad (2.112)$$

Según [78] el desarrollo obtenido por Chebyshev (2.112), coincide con el obtenido por Euler, véase a [74], sin embargo, Chebyshev se apoya en argumentos matemáticos para demostrar que encontrar una x^* tal que $f(x^*) = 0$, es lo mismo que encontrar a $f^{-1}(0)$. En síntesis, tanto

Euler como Chebyshev aplican un método de aproximación de la función inversa mediante funciones polinomiales.

En la Sección VI del *Cálculo de las raíces de una ecuación*, Chebyshev trunca en el tercer término la fórmula dada en (2.112), obteniendo

$$x^* \approx x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} - \left(\frac{f(x_0)}{f'(x_0)}\right)^2 \frac{f''(x_0)}{2f'(x_0)},$$

y la estimación del error

$$|x^* - x_1| \leq \text{máx} \left| f(x_0)^3 \left(\frac{f'(x_0)^2}{6f(x_0)^2} - \frac{f''(x_0)}{6f(x_0)^2} \right) \right|$$

Es en esta sección en donde Chebyshev estudia el método definido por (2.107) tal y como lo conocemos hoy en día. Además, Chebyshev comenta que la aproximación a la raíz será mayor, si se consideran 5, 6 o más términos de la serie (2.112), advirtiéndole que de hacerse así, los cálculos podrían resultar más costosos. Con estos comentarios, Chebyshev intuitivamente se está refiriendo a los conceptos de orden de convergencia y de eficiencia computacional, aspectos que posteriormente fueron formalizados por Schröder. Para mayores detalles de estas formalizaciones, véase a [162] y [?].

2.4.2. Descripción del método de Chebyshev

El método de Chebyshev, junto con sus mejoras y modificaciones ha llamado la atención de muchos investigadores. Por ejemplo, desde el punto de vista histórico, podemos encontrar una forma equivalente para escribir (2.107) conocida como fórmula de Schröder [162]. Como una muestra de recientes publicaciones sobre este tema, podemos citar [5, 6], [42, 43, 44], [90], [120], [132].

En el presente capítulo, nos dedicaremos a mostrar la construcción del método de Chebyshev de distintas maneras. En primer lugar, usaremos una forma clásica de construcción de métodos iterativos, que consiste en la aproximación de la función inversa de $f(x)$. Esta forma se conoce como «interpolación cuadrática inversa», la misma a sido estudiada por autores como [4], [132] y [180]. Otra forma de deducción del método de Chebyshev que presentamos, es el método de la parábola tangente, véase [5], [157], [158] y [193]. Dicho método consiste en tomar una parábola que tenga algún punto de tangencia respecto a la función $f(x)$, y a continuación con un poco de manipulación algebraica se obtiene la función iterativa de Chebyshev.

El capítulo finaliza con una forma muy particular de deducir una familia de métodos iterativos, entre los cuales está el método de Chebyshev, para encontrar ceros de ecuaciones, que tiene como base, los métodos de Obreshkov, (refiérase a [92], [125] y [140]), que son métodos para resolver ecuaciones diferenciales.

Interpolación cuadrática inversa

A través del tiempo se han desarrollado una gran variedad de técnicas que permiten construir un algoritmo iterativo numérico. Algunas de estas técnicas, como la interpretación geométrica, el desarrollo de polinomios en serie de Taylor, ya han sido aplicadas en este sentido. Otra técnica es la de aproximar una función inversa usando el desarrollo de Taylor hasta un polinomio de segundo grado, la cual utilizaremos para encontrar la expresión del método de Chebyshev, conocido también como método de la interpolación cuadrática inversa, (véase nuevamente a [4], [132] y [180]).

Dada la función $y = f(x)$, llamaremos $x = \phi(y)$ a la función inversa de f . Si la aproximamos en un punto y_0 , entonces

$$\phi(y) \approx \phi(y_0) + \phi'(y_0)(y - y_0) + \frac{1}{2}\phi''(y_0)(y - y_0)^2.$$

Encontrar la raíz α de la ecuación $f(x) = 0$, es lo mismo que encontrar la imagen del cero por medio de la función ϕ , es decir a $\phi(0)$. Así, si x_0 es un valor próximo a α y $f(x_0) = y_0$, entonces,

$$\phi(0) \approx \phi(y_0) - \phi'(y_0)y_0 + \frac{1}{2}\phi''(y_0)(y_0)^2.$$

De aquí podemos deducir una nueva aproximación de α , que denotamos x_1 :

$$x_1 = x_0 - \phi'(y_0)f(x_0) + \frac{1}{2}\phi''(y_0)f(x_0)^2. \quad (2.113)$$

Ahora calculamos $\phi'(y)$ y $\phi''(y)$ para sustituirlo en (2.113). De antemano sabemos que $x = \phi(y)$, $y' = f'(x)$ e $y'' = f''(x)$, de modo que ahora, por el teorema de la función inversa,

$$\phi'(y) = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{y'} = \frac{1}{f'(x)}, \quad (2.114)$$

y

$$\phi''(y) = \frac{d\phi'(y)}{dy} = \frac{d(1/f'(x))}{dx} \cdot \frac{dx}{dy} = -\frac{y''}{(y')^3}. \quad (2.115)$$

Sustituyendo (2.115) y (2.114) en (2.113) se sigue

$$x_1 = x_0 - \left(1 + \frac{1}{2}L_f(x_0)\right) \frac{f(x_0)}{f'(x_0)}.$$

Repitiendo para este punto x_1 el mismo proceso anterior conseguiremos la nueva aproximación x_2 , luego por recurrencia obtenemos a (2.107).

Método de las parábolas tangentes

Una forma sencilla de obtener el método iterativo de Chebyshev, es por medio de una derivación geométrica a partir de una parábola tangente a la gráfica de la función cuya raíz se desea aproximar, como ilustramos en la Figura 2.12.

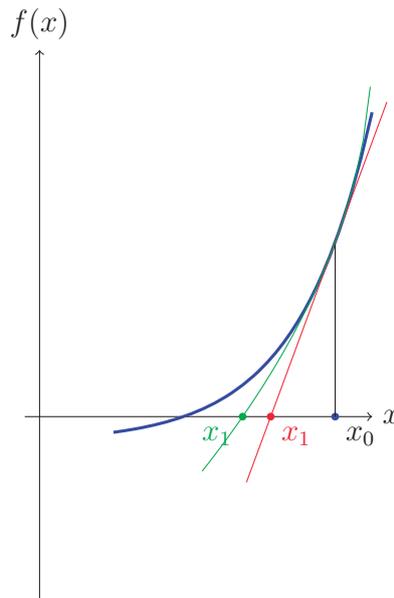


Figura 2.12: Interpretación geométrica del método de Chebyshev.

Consideramos una parábola de la forma

$$ay^2 + y + bx + c = 0. \quad (2.116)$$

La idea consiste en aproximar una función $f(x)$ tomando un punto x_n próximo a la raíz $f(x) = 0$. Si ahora se tomamos en cuenta que la función debe cumplir las condiciones de tangencia siguientes

$$y(x_n) = f(x_n), \quad (2.117)$$

$$y'(x_n) = f'(x_n), \quad (2.118)$$

$$y''(x_n) = f''(x_n). \quad (2.119)$$

Podemos escribir la parábola (2.116) de la forma conveniente

$$a(y - f(x_n))^2 + (y - f(x_n)) + b(x - x_n) = 0 \quad (2.120)$$

de modo que si $x = x_n$, entonces $y(x_n) = f(x_n)$, este hecho cumple la primera condición de tangencia, ahora determinamos los parámetros a y b para que $y'(x_n) = f'(x_n)$, $y''(x_n) = f''(x_n)$. Si derivamos la expresión (2.120), nos queda

$$2a(y(x_n) - f(x_n))y'(x_n) + y'(x_n) + b = 0 \quad (2.121)$$

de modo que en $(x_n, f(x_n))$ nos queda sólo

$$y'(x_n) + b = 0$$

o lo que es lo mismo

$$b = -y'(x_n) = -f'(x_n),$$

como podemos ver, sólo falta derivar la expresión (2.121) y despejar a , como sigue

$$2a(y'(x_n))^2 + 2a(y(x_n) - f(x_n))y''(x_n) + y''(x_n) = 0.$$

Esta expresión en el punto de intersección $(x_n, f(x_n))$ es

$$2a(y'(x_n))^2 + y''(x_n) = 0.$$

Si ahora despejamos a , se tiene que

$$a = -\frac{y''(x_n)}{2y'(x_n)^2} = -\frac{f''(x_n)}{2f'(x_n)^2}.$$

Luego de sustituir los valores a y b encontrados en la expresión (2.120), la podemos escribir como

$$-\frac{f''(x_n)}{2f'(x_n)^2}(y - f(x_n))^2 + y - f(x_n) - f'(x_n)(x - x_n) = 0.$$

La intersección de esta curva con el eje OX, es decir cuando $y = 0$ nos daría la siguiente iteración del método de Chebyshev:

$$\begin{aligned} &-\frac{f(x_n)^2 f''(x_n)}{2f'(x_n)^2} - f(x_n) - f'(x_n)(x - x_n) = 0 \\ f'(x_n)(x - x_n) &= -\frac{f(x_n)^2 f''(x_n)}{2f'(x_n)^2} - f(x_n) \\ x - x_n &= -\frac{f(x_n)^2 f''(x_n)}{2f'(x_n)^3} - \frac{f(x_n)}{f'(x_n)} \\ x &= x_n - \left(1 + \frac{f(x_n) f''(x_n)}{2f'(x_n)^2}\right) \frac{f(x_n)}{f'(x_n)}, \end{aligned}$$

como puede verse, esta es la expresión del método de Chebyshev.

La Figura 2.12 es la interpretación geométrica del método de Chebyshev por medio de la parábola tangente, donde la línea recta en color rojo representa la recta tangente de la primera iteración del método de Newton, la curva de color verde es la parábola aproximante que emplea el método de Chebyshev y la curva de color azul es la función cuyas raíces queremos aproximar.

Para ilustrar esta técnica tomamos como ejemplo la situación que se refleja en la Figura 2.12. En este caso, se pretende aproximar la raíz de la función $f(x) = e^x - 7$. Luego de hacer los cálculos apropiados, aplicando tanto el método de Newton como el método de Chebyshev, y tomando como punto de partida $x_0 = 4$, se han obtenido la recta tangente

$$y = f(x_0) + f'(x_0)(x - x_0) = 47.5982 + 54.5982(x - 4)$$

y la parábola

$$\begin{aligned} & -\frac{f''(x_0)}{2f'(x_0)^2}(y - f(x_0))^2 + y - f(x_0) - f'(x_0)(x - x_0) \\ & = -\frac{54.5982}{2 \cdot 54.5982^2}(y - 47.5982)^2 + y - 47.5982 - 54.5982(x - 4) = 0, \end{aligned}$$

respectivamente. Calculando las intersecciones de ambas curvas con el eje $y = 0$, se obtienen las aproximaciones $x_1 = 3.12821$ para el método de Newton y $x_1 = 2.7482$ para el método de Chebyshev. Como vemos tanto gráfica como numéricamente, en este caso, la aproximación dada por el método de Chebyshev está más cercana a la solución exacta del problema que es $\log 7 \approx 1.94591$.

Otra interpretación geométrica: interpolación exponencial

Siguiendo un procedimiento similar al aplicado por Kanwar en [23] y [112], pero considerando ahora la curva de aproximación

$$y(x) = e^{a(y-y(x_n))}(b(x-x_n) + c), \quad (2.122)$$

donde a , b y c son tres parámetros a determinar, llegamos a lo siguiente:

$$\begin{aligned} y(x_n) &= f(x_n) \Rightarrow c = f(x_n) \\ y'(x_n) &= f'(x_n) \Rightarrow b = f'(x_n)(1 - af(x_n)) \end{aligned}$$

Los nuevos métodos iterativos de la forma $x_{n+1} = x_n - c/b$ se convierten ahora en

$$x_{n+1} = x_n - \frac{1}{1 - af(x_n)} \frac{f(x_n)}{f'(x_n)} \quad (2.123)$$

Si buscamos ahora que los métodos de (2.122) tengan convergencia cúbica, debemos exigir que $y''(x_n) = f'(x_n)$, es decir,

$$a^2 f(x_n) f'(x_n)^2 - a(f(x_n) f''(x_n) + 2f'(x_n)^2) + f''(x_n) = 0 \quad (2.124)$$

Para valores de a muy próximos a cero y despreciando el término en a^2 de (2.124), se obtiene que

$$a = \frac{f''(x_n)}{f(x_n) f''(x_n) + 2f'(x_n)^2} \quad (2.125)$$

Sustituyendo (2.125) en (2.124) se obtiene el método de Chebyshev

$$x_{n+1} = x_n - \left[1 + \frac{1}{2} L_f(x_n)\right] \frac{f(x_n)}{f'(x_n)}$$

Siguiendo los métodos de Obreshkov

Los métodos de Obreshkov fueron introducidos en [92] con el objetivo de combinar los atributos de los métodos multipaso, métodos multietapas y métodos de derivadas múltiples para la resolución de ecuaciones diferenciales. En efecto, para una ecuación autónoma y para $m \geq 1$, definimos los operadores diferenciales lineales \mathcal{L} y \mathcal{M} :

$$\begin{aligned} \mathcal{L}(x(y); h, h_1, \dots, h_{m-1}) = & x(y+h) - x(y) - h \sum_{k=1}^m [\beta_k D(x)(y_k) + \\ & + \gamma_k H_k D^2(x)(y_k)] \end{aligned}$$

y

$$\begin{aligned} \mathcal{M}(x(y); h, h_1, \dots, h_{m-1}) = & x(y+h) - x(y) - h \left[\sum_{k=1}^m \beta_k D(x)(y_k) + \right. \\ & \left. + \sum_{k=1}^{m-1} \gamma_k H_k D^2(x)(y_k) \right], \end{aligned}$$

donde

$$y_k = y - \sum_{\ell=1}^{k-1} h_\ell, \beta_k \quad \text{y} \quad \gamma_k, 1 \leq k \leq m$$

son constantes reales a determinar, y H_k está definida por

$$H_k = \begin{cases} h, & \text{si } k = 1, \\ \sum_{\ell=1}^{k-1} h_\ell, & \text{si } k > 1. \end{cases}$$

Consideremos una función $x(y)$ suficientemente diferenciable en torno a y . Entonces al expandir $x(y+h)$, $x'(y_k)$ y $x''(y_k)$ en torno a y , obtenemos el siguiente tipo de serie:

$$\sum_{q \geq 0} C_q x^{(q)}(y) \mathcal{H}^q,$$

donde la C_q es una constante real y $\mathcal{H}^q = h^{q_0} h_1^{q_1} \dots h_{m-1}^{q_{m-1}}$ con $q_0 + q_1 + \dots + q_{m-1} = q$. Decimos que el proceso desarrollado es de orden d cuando $C_0 = C_1 = \dots = C_d = 0$ y $C_{d+1} \neq 0$.

Decimos que una función que depende de m variables x_0, \dots, x_{m-1} es de la forma

$$O_q(x_0, \dots)$$

si resulta que

$$O(x_0^{q_0} x_1^{q_1} \dots x_{m-1}^{q_{m-1}}),$$

donde

$$q_0 + q_1 + \dots + q_{m-1} = q, \quad q_i \geq 0, i = 0, 1, \dots, m-1.$$

Entonces, una elección adecuada de β y γ nos permite encontrar el máximo orden de convergencia.



Figura 2.13: N. Obreshkov, (1896 – 1963)

Nuestro objetivo en este apartado, es presentar una nueva técnica para obtener el método de Chebyshev teniendo como base la esencia de los métodos de Obreshkov (ver nuevamente a [125], [140]), que fueron ideados, en principio, para la solución numérica de problemas de valor inicial.

La clave de la idea es la siguiente. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función con una raíz simple α , es decir, $f(\alpha) = 0$ y $f'(\alpha) \neq 0$ en una vecindad I_α de α . Entonces, la función inversa $x = g(y)$ está bien definida en $J_0 \subseteq f(I_\alpha)$ y satisface la relación $g(0) = \alpha$. Al derivar la función $x(y)$ con respecto a la variable y , obtenemos

$$\frac{dx}{dy} = g'(y) = \frac{1}{f'(x)}. \quad (2.126)$$

La ecuación (2.126) puede ser escrita como

$$D(x)(y) = x'(y) = F(x), \quad (2.127)$$

donde

$$F(x) = \frac{1}{f'(x)}. \quad (2.128)$$

Si x_0 es una primera aproximación de α , a la raíz simple de f , y denotamos $y_0 = f(x_0)$.

Definimos ahora el operador \mathcal{L} como

$$\begin{aligned} \mathcal{L}(x(y); h, h_1, \dots, h_{m-1}) &= x(y+h) - x(y) - h \sum_{k=1}^m [\beta_k D(x)(y_k) + \\ &+ \gamma_k H_k D^2(x)(y_k)]. \end{aligned} \quad (2.129)$$

Ahora, siguiendo esta técnica, deducimos el método de Chebyshev en términos del operador \mathcal{L} definido en (2.129). Para $m = 1$ la ecuación (2.129) se convierte en

$$\mathcal{L}(x(y); h) = x(y+h) - x(y) - h \left(\beta_1 D(x)(y) + h\gamma_1 D^{(2)}(x)(y) \right), \quad (2.130)$$

donde β_1 y γ_1 son las constantes reales a determinar. Desarrollando $\mathcal{L}(x(y); h)$ en series de Taylor en torno a y se obtiene

$$\mathcal{L}(x(y); h) = (1 - \beta)hD(x)(y) + \left(\frac{1}{2} - \gamma_1\right)h^2D^{(2)}(x)(y)O_3(h).$$

Si en la ecuación anterior hacemos $\beta = 1$ y $\gamma = \frac{1}{2}$, entonces encontramos el máximo orden de aproximación de $x(y+h)$. Sustituyendo este valor en (2.130) llegamos al método iterativo de la forma

$$x^+ = x_n + h \left(D(x)(y) + \frac{h}{2}D^{(2)}(x)(y) \right).$$

En otras palabras, iniciando con x_0 , la secuencia obtenida es definida por

$$x_{n+1} = x_n + h \left(F(x_n) + \frac{h}{2}F(x_n)F'(x_n) \right).$$

Hemos tomado en cuenta el hecho de que $D(x)(y) = F(x)$ y $D^{(2)}(x)(y) = F'(x)F(x)$, con $F(x)$ definida en (2.128).

Es importante observar que la clave de este aporte consiste en la elección del paso de integración h . En este caso, consideramos $h = -f(x_n)$. Como $F(x_n) = \frac{1}{f'(x_n)}$ y $F'(x_n) = -\frac{f''(x_n)}{f'(x_n)^2}$, el método derivado correspondiente es el método de Chebyshev (2.107).

2.4.3. Convergencia del método de Chebyshev

Convergencia local del método de Chebyshev

Como en el caso de los métodos de Newton y Halley, mostramos a continuación un teorema de convergencia local para el método de Chebyshev.

Teorema 2.20. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz simple de $f(x) = 0$, ($f(x^*) = 0$ y $f'(x^*) \neq 0$). Supongamos que f es tres veces derivable con continuidad en I . Entonces el método de Chebyshev (2.107) definido por $x_{n+1} = C_f(x_n)$,*

$$C_f(x) = x - \left(1 + \frac{1}{2}L_f(x)\right) \frac{f(x)}{f'(x)},$$

tiene al menos convergencia cúbica localmente. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = (2A_2^2 - A_3)e_n^3 + O_4.$$

donde A_i , e_n y O_N se han introducido en (2.12), (2.13) y (2.14) respectivamente.

Demostración. (Usando el Teorema de Schröder) El cálculo de las derivadas de C_f nos permite deducir que

$$C_f(x^*) = x^*, \quad C'_f(x^*) = C''_f(x^*) = 0, \quad C'''_f(x^*) = \frac{3f''(x)^2 - f'(x)f^{(3)}(x)}{f'(x)^2} \neq 0.$$

Por lo tanto,

$$e_{n+1} = x_{n+1} - x^* = C_f(x_n) - x^* = \frac{C_f'''(x^*)}{3!}(x_n - x^*)^3 + O_4,$$

de donde se deduce que

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^3} = \frac{C_f'''(x^*)}{3!} = 2A_2^2 - A_3.$$

Por lo tanto, el método de Chebyshev es un método con orden de convergencia local igual a 3 y contante de error asintótico $2A_2^2 - A_3$. \square

Demostración. (Usando desarrollos de Taylor) De nuevo, podemos obtener las mismas conclusiones del teorema siguiendo usando desarrollos en serie de Taylor. En concreto, teniendo en cuenta las notaciones A_i , e_n y O_N introducidas en (2.12), (2.13) y (2.14) respectivamente, deducimos que

$$\frac{f(x_n)}{f'(x_n)} = e_n - A_2 e_n^2 + 2(A_2^2 - A_3)e_n^3 + O_4.$$

$$\frac{L_f(x_n)}{2} = A_2 e_n + 3(A_3 - A_2^2)e_n^2 + (14A_2 A_3 - 8A_2^2)e_n^3 + O_4.$$

En consecuencia,

$$\frac{1}{2}L_f(x_n)\frac{f(x_n)}{f'(x_n)} = A_2 e_n^2 + (3A_3 - 4A_2^2)e_n^3 + O_4.$$

Por lo tanto,

$$\begin{aligned} e_{n+1} &= x_{n+1} - x^* = x_n - x^* - \frac{f(x_n)}{f'(x_n)} - \frac{1}{2}L_f(x_n)\frac{f(x_n)}{f'(x_n)}. \\ &= e_n - \left(e_n - A_2 e_n^2 + 2(A_2^2 - A_3)e_n^3 + O_4 \right) - \left(A_2 e_n^2 + (3A_3 - 4A_2^2)e_n^3 + O_4 \right). \\ &= (2A_2^2 - A_3)e_n^3 + O_4, \end{aligned}$$

con lo que hemos demostrado el resultado. \square

Convergencia semilocal del método de Chebyshev

En esta sección demostraremos un teorema de convergencia semilocal del método de Chebyshev que aparece en [61], como una versión en \mathbb{R} que podrá ser extendida a espacios más generales.

A partir del estudio de la convergencia semilocal del método de Chebyshev, podemos ver que el tema ha sido ampliamente tratado, véase a [11], [14], [30], [68] y [72]. En esta sección aplicaremos nuevamente el análisis de Kantorovich [110], para demostrar la convergencia semilocal del método de Chebyshev. En ese sentido, en esta sección asumiremos las siguientes condiciones:

(i) Existe un punto $x_0 \in \mathbb{R}$ tal que $f'(x_0) \neq 0$.

(ii) $|\frac{f''(x)-f''(y)}{f'(x_0)}| \leq k|x-y|$, $x, y \in \mathbb{R}$, $k > 0$.

(iii) $|\frac{f(x_0)}{f'(x_0)}| \leq a$, $|\frac{f''(x_0)}{f'(x_0)}| \leq b$.

(iv) El polinomio

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a = 0, \quad (2.131)$$

tiene una raíz negativa y dos raíces positivas, r_1 y r_2 , con $r_1 \leq r_2$.

(v) $S = \{x : |x - x_0| \leq r_1\} \subseteq \mathbb{R}$.

Previo a la demostración, se presentaran los siguientes lemas técnicos.

Lema 2.21. *Definamos la sucesión escalar $\{t_n\}$ mediante*

$$t_0 = 0, \quad t_{n+1} = H(t_n) = t_n - (1 + \frac{1}{2}L_p(t_n))\frac{p(t_n)}{p'(t_n)}, \quad n \geq 0, \quad (2.132)$$

donde $p(t)$ está definido en (2.131) y $L_p(t)$ se ha definido en (2.11). Entonces $\{t_n\}$ es monótonamente creciente y converge a r_1 , la cual es la raíz positiva más pequeña de $p(t)$.

Demostración. Dado que $p(t_0) = a > 0$, sabemos que $t_0 < r_1$. Luego, en virtud del Teorema del Valor Medio tenemos que para algún $s_0 \in (t_0, r_1)$ se cumple que

$$t_1 - r_1 = H'(s_0)(t_0) - r_1.$$

Notemos que

$$H'(t) = \frac{L_{p'}(t)^2}{2(1 - L_{p'}(t))^2}(1 - L_{p'}(t))^2(3 - L_{p'}(t)),$$

donde

$$L_{p'}(t) = \frac{p'(t)p'''(t)}{p''(t)^2}.$$

Luego, tomando en cuenta que $p(t)$ es positivo, decreciente y convexo en el intervalo $[0, r_1]$, sigue que $H'(t) \geq 0$ en $[0, r_1]$. Por lo tanto $t_1 \leq r_1$.

Por otro lado, tenemos que

$$t_1 - t_0 = -\frac{p(t_0)}{p'(t_0)}(1 + \frac{1}{2}L_p(t_0)) \geq 0.$$

Luego, para todo $n \geq 1$, por inducción matemática se obtiene que $t_n \leq r_1$ y $t_{n-1} \geq t_n$, ya que $(t_{n-1}, r_1) \subset (t_0, r_1)$. Luego la sucesión (2.132) converge a r_1 .

□

Lema 2.22. Sea ρ la única raíz positiva de $p'(t) = 0$, donde $p(t)$ es el polinomio definido por (2.131). Bajo las condiciones (i) – (v) definidas anteriormente, si

$$|x - x_0| \leq \rho,$$

entonces $1/f'(x)$ existe y se cumple que

$$\begin{aligned} \left| \frac{f'(x_0)}{f'(x)} \right| &\leq \frac{1}{p'(|x - x_0|)}, \\ \left| \frac{f''(x_0)}{f'(x_0)} \right| &\leq p''(|x - x_0|) \end{aligned}$$

Demostración. Notemos que

$$p'(t) = \frac{k}{2}t^2 + bt - 1$$

tiene dos raíces reales dadas por

$$t_+ = \frac{-b + \sqrt{b^2 - 2k}}{k}, \quad t_- = \frac{-b - \sqrt{b^2 - 2k}}{k}.$$

Entonces, cuando $|x - x_0| < t_+$ sigue que

$$p'(|x - x_0|) = \frac{k}{2}|x - x_0|^2 + b|x - x_0| - 1 < 0.$$

Así, bajo las condiciones del teorema de Kantorovich tenemos que

$$\left| \frac{f''(x_0)(x - x_0)}{f'(x_0)} + \int_0^1 \frac{f''(x_0 + t(x - x_0)) - f''(x_0)}{f'(x_0)} dt(x - x_0) \right| \leq b|x - x_0| + \frac{k}{2}|x - x_0|^2 < 1.$$

Tomando en cuenta el Teorema del Valor Medio, sigue que

$$f'(x) = f'(x_0) + \int_0^1 f''(x_0 + t(x - x_0)) dt(x - x_0),$$

de donde

$$\frac{f'(x)}{f'(x_0)} = 1 + \frac{f''(x_0)(x - x_0)}{f'(x_0)} + \int_0^1 \frac{f''(x_0 + t(x - x_0)) - f''(x_0)}{f'(x_0)} dt(x - x_0).$$

Luego, en virtud del Lema de Neumann, $\frac{f'(x_0)}{f'(x)}$ existe. Entonces

$$\left| \frac{f'(x_0)}{f'(x)} \right| \leq \frac{1}{1 - b|x - x_0| - \frac{k}{2}|x - x_0|^2} = \frac{1}{p'(t)}$$

y

$$\left| \frac{f''(x)}{f'(x_0)} \right| \leq \left| \frac{f''(x_0)}{f'(x_0)} \right| + \left| \frac{f''(x_0) - f''(x)}{f'(x_0)} \right| \leq b + k|x - x_0| = p''(|x - x_0|).$$

Luego el Lema está probado. □

Lema 2.23. *Con la misma notación y supuestos anteriores, podemos escribir $f(x_{n+1})$ de la manera siguiente*

$$f(x_{n+1}) = \frac{1}{8}f''(x_n)y_n^2 + \frac{1}{2}\frac{f''(x_n)f(x_n)y_n}{f'(x_n)} + \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)](x_{n+1} - x)dx, \quad (2.133)$$

donde

$$y_n = \frac{L_f(x_n)f(x_n)}{f'(x_n)}.$$

Demostración. Por la fórmula de Taylor y tomando en cuenta a ((2.11)) deducimos que

$$\begin{aligned} f(x_{n+1}) &= f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{1}{2}f''(x_n)(x_{n+1} - x_n)^2 + \int_{x_n}^{x_{n+1}} (f''(x) - f''(x_n))(x_{n+1} - x)dx \\ &= -\frac{1}{2}\frac{f''(x_n)(f(x_n))^2}{(f'(x_n))^2} + \frac{1}{2}\frac{f''(x_n)(f(x_n))^2}{(f'(x_n))^2} + \frac{1}{8}f''(x_n)y_n^2 + \frac{1}{2}\frac{f''(x_n)f(x_n)y_n}{f'(x_n)} \\ &\quad + \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)](x_{n+1} - x)dx, \end{aligned}$$

cumpléndose así el Lema. □

Veamos a continuación el teorema

Teorema 2.24. *Asumamos que se cumplen las condiciones (i) – (v). Entonces la sucesión $\{x_n\}$ definida por el método de Chebyshev, (2.107), está bien definida, pertenece a S y converge a x^* solución de $f(x) = 0$. Además se satisface la siguiente estimación de error*

$$|x^* - x_n| \leq r_1 - t_n,$$

donde $\{t_n\}$ es la sucesión definida en (2.132) y r_1 es la raíz positiva más pequeña de $p(t)$.

Demostración. Siguiendo a [200] probaremos que para todo entero no negativo n , se cumplen las siguientes condiciones:

1. $x_n \in S = \{x : |x - x_0| \leq r_1\}$.
2. $1/f'(x_n)$ existe.
3. $|\frac{f(x_n)}{f'(x_0)}| \leq p(t_n)$
4. $|\frac{f'(x_0)}{f'(x_n)}| \leq -\frac{1}{p(t_n)}$
5. $|\frac{f''(x_n)}{f'(x_0)}| \leq p''(t_n)$
6. $|x_{n+1} - x_n| \leq t_{n+1} - t_n$.

Notemos que la condición 6 es inmediata de las otras cinco, además, las condiciones (i) – (v) garantizan que las condiciones 1 – 5 se cumplan para $n = 0$. Asumamos que las condiciones 1 – 5 se cumplen para $n \leq k$, luego la condición 6 también se cumple, por tanto

$$|x_{n+1} - x_0| \leq t_{n+1} - t_0 < r_1 \leq \rho,$$

en consecuencia $x_{n+1} \in S$.

Por el Lema 2.22 tenemos que $1/f'(x_{n+1})$ existe, luego las condiciones 4 y 5 se cumplen para $n = k + 1$.

Por otro lado, en virtud del Lema 2.23 sigue que

$$\begin{aligned} \left| \frac{f(x_{k+1})}{f'(x_0)} \right| &= \frac{1}{8} \left| \frac{f''(x_k)y_k^2}{f'x_0} \right| + \frac{1}{2} \left| \frac{f''(x_k)f(x_k)y_k}{f'x_0f'(x_k)} \right| \\ &+ \left| \int_0^1 \frac{f''(x_k + u(x_{k+1} - x_k) - f''(x_k))}{f'(x_0)} (1 - u) du (x_{k+1} - x_k) \right| \\ &\leq \frac{1}{8} L_p(t_k)^3 p(t_k) + \frac{1}{2} L_p(t_k)^2 p(t_k) + \frac{k}{6} (t_{k+1} - t_k)^3 = p(t_{k+1}), \end{aligned}$$

por tanto la condición 3 también se cumple para $n = k + 1$. Luego las condiciones 1 – 5 se cumplen para todo $n \geq 0$, y como consecuencia de esto, la condición 6 también se cumple para todo $n \geq 0$.

Como $\{t_n\}$ es una sucesión convergente, la condición 6 implica que $\{x_n\}$ también es convergente. Llamemos x^* al límite de $\{x_n\}$, luego, por la condición 3 se deduce que cuando $n \rightarrow \infty$ se cumple que $f(x^*) = 0$, por tanto x^* es solución de $f(x) = 0$.

Finalmente, la estimación del error se deduce aplicando recursivamente la condición

$$|x_{n+j} - x_n| \leq t_{n+j} - t_n, \quad j \geq 1,$$

cuando hacemos que $j \rightarrow \infty$.

□

Convergencia global del método de Chebyshev

En [6] encontramos un novedoso teorema de convergencia global. Parte del mismo se presenta a continuación.

Teorema 2.25. *Sea $f(x)$ una función tal que $f'''(x)$ es continua en un intervalo I que contiene la raíz x^* de $f(x)$. Sea $f'(x) \neq 0$, $L_f(x) > -2$ y $\left((n/f'(x))^2 \right)'' \geq 0$ en I , con $n = \text{sgn}(f'(x))$. Entonces (2.107), desde cualquier punto de I , converge monótonamente a la solución x^**

Demostración. La demostración estará limitada al caso en que $f'(x) > 0$ en I , pues para $f'(x) < 0$ la demostración es similar.

En primer lugar partiremos de un punto que está a la izquierda de x^* , es decir, asumiremos que $x_0 \leq x^*$, para demostrar que la intersección \hat{x} de la parábola $y(x)$ dada en (2.120) con el eje OX estará en $[x_0, x^*]$.

Por hipótesis sabemos que $L_f(x_0) > -2$, luego

$$x_0 - \hat{x} = x_0 - \left(1 + \frac{1}{2}L_f(x_0)\frac{f(x_0)}{f'(x_0)}\right) \leq 0 \Leftrightarrow x_0 \leq \hat{x}.$$

Por tanto, será suficiente que $x \geq x_0$ para que se cumpla

$$y(x) = \frac{-1 + \sqrt{1 - 4a(bx + c)}}{2a} \geq f(x). \quad (2.134)$$

En ese caso, obtendremos una secuencia monótona creciente, acotada por arriba por x^* , luego converge en el límite $\gamma \leq x^*$. Por lo tanto, debido a la construcción del método y la continuidad de $f(x)$ se obtiene la convergencia $\gamma = x^*$.

Por otro lado, notemos que (2.134) es equivalente a

$$\frac{-1 + \sqrt{1 - 4a(bx + c)}}{2a} - \frac{-1 + \sqrt{1 - 4a(bx_0 + c)}}{2a} \geq f(x) - f(x_0)$$

o

$$\int_{x_0}^x \frac{-b}{\sqrt{1 - 4a(bt + c)}} dt \geq \int_{x_0}^x f'(t) dt. \quad (2.135)$$

Como $f'(x) > 0$, entonces, por hipótesis se cumple que $((1/f'(x))^2)'' \geq 0$ en I , luego $(1/f'(x))^2$ es convexa, por tanto

$$\left(\frac{1}{f'(x)}\right)^2 \geq \frac{1 - 4a(bx + c)}{(-b)^2},$$

porque $(1 - 4a(bx + c))/(-b)^2$ aproxima a $(1/f'(x))^2$ hasta el segundo orden.

Así,

$$\frac{-b}{\sqrt{1 - 4a(bx + c)}} \geq f'(x) > 0,$$

y en consecuencia la expresión (2.135) se cumple.

Por último, si comenzamos desde cualquier punto ubicado a la derecha de la raíz, se obtendrá que

$$\frac{-1 + \sqrt{1 - 4a(bx + c)}}{2a} \leq f(x),$$

luego la convergencia será monótona por la derecha. \square

Otro teorema de convergencia global de la autoría de M. Hernández y A. Salanova aparece en [98]. Parte del mismo ha sido tomado para demostrar el siguiente teorema.

Teorema 2.26. *Sea $f(x)$ una función tal que en $[a, b]$ se cumple que $L_{f'}(x) \leq 3$. Si $f(x) > 0$, entonces la sucesión $\{x_n\}$ definida por (2.107) es decreciente y converge a x^* . Además, si $f(x_0) < 0$ y $L_f(x_0) \geq -2$ en $[a, b]$, entonces la sucesión $\{x_n\}$ es creciente y converge a x^* .*

Demostración. La demostración estará limitada al caso en que $f(x_0) > 0$, pues para $f(x_0) < 0$ la demostración es análoga.

Asumamos que $x_0 \in [a, b]$ tal que $f(x_0) > 0$, probaremos que $x_n \leq x^*$ para todo $n \in \mathbb{N}$.

Notar que $x_0 - x^* \geq 0$, pues por hipótesis $f(x_0) > 0$, además

$$x_1 - x^* = F(x_0) - F(x^*) = F'(\delta_0)(x_0 - x^*), \quad \text{con } \delta_0 \in (x^*, x_0),$$

y $F(x) = x_{n+1}$ definida por (2.107).

Por otro lado, como $L_{f'}(x) \leq 3$ y

$$F'(x) = L_f(x)^2 \frac{3 - L_{f'}(x)}{2}, \quad (2.136)$$

luego $F'(x) \geq 0$ en $[a, b]$. Entonces $x_1 \geq x^*$, y por inducción sigue que $x_n \geq x^*$ para todo $n \in \mathbb{N}$.

Por otra parte, como $f(x_n) > 0$ para todo $n \in \mathbb{N}$, luego

$$x_n - x_{n-1} = -\frac{f(x_{n-1})}{f'(x_{n-1})} \left[1 + \frac{L_f(x_{n-1})}{2} \right] \leq 0 \Leftrightarrow x_{n-1} \geq x_n,$$

luego $\{x_n\}$ es decreciente y pasando al límite de (2.107), sigue que converge a la única raíz x^* de $f(x) = 0$. □

El teorema que sigue, es una adaptación de un resultado que aparece en [88] para una familia de procesos iterativos que incluye como caso particular al método de Chebyshev. Se trata de un resultado de convergencia global para dicho método, con condiciones diferentes a las dadas en el teorema anterior.

Teorema 2.27. *Supongamos que $f(x)$ es una función decreciente y convexa en un intervalo $[0, x^*]$, con $f'''(x) > 0$ en $[0, x^*]$. En estas condiciones, la sucesión dada por (2.107), empezando en $x_0 = 0$ es creciente y converge a x^* , solución de $f(x) = 0$.*

Demostración. Primero comprobaremos que

$$C_f(x_n) = x_{n+1} = x_n - \left[1 + \frac{1}{2}L_f(x_n) \right] \frac{f(x_n)}{f'(x_n)}, \quad (2.137)$$

es creciente en $[0, x^*]$. En efecto,

$$\begin{aligned} C_f'(x_n) &= -\frac{1}{2} \frac{f(x_n)^2 f'''(x_n)}{f'(x_n)^3} + \frac{3}{2} \frac{f(x_n)^2 f''(x_n)^2}{f'(x_n)^4} \\ &= -\frac{1}{2} \frac{f(x_n)^2 f'''(x_n)}{f'(x_n)^3} + \frac{3}{2} L_f^2(x_n) \end{aligned}$$

El primer sumando es positivo ya que $f'''(x_n) \geq 0$ y $f'(x_n) < 0$ para $x \in [0, x^*]$. El segundo sumando también es positivo por tener un exponente par. En consecuencia tenemos que $C'_f(x_n) > 0$. Luego es una función creciente en el intervalo $[0, x^*]$. De aquí se sigue que la sucesión 2.137 está acotada por x^* . En efecto, $x_1 = C(x_0) = C(0) < C(x^*) = x^*$, y por inducción se sigue que $x_n < x^*, \forall n \in \mathbb{N}$. Como además

$$x_{n+1} - x_n = - \left[1 + \frac{1}{2} L_f(x_n) \right] \frac{f(x_n)}{f'(x_n)} \geq 0$$

se tiene que $\{x_n\}$ también es creciente, y por tanto convergente. \square

2.5. El método super-Halley

La aceleración convexa del método de Newton, también conocida como método super-Halley, es un novedoso algoritmo iterativo que también se aplica para resolver una ecuación no lineal de la forma (1.1). Aunque en los trabajos iniciales de Hernández et al. ([69], [85], [87], [88]) el método recibe el nombre de aceleración convexa del método de Newton, a partir de la publicación de Argyros, Chen y Qian, [41], que data del año 1994, dicho método ha sido estudiado por varios autores y recibe mayoritariamente el nombre de método super-Halley. Por ejemplo, este método ha sido ampliamente estudiado en [13], [18], [45], [121] y [122]. Por último, destacar que el estudio de la convergencia del método super-Halley o aceleración convexa del método de Newton ha sido una parte fundamental de la tesis doctoral del Dr. Gutiérrez [83].

2.5.1. Notas históricas sobre el método super-Halley

En [98], M. A. Hernández y M. A. Salanova introducen el grado de convexidad logarítmico como una medida puntual de la convexidad de una función y lo utilizan para un problema de localización de neuronas.

Esta forma de medir la convexidad surge del Teorema de Bohr-Mollerup [29], para la definición de la función gamma. En este resultado aparece el concepto de función logarítmicamente convexa, es decir, una función cuyo logaritmo es una función convexa. El grado de convexidad logarítmico es una medida puntual de la convexidad, dada por la resistencia de una función a ser “concavizada” por el operador logarítmico, es decir, el número de veces que hay que aplicar el operador logarítmico a una función convexa para obtener como resultado una función cóncava. En [97] se relaciona dicho grado de convexidad con la velocidad de convergencia del método de Newton.

Si bien el concepto de grado de convexidad logarítmico puede considerarse como reciente, y su aplicación al estudio de la velocidad de convergencia de procesos iterativos tenga como origen [102], podemos encontrar algún antecedente en la bibliografía matemática. Así, por

ejemplo, B. Neta [136] da la expresión de diversos métodos iterativos de tercer orden en función del grado de convexidad logarítmico, y en el trabajo de M. Altman [4] nos encontramos con la expresión

$$L_f(t) = \frac{f(t)f''(t)}{f'(t)^2}, \quad (2.138)$$

que el autor utiliza para probar la convergencia de un proceso iterativo de orden tres. Además, en dicho trabajo se dan unas propiedades sobre el grado de convexidad logarítmico que destacamos por su utilidad posterior.

Lema 2.28 (Lema de Altman). *Sea f una función que satisface las condiciones $f(t) \geq 0$, $f'(t) \neq 0$ y $f'''(t) \geq 0$ en $[0, x^*]$ y $f(x^*) = 0$. Entonces*

$$L_f(t) < \frac{1}{2}, \quad t \in [0, x^*].$$

Corolario 2.29. *Si exigimos que $f'(t) \neq 0$ en $[0, x^*)$, entonces se tiene que*

$$L_f(t) \leq \frac{1}{2}, \quad t \in [0, x^*].$$

2.5.2. Descripción del método super-Halley

El método super-Halley es menos conocido que los anteriormente considerados en esta memoria y, hasta donde nosotros conocemos, no existen tantas formas de obtenerlo como tienen, por ejemplo, los métodos de Newton o Halley. No obstante, mostramos a continuación tres construcciones que dan lugar al método super-Halley.

Como aceleración convexa del método de Newton

El método super-Halley surge al estudiar la influencia de la convexidad en el método de Newton $\{t_n\}$ para resolver una ecuación $f(t) = 0$.

Dada la importancia que tiene para esta sección, damos a continuación la siguiente

Definición 2.8 (Función convexa o cóncava hacia arriba). *Sea $S \subseteq \mathbb{R}$ un conjunto convexo no vacío, y sea $f : S \rightarrow \mathbb{R}$, f es una función convexa en S si y sólo si*

$$\forall t \in [0, 1] \wedge \forall x, y \in S : f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Consideramos, sin pérdida de generalidad, una función f decreciente y convexa en un intervalo $[a, b]$ tal que $f(a)f(b) < 0$. En estas condiciones, tomando $t_0 \in [a, b]$ tal que $f(t_0) > 0$, se garantiza la convergencia de la sucesión $\{t_n\}$ a x^* , la única raíz de $f(t) = 0$ en $[a, b]$. Además, la sucesión $\{t_n\}$ es creciente a x^* .

Sea g una función en las mismas condiciones que f en $[a, b]$, tal que $g(x^*) = 0$. Consideramos la sucesión $\{s_n\}$ obtenida aplicando el método de Newton a $g(t) = 0$, con $s_0 = t_0$.

Hernández y Salanova, [97], [102], utilizan el grado de convexidad logarítmico para comparar las sucesiones $\{t_n\}$ y $\{s_n\}$. En concreto, prueban que si $L_g(t) < L_f(t)$ para $t \in [t_0, x^*)$, entonces la sucesión $\{s_n\}$ converge a x^* más rápidamente que $\{t_n\}$. Es más, se tiene que $t_n < s_n < x^*$ para $n \geq 1$.

Utilizando este resultado intentaremos obtener una aceleración punto a punto para el método de Newton-Raphson. Es decir, dada $\{t_n\}$ tratamos de definir una aceleración $\{s_n\}$ tal que $s_{n+1} = G(t_n)$, a partir de un operador G . Este tipo de aceleraciones tienen gran interés, ya que una vez construido G nos permiten definir nuevos procesos iterativos de la forma

$$z_{n+1} = G(z_n)$$

independientemente de la sucesión $\{t_n\}$.

Con esta idea en la mente, obsérvese que el grado de convexidad logarítmico de una recta es cero y, por tanto, es el menor grado de convexidad logarítmico que puede tener una función convexa. Así, la recta $g(t) = f'(x^*)(t - x^*)$ cumple las condiciones exigidas para f y su grado de convexidad logarítmico es cero. Obviamente, la sucesión obtenida por el método de Newton-Raphson para esta recta converge más rápidamente que la sucesión obtenida para la función f .

Teniendo en cuenta el desarrollo de Taylor para la función f , tomaremos la cantidad

$$f(t) - \frac{f''(x^*)}{2}(t - x^*)^2$$

como aproximación de la recta $f'(x^*)(t - x^*)$.

El problema que surge es que x^* , en principio, es desconocido. Por ello hacemos las siguientes estimaciones.

Si consideramos $g(t) = f(t) - \frac{f''(x^*)}{2}(t - x^*)^2$ entonces podemos aproximar

$$g(t_n) \sim f(t_n) - \frac{f''(t_n)}{2}(t_n - t_{n+1})^2,$$

ya que $f''(t_n) \sim f''(x^*)$, la sucesión $\{t_n\}$ converge a x^* y $(t_n - t_{n+1})^2 \sim (t_n - x^*)^2$ debido a que

$$\lim_{n \rightarrow \infty} \frac{t_n - x^*}{t_n - t_{n+1}} = 1,$$

como se comprueba fácilmente.

Por otra parte, como $g'(t) = f'(t) - f''(x^*)(t - x^*)$ entonces

$$g'(t_n) \sim f'(t_n) - f''(t_n)(t_n - t_{n+1}).$$

Con estas aproximaciones tratamos de construir una aceleración para el método de Newton

$$s_{n+1} = t_n - \frac{g(t_n)}{g'(t_n)} = t_n - \frac{f(t_n) - \frac{f''(t_n)}{2} \left(\frac{f(t_n)}{f'(t_n)} \right)^2}{f'(t_n) - f''(t_n) \frac{f(t_n)}{f'(t_n)}}$$

$$s_{n+1} = t_n - \frac{f(t_n)}{2f'(t_n)} \frac{2 - L_f(t_n)}{1 - L_f(t_n)}. \quad (2.139)$$

Es una comprobación inmediata ver que si f es una función que satisface las condiciones anteriores, entonces la sucesión $\{s_n\}$ definida por (2.139) es una aceleración de la sucesión de Newton. En efecto, denotando

$$F(t) = t - \frac{f(t)}{f'(t)}$$

y

$$G(t) = t - \frac{f(t)}{2f'(t)} \frac{2 - L_f(t)}{1 - L_f(t)} = t - \frac{f(t)}{2f'(t)} \left[1 + \frac{1}{1 - L_f(t)} \right], \quad (2.140)$$

tenemos que

$$F'(t) = L_f(t), \quad G'(t) = \frac{L_f^2(t)(L_f(t) - L_{f'}(t))}{2(1 - L_f(t))^2}.$$

Por lo tanto,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{x^* - s_{n+1}}{x^* - t_{n+1}} &= \lim_{n \rightarrow \infty} \frac{x^* - G(t_n)}{x^* - F(t_n)} \\ &= \lim_{t \rightarrow x^*} \frac{x^* - G(t)}{x^* - F(t)} = \lim_{t \rightarrow x^*} \frac{G'(t)}{F'(t)} = \lim_{t \rightarrow x^*} \frac{L_f(t)(L_f(t) - L_{f'}(t))}{2(1 - L_f(t))^2} = 0. \end{aligned}$$

La aceleración anterior nos permite definir un nuevo proceso iterativo de la forma

$$t_{n+1} = G(t_n)$$

con $t_0 \in [a, b]$ y G definida en (2.140). Notemos que la sucesión obtenida no es otra que la conocido como método super-Halley.

Construcción geométrica del método super-Halley

Una forma elegante de obtener el método iterativo super-Halley es mediante su derivación a partir de la tangente de una hipérbola con el gráfico de la función en el punto $(x_n, f(x_n))$. Esta idea a sido aplicada en los artículos de diversos autores, [5], [132], [165], teniendo en común la obtención de una familia de métodos iterativos de convergencia cúbica.

Consideremos una hipérbola de la forma

$$ay^2 + bxy + y + cx + d = 0, \quad (2.141)$$

En lo adelante deberemos encontrar la forma de los parámetros a , b y c , para los cuales se cumplen las condiciones de tangencias dadas en (2.117), (2.118) y (2.119). Para tal fin, escribimos a (2.141) de la forma

$$a_n(y - f(x_n))^2 + b_n(x - x_n)(y - f(x_n)) + (y - f(x_n)) + c_n(x - x_n) = 0. \quad (2.142)$$

De donde

$$x - x_n = -(y - f(x_n)) \frac{1 + a_n(y - f(x_n))}{b_n(y - f(x_n)) + c_n} \quad (2.143)$$

En la intersección de esta curva con el eje OX , se cumple que $y = 0$, luego (2.143) se transforma en

$$x - x_n = f(x_n) \frac{1 - a_n f(x_n)}{-b_n f(x_n) + c_n} \quad (2.144)$$

Fijando el valor de

$$c_n = -f'(x_n), \quad (2.145)$$

y derivando 2 veces a (2.142) se obtiene

$$a_n = -\frac{f''(x_n)}{2f'(x_n)^2} - \frac{b_n}{f'(x_n)} \quad (2.146)$$

Sustituyendo a (2.145) y (2.146) en (2.144) sigue

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 + b_n \left(\frac{f(x_n)}{f'(x_n)} \right)} \right) \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0 \quad (2.147)$$

La expresión (2.147) representa una familia de métodos iterativos de orden tres. Tomando a $b_n = -\frac{f''(x_n)}{f'(x_n)}$, obtenemos el método

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 - L_f(x_n)} \right) \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0$$

el cual representa al método de super-Halley.

Otra construcción geométrica: usando funciones exponenciales

En la sección 2.3.2, siguiendo a [5], [23] y [112], describimos el método de Halley a partir de una construcción geométrica que se obtuvo a partir de la función osculatriz definida por (2.66), la cual sirvió para obtener la familia uniparamétrica

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n) - af(x_n)}. \quad a \in \mathbb{R}. \quad (2.148)$$

Si ahora aplicamos el método de Newton a la función de iteración

$$G(x) = x - \frac{f(x)}{f'(x) - af(x)},$$

introducida en (2.148,) obtenemos una nueva función de iteración:

$$x_{n+1} = x_n - \frac{G(x_n)}{G'(x_n)} = x_n - \frac{1 - aU_f(x_n)}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}, \quad (2.149)$$

donde $U_f(x_n) = \frac{f(x)}{f'(x)}$.

De nuevo, para lograr que (2.148) tenga convergencia cúbica, se exige que

$$a^2 f(x_n) - 2af'(x_n) + f''(x_n) = 0. \quad (2.150)$$

Siguiendo los mismos pasos que en el método de Halley, considerando que para valores de $a \neq 0$, muy próximos a cero, el término a^2 dado en (2.150) puede ser despreciado, se obtiene la solución aproximada

$$a = \frac{f''(x_n)}{2f'(x_n)}. \quad (2.151)$$

Sustituyendo el valor de a dado en (2.151) en la fórmula (2.149) se obtiene el método super-Halley:

$$\begin{aligned} x_{n+1} &= x_n - \frac{1 - 1/2L_f(x_n)}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \left[1 + \frac{1}{2} \frac{1}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)} \right] \end{aligned}$$

La misma nota realizada para el método de Halley acerca de los valores de a dados por (2.150) es válida. Teóricamente deben ser muy próximos a 0, pero, en general, el proceso obtenido tiene convergencia cúbica para funciones $f(x)$ con raíces simples.

2.5.3. Convergencia del método super-Halley

Convergencia local del método super-Halley

Para demostrar la convergencia local de un método existen diversos teoremas tales como el de Teorema de Gander, [76], Schröder, [162] o Taylor, [179]. Mostramos a continuación un teorema de convergencia local para el método super-Halley a partir [162] y [179].

Teorema 2.30. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz simple de $f(x) = 0$, ($f(x^*) = 0$ y $f'(x^*) \neq 0$). Supongamos que f es tres veces derivable con continuidad en I . Entonces el método de super-Halley (2.140) definido por $x_{n+1} = G_f(x_n)$,*

$$G_f(x) = x - \frac{f(x)}{2f'(x)} \left[1 + \frac{1}{1 - L_f(x)} \right], \quad (2.152)$$

tiene al menos convergencia cúbica localmente. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = (-A_3)e_n^3 + O_4.$$

donde A_i , e_n y O_N se han introducido en (2.12), (2.13) y (2.14) respectivamente.

Demostración. (Usando el Teorema de Schröder) El cálculo de las derivadas de G_f nos permite deducir que

$$G_f(x^*) = x^*, \quad G'_f(x^*) = G''_f(x^*) = 0, \quad G'''_f(x^*) = -\frac{f'''(x)}{f'(x)} \neq 0.$$

Por lo tanto,

$$e_{n+1} = x_{n+1} - x^* = G_f(x_n) - x^* = -\frac{G'''_f(x^*)}{3!}(x_n - x^*)^3 + O_4,$$

de donde se deduce que

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^3} = -\frac{G'''_f(x^*)}{3!} = -A_3.$$

Por lo tanto, el método super-Halley es un método con orden de convergencia local igual a 3 y contante de error asintótico $-A_3$. \square

Demostración. (Usando desarrollos de Taylor) A partir de aplicar el desarrollo en serie de Taylor, podemos llegar a la conclusión anterior. En concreto, teniendo en cuenta las notaciones A_i , e_n y O_N introducidas en (2.12), (2.13) y (2.14) respectivamente, deducimos que

$$\frac{f(x_n)}{f'(x_n)} = e_n - A_2 e_n^2 + 2(A_2^2 - A_3)e_n^3 + O_4.$$

$$\frac{1}{2} \frac{L_f(x_n)}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)} = A_2 e_n^2 + (2A_2 + 3A_3 - 4A_2^2)e_n^3 + O_4.$$

Por lo tanto,

$$\begin{aligned} e_{n+1} &= x_{n+1} - x^* = x_n - x^* - \frac{f(x_n)}{f'(x_n)} - \frac{f(x_n)}{2f'(x_n)} \left(\frac{L_f(x_n)}{1 - L_f(x_n)} \right) \\ &= e_n - \left(e_n - A_2 e_n^2 + 2(A_2^2 - A_3)e_n^3 \right) - \left(A_2 e_n^2 + (2A_2 + 3A_3 - 4A_2^2)e_n^3 \right) + O_4 \\ &= (-A_3)e_n^3 + O_4, \end{aligned}$$

con lo que hemos demostrado el resultado. \square

Convergencia semilocal del método super-Halley

En este acápite presentaremos un teorema de convergencia semilocal que aparece en la tesis doctoral de [83]. En dicha tesis, el teorema aparece para funciones definidas en espacios de Banach. En lo posterior daremos una versión para \mathbb{R} . Al igual que en ocasiones anteriores, para la convergencia semilocal del método super-Halley asumiremos condiciones tipo Kantorovich. En ese sentido, supondremos que $f(x)$ es una función convexa dos veces diferenciable en un conjunto convexo $S \in \mathbb{R}$ y además que se cumplen las siguientes condiciones:

[C₁] Existe un punto $x_0 \in S$ tal que $1/f'(x_0) \neq 0$.

[C₂] $|\frac{f''(x)-f''(y)}{f'(x_0)}| \leq k|x-y|$, $x, y \in S$, $k \geq 0$.

[C₃] El polinomio

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a, \quad a = |\frac{f(x_0)}{f'(x_0)}|, \quad b = |\frac{f''(x_0)}{f'(x_0)}|,$$

tiene una raíz negativa y 2 raíces positivas, r_1, r_2 con $r_1 < r_2$.

Teorema 2.31. Sea $\{t_n\}$ definida por

$$t_0 = 0, \quad t_{n+1} = t_n - \frac{f(t_n)}{2f'(t_n)} \frac{2 - L_f(t_n)}{1 - L_f(t_n)},$$

la sucesión construida al aplicar el método super-Halley al polinomio $p(t)$. Si bajo las condiciones 1) – 3) se verifica que

$$|x_{n+1} - x_n| \leq t_{n+1} - t_n, \quad n \geq 0,$$

entonces el método super-Halley está bien definido y converge a x^* solución de $f(x) = 0$.

Demostración. En lo adelante probaremos que para todo $n \geq 0$ se cumple que

[I_n] Existe un punto $x_n \in S$ tal que $f'(x_n) \neq 0$.

[II_n] $|\frac{f''(x_n)}{f'(x_0)}| \leq -\frac{p''(t_n)}{p'(t_0)}$.

[III_n] $|\frac{f'(x_0)}{f'(x_n)}| \leq \frac{p'(t_0)}{p'(t_n)}$.

[IV_n] $|\frac{f(x_n)}{f'(x_0)}| \leq -\frac{p(t_n)}{p'(t_0)}$.

La demostración se hará por inducción. Notemos que para $n = 0$ se verifican inmediatamente las hipótesis. Asumamos que [I_n], ..., [IV_n] se cumplen, luego debemos probar que [I_{n+1}], ..., [IV_{n+1}] también se verifican.

Tomando en cuenta que $L_f(x_n) \neq 0$, en virtud del Lema de Altman sigue que

$$|L_f(x_n)| \leq |\frac{f'(x_0)}{f'(x_n)}| |\frac{f''(x_n)}{f'(x_0)}| |\frac{f'(x_0)}{f'(x_n)}| |\frac{f(x_n)}{f'(x_0)}| \leq \frac{p''(t_n)p(t_n)}{p'(t_n)^2} = L_p(t_n) < \frac{1}{2}.$$

En consecuencia, por el Lema de Banach se cumple $\frac{1}{1-L_f(x_n)} \neq 0$ y que

$$|1 - L_f(x_n)| \leq \frac{1}{1 - L_p(t_n)}.$$

Esto garantiza la existencia del término siguiente, x_{n+1} , en la sucesión.

Comprobemos ahora que $f'(n+1) \neq 0$. Notar que

$$1 - \frac{f'(n+1)}{f'(x_0)} = - \int_{x_0}^{x_{n+1}} \frac{f''(x)}{f'(x_0)} dx = - \int_{x_0}^{x_{n+1}} \frac{f''(x) - f''(x_0)}{f'(x_0)} dx - \frac{f''(x_0)(x_{n+1} - x_0)}{f'(x_0)}.$$

Luego, considerando a $[\mathbf{C}_1]$, $[\mathbf{C}_2]$ y $[\mathbf{C}_3]$, sigue que

$$\begin{aligned} \left| 1 - \frac{f'(x_{n+1})}{f'(x_0)} \right| &\leq \left| \int_0^1 \frac{f''(x_0 + t(x_{n+1} - x_0)) - f''(x_0)}{f'(x_0)} (x_{n+1} - x_0) dt \right| + \left| \frac{f''(x_0)}{f'(x_0)} \right| |x_{n+1} - x_0| \\ &\leq \frac{k}{2} |x_{n+1} - x_0|^2 + b |x_{n+1} - x_0| \leq \frac{k}{2} t_{n+1}^2 + b t_{n+1} = p'(t_{n+1}) + 1 < 1. \end{aligned}$$

Entonces, en virtud del Lema de Banach, se cumple que $\frac{f'(x_0)}{f'(x_{n+1})} \neq 0$ y además

$$\left| \frac{f'(x_0)}{f'(x_{n+1})} \right| \leq \frac{1}{1 - \left| 1 - \frac{f'(x_{n+1})}{f'(x_0)} \right|} \leq -\frac{1}{p'(t_{n+1})}. \quad (2.153)$$

Luego sigue que

$$\frac{1}{f'(x_{n+1})} = \frac{f'(x_0)}{f'(x_{n+1})} f'(x_0) \neq 0.$$

Además, tomando en cuenta a (2.153) sigue que

$$\left| \frac{f'(x_0)}{f'(x_{n+1})} \right| = \left| \frac{f'(x_0)}{f'(x_{n+1})} \right| \leq -\frac{1}{p'(t_{n+1})} = \frac{p'(t_0)}{p'(t_{n+1})},$$

luego se verifica $[\mathbf{III}_{n+1}]$.

Finalmente, veamos que $[\mathbf{IV}_{n+1}]$ también se cumple. Aplicando la técnica de Altman y aplicando el desarrollo de Taylor a f en relación con x_n sigue que

$$f(x_{n+1}) = f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{1}{2} f''(x_n)(x_{n+1} - x_n)^2 + \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)](x_{n+1} - x_n) dx.$$

Como

$$x_{n+1} - x_n = - \left[1 + \frac{1}{2} L_f(x_n) \frac{1}{1 - L_f(x_n)} \right] \frac{f(x_n)}{f'(x_n)},$$

se sigue

$$\begin{aligned} f'(x_n)(x_{n+1} - x_n) &= -f(x_n) - \frac{1}{2} f'(x_n) L_f(x_n) \frac{1}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)} \\ &= -f(x_n) - \frac{1}{2} \frac{f''(x_n) f(x_n) f(x_n)}{f'(x_n) f'(x_n)} \frac{1}{1 - L_f(x_n)}. \end{aligned}$$

Por tanto

$$\begin{aligned} f(x_{n+1}) &= -\frac{1}{2} \frac{f''(x_n) f(x_n) f(x_n)}{f'(x_n) f'(x_n)} \frac{1}{1 - L_f(x_n)} + \frac{1}{2} f''(x_n) (x_{n+1} - x_n)^2 \\ &+ \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)] (x_{n+1} - x_n) dx = -\frac{1}{2} \frac{f''(x_n) f(x_n) f(x_n)}{f'(x_n) f'(x_n)} \frac{1}{1 - L_f(x_n)} \\ &+ \frac{1}{2} f''(x_n) \left(\frac{f(x_n)}{f'(x_n)} \right)^2 + \frac{1}{8} f''(x_n) \left(\frac{L_f(x_n) f(x_n)}{f'(x_n)} \frac{1}{1 - L_f(x_n)} \right)^2 \\ &+ \frac{1}{2} \frac{f''(x_n) f(x_n) L_f(x_n) f(x_n)}{f'(x_n) f'(x_n)} \frac{1}{1 - L_f(x_n)} + \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)] (x_{n+1} - x_n) dx. \end{aligned}$$

Haciendo $\Phi_n = \frac{1}{1-L_f(x_n)}$ y $y_n = \frac{L_f(x_n)f(x_n)}{f'(x_n)}\Phi_n$, sigue que

$$f(x_{n+1}) = \frac{1}{8}f''(x_n)(y_n)^2 + \int_{x_n}^{x_{n+1}} [f''(x) - f''(x_n)](x_{n+1} - x_n)dx.$$

Así, usando que

$$|y_n| \leq |L_f(x_n)| |\Phi_n| \left| \frac{f(x_n)}{f'(x_n)} \right| \leq -\frac{a_n p(t_n)}{(1-a_n)p'(t_n)},$$

donde $a_n = L_p(t_n)$ se tiene que

$$\begin{aligned} \left| \frac{f(x_{n+1})}{f'(x_0)} \right| &\leq \frac{1}{8}p''(t_n) \frac{a_n^2 p(t_n)^2}{(1-a_n)^2 p'(t_n)^2} + \left| \int_0^1 \frac{f''(x_n + t(x_{n+1} - x_n)) - f''(x_n)}{f'(x_0)} (x_{n+1} - x_0)^2 (1-t) dt \right| \\ &\leq \frac{1}{8} \frac{a_n^3 p(t_n)}{(1-a_n)^2} + \frac{k}{6} |x_{n+1} - x_n|^3 \leq \frac{1}{8} \frac{a_n^3 p(t_n)}{(1-a_n)^2} + \frac{k}{6} (t_{n+1} - t_n)^3. \end{aligned}$$

Notemos que procediendo de forma análoga con el polinomio $p(t)$ se obtiene que

$$p(t_{n+1}) = \frac{1}{8} \frac{a_n^3 p(t_n)}{(1-a_n)^2} + \frac{k}{6} (t_{n+1} - t_n)^3,$$

luego

$$\left| \frac{f(x_{n+1})}{f'(x_0)} \right| \leq p(t_{n+1}),$$

por tanto se verifica a $[\mathbf{IV}_{n+1}]$, quedando demostrado teorema. □

Convergencia global del método super-Halley

En esta sección nos referiremos nuevamente al teorema de convergencia global que aparece en [88]. Notar que al sustituir en (2.106) el parámetro α por la constante uno, $\alpha = 1$, obtenemos la función de iteración del método super-Halley.

Teorema 2.32. *Supongamos que $f(x)$ es una función decreciente y convexa en un intervalo $[0, x^*]$, con $f'''(x) > 0$ en $[0, x^*]$. En estas condiciones, la sucesión dada por (2.152), empezando en $x_0 = 0$ es creciente y converge a x^* , solución de $f(x) = 0$.*

Demostración. Llamamos

$$G(x) = x - \left[1 + \frac{1}{2} \frac{L_f(t)}{1 - L_f(x)} \right] \frac{f(x)}{f'(x)}.$$

Se comprueba que $G(x)$ es creciente en $[0, x^*]$. En efecto,

$$G'(x) = L_f(x) - \frac{1}{2} \frac{L_f'(x)}{(1 - L_f(x))^2} \frac{f(x)}{f'(x)} - \frac{1}{2} \frac{L_f(x)}{1 - L_f(x)} (1 - L_f(x))$$

$$\begin{aligned}
&= L_f(x) - \frac{1}{2} \frac{1}{(1 - L_f(x))^2} \left(L_f(x) + \frac{f^2(x)f'''(x)}{(f'(x))^3} - 2L_f^2(x) \right) - \frac{1}{2} \frac{L_f(x)}{1 - L_f(x)} + \frac{1}{2} \frac{L_f^2(x)}{1 - L_f(x)} \\
&= -\frac{1}{2} \frac{1}{(1 - L_f(x))^2} \frac{f^2(x)f'''(x)}{(f'(x))^3} + \frac{1}{2} \frac{L_f^2(x)}{(1 - L_f(x))^2} (2L_f(x) - L_f^2(x)).
\end{aligned}$$

El primer sumando es positivo ya que $f'''(x) \geq 0$ y $f'(x) < 0$ para $x \in [0, x^*]$. El segundo sumando es positivo si lo es

$$2L_f(x) - L_f^2(x) = L_f(x).$$

Por el lema de Altman, si f es decreciente y convexa, y además verifica $f'''(x) \geq 0$ para $x \in [0, x^*]$, tenemos que $0 \leq L_f(x) < 1/2$ si $x \in [0, x^*]$.

En consecuencia tenemos que $G'(x) > 0$.

Luego, $G(x)$ es una función creciente en el intervalo $[0, x^*]$. De aquí se sigue que la sucesión $\{x_n\}$ está acotada por x^* . En efecto, $x_1 = G(x_0) = G(0) < G(x^*) = x^*$, y por inducción se sigue que $x_n < x^*, \forall n \in \mathbb{N}$. Como además

$$x_{n+1} - x_n = - \left[1 + \frac{1}{2} \frac{L_f(x_n)}{1 - L_f(x_n)} \right] \frac{f(x_n)}{f'(x_n)} \geq 0$$

se tiene que $\{x_n\}$ también es creciente, y por tanto convergente. □

2.6. El método de la Secante

Uno de los métodos más utilizados para resolver ecuaciones no lineales es el método de Newton:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, 3, \dots \quad (2.154)$$

Sin embargo, si la función no es diferenciable o es difícil de calcular su derivada, este método no es favorable. Una alternativa para solucionar este inconveniente es considerar las diferencias divididas. La definición de diferencia dividida de un operador fue introducida por J. Schröder, consúltese [162] y representa una generalización de la definición usual de diferencia dividida de una función escalar. De acuerdo con [162], una función lineal y acotada $[x, y; f] : I \rightarrow \mathbb{R}$ se dice que es una diferencia dividida de primer orden para la función f en los puntos x e y , $x \neq y$, si cumple la igualdad siguiente:

$$[x, y; f] = \frac{f(x) - f(y)}{(x - y)}.$$

Para conocer más sobre la existencia de las diferencias divididas, véase también a [21] y [181]. Esta función se ha aplicado para estudiar el método de la Secante, el cual parte de dos puntos y estima su tangente de acuerdo a la ecuación

$$f'(x) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \quad (2.155)$$

Sustituyendo (2.155) en la ecuación (2.5) del método de Newton-Raphson, se obtiene como resultado la recurrencia

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n), \quad n \geq 1. \quad (2.156)$$

Continuando con los esquemas anteriores, en lo adelante revisaremos la evolución histórica del método de la Secante, la técnica para deducirlo y analizaremos algunos resultados que garanticen su convergencia.

2.6.1. Notas históricas sobre el método de la Secante

En [195] encontramos que este método fue utilizado en el siglo *XVI* por Viète (1540–1603) y más tarde por Descartes (1596 – 1650).



Figura 2.14: Francis Viète (1540 – 1603) y René Descartes (1596 – 1650).

De acuerdo a [195], los trabajos de Viète, concernientes a la solución numérica de una ecuación polinomial mónica no lineal, fueron publicados por primera vez en 1600, en el libro *De numerosa potestatum*, en la ciudad de París. Luego, en 1646 apareció en una colección de trabajos de Viète recopilada por Fran Van Schooten bajo el título *Francisci Vietae Opera Mathematica* y más tarde, en 1647, fueron publicados en varias ediciones de los trabajos de Oughtred bajo el título *Clavis Mathematicae*. Newton tuvo acceso, tanto a la colección de Schooten como a la tercera edición del libro de Oughtred, publicada en Oxford en 1652, lo cual constituye una primera señal de que él estaba muy interesado en el estudio de la solución numérica de una ecuación no lineal.

No fue sino hasta el año 1665 cuando en una colección de notas no publicadas tituladas *El libro de los residuos de Newton*, que apareció la demostración de una técnica dada por Newton que podría ser identificada como el método de la Secante para resolver ecuaciones

no lineales. La misma se basa en la semejanza de los triángulos que se forman cuando una Secante corta a la curva de una función.

Por otro lado, existe una información, documentada por Carlos Suárez Alemán [176], en la cual se establece la invención del método de la Secante al matemático español José Mariano Vallejo y Ortega (1779–1846). Vallejo nació en Albuñuelas, Granada. De acuerdo a esta fuente, Vallejo, en su obra *Compendio de Matemáticas puras y mixtas*, publicada en 1819, en concreto, en la edición de 1835 formula el método de manera completa.



Figura 2.15: Sir Isaac Newton (1643 – 1727) y José Mariano Vallejo (1779 – 1846).

Tras un breve recorrido histórico de los diferentes métodos de resolución numérica de ecuaciones y sus variantes aportados por Viète, Harriot, Oughtred, Pell, Newton, Halley, Lagrange, Raphson, Jacques y Johann Bernoulli, Daniel Bernoulli, Euler, Lagrange y Budan de Boislaurent, Vallejo indica que *estos métodos necesitan de los conocimientos más sublimes de las Matemáticas* mientras que su propio método sólo necesita de conocimientos elementales de la Aritmética. Tomando esto en cuenta, Suárez [176] considera a José Mariano Vallejo y Ortega como el primero que enuncia el método de la Secante para resolver numéricamente ecuaciones polinómicas de cualquier grado, ya que no se ha encontrado ningún otro autor que incluya este método en alguna obra de fecha anterior a 1835.

El método de Vallejo está basado en la combinación de dos importantes ideas: el método de la Falsa Posición y la iteración. De acuerdo a [176], antes de dar inicio a la demostración de su método, Vallejo [184] enuncia los siguientes resultados, los cuales eran conocidos a esa fecha:

1. Toda ecuación de grado impar, tiene al menos una raíz real de signo contrario al de su último término.

2. Toda ecuación de grado par, cuyo último término es negativo, tiene al menos dos raíces reales, una positiva y otra negativa.
3. La ecuación en que todos los exponentes de la incógnita sean números pares, y su último término sea negativo, tendría al menos dos raíces reales de igual valor numérico; una positiva y otra negativa.
4. La ecuación en que todos los exponentes sean números pares y todos los coeficientes sean números positivos, no tiene ninguna raíz real; y todas serán imaginarias.

De acuerdo a [176], Vallejo [184], para resolver una ecuación polinómica con coeficientes reales de cualquier grado, establece el siguiente procedimiento:

1. Sea $f(x) = 0$, siendo $f(x)$ una ecuación polinómica con coeficientes reales. Considérese como primeras aproximaciones de la raíz 1 y 2. En caso de que sus imágenes sean iguales, se toman como aproximaciones a 1 y 3 o 1 y 4, etc. Así hasta que las dos primeras tengan imágenes distintas.
2. Procedimiento para determinar una nueva aproximación a partir de las dos últimas aproximaciones a_i y b_i .

a) Si $|f(a_i)| \leq |f(b_i)|$, entonces

$$c_i = a_i + \frac{f(a_i)(b_i - a_i)}{f(a_i) - f(b_i)}.$$

b) Si $|f(a_i)| > |f(b_i)|$, entonces

$$c_i = b_i + \frac{f(b_i)(a_i - b_i)}{f(b_i) - f(a_i)}.$$

3. Comprobación de la raíz.

Se considera que c_i es una raíz si $|f(c_i)| < 10^{-n}$, $n \in \mathbb{N}$, en cuyo caso se divide a $f(x)$ por $x - c_i$ y se inicia el procedimiento de nuevo con el cociente.

4. Elección del nuevo intervalo.

- a) En el caso de que $f(c_i) \times f(a_i) < 0$ y $f(c_i) \times f(b_i) > 0$, el nuevo intervalo es $[a_i, c_i]$, esto es $a_{i+1} = a_i$ y $b_{i+1} = c_i$.
- b) En el caso de que $f(c_i) \times f(a_i) > 0$ y $f(c_i) \times f(b_i) < 0$, el nuevo intervalo es $[c_i, b_i]$, esto es $a_{i+1} = c_i$ y $b_{i+1} = b_i$.
- c) En el caso de que $f(c_i) \times f(a_i) < 0$ y $f(c_i) \times f(b_i) < 0$, puede adoptarse uno cualquiera de los casos siguientes:

- $a_{i+1} = c_i$ y $b_{i+1} = a_i$, si $c_i < a_i$.
- $a_{i+1} = a_i$ y $b_{i+1} = c_i$, si $a_i < c_i < b_i$ y $|c_i - a_i| < |b_i - c_i|$.
- $a_{i+1} = c_i$ y $b_{i+1} = b_i$, si $a_i < c_i < b_i$ y $|c_i - a_i| > |b_i - c_i|$.
- $a_{i+1} = b_i$ y $b_{i+1} = c_i$, si $b_i < c_i$.

d) En el caso de que $f(c_i) \times f(a_i) > 0$ y $f(c_i) \times f(b_i) > 0$, es decir, los tres valores tienen el mismo signo:

1) Si $|f(c_i)| \leq \min\{|f(a_i)|, |f(b_i)|\}$

(a') Si existe N , con $N = \pm 10^k$ con $k \in \mathbb{N}$ tal que:

- $f(N) \times f(a_i) < 0$, entonces se toma como nuevo límite del intervalo $a_{i+1} = a_i$ y $b_{i+1} = N$.
- $f(N) \times f(b_i) < 0$, entonces se toma como nuevo límite del intervalo $a_{i+1} = N$ y $b_{i+1} = b_i$.

(b') En caso contrario se busca una raíz real r de $f'(r)$ tal que $f(r) \times f(a_i) < 0$ y se toma como nuevo intervalo el formado por r y el valor más cercano a él de todos los a_i, b_i, c_i encontrados anteriormente.

(c') Si no existe raíz real, r , de $f'(x)$ tal que $f(r) \times f(a_i) < 0$, el polinomio $f(x)$ no tiene raíces reales.

2) Si $|f(c_i)| < \min\{|f(a_i)|, |f(b_i)|\}$, se toma como nuevo intervalo el formado por:

- Si $|f(a_i)| < |f(b_i)|$, entonces $a_{i+1} = a_i$ y $b_{i+1} = c_i$.
- Si $|f(a_i)| > |f(b_i)|$, entonces $a_{i+1} = c_i$ y $b_{i+1} = b_i$.

Para un análisis más profundo del método de la Secante, según Vallejo, léase nuevamente a [176] y [184]

2.6.2. Descripción del método de la Secante

El método de la Secante, es una relación recurrente que se aplica para resolver ecuaciones no lineales, teniendo como principal ventaja el hecho de que tan sólo necesita que la función sea de variable real. En lo que sigue mostraremos algunas construcciones que desembocan en la obtención del método de la Secante.

Construcción analítica del método de la Secante

El método de la Secante puede ser obtenido analizando la fórmula que se obtiene cuando una recta toca dos puntos de la gráfica de una función. Existen una inmensa cantidad de libros y artículos en los que se presenta esta forma de deducir la expresión (2.156). Parte de la deducción siguiente aparece en [77], [107], [113], [133], [135], [138], [141], [153] y [163].

Sean x_0 y x_1 dos puntos, entonces la ecuación de la secante que pasa a través de esos puntos a la curva $y = f(x)$ es

$$y - f(x) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1).$$

Si $y = 0$, entonces obtenemos una nueva aproximación dada por

$$x_2 = x_1 - \frac{x_1 - x_0}{f(x_1) - f(x_0)}f(x_1).$$

Repitiendo el proceso, se llega a la fórmula general de las aproximaciones sucesivas conocida como el método de la Secante:

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}f(x_n), \quad n \geq 1. \quad (2.157)$$

Construcción geométrica del método de la Secante

El método de la Secante, tal y como se muestra en la Figura 2.16, puede ser derivado geoméricamente.

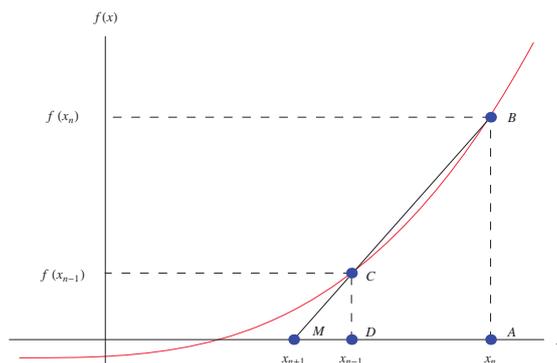


Figura 2.16: Interpretación geométrica del método de la Secante

Tomando 2 puntos iniciales, x_{n-1} y x_n , dibujamos una línea recta entre $f(x_n)$ y $f(x_{n-1})$ pasando a través del eje X en x_{n+1} , obteniéndose que $\triangle ABM \sim \triangle DCM$, por lo tanto

$$\frac{AB}{AM} = \frac{DC}{DM} \Leftrightarrow \frac{f(x_n)}{x_n - x_{n+1}} = \frac{f(x_{n-1})}{x_{n-1} - x_{n+1}}.$$

De donde se obtiene a (2.157)

Aproximando la derivada del método de Newton

Al principio de esta sección señalamos que una de las mayores desventajas que presenta el método de Newton es que aplica la derivada de la función que se plantea resolver. Con

la intención superar esta desventaja se han propuesto una serie de métodos, como el de Steffensen [172], cuya función de iteración viene dada por

$$x_{n+1} = x_n - \frac{f(x_n)^2}{f(x_n + f(x_n)) - f(x_n)}.$$

Pero el método de Steffensen es de limitada aplicación, pues su extensión a operadores definidos en espacios de Banach sólo es posible cuando el espacio donde está contenido el dominio es el mismo que el de llegada.

Otra manera de mejorar el método de Newton es aproximar la derivada de la función mediante una fórmula numérica dada por

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}},$$

que al ser sustituida en (2.5) se obtiene el algoritmo del método de la Secante.

Por otro lado, notemos que podemos derivar el método de la Secante a partir del análisis de una familia uniparamétrica de procesos iterativos definidos por

$$\begin{cases} S_n = \lambda x_n + (1 - \lambda)x_{n-1}, & \lambda \in [0, 1], \quad n \geq 0. \\ x_{n+1} = x_n - \frac{S_n - x_n}{f(S_n) - f(x_n)} f(x_n), & n \geq 0. \end{cases}$$

Notemos que para $\lambda = 0$ obtenemos el método de la Secante, mientras que para $\lambda = 1$, el método de Newton.

2.6.3. Convergencia del método de la Secante

Dado el hecho de que el método de la secante no siempre converge, se hace necesario establecer una serie de condiciones que aseguren dicha convergencia. En este acápite mostraremos algunos ejemplos de convergencia local y semilocal y global. No es de nuestro interés hacer un estudio profundo de estos ejemplos, sino dar una visión básica de los mismos. Para un estudio profundo de los mismos ver [15], [16], [17], [36], [59], [77], [107], [113], [133], [135], [138], [141], [153], [163], [173] [177] y [194].

Convergencia local del método de la Secante

La convergencia local del método de la Secante puede ser vista como una variación del método de Newton-Raphson, en donde la derivada de la función es reemplazada por su incremento:

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \quad (2.158)$$

El análisis de convergencia de un método está basado en la caracterización de la evolución de la sucesión del error. $\{e_0, e_1, e_2, \dots\}$, donde $e_n = x^* - x_n$. Recordar que un método es convergente si cumple que

$$\lim_{n \rightarrow \infty} |e_n| = 0. \quad (2.159)$$

En términos más precisos, decimos que un método tiene un orden de convergencia p , si existe C_n tal que

$$\lim_{n \rightarrow \infty} |e_{n+1}| \geq C_n, \quad (2.160)$$

donde C_n representa la constante de error asintótico.

Asumiendo que el método es convergente, el orden de convergencia se analiza aplicando los siguientes polinomios de Taylor:

$$f(x_n) = -f'(x^*)e_n + \frac{1}{2}f''(x^*)(e_n)^2 + O(|e_n|^3) \quad (2.161)$$

$$f'(x_n) = f'(x^*) - f''(x^*)(e_n) + O(|e_n|^2). \quad (2.162)$$

Parte del teorema que presentamos a continuación aparece en [59].

Teorema 2.33. Sean x_0 y x_1 dos aproximaciones a una raíz simple x^* de $f(x) = 0$, tales que $f(x_1)f(x_2) > 0$, entonces

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n) \quad \text{para } n \geq 1, \quad (2.163)$$

converge a x^* con una constante de error asintótico igual a

$$\left| \frac{f''(x^*)}{2f'(x^*)} \right|^{1/p},$$

donde $p = \frac{1+\sqrt{5}}{2}$ es el orden de convergencia.

Demostración. De (2.155) obtenemos

$$e_{n+1} = \frac{(e_{n-1})f(x_n) - (e_n)f(x_{n-1})}{f(x_n) - f(x_{n-1})}. \quad (2.164)$$

Luego, sustituyendo (2.161) y (2.162) en (2.164) sigue

$$e_{n+1} = \frac{\frac{1}{2}f''(x^*)e_n e_{n-1} + O(|e_n|^4)}{f'(x^*) + \frac{1}{2}f''(x^*)(e_n + e_{n-1}) + O(|e_n|^3)}. \quad (2.165)$$

Notar que $O(|e_{n-1}|^r)$ mayoriza al término $O(|e_n|^r)$ para $r = 3$ y $r = 4$. Luego, la forma asintótica de (2.165), cuando $n \rightarrow \infty$ es

$$e_{n+1} \cong \frac{e_{n-1}e_n f''(x^*)}{2! f'(x^*)}, \quad (2.166)$$

luego

$$|e_{n+1}| = C_n |e_n e_{n-1}|, \quad \text{donde } C = \lim_{n \rightarrow \infty} C_n = 1/2 |f''(x^*) f'(x)|,$$

por tanto (2.163) converge a x^* .

Para determinar el orden de convergencia necesitamos determinar un número p tal que

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} = C,$$

para cualquier constante C no nula. Como

$$\frac{|e_{n+1}|}{|e_n|^p} = C_n |e_n|^{1-p} |e_{n-1}| = C_n \left(\frac{|e_n|}{|e_{n-1}|^p} \right)^\beta, \quad (2.167)$$

haciendo a $\beta = 1 - p$ ó $\beta p = -1$, esto es, haciendo que $p^2 - p - 1 = 0$.

La ecuación $p^2 - p - 1 = 0$ tiene como raíz simple a $p = \frac{1+\sqrt{5}}{2} = 1.618\dots$

Esto muestra que

$$\frac{|e_{n+1}|}{|e_n|^p} \cong C^{\frac{1}{p}} = \left| \frac{f''(x^*)}{2f'(x^*)} \right|^{\frac{1}{p}},$$

con n lo suficientemente grande y $p = 1.618\dots$

Por tanto el método de la Secante exhibe una convergencia superlineal y su constante de error asintótica es

$$\left| \frac{f''(x^*)}{2f'(x^*)} \right|^{1/p}.$$

□

En [191] aparece un interesante teorema de convergencia local en el cual se aplica la γ -teoría. Dado lo novedoso del mismo, lo presentamos a continuación, pero para funciones reales.

Previo a la demostración del teorema de convergencia local aplicando la γ -teoría, Yakoubsohn demuestra la siguiente

Proposición 2.34. *Sea x^* una raíz de f . Para $y, z \in \mathbb{R}$ consideremos las cantidades*

$$u = \gamma(f, f'(x^*), x^*) |x^* - y|.$$

$$v = \gamma(f, f'(x^*), x^*) |x^* - z|.$$

Supongamos que $\psi(u, v) > 0$, entonces, el punto

$$z' = z - \frac{f(z)(z - y)}{f(z) - f(y)},$$

está bien definido y

$$|z' - x^*| \leq S(u, v) |z - x^*|$$

Además toma en cuenta las siguientes cantidades:

$$u = \gamma(f, f'(x^*), x^*) |x_0 - x^*| \quad y \quad v = \gamma(f, f'(x^*), x^*) |x_1 - x^*|,$$

donde x^* es una raíz de f y x_0, x_1 son dos puntos en \mathbb{R} . También aplica la función

$$S(u, v) = \frac{u}{(1 - 2u - 2v + 2uv)}, \quad \psi(u, v) = 1 - 2u - 2v + 2uv$$

y la sucesión de Fibonacci

$$i_{-1} = 0, \quad i_0 = 1, \quad i_{n+1} = i_n + i_{n-1}, \quad n \geq 0.$$

Teorema 2.35 (Teorema de convergencia local de Yakoubsohn, aplicando la γ -teoría). Sean $x_0, x_1 \in \mathbb{R}$ dos puntos tales que para

$$u = \gamma(f, f'(x^*), x^*)|x_0 - x^*| \quad y \quad v = \gamma(f, f'(x^*), x^*)|x_1 - x^*|,$$

se cumple que

$$1 - 3u - 2v + 2uv > 0, \quad y \quad 1 - 2u - 3v + 2uv > 0.$$

Entonces la sucesión del método de la Secante definida por

$$x_{n+1} = x_n - \frac{f(x_n)}{[x_n, x_{n-1}]}, \quad n \geq 1,$$

está bien definida y para todo $n \geq 2$ se cumple que

$$\begin{aligned} |x_n - x^*| &\leq S(u, v)^{i_{n-1}-1} S(u, v)^{i_{n-2}} |x_1 - x^*| \\ &\leq S(u, v)^{i_{n-1}} S(u, v)^{i_{n-2}-1} |x_0 - x^*|. \end{aligned} \quad (2.168)$$

Demostración. La demostración se hará por inducción.

Notemos que la desigualdad $1 - 3u - 2v + 2uv > 0$ implica que $S(u, v) < 1$ y $\psi(u, v) > 0$. Asimismo, la desigualdad $1 - 2u - 3v + 2uv > 0$ implica que $S(u, v) < 1$.

Por otro lado tenemos que

$$|x_1 - x^*| = S(u, v)^{i_0-1} S(u, v)^{i_0-1} |x_1 - x^*| = S(u, v)^{i_0} S(u, v)^{i_0-1} |x_0 - x^*|,$$

de ahí se cumple que para $n = 1$.

Supongamos que la desigualdad 2.168 se cumple $\forall i, 2 \leq i \leq n$. En consecuencia tenemos que

$$|x_i - x^*| \leq |x_1 - x^*| \quad y \quad |x_i - x^*| \leq |x_0 - x^*|.$$

Por lo tanto el punto

$$\psi(\gamma(f, f'(x^*), x^*)|x_{i-1} - x^*|, \gamma(f, f'(x^*), x^*)|x_i - x^*|) \geq \psi(u, v) > 0.$$

y por la Proposición 2.34, el punto x_{n+1} está bien definido y cumple que

$$|x_{n+1} - x^*| \leq S(\gamma(f, f'(x^*), x^*)|x_{n-1} - x^*|, \gamma(f, f'(x^*), x^*)|x_n - x^*|) |x_n - x^*|.$$

Por último, aplicando la hipótesis de inducción y la definición de sucesión de Fibonacci, sigue

$$\begin{aligned} |x_{n+1} - x^*| &\leq \frac{\gamma(f, f'(x^*), x^*)}{\psi(\gamma(f, f'(x^*), x^*)|x_{n-1} - x^*|, \gamma(f, f'(x^*), x^*)|x_n - x^*|)} |x_{n-1} - x^*| |x_n - x^*|. \\ &\leq \frac{\gamma}{\psi(u, v)} S(v, u)^{i_{n-1}-1} S(u, v)^{i_{n-2}} |x_1 - x^*| S(v, u)^{i_{n-2}-1} S(u, v)^{i_{n-3}} |x_1 - x^*|. \\ &\leq S(v, u)^{i_n-1} S(u, v)^{i_{n-1}} |x_1 - x^*|. \end{aligned}$$

La desigualdad

$$|x_{n+1} - x^*| \leq S(v, u)^{i_n} S(u, v)^{i_{n-1}-1} |x_0 - x^*|,$$

se obtiene de forma similar a la anterior y con ella se completa la demostración. \square

Convergencia semilocal del método de la Secante

El estudio de la convergencia semilocal del método de la Secante ha sido ampliamente desarrollado, para muestra, véase a [100], [101], [155], [191]. En esta sección demostraremos algunos teoremas de convergencia semilocal del método de la Secante que aparecen en [191], como una versión en \mathbb{R} que podrán ser extendida a espacios generales o ecuaciones definidas en espacios de Banach.

En lo que sigue aplicaremos el análisis de Kantorovich y la α -teoría de Smale tal y como se presenta en [111] y [169]. Recordemos que las condiciones de convergencia semilocal tipo Kantorovich están condicionadas por la convergencia de una sucesión real $\{s_n\}_{n=0}^{\infty}$ asociada a la función mayorizante $h(t)$ definida por

$$h(t) = b - 2t + \frac{t}{1 - gt}, \quad (2.169)$$

donde b y g son dos número positivos y reales. Por, conveniencia representaremos a (2.169) mediante una función definida por

$$h^*(\tau) = gh\left(\frac{\tau}{g}\right) = a - 2\tau + \frac{\tau}{1 - \tau}, \quad (2.170)$$

donde $a = bg$.

Notar que (2.170) es una función convexa en el intervalo $[0, 1[$. A lo largo de esta sección asumiremos que la desigualdad $a < 3 - 3\sqrt{2}$ se cumple. Notemos también que bajo estas condiciones, la función (2.170) tiene dos raíces $\tau_1 < \tau_2$ dadas por

$$\tau_i = \frac{a + 1 \pm \sqrt{(a + 1)^2 - 8a}}{4}, \quad i = 1, 2, \dots \quad (2.171)$$

Recordar también que la la función de iteración del método de la Secante, en términos de diferencia dividida está dada por

$$x_{n+1} = x_n - \frac{f(x_n)}{[x_n, x_{n-1}]}, \quad [x_n, x_{n-1}] = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}, \quad (2.172)$$

luego, el método de la Secante aplicado a la función mayorizante (2.170) viene dado a partir de

$$\begin{cases} x_0 = 0, & 0 < x_1 < \tau_1, \\ x_{n+1} = x_n - \frac{(x_n - x_{n-1})h(x_n)}{h(x_n) - h(x_{n-1})}, \end{cases} \quad (2.173)$$

Teorema 2.36 (Teorema de Yakoubsohn bajo condiciones tipo Kantorovich). *Consideremos la sucesión del método de la Secante $\{x_n\}_{n \geq 0}$ asociada a la función $h^*(\tau)$:*

$$x_0 = 0, \quad 0 < x_1 < \tau_1,$$

$$x_{n+1} = x_n - \frac{h^*(x_n)}{[x_n, x_{n-1}]h^*}. \quad (2.174)$$

Definamos también para todo $n \geq 0$ las sucesiones

$$u_n = \frac{(x_n - \tau_1)}{(x_n - \tau_2)} \quad (2.175)$$

$$q = \frac{(1 - \tau_2)\tau_1}{(1 - \tau_1)\tau_2}. \quad (2.176)$$

Además consideremos la sucesión de Fibonacci:

$$i_0 = i_1 = 1, \quad y \quad i_{n+1} = i_n + i_{n-1}, \quad n \geq 1.$$

Entonces la sucesión $\{x_n\}_{n>0}$ es estrictamente creciente y converge hacia τ_1 . La velocidad de convergencia de esta sucesión está dada por:

$$u_{n+1} = \left(\frac{q}{u_0}\right)u_n u_{n-1} \leq \left(\frac{u_0}{q}\right)q^{i_{n+1}}, \quad n \geq 1. \quad (2.177)$$

$$\tau_1 - x_{n+1} = \left(\frac{(\tau_2 - \tau_1)}{(1 - u_{n+1})}\right)u_{n+1} \leq \left(\frac{\tau_2 - \tau_1}{1 - u_1}\right)u_1 q^{i_{n+1}-1} = (\tau_1 - x_1)q^{i_{n+1}-1}, \quad n \geq 0 \quad (2.178)$$

$$x_{n+1} - x_n = \left(\frac{(\tau_2 - \tau_1)}{(1 - u_n)(1 - u_{n+1})}\right)\left(1 - \left(\frac{q}{u_0}\right)u_{n-1}\right)u_n \leq aq^{i_n-1}, \quad n \geq 1 \quad (2.179)$$

$$s_n = \frac{x_n}{g}. \quad (2.180)$$

Demostración. Sea $h^*(\tau) = \frac{2(\tau - \tau_1)(\tau - \tau_2)}{1 - \tau}$. Sustituyendo en (2.174) se sigue

$$\begin{aligned} x_{n+1} &= x_n - \frac{\frac{2(x_n - \tau_1)(x_n - \tau_2)}{1 - x_n}}{\frac{h^*(x_n) - h^*(x_{n-1})}{x_n - x_{n-1}}} \\ &= x_n - \frac{(x_n - \tau_1)(x_n - \tau_2)(x_n - x_{n-1})(1 - x_{n-1})}{(x_n - \tau_1)(x_n - \tau_2)(1 - x_{n-1}) - (x_{n-1} - \tau_1)(x_{n-1} - \tau_2)(1 - x_n)}. \\ &= \frac{x_{n-1}(x_n - \tau_1)(x_n - \tau_2)(1 - x_{n-1}) - x_n(x_{n-1} - \tau_1)(x_{n-1} - \tau_2)(1 - x_n)}{(x_n - \tau_1)(x_n - \tau_2)(1 - x_{n-1}) - (x_{n-1} - \tau_1)(x_{n-1} - \tau_2)(1 - x_n)} \end{aligned}$$

Luego

$$x_{n+1} - \tau_1 = \frac{(x_n - \tau_1)(x_{n-1} - \tau_1)[(x_{n-1} - \tau_2)(1 - x_n) - (x_n - \tau_2)(1 - x_{n-1})]}{(x_{n-1} - \tau_1)(x_{n-1} - \tau_2)(1 - x_n) - (x_n - \tau_1)(x_n - \tau_2)(1 - x_{n-1})} \quad (2.181)$$

Notar que

$$\begin{aligned} [(x_{n-1} - \tau_2)(1 - x_n) - (x_n - \tau_2)(1 - x_{n-1})] &= x_{n-1} + x_n \tau_2 - x_n - x_{n-1} \tau_2. \\ &= x_{n-1}(1 - \tau_2) - x_n(1 - \tau_2). \\ &= (1 - \tau_2)(x_{n-1} - x_n). \end{aligned} \quad (2.182)$$

Por otro lado, del denominador de la parte derecha de (2.181) se obtiene

$$\begin{aligned}
&= (x_{n-1}^2 - x_{n-1}\tau_2 - x_{n-1}\tau_1 + \tau_1\tau_2)(1 - x_n) - (x_n^2 - x_n\tau_2 - x_n\tau_1 + \tau_1\tau_2)(1 - x_{n-1}). \\
&= x_{n-1}^2 - x_n^2 - x_{n-1}\tau_2 + x_n\tau_2 - x_{n-1}\tau_1 + x_n\tau_1 - x_{n-1}x_n + x_n^2x_{n-1} - x_n\tau_1\tau_2 + x_{n-1}\tau_1\tau_2. \\
&= (x_{n-1} - x_n)(x_n + x_{n-1} - x_nx_{n-1} + \tau_1\tau_2 - \tau_1 - \tau_2). \tag{2.183}
\end{aligned}$$

Luego, sustituyendo (2.182) y (2.183) en (2.181), sigue que

$$x_{n+1} - \tau_1 = \frac{(x_n - \tau_1)(x_{n-1} - \tau_1)(1 - \tau_2)}{x_n + x_{n-1} - x_nx_{n-1} + \tau_1\tau_2 - \tau_1 - \tau_2}. \tag{2.184}$$

De forma similar se obtiene para $x_{n+1} - \tau_2$ la expresión:

$$x_{n+1} - \tau_2 = \frac{(x_n - \tau_2)(x_{n-1} - \tau_2)(1 - \tau_1)}{x_n + x_{n-1} - x_nx_{n-1} + \tau_1\tau_2 - \tau_1 - \tau_2}. \tag{2.185}$$

De (2.175) sigue que

$$\begin{aligned}
u_{n+1} &= \frac{x_{n+1} - \tau_1}{x_{n+1} - \tau_2} \\
u_{n-1} &= \frac{x_{n-1} - \tau_1}{x_{n-1} - \tau_2} \\
u_0 &= \frac{\tau_1}{\tau_2}
\end{aligned} \tag{2.186}$$

Sustituyendo a (2.184) y (2.185) en (2.186) sigue que

$$u_{n+1} = cu_n u_{n-1}, \tag{2.187}$$

donde $c = q/u_0$.

Si $u_1 < u_0$, luego en (2.187) se cumple que

$$u_{n+1} = \frac{1}{c}(cu_1)^{in}(cu_0)^{i_{n-1}} \leq \frac{1}{c}(cu_0)^{i_{n+1}} = \left(\frac{u_0}{q}\right)q^{i_{n+1}}.$$

Con esto se comprueba la primera parte del teorema.

Notemos que

$$\begin{aligned}
\frac{(\tau_2 - \tau_1)}{(1 - u_{n+1})}u_{n+1} &= \frac{\tau_2 - \tau_1}{1 - \frac{x_{n+1} - \tau_1}{x_{n+1} - \tau_2}} \frac{x_{n+1} - \tau_1}{x_{n+1} - \tau_2} \\
&= \frac{(\tau_2 - \tau_1)(x_{n+1} - \tau_1)}{(\tau_1 - \tau_2)} \\
&= \tau_1 - x_{n+1}.
\end{aligned}$$

Recordar que para $u_1 < u_2$, se cumple que $(cu_1)^{in}(cu_0)^{i_{n-1}} \leq (cu_0)^{i_{n+1}} = q^{i_{n+1}} \geq q^{i_{n+1}}/q$

Luego,

$$\tau_1 - x_{n+1} \leq \frac{\tau_2 - \tau_1}{1 - u_1}u_1 q^{i_{n+1}-1} = (\tau_1 - x_1)q^{i_{n+1}-1}, \quad n \geq 0.$$

Luego la segunda parte del teorema está demostrada.

La tercera parte del teorema se deriva de cálculos sencillos. Vamos a probar la desigualdad de la tercera parte.

Dado que la sucesión $\{u_n\}$ es decreciente y que para $c = q/u_0$ la función $u \rightarrow (1 - cu/(1 - u))$ es creciente, en consecuencia

$$\frac{\tau_2 - \tau_1}{(1 - u_n)(1 - u_{n+1})}(1 - cu_{n-1}) \leq \frac{\tau_2 - \tau_1}{(1 - u_n)(1 - u_{n+1})}(1 - cu_n) \leq \frac{\tau_2 - \tau_1}{(1 - u_0)(1 - qu_0)}(1 - q) = a$$

□

Por otra parte, tenemos que Yakoubsohn, [191], también nos presenta un novedoso teorema de convergencia semilocal a partir de las condiciones de la α -teoría de Smale, [170]. Recordar que los criterios de convergencia semilocal de la α -teoría de Smale vienen dados a partir de 3 invariantes de la forma:

$$\begin{aligned} \beta(f, f'(x_0), x) &= \|f'(x_0)^{-1}f(x)\|, \\ \gamma(f, f'(x_0), x) &= \sup_{n \geq 2} \left(\left\| \frac{f'(x_0)^{-1}f^n(x)}{n!} \right\| \right)^{\frac{1}{(n-1)}}, \\ \alpha(f, f'(x_0), x) &= \beta(f, f'(x_0), x)\gamma(f, f'(x_0), x). \end{aligned}$$

donde $x \in E$ y $f'(x_0) : E \rightarrow F$ es un operador lineal invertible y acotado definido entre dos espacios de Banach.

Previo a la demostración de dicho teorema, Yakoubsohn demuestra los siguientes

Lema 2.37. *Sea*

$$b = \frac{1 - 2\gamma(\beta_0 - \beta_1)}{1 - \gamma(\beta_0 - \beta_1)}\beta_0, \quad (2.188)$$

$$g = \gamma(f, A, x_0). \quad (2.189)$$

Supongamos que $bg < 3 - 2\sqrt{2}$. Si $0 < t < t_1$, entonces:

$$a) \quad -1 = h'(0) < \frac{h(t) - h(0)}{t - 0} < 0.$$

$$b) \quad 0 < \frac{h(t) - h(0)}{t - 0} + 1 < 1.$$

Lema 2.38. *Sea $x_0 \in \mathbb{R}$, luego*

1. *Para algún $x_1 \in \mathbb{R}$ supongamos que*

$$A := \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

es invertible.

Para $i \in \{0, 1\}$ definimos a $\gamma_i = \gamma(f, A, x_i)$, y asumamos que

$$1 - \gamma_i|x_0 - x_1| > 0, \quad 1 - \gamma_i|y - x_i| > 0,$$

para algún $y \in \mathbb{R}$. Además, supongamos que

$$B := \frac{f(y) - f(x_0)}{y - x_0},$$

para algún $y \in \mathbb{R}$. Entonces

$$\left| \frac{B}{A} - 1 \right| \leq \frac{\gamma_i|y - x_1|}{(1 - \gamma_i|x_0 - x_1|)(1 - \gamma_i|y - x_i|)}.$$

En adicción, si

$$(1 - \gamma_i|x_0 - x_1|)(1 - \gamma_i|y - x_i|) > \gamma_i|y - x_1|,$$

entonces B es invertible y se cumple que

$$\left| \frac{A}{B} \right| \leq \frac{(1 - \gamma_i|x_0 - x_1|)(1 - \gamma_i|y - x_i|)}{(1 - \gamma_i|x_0 - x_1|)(1 - \gamma_i|y - x_i|) - \gamma_i|y - x_1|}.$$

2. Supongamos que $f'(x_i)$ es invertible para $i \in \{0, 1\}$, y definamos a $\gamma = \gamma(f, f'(x_i), x_i)$. Asumamos que para algunos $x, y \in \mathbb{R}$ se define $\hat{B} := \frac{f(y) - f(x)}{y - x}$ y se cumple que

$$1 - \gamma|x - x_i| > 0, \quad 1 - \gamma|y - x_i| > 0.$$

Entonces

$$\left| \frac{\hat{B}}{f'(x_i)} - 1 \right| \leq \frac{1}{(1 - \gamma|x - x_i|)(1 - \gamma|y - x_i|)} - 1.$$

Además, si

$$2(1 - \gamma|x - x_i|)(1 - \gamma|y - x_i|) - 1 > 0,$$

entonces \hat{B} tiene inversa y se cumple que

$$\left| \frac{f'(x_i)}{\hat{B}} \right| \leq \frac{(1 - \gamma|x - x_i|)(1 - \gamma|y - x_i|)}{2(1 - \gamma|x - x_i|)(1 - \gamma|y - x_i|) - 1}$$

3. En particular, si $f'(x_0)$ es invertible, entonces

$$\check{B} := \frac{f(y) - f(x_0)}{y - x_0}$$

es invertible para todo y tal que

$$2\gamma|y - x_0| < 1, \quad \gamma = \gamma(f, f'(x_0), x_0).$$

Por otra parte tenemos que

$$\left| \frac{f'(x_0)}{\check{B}} \right| \leq \frac{1 - \gamma|y - x_0|}{1 - 2\gamma|y - x_0|}.$$

4. Recíprocamente, si A es invertible y si

$$2\gamma_0|x_0 - x_1| < 1, \quad \gamma_0 = \gamma(f, A, x_0),$$

entonces $f'(x_0)$ es invertible y se cumple que

$$\left| \frac{A}{f'(x_0)} \right| \leq \frac{1 - \gamma_0|x_1 - x_0|}{1 - 2\gamma_0|x_1 - x_0|}.$$

Lema 2.39. Sea A un operador lineal acotado de \mathbb{R} y $\gamma_i = \gamma(f, A, x_i)$, $i \in \{0, 1\}$. Asumamos que

$$x_2 = x_1 - \frac{(x_1 - x_0)f(x_1)}{f(x_1) - f(x_0)},$$

entonces se cumple que

$$\left| \frac{f(x_2)}{A} \right| \leq \frac{\gamma_i|x_2 - x_1||x_2 - x_0|}{(1 - \gamma_i|x_2 - x_i|)(1 - \gamma_i|x_1 - x_i|)(1 - \gamma_i|x_0 - x_i|)}, \quad i \in \{0, 1\}.$$

En esta parte seguiremos considerando la función $h(t)$ dada en (2.169) con b, g dados en (2.188) y (2.189). Además, consideraremos la sucesión $\{s_n\}_{n \geq 0}$ con $s_0 = 0$ y $s_1 = \beta_0 - \beta_1$. Veamos a continuación el teorema

Teorema 2.40 (Teorema de Yakoubsohn bajo condiciones de Smale). Sean x_0 y x_1 dos puntos dados en \mathbb{R} tales que $f'(x_0)$ es invertible. Si

$$2\gamma|x_1 - x_0| < 1. \quad (2.190)$$

$$b\gamma < 3 - 2\sqrt{2}. \quad (2.191)$$

$$0 \leq \beta_0 - \beta_1 \leq t_1. \quad (2.192)$$

Entonces

1. La función $h(t)$ tiene dos raíces positivas, $t_1 < t_2$ y la sucesión $\{s_n\}_{n \geq 0}$ converge a t_1 .
2. La función analítica f tiene una solución $x^* \in (x_0, t_1)$.
3. La sucesión del método de la Secante está bien definida y converge a x^* . Además, para toda $n \geq 1$

$$|x_{n+1} - x_n| \leq s_{n+1} - s_n \quad \text{y} \quad |x_n - x^*| \leq t_1 - s_n$$

Vamos a dar la significación de las hipótesis del teorema anterior. La condición (2.190) asegura que la diferencia dividida

$$[x_1, x_0] := \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

es invertible. La desigualdad dada a partir de (2.191) implica que la función $h(t)$ tiene dos raíces positivas, por tanto se pueden asumir las condiciones de Kantorovich. La condición (2.192) nos permite establecer que $\{s_n\}_{n \geq 0}$ es una sucesión creciente y convergente a t_1

Demostración. Para demostrar que

$$|x_{n+1} - x_n| \leq s_{n+1} - s_n,$$

procederemos por inducción. Por la hipótesis sabemos que la derivada en el punto x_0 es invertible, y por el Lema 2.38, parte 3, y tomando en cuenta que

$$1 - 2\gamma|x_1 - x_0| > 0,$$

podemos afirmar que $[x_1, x_0]$ es invertible y que el punto x_2 está bien definido. Notemos que la función $h(t)$ y la sucesión $\{s_n\}_{n \geq 0}$ nos permiten establecer que

$$|x_2 - x_0| = \beta_0 = s_2 - s_0 = s_2 \quad y \quad |x_2 - x_1| = \beta_1 = s_2 - s_1.$$

La desigualdad (2.191) implica que la función $h(t)$ tiene dos raíces positivas, $t_1 < t_2$. Luego, de $0 \leq s_1 = \beta_0 - \beta_1 < t_1$, sigue que la sucesión $\{s_n\}_{n \geq 0}$ es creciente y converge a t_1 .

Supongamos ahora que para toda j , con $2 \leq j \leq n$ se cumple que

$$|x_j - x_{j-1}| \leq s_j - s_{j-1}.$$

Primero demostraremos que x_{n+1} está bien definido, es decir, que $[x_n, x_0]$ es invertible. En ese sentido tenemos que

$$|x_n - x_0| \leq \sum_{j=3}^n |x_j - x_{j-1}| + |x_2 - x_0|. \quad (2.193)$$

$$\leq \sum_{j=3}^n s_j - s_{j-1} + s_2 - s_0 = s_n. \quad (2.194)$$

Esto implica que

$$|x_n - x_0| < t_1 < \frac{(b\gamma + 1)}{4\gamma}.$$

De $b\gamma < 3 - 1\sqrt{2}$ sigue que

$$\gamma|x_n - x_0| < 1 - \frac{\sqrt{2}}{2}.$$

Esta desigualdad también se cumple para $n - 1$. Por tanto

$$\gamma|x_n - x_0| < 1 < 1 - \frac{1}{2(1 - \gamma|x_{n-1} - x_0|)}.$$

de ahí que la desigualdad

$$2(1 - \gamma|x_n - x_0|)(1 - \gamma|x_{n-1} - x_0|) - 1 > 0$$

se cumple.

Por el Lema 2.38, parte 2 aplicado a los puntos x_n, x_{n-1}, x_0 , la diferencia dividida $[x_n, x_{n-1}]$ es invertible. Por lo tanto el punto x_{n+1} está bien definido y se cumple que

$$\begin{aligned} \left| \frac{f'(x_0)}{[x_n, x_{n-1}]} \right| &\leq \frac{(1 - \gamma|x_n - x_0|)(1 - \gamma|x_{n-1} - x_0|)}{2(1 - \gamma|x_n - x_0|)(1 - \gamma|x_{n-1} - x_0|) - 1} \\ &\leq \frac{(1 - \gamma s_n)(1 - \gamma s_{n-1})}{2(1 - \gamma s_n)(1 - \gamma s_{n-1}) - 1} \\ &\leq \frac{-1(s_n - s_{n-1})}{h(s_n) - h(s_{n-1})} = \frac{-1}{[s_n, s_{n-1}]h}. \end{aligned} \quad (2.195)$$

Ahora probaremos que $|x_{n+1} - x_n| \leq s_{n+1} - s_n$. Notemos que

$$|x_{n+1} - x_n| = \frac{f(x_n)}{[x_n, x_{n-1}]} \leq \left| \frac{f'(x_0)}{[x_n, x_{n-1}]} \right| \frac{f(x_n)}{f'(x_0)}$$

Hemos obtenido previamente una cota superior para $f'(x_0)/[x_n, x_{n-1}]$. Para $n \geq 2$, el Lema 2.39 da una cota superior para $|f(x_n)/f'(x_0)|$ definida por

$$\begin{aligned} |f(x_n)/f'(x_0)| &\leq \frac{\gamma|x_n - x_{n-1}||x_n - x_{n-2}|}{(1 - \gamma|x_n - x_0|)(1 - \gamma|x_{n-1} - x_0|)(1 - \gamma|x_{n-2} - x_0|)} \\ &\leq \frac{\gamma(s_n - s_{n-1})(s_n - s_{n-2})}{(1 - \gamma s_n)(1 - \gamma s_{n-1})(1 - \gamma s_{n-2})}. \end{aligned} \quad (2.196)$$

Como $s_{n-2} = s_0 = 0$ y considerando el Lema 2.37, sigue

$$\left| \frac{f(x_n)}{f'(x_0)} \right| \leq \frac{\gamma(s_n - s_{n-1})s_n}{(1 - \gamma s_n)(1 - \gamma s_{n-1})} = h(s_n).$$

Por último, las estimaciones puntuales dadas por (2.195) y (2.196) implican que

$$|x_{n+1} - x_n| \leq \frac{-h(s_n)}{[s_n, s_{n-1}]h} = s_{n+1} - s_n.$$

Luego el teorema se deduce fácilmente. □

Convergencia global del método de la Secante

En lo adelante presentamos un teorema de convergencia global, parte del mismo aparece en [107] y [194] respectivamente. Recordemos que la convergencia global impone condiciones sobre el intervalo donde aparece la solución x^* de la ecuación. Recordemos también que e_n es el error que se dá entre el valor exacto x_n de la sucesión $\{x_n\}_{n=0}^{\infty}$ y x^* .

Teorema 2.41. *Sea x^* una raíz simple de $f(x) = 0$, donde $f'(x)$ y $f''(x)$ existen y son continua en un intervalo apropiado alrededor de x^* . Asumamos también que para ese intervalo*

$$|f'(x)| \geq m_1 > 0 \quad \text{y} \quad |f''(x)| \leq M_2,$$

siendo m_1 y M_2 el mínimo y el máximo respectivamente del intervalo en cuestión, entonces

$$x_n = x_{n-1} - \frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})} f(x_{n-1}) \quad \text{para } n \geq 2,$$

converge a x^* con un orden de convergencia superlineal, pues su orden es igual a $\frac{1+\sqrt{5}}{2}$.

Demostración. Por (2.156) obtenemos

$$e_{n+1} = \frac{e_{n-1}f(x_n) - e_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}. \quad (2.197)$$

Ahora estudiemos el comportamiento de la expresión

$$(y - x^*)f(x) - (x - x^*)f(y), \quad (2.198)$$

para todo x y y próximos a x^* . Si ha esta expresión le aplicamos el teorema del resto de Taylor en su forma integral, se obtiene que

$$\begin{cases} f(x) = f(x^*) + (x - x^*)f'(x^*) + \int_{x^*}^x (x - t)f''(t)dt \\ f(y) = f(x^*) + (y - x^*)f'(x^*) + \int_{x^*}^y (y - t)f''(t)dt. \end{cases} \quad (2.199)$$

En lo adelante asumamos primeramente que $x^* \leq y \leq x$ y si $f(x^*) = 0$, entonces (2.198) se convierte en

$$(y - x^*)f(x) - (x - x^*)f(y) = \int_{x^*}^y f''(t)(x - y)(t - x^*)dt + \int_y^x f''(t)(x - t)(y - x^*)dt. \quad (2.200)$$

Si $(x - y)(t - x^*) \geq 0$ y $(x - t)(y - x^*) \geq 0$ en el intervalo $x^* \leq t \leq x$ entonces (2.200) es igual a

$$\begin{aligned} f''(c) & \left[\int_{x^*}^y (x - y)(t - x^*)dt + \int_y^x (x - t)(y - x^*)dt \right] = \\ & = f''(c) \left[(x - y) \frac{(y - x^*)^2}{2} + \frac{(x - y)^2}{2} (y - x^*) \right] \\ & = \frac{1}{2} f''(c) (x - y)(y - x^*)(x - x^*). \end{aligned}$$

Luego

$$(y - x^*)f(x) - (x - x^*)f(y) = \frac{1}{2} f''(c) (x - y)(y - x^*)(x - x^*).$$

Ahora, por el teorema del valor medio sabemos que

$$f(x_n) - f(x_{n-1}) = (x_n - x_{n-1})f'(\varphi), \quad (2.201)$$

donde φ está entre x_n y x_{n-1} . Por lo tanto, de (2.197) obtenemos que

$$e_{n+1} = \frac{1}{2} \frac{f''(c)}{2f'(\varphi)} e_n e_{n-1},$$

que al tomar en cuenta las hipótesis del inicio del planteamiento se transforma en

$$|e_{n+1}| \leq \frac{M_2}{2m_1} |e_n| |e_{n-1}|. \quad (2.202)$$

Asumamos ahora, como segunda opción, que $x \geq x^*$ y que $y \geq x^*$. De (2.199) obtenemos

$$\begin{aligned} (y - x^*)f(x) - (x - x^*)f(y) &= \int_{x^*}^x (y - x^*)(x - t)f''(t)dt - \int_{x^*}^y (x - x^*)(y - t)f''(t)dt \\ &= \int_y^x f''(t)g(t)dt, \end{aligned}$$

donde

$$g(t) = \begin{cases} (y - x^*)(x - t), & x^* \leq t \leq x, \\ (x - x^*)(y - t), & y \leq t \leq x^*. \end{cases}$$

Es claro que para $y \leq t \leq x^*$ se cumple que $g(t) \leq 0$, luego, por el teorema del valor medio para integrales obtenemos que

$$\begin{aligned} (y - x^*)f(x) - (x - x^*)f(y) &= f''(c) \int_y^x g(t)dt \\ &= f''(c) \left[(y - x^*) \frac{(x - x^*)^2}{2} - (x - x^*) \frac{(y - x^*)^2}{2} \right] \\ &= -f''(c) \frac{(x^* - y)(x - x^*)(x - y)}{2}, \end{aligned}$$

de ahí que (2.201) se cumple.

Ahora resolvamos la siguiente ecuación recurrente

$$q_{n+1} = q_n q_{n-1},$$

si tomamos a q_0 y q_1 arbitrariamente e introducimos un cambio de variable $z_n = \log q_n$ se obtiene

$$z_{n+1} = z_n + z_{n-1}, \quad (2.203)$$

la cual es una ecuación en recurrencia lineal cuya solución se obtiene a partir de

$$z_n = c_1 \lambda_1^n + c_2 \lambda_2^n, \quad (2.204)$$

donde

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}$$

$$\lambda_2 = \frac{1 - \sqrt{5}}{2}$$

son las raíces de la ecuación característica

$$\lambda^2 - \lambda - 1 = 0,$$

que se construye a partir de (2.204). Para determinar a c_1 y c_2 hacemos a $z_0 = \log q_0$ y $z_1 = \log q_1$ y obtenemos que

$$z_0 = c_1 + c_2, \quad z_1 = c_1 \lambda_1 + c_2 \lambda,$$

por tanto

$$c_1 = \frac{\lambda_2 z_0 - z_1}{\lambda_2 - \lambda_1}, \quad c_2 = \frac{z_1 - \lambda_1 z_0}{\lambda_2 - \lambda_1},$$

luego

$$z_n = \frac{1}{\lambda_2 - \lambda_1} \left(\lambda_1^n \log \frac{q_0^{\lambda_2}}{q_1} + \lambda_2^n \log \frac{q_1}{q_0^{\lambda_1}} \right),$$

como

$$\lambda_2 - \lambda_1 = -\sqrt{5}, \quad \lambda_1 \lambda_2 = -1,$$

entonces

$$q_n = \left[\left(\frac{q_0^{\lambda_2}}{q_1} \right)^{\lambda_1^n} \left(\frac{q_1}{q_0^{\lambda_1}} \right)^{\lambda_2^n} \right]^{-1/\sqrt{5}} = \left(\frac{q_0^{-\lambda_1^{n-1} + \lambda_2^{n-1}}}{q_1^{\lambda_1^n - \lambda_2^n}} \right)^{-1/\sqrt{5}},$$

luego

$$q_n = q_0^{(\lambda_1^{n-1} - \lambda_2^{n-1})^{-1/\sqrt{5}}} q_1^{(\lambda_1^n - \lambda_2^n)^{1/\sqrt{5}}}, \quad (2.205)$$

como $\lambda_1 \doteq 1.618$ y $\lambda_2 \doteq -0.618$, luego, para n lo suficientemente grande se obtiene que

$$q_n \sim \left[(q_0^{1/\lambda_1} q_1)^{1/\sqrt{5}} \right]^{\lambda_1^n}, \quad q_{n+1} \sim q_n^{\lambda_1},$$

Tomando en cuenta a (2.202) y (2.205) sigue que

$$|e_{n+1}| \leq \frac{2m_1}{M_2} \left[\left[\left(\frac{|e_0| M_2}{2m_1} \right)^{1/\lambda_1} \left(\frac{|e_1| M_2}{2m_1} \right) \right]^{1/\sqrt{5}} \right]^{\lambda_1^{n+1}},$$

por tanto, para $|e_0|$ y e_1 lo suficientemente pequeño, se sigue que

$$\left| \frac{|e_n| M_2}{2m_1} \right| < 1, \quad n = 0, 1. \quad (2.206)$$

De esta forma, si (2.206) se cumple, entonces el método convergerá con un orden

$$p = \frac{1 + \sqrt{5}}{2}.$$

□

2.7. Otros métodos iterativos

En la literatura matemática se pueden encontrar un gran número de procesos iterativos de diversos tipos (punto a punto, multipunto, con o sin memoria, usando derivadas o diferencias divididas, híbridos, etc.) para resolver ecuaciones no lineales. Por lo tanto, somos conscientes de que un estudio detallado de estos métodos al nivel de lo realizado en las secciones anteriores sería una labor ardua y prolongada en el tiempo. Dicha tarea escapa, por tanto de los objetivos de esta memoria. El lector interesado puede consultar algunas referencias especializadas en este ámbito, como por ejemplo [8], [19] o [180]. No obstante, en esta sección se presentan un par de métodos iterativos que se emplearán en la resolución de la ecuación de Kepler. Se realiza un estudio de su construcción y de su convergencia, aunque sin seguir el esquema desarrollado para los métodos de las secciones anteriores. La elección de estos dos métodos y no la de muchos otros (Steffensen, Schröder, Stirling, etc.) ha sido motivada por los siguientes hechos:

- Se ha tomado el método de Yun-Petković [197] porque es una generalización directa de otros de los métodos más conocidos para resolver ecuaciones no lineales, como es el método de la bisección. De hecho, lo que el método de Yun-Petković consigue es mejorar el orden de convergencia del método de la bisección.
- El método de Danby (también conocido como de Danby-Burkardt) se ha elegido como representante de los muchos métodos con orden de convergencia mayor que tres. Se ha tomado porque ha sido empleado por otros autores ([51], [147]) en la resolución de la ecuación de Kepler. De hecho, el propio Danby, recomienda el uso de métodos de alto orden para resolver la ecuación de Kepler debido a que, en este caso, es sencillo el cálculo de las derivadas de la función que define dicha ecuación.

2.7.1. El método de Yun-Petković

El método de Yun-Petković es un método para hallar las raíces de una ecuación no lineal de convergencia cuadrática que tiene como ventaja principal el hecho de que no necesita de la

derivada de la función, ni del esfuerzo que implica elegir un valor inicial adecuado. El mismo viene dado a partir de la expresión

$$x_{n+1} = x_n - \left[\frac{2h_n}{f(b_n) - f(a_n)} \right] f(x_n), \quad n = 0, 1, 2, \dots \quad (2.207)$$

donde, para $n \geq 1$,

$$a_n = x_n - h_n \quad (2.208)$$

$$b_n = x_n + h_n \quad (2.209)$$

$$h_n = x_n - x_{n-1}, \quad (2.210)$$

El método de Yun-Petković parece ser una variante del método de la Secante, (2.156), que requiere una evaluación de la función adicional $b_n = 2x_n - x_{n-1}$ en cada iteración, en comparación del método de la Secante tradicional. Recordar que el método de la Secante, para raíces simples posee un orden de convergencia superlineal, $(1 + \sqrt{5})/2$. Ahora bien, el presente método, con el paso $h_n := x_n - x_{n-1}$ tendrá, como límite inferior de convergencia a la convergencia del método de la Secante. Para mayores detalles de este método, véase a [197].

Notas históricas sobre el método de Yun-Petković

Hoy en día, la mayoría de los métodos iterativos que se conocen tienen como base al método de Newton-Raphson o en al método de la Secante. Cabe señalar que la fortaleza del método de Newton-Raphson depende principalmente de dos condiciones: una estimación inicial adecuada y el comportamiento de la función cerca de la raíz. Además, es necesario conocer explícitamente a la derivada de la función. Aunque el método de la Secante supera las dificultades mencionadas anteriormente, la aplicación del mismo requiere un coste de más lenta velocidad de convergencia.

Para superar estas dificultades, Yun, [196], desarrolló en el 2008 un nuevo método iterativo, el cual está basado en la transformada de la función $f(x)$ a través de una función tangente hiperbólica. Se debe destacar que la evaluación numérica de la integración de la transformada de la función para encontrar una aproximación inicial, se puede realizar directamente. El error de este método sólo depende de la exactitud de la integración numérica. Sin embargo, para obtener una integración numérica lo suficientemente precisa, se requiere un gran número de puntos de integración porque el integrando o función transformada es una función escalonada.

Tres años después, Yun y Petković, [197], unieron esfuerzos con el propósito de desarrollar un nuevo y sencillo método iterativo que remediara los inconvenientes del método de Newton-Raphson y el de la Secante. Como resultado de sus esfuerzos obtuvieron el método iterativo definido por (2.215), de convergencia cuadrática y que además puede aumentar su velocidad



Figura 2.17: Miodrag S. Petković, (1948).

de convergencia, reemplazando en dicho método un intervalo inicial de los que se utilizan en el método iterativo planteado por Yun, [196].

Descripción del método de Yun-Petković

Para desarrollar su método, Yun y Petković, asumieron que la función $f(x)$ sea continua con solución única en un intervalo $[a, b]$ con $f(a)f(b) < 0$. Además que $x_n \in (a, b)$ sea un punto inicial próximo a la raíz x^* de la función $f(x^*)=0$, de forma tal que la raíz x^* esté incluida en el subintervalo $[a_n, b_n]$ o $[b_n, a_n]$ de $[a, b]$ para algún $h_n \neq 0$.

Luego definieron a $Y(x^*, x)$ y $P(x_n, x)$ como dos funciones lineales a trozos de la forma

$$Y(x^*, x) := \begin{cases} \frac{f(a_n)}{a_n - x^*}(x - x^*), & x \leq x^*; \\ \frac{f(b_n)}{b_n - x^*}(x - x^*), & x > x^*. \end{cases} \quad (2.211)$$

y

$$P(x_n, x) := \begin{cases} \frac{f(a_n) - f(x_n)}{a_n - x_n}(x - x_n) + f(x_n), & x \leq x_n; \\ \frac{f(b_n) - f(x_n)}{b_n - x_n}(x - x_n) + f(x_n), & x > x_n. \end{cases} \quad (2.212)$$

Luego definieron dos integrales de la forma

$$\begin{aligned} I_n(x^*) &:= \int_{a_n}^{b_n} Y(x^*, x) dx = \frac{1}{2} \{ (x^* - a_n)f(a_n) + (b_n - x^*)f(b_n) \} \\ &= \frac{1}{2} \{ (x^* - x_n)[f(a_n) - f(b_n)] + h_n[f(a_n) + f(b_n)] \} \end{aligned} \quad (2.213)$$

y

$$J_n(x_n) := \int_{a_n}^{b_n} P(x_n, x) dx = \frac{h_n}{2} \{ 2f(x_n) + [f(a_n) + f(b_n)] \}, \quad (2.214)$$

Donde a , b y h están definido por (2.208), (2.209) y (2.210) respectivamente.

Notar que las integrales $I_n(x^*)$ y $J_n(x_n)$, las cuales inician muy próximas al radio $|h_n|$ del intervalo $[a_n, b_n]$ o $[b_n, a_n]$, son decrecientes, por lo que reemplazaron a la raíz x^* por x_{n+1} en (2.213). Luego, resolvieron la ecuación $I_n(x_{n+1}) = J_n(x_n)$, obteniendo, para el punto inicial $x_0 = (a + b)/2$ la fórmula de iteración

$$x_{n+1} = x_n - \left[\frac{2h_n}{f(b_n) - f(a_n)} \right] f(x_n), \quad n = 0, 1, 2, \dots \quad (2.215)$$

donde a_n , b_n y h_n se definen en (2.208), (2.209) y (2.210) respectivamente. La función de iteración dada por (2.215) define al método iterativo de Yun-Petković.

Es importante resaltar que este método incluye un punto de aproximación más preciso que el de la Secante, dado por

$$\left[\frac{2h_n}{f(b_n) - f(a_n)} \right]^{-1} = \frac{1}{2} \left[\frac{f(b_n) - f(x_n)}{h_n} + \frac{f(x_n) - f(a_n)}{h_n} \right] \approx f'(x_n),$$

el cual es un promedio de las pendientes de dos líneas rectas que pasan respectivamente a través de los puntos definidos por $(a_n, f(a_n))$, $(x_n, f(x_n))$ y $(b_n, f(b_n))$, $(x_n, f(x_n))$, mientras que el método de la Secante tiene una aproximación dada por la pendiente de una línea que pasa por los puntos $(a_n, f(a_n))$ y $(x_n, f(x_n))$.

Para demostrar la efectividad de su método, Yun y Petković lo aplicaron a diferentes funciones. A modo de ilustración presentamos el caso en que

$$f(x) = x^4 + x - 1, \quad 0 \leq x \leq 2. \quad (2.216)$$

A diferencia del resto de métodos analizados en este capítulo, para los cuales existen numerosas referencias a códigos en diversos lenguajes de programación y programas de cálculo, para el método de Yun-Petković no hemos encontrado una implementación adecuada. Por este motivo, en el Cuadro ?? presentamos el código Mathematica del método de Yun-Petković que hemos desarrollado nosotros y que emplearemos en los experimentos numéricos que se desarrollan.

Los resultados de aplicar el método de Yun-Petković a (2.216), están representados en el Cuadro 2.1. Notar que para (2.216), $a = 0$, $b = 2$, $x_0 = \frac{a+b}{2} = 1$ y $h_0 = 1$. Luego, para la primera iteración se obtiene que

$$x_1 = x_0 - \frac{2h_0}{f(b) - f(a)} f(x_0) = 1 - \frac{2}{18} = \frac{8}{9}, \quad \text{luego } f(x_1) \approx 0.513184$$

Para la segunda iteración se tiene que $h_1 = x_1 - x_0 = \frac{8}{9} - 1 = -\frac{1}{9}$, $a_1 = x_1 - h_1 = \frac{8}{9} + \frac{1}{9} = 1$, $b_1 = x_1 + h_1 = \frac{8}{9} - \frac{1}{9} = \frac{7}{9}$. Por tanto

$$x_2 = x_1 - \frac{2h_1}{f(b_1) - f(a_1)} f(x_1) = \frac{8}{9} + \frac{2/9}{0.143 - 1} \times 0.513184 \approx 0.755706. \quad \text{Así, } f(2) \approx 0.081852$$

Cuadro 2.1: Iteraciones obtenidas al aplicar el método de Yun-Petković a (2.216).

n	x_n	a_n	b_n	$f(x_n)$
0	1	0.	2	1
1	0.88888	1.	0.77777	0.51318
2	0.7557	0.8888	0.62252	0.08185
3	0.7262	0.7557	0.6968	0.0044
4	0.7244	0.7262	0.7227	0.00001

En el Cuadro 2.2 se presentan los diversos valores de a_n , b_n , x_n y $f(x_n)$ obtenidos al aplicar dicho método a la función

$$f(x) = x - 0.9 \operatorname{sen}(x) - 3 \quad (2.217)$$

en el intervalo $[0, \pi]$.

Cuadro 2.2: Iteraciones obtenidas al aplicar el método de Yun-Petković a (2.217).

n	x_n	a_n	b_n	$f(x_n)$
0	1.5707963	0.	3.1415926	-1.5192036
1	3.09	1.5707963	4.6092036	0.0853587
2	3.0094032	3.09	2.9288065	-0.0024591
3	3.0116612	3.0094032	3.0139191	2.4808637
4	3.0116609	3.0116612	3.0116607	-1.6818144

Con la finalidad de comparar los métodos de la Secante, Bisección y Yun-Petković, en el Cuadro 2.3 y el Cuadro 2.4 presentamos los resultados obtenidos por el método de Bisección aplicados a (2.216) y (2.217). Asimismo, en el Cuadro 2.5 y el Cuadro 2.6 aparecen los resultados obtenidos por el método de la Secante aplicados a ecuaciones citadas anteriormente.

Convergencia local del método de Yun-Petković

El teorema siguiente demuestra que el orden de convergencia del método iterativo de Yun-Petković es cuadrático. Parte del mismo a sido tomado de [197].

Teorema 2.42 (Teorema de Yun-Petković). *Sea $f(x)$ una función dos veces continuamente diferenciable sobre un intervalo $[a, b]$ y que $f'''(x)$ existe en el intervalo (a, b) . Supongamos que $f(x)$ tiene a x^* como solución única y que $f'(x) \neq 0$ en el intervalo $[a, b]$. Entonces el método iterativo definido por (2.215) es cuadráticamente convergente una vez que una iteración x_n para algún n está lo suficientemente próximo a x^* .*

Cuadro 2.3: Iteraciones obtenidas al aplicar el método de Bisección a (2.216).

n	a_n	x_n	b_n	$f(x_n)$
0	0	1	2	1
1	0	0.5	1	-0.4375
2	0.5	0.75	1	0.06640
3	0.5	0.625	0.75	-0.2224
4	0.625	0.6875	0.75	-0.0890
5	0.6875	0.7187	0.75	-0.0143
6	0.7187	0.7343	0.75	0.0252
7	0.7187	0.7265	0.7343	0.00523
8	0.7187	0.7226	0.7265	-0.0046

Cuadro 2.4: Iteraciones obtenidas al aplicar el método de Bisección a (2.217).

n	a_n	x_n	b_n	$f(x_n)$
0	0	1	2	-2.7573
1	0	0.5	1	-2.9314
2	0	0.25	0.5	-2.9726
3	0	0.125	0.25	-2.9872
4	0	0.625	0.125	-2.9937
5	0	0.03125	0.0625	-2.9968
6	0	0.0156	0.03125	-2.9984
7	0	0.0078	0.0156	-2.9992
8	0	0.0039	0.0078	-2.9996

Cuadro 2.5: Iteraciones obtenidas al aplicar el método de la Secante a (2.216).

n	a_n	b_n	$f(a_n)$	$f(b_n)$
0	0	2	-1	17
1	0.1111	0.2049	-0.8887	-0.7932
2	0.9848	0.5648	0.9255	-0.3332
3	0.6760	0.7346	-0.1150	0.0259
4	0.7238	0.7244	-0.0015	-0.00001

Demostración. De la ecuación (2.215), usando la notación de la primera diferencia dividida

Cuadro 2.6: Iteraciones obtenidas al aplicar el método de la Secante a (2.216).

n	a_n	b_n	$f(a_n)$	$f(b_n)$
0	0	2	-3	-1.8183
1	5.0777	3.1815	2.9183	0.2174
2	3.0288	3.0669	-0.0723	-0.00009
3	3.0670	3.0670	7.6×10^{-8}	-6.8×10^{-14}
4	3.0670	3.0670	5.5×10^{-17}	5.5×10^{-17}

de Newton y la ecuación $f(x^*) = 0$, obtenemos que

$$\begin{aligned}
 x^* - x_{n+1} &= x^* - x_n + \left[\frac{b_n - a_n}{f(b_n) - f(a_n)} \right] f(x_n) \\
 &= x^* - x_n + \frac{f(x_n)}{f[a_n, b_n]} \\
 &= \frac{x^* - x_n}{f[a_n, b_n]} \left[f[a_n, b_n] - \frac{f(x^*) - f(x_n)}{x^* - x_n} \right] \\
 &= -\frac{(x^* - x_n)(x^* - b_n)}{f[a_n, b_n]} \{ f[x_n, x^*] - f[a_n, b_n] \} / (x^* - b_n) \\
 &= -\frac{(x^* - x_n)(x^* - b_n)}{f[a_n, b_n]} \left\{ \frac{f[x_n, x^*] - f[b_n, x_n]}{x^* - b_n} + \frac{f[b_n, x_n] - f[a_n, b_n]}{x^* - b_n} \right\}
 \end{aligned}$$

Introduciendo la diferencia dividida de segundo orden de Newton, sigue

$$\begin{aligned}
 x^* - x_{n+1} &= -\frac{(x^* - x_n)(x^* - b_n)}{f[a_n, b_n]} \left\{ f[b_n, x_n, x^*] - \frac{f[b_n, a_n] - f[x_n, b_n]}{a_n - x_n} \times \frac{a_n - x_n}{x^* - b_n} \right\} \\
 &= -\frac{(x^* - x_n)(x^* - b_n)}{f[a_n, b_n]} \left\{ f[b_n, x_n, x^*] - f[x_n, b_n, a_n] \times \frac{a_n - x_n}{x^* - b_n} \right\} \\
 &= -\frac{(x^* - x_n)(x^* - b_n)}{2f'(\xi_n)} \left\{ f''(\eta_{n,1}) - f''(\eta_{n,2}) \times \frac{a_n - x_n}{x^* - b_n} \right\} \quad (2.218)
 \end{aligned}$$

para algún ξ_n , $\eta_{n,1}$ y $\eta_{n,2}$ que están en un entorno de x^* que contiene el intervalo $[a_n, b_n]$.

En (2.218) se aplicaron las siguientes identidades:

$$f[\alpha, \beta] = f'(\xi), \quad f[\alpha, \beta, \gamma] = \frac{1}{2}f''(\eta)$$

para algún ξ entre α y β , y η entre el máximo y el mínimo de α , β y γ .

Por otro lado, por el teorema del valor medio para f'' , sigue

$$f''(\eta_{n,2}) = f''(\eta_{n,1}) + (\eta_{n,2} - \eta_{n,1})f'''(w_n)$$

para algún w_n entre $\eta_{n,1}$ y $\eta_{n,2}$.

Así obtenemos que

$$\begin{aligned} f''(\eta_{n,1}) - f''(\eta_{n,2}) \times \frac{a_n - x_n}{x^* - b_n} &= f''(\eta_{n,1}) - [f''(\eta_{n,1}) + (\eta_{n,2} - \eta_{n,1})f'''(w_n)] \times \frac{a_n - x_n}{x^* - b_n} \\ &= f''(\eta_{n,1}) \frac{x^* - x_n}{x^* - b_n} - (\eta_{n,2} - \eta_{n,1})f'''(w_n) \times \frac{a_n - x_n}{x^* - b_n}, \end{aligned}$$

De modo que (2.218) se convierte en

$$x^* - x_{n+1} = -\frac{(x^* - x_n)^2}{2f'(\xi_n)} f''(\eta_{n,1}) - \frac{(x^* - x_n)}{2f'(\xi_n)} \{(a_n - x_n)(\eta_{n,2} - \eta_{n,1})\} f'''(w_n). \quad (2.219)$$

Estableciendo que $e_n = |x^* - x_n|$, a continuación podemos ver que

$$|a_n - x_n| = |h_n| = |x_n - x_{n-1}| = |x_n - x^* + x^* - x_{n-1}| \leq e_n + e_{n-1}. \quad (2.220)$$

Además, suponiendo, sin pérdida de generalidad, que $h_n > 0$ para x_n próximo a x^* en la vecindad de U_n de x^* , sigue que

$$U_n := [a_n - m_n h_n, b_n + m_n h_n],$$

donde $m_n = O(1)$ es un multiplicador positivo, de tal manera que

$$x^* \in U_n, \quad \eta_{n,1}, \eta_{n,2} \in U_n, \quad [a_n, b_n] \subset U_n.$$

Entonces

$$\begin{aligned} |\eta_{n,2} - \eta_{n,1}| &\leq (b_n + m_n h_n) - (a_n - m_n h_n) \\ &= (b_n - a_n) + 2m_n h_n = 2h_n + 2m_n h_n \\ &\leq (2m_n + 2)(e_n + e_{n-1}). \end{aligned} \quad (2.221)$$

Sustituyendo (2.220) y (2.221) en (2.219), sigue

$$e_{n+1} \leq c_1 e_n^2 + c_2 e_n (e_n^2 + 2e_n e_{n-1} + e_{n-1}^2) \quad (2.222)$$

para algunas constantes positivas c_1 y c_2 . Si los errores e_n y e_{n-1} son lo suficiente pequeños, se deduce que

$$e_{n+1} \leq d_1 e_n^2 + d_2 e_n e_{n-1}^2 \quad (2.223)$$

para algunas constantes positivas d_1 y d_2 .

Supongamos que el orden de convergencia de (2.215) es $r > 0$, esto es que para algunas constantes positivas q_1 y q_2 se cumple que

$$e_{n+1} = q_1 e_n^r, \quad e_n = q_2 e_{n-1}^r.$$

Entonces (2.223) se convierte en

$$(q_1 q_2^r) e_{n-1}^{r^2} \leq (d_1 q_2^2) e_{n-1}^{2r} + (d_2 q_2) e_{n-1}^{r+2},$$

lo que implica que

$$r^2 = \min\{2r, r + 2\}.$$

Por lo tanto tenemos que $r = 2$, luego la prueba de la convergencia cuadrática está completa. □

A continuación presentamos el código Mathematica del método de Yun-Petković. Está desarrollado para el caso particular de la función $f(x) = x^4 + x - 1$ en el intervalo $[0, 2]$.

```
yun[a0_, b0_, m_] :=
Module[{},
  a = N[a0]; b = N[b0]; x0 = (a + b)/2; h =
    (a + b)/2;
  k = 0;
  output = {{k, x0, a, b, f[x0]}};
  While[k < m,
    x1 = x0 - 2 h*f[x0]/(f[b] - f[a]);
    h = x1 - x0;
    a = x1 - h;
    b = x1 + h;
    x0 = x1;
    k = k + 1;
    output = Append[output, {k, x0, a, b, f[x0]}];];
  Print[NumberForm[TableForm[output,
    TableHeadings -> {None, {"k", "xk", "a", "b", "f[xk]"}}, 16]];
]
f[x_] = x^4 + x - 1; yun[0, 2, 5]
```

2.7.2. El método de Danby

Recordemos que la ecuación de Kepler da la relación entre las coordenadas polares de un cuerpo celeste y el tiempo transcurrido desde un punto inicial dado. Cuando la anomalía media M , no es un múltiplo de π , la ecuación de Kepler tiene una solución única, pero dado que es una ecuación trascendental, la misma no puede ser resuelta directamente. Sin embargo, como resultado de su importancia en la mecánica celeste, se han derivado muchos métodos para resolver dicha ecuación.

El método de Danby, también conocido como método de Danby-Burkardt, es un proceso iterativo presentado en [51] como alternativa a otros métodos iterativos (Newton, Halley, Steffensen) para resolver la ecuación de Kepler (véase también [147]). A diferencia de estos

últimos, el método de Danby es un algoritmo especialmente diseñado para resolver la ecuación en cuestión. Incluso, junto con la descripción de la función de iteración, se sugiere una elección adecuada de puntos de partida, en función de la excentricidad media M que aparece en la ecuación de Kepler.

En este capítulo, vamos a considerar la formulación general del método de Danby para resolver una ecuación lineal cualquiera $f(x) = 0$. Teniendo en cuenta las notaciones

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}, \quad L_{f'}(x) = \frac{f'(x)f'''(x)}{f''(x)^2}$$

introducidas en secciones anteriores, el método de Danby y Burkardt puede expresarse como sigue:

$$x_{n+1} = x_n - \frac{3}{2} \frac{(2 - L_f(x_n))^2}{6 - 9L_f(x_n) + 3L_f(x_n)^2 + L_f(x_n)^2 L_{f'}(x_n)} \frac{f(x_n)}{f'(x_n)}. \quad (2.224)$$

Notas históricas y descripción del método de Danby

De acuerdo con [201], Danby nació en 1929 y murió en 2009 en Londres, pero pasó sus primeros 6 años en Palestina. Realizó un doctorado en Astronomía en la Universidad de Manchester. En 1957 aceptó una oferta para enseñar astronomía en la Universidad de Minnesota, y cuatro años después ingresó al Departamento de Astronomía de la Universidad de Yale, donde fue compañero de Brouwer Dirk. Luego, en 1967, se incorporó, como profesor de Matemáticas, de la Universidad Estatal de Carolina del Norte, donde permaneció hasta su jubilación en 1998.



Figura 2.18: J. M. A. Danby (1929 – 2009).

El método de Danby es un método de uso bastante específico y de creación relativamente reciente, así que todo el apartado histórico y su desarrollo posterior es el que se puede

encontrar en el artículo de Danby y Burkardt [51]. El mismo fue publicado en 1983 por la revista *Celestial Mechanics*, apareciendo en las páginas 95 – 107 del volumen 31.

En realidad, lo que hace Danby en [51], es obtener una familia de métodos iterativos con orden k para resolver una ecuación no lineal $f(x) = 0$.

En concreto, sea x^* una solución de $f(x) = 0$, x_n una aproximación de x^* y $\varepsilon_n = x^* - x_n$ el error cometido. Entonces

$$0 = f(x^*) = f(x_n + \varepsilon_n) = f(x_n) + f'(x_n)\varepsilon_n + \cdots + \frac{f^{(k)}(x_n)}{k!}\varepsilon_n^k + \cdots$$

Supongamos que el desarrollo de Taylor truncado en el término k -ésimo tiene una raíz, que denotaremos δ_n , es decir

$$0 = f(x_n) + f'(x_n)\delta_n + \cdots + \frac{f^{(k)}(x_n)}{k!}\delta_n^k. \quad (2.225)$$

Entonces, Danby define el proceso iterativo

$$x_{n+1} = x_n + \delta_n, \quad (2.226)$$

con δ_n solución de (2.225), y comprueba que converge a x^* con orden $k + 1$.

Notemos que para $k = 1$, el método obtenido en (2.226) no es otro que el método de Newton. Cuando $k = 2$, se obtiene el método de Euler o método de Halley irracional ([132], [180]).

$$x_{n+1} = x_n - \frac{2}{1 + \sqrt{1 - 2L_f(x_n)}} \frac{f(x_n)}{f'(x_n)}$$

Para evitar el cálculo de radicales, se puede seguir la idea vista en la construcción del método de Halley, (véase también a [70])

$$0 = f(x_n) + f'(x_n)\delta_n + \frac{1}{2}f''(x_n)\delta_n^2,$$

de donde

$$\delta_n = -\frac{f(x_n)}{f'(x_n) + \frac{1}{2}f''(x_n)\delta_n}.$$

Como vemos, tomando como aproximación del δ_n que aparece en el segundo término de la igualdad $-\frac{f(x_n)}{f'(x_n)}$, se llega al método de Halley.

La idea de Danby consiste en dar un paso más en este proceso. En concreto, para $k = 3$, la ecuación (2.225) puede escribirse como

$$\delta_n = -\frac{f(x_n)}{f'(x_n) + \frac{1}{2}f''(x_n)\delta_n + \frac{1}{6}f'''(x_n)\delta_n^2}. \quad (2.227)$$

La clave ahora es elegir las aproximaciones δ_n de la parte derecha de la fórmula (2.227). Danby [51], propone una elección en dos pasos:

$$\delta_{n_1} = -\frac{f(x_n)}{f'(x_n)} \quad (2.228)$$

$$\delta_{n_2} = -\frac{f(x_n)}{f'(x_n) + \frac{1}{2}f''(x_n)\delta_{n_1}}, \quad (2.229)$$

y así

$$\delta_n = -\frac{f(x_n)}{f'(x_n) + \frac{1}{2}f''(x_n)\delta_{n_2} + \frac{1}{6}f'''(x_n)\delta_{n_2}^2}. \quad (2.230)$$

El método de hacer

$$x_{n+1} = x_n + \delta_n, \quad (2.231)$$

con δ_n definido por (2.228), (2.229) y (2.230), es el que toma Palacios [147] como base para artículo, comparándolo con otros procesos iterativos.

El método (2.231), teniendo en cuenta las notaciones L_f y $L_{f'}$, puede escribirse como (2.224). Notemos que (2.224) es un proceso iterativo de orden de convergencia 4. El propio Danby justifica el uso de métodos con alto nivel de convergencia para resolver la ecuación de Kepler, ya que en este caso, el cálculo de la derivada de la función es relativamente sencillo.

Siguiendo a Palacios [147], nosotros también consideraremos el método (2.224) como referente para el estudio de la ecuación de Kepler y nos referiremos a él como el método de Danby.

No obstante, Danby también propone el uso de métodos con orden de convergencia 5, incrementando un poco más el proceso visto anteriormente. También propone usar el desarrollo en serie de la función

$$\frac{3}{2} \frac{(2-x)^2}{6-9x+3x^2+x^2y} = 1 + \frac{x}{2} + \frac{x^2}{2} - \frac{x^2y}{6} + O(x^3).$$

Haciendo $x = L_f(x_n)$, $y = L_{f'}(x_n)$ y despreciando los términos que contengan a $f^4(x_n)$ se obtiene una variación de (2.224) de la forma

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}L_f(x_n) + \frac{1}{2}L_f(x_n)^2 - \frac{1}{6}L_f(x_n)^2L_{f'}(x_n)\right) \frac{f(x_n)}{f'(x_n)},$$

que también tiene orden 4.

Convergencia local del método de Danby

Teorema 2.43. *Sea $f : I \rightarrow \mathbb{R}$ una función de variable real definida en un entorno I de x^* , una raíz simple de $f(x) = 0$, ($f(x^*) = 0$ y $f'(x^*) \neq 0$). Supongamos que f es cuatro veces derivable con continuidad en I . Entonces el método de Danby (2.224) definido por $x_{n+1} = D_f(x_n)$,*

$$D_f(x) = x - \frac{3}{2} \frac{(2 - L_f(x))^2}{6 - 9L_f(x) + 3L_f(x)^2 + L_f(x)^2L_{f'}(x)} \frac{f(x)}{f'(x)}, \quad (2.232)$$

tiene convergencia local de, al menos, orden 4. Además, se cumple la siguiente ecuación del error:

$$e_{n+1} = A_2^3 - A_2A_3 + A_4 + O_5,$$

donde A_i , e_n y O_N se han introducido en (2.12), (2.13) y (2.14) respectivamente.

Demostración. (Usando el Teorema de Schröder) El cálculo de las derivadas de D_f nos permite deducir que

$$D_f(x^*) = x^*, \quad D_f^{(i)}(x^*) = 0, \quad i = 1, 2, 3,$$

$$D_f^{(4)}(x^*) = \frac{3f''(x^*)^3 - 2f'(x^*)f^{(3)}(x^*)f''(x^*) + f'(x^*)^2f^{(4)}(x^*)}{f'(x^*)^3} \neq 0.$$

Por lo tanto,

$$e_{n+1} = x_{n+1} - x^* = D_f(x_n) - x^* = \frac{D_f^{(4)}(x^*)}{4!}(x_n - x^*)^4 + O_5,$$

de donde se deduce que

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^4} = \frac{D_f^{(4)}(x^*)}{4!} = 27A_3^3 - A_2A_3 + A_4.$$

Por lo tanto, el método de Danby es un método con orden de convergencia local igual a 4 y constante de error asintótico $A_2^3 - A_2A_3 + A_4$. \square

Capítulo 3

Métodos iterativos y la ecuación de Kepler

3.1. Introducción

Fue alrededor del año 1609, cuando Kepler, en su libro *Astronomia Nova*, enunció la ley horaria del movimiento elíptico, que requiere la resolución de una ecuación algebraica no lineal de la forma

$$f(E) = 0, \quad \text{con} \quad f(E) = E - e \sin E - M. \quad (3.1)$$

Para la solución de esta ecuación, es necesario conocer los valores de la anomalía media, M , y de la excentricidad, e , de una órbita elíptica.

Aunque la ecuación anterior podría formularse en un lenguaje puramente algebraico, $f(x) = x - a \sin x - b$, con a y b constantes arbitrarias, en este apartado mantendremos la formulación inicial (3.1), donde los parámetros e y M tienen claras interpretaciones astronómicas (véase [49] para más información). En concreto, tanto E como M son ángulos entre 0 y 2π radianes, mientras que la excentricidad es un valor comprendido entre 0 y 1. Además, debido a la simetría de la elipse respecto a su eje mayor, podemos reducir nuestro estudio a anomalías medias comprendidas entre 0 y π .

Existen muchas maneras de resolver la ecuación de Kepler (véase de nuevo el texto de [49] para una información más detallada a este respecto), pero en este capítulo nos centraremos en su resolución mediante métodos iterativos, tal y como se hace, por ejemplo, en los artículos de Charles y Tatum [35], Danby y Burkardt [51], [52] o Palacios [147]. En concreto, vamos a analizar el comportamiento de diversos procesos iterativos para resolver la ecuación de Kepler (3.1): método de Newton, método de Halley, método de Chebyshev, método de Danby, etc.

En consecuencia, en este acápite consideraremos la famosa ecuación de Kepler (3.1) como función test para comparar el comportamiento de diversos métodos iterativos. En concreto,

junto con el conocido método de Newton

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0 \quad (3.2)$$

se consideran también los métodos de la familia uniparamétrica

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \alpha L_f(x_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R} \quad (3.3)$$

introducida en [5]. En la fórmula anterior, y durante todo este acápite, se usa la notación

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}. \quad (3.4)$$

La familia (3.3) incluye alguno de los métodos iterativos de tercer orden más conocidos, como el método de Chebyshev ($\alpha = 0$), el método de Halley ($\alpha = 1/2$) o el método super-Halley ($\alpha = 1$). Además, como un caso límite cuando el parámetro α tiende a ∞ (ó a $-\infty$), se obtiene el método de Newton (3.2).

Por último, también consideraremos el método recomendado por Danby y Burkardt [51] para la resolución de la ecuación de Kepler (3.1). Con la notación introducida en (3.4), el método de Danby y Burkardt puede expresarse como sigue:

$$x_{n+1} = x_n - \frac{3}{2} \frac{(2 - L_f(x_n))^2}{6 - 9L_f(x_n) + 3L_f(x_n)^2 + L_f(x_n)^2 L_{f'}(x_n)} \frac{f(x_n)}{f'(x_n)}. \quad (3.5)$$

Nótese que en la expresión anterior hemos usado la notación $L_{f'}(x)$ que consiste en aplicar la fórmula (3.4) a $f'(x)$ en lugar de a $f(x)$, es decir,

$$L_{f'}(x) = \frac{f'(x)f'''(x)}{f''(x)^2}.$$

Llegados a este punto nos gustaría destacar que el método de Danby y Burkardt (3.5) no está contenido en la familia (3.3) para ningún valor de α . De hecho, (3.5) es un método con orden de convergencia 4, mientras que los métodos de la familia (3.3) tienen orden de convergencia 3 y el método de Newton (3.2) tiene convergencia cuadrática.

Para determinar la efectividad de los métodos señalados anteriormente, tomaremos en cuenta diversos puntos iniciales, los cuales estarán relacionados con la excentricidad e y la desviación media M .

En el Cuadro 3.1, presentamos las funciones de iteración de los métodos que vamos a considerar en esta sección para estudiar la ecuación de Kepler (3.1), así como algunas de sus características principales. Notemos que la constante asintótica de error, β , de cada método esta expresada en términos de las constantes A_k , definidas por:

$$A_k = \frac{f^{(k)}(x^*)}{k!f'(x^*)}. \quad (3.6)$$

Cuadro 3.1: Función de iteración, orden de convergencia (p), constante asintótica del error (β) e índice de eficiencia (δ) de los distintos métodos iterativos que vamos a aplicar a la ecuación de Kepler, donde las constantes A_k están definidas en (3.6).

Método	Función de iteración	p	β	δ
Newton	$N(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$	2	A_2	$\sqrt{2}$
Chebyshev	$C(x_n) = x_n - \left(1 + \frac{1}{2}L_f(x_n)\right) \frac{f(x_n)}{f'(x_n)}$	3	$2A_2^2 - A_3$	$\sqrt[3]{3}$
Halley	$H(x_n) = x_n - \left(\frac{2}{2-L_f(x_n)}\right) \frac{f(x_n)}{f'(x_n)}$	3	$A_2^2 - A_3$	$\sqrt[3]{3}$
Super-Halley	$SH(x_n) = x_n - \left[1 + \frac{1}{2}\left(\frac{L_f(x_n)}{1-L_f(x_n)}\right)\right] \frac{f(x_n)}{f'(x_n)}$	3	$-A_3$	$\sqrt[3]{3}$
$\alpha = 3/2$	$F_{3/2}(x_n) = x_n - \left[1 + \frac{1}{2}\left(\frac{L_f(x_n)}{1-\frac{3}{2}L_f(x_n)}\right)\right] \frac{f(x_n)}{f'(x_n)}$	3	$-A_2^2 - A_3$	$\sqrt[3]{3}$
$\alpha = 2$	$F_2(x_n) = x_n - \left[1 + \frac{1}{2}\left(\frac{L_f(x_n)}{1-2L_f(x_n)}\right)\right] \frac{f(x_n)}{f'(x_n)}$	3	$-2A_2^2 - A_3$	$\sqrt[3]{3}$
Danby	$D(x_n) = x_n - \frac{3}{2} \frac{(2-L_f(x_n))^2}{6-9L_f(x_n)+3L_f(x_n)^2+L_f(x_n)^2L_{ff}(x_n)} \frac{f(x_n)}{f'(x_n)}$	4	$-9A_2^3 + 10A_3A_2 + A_4$	$\sqrt[4]{4}$

3.2. Métodos iterativos punto a punto aplicados a la ecuación de Kepler

Con el objetivo de comparar las funciones de iteración presentadas en el Cuadro 3.1 cuando se aplican a la ecuación de Kepler (3.1), hemos seleccionado (véase el Cuadro 3.2) algunos valores de los parámetros e y M que han sido utilizados por otros autores: [49], [51], [62], [147] y [198]. Además, hemos complementado esta selección con una muestra formada por valores reales asociados a los asteroides Amor, Apollo, Chiron, Damocles, Hidalgo, Hungaria, Pholus y al Planeta Mercurio [208]. Debemos resaltar que los valores considerados en [147] corresponden a situaciones extremas para el método de Danby. Así, el caso (2), en el que M está próximo a cero, e está próximo a uno y $f'(E)$ es casi nula, puede presentar dificultad cuando el valor inicial es menor que la solución alcanzada por el método.

Como nota importante, debemos señalar que todos los métodos iterativos fueron aplicados utilizando el paquete informático Mathematica, versión 7, los cálculos se realizaron tomando 10 cifras significativas.

Un aspecto clave a la hora de estudiar la convergencia de un proceso iterativo a la solución de una ecuación no lineal, es la elección del punto de partida. En el caso de la ecuación de Kepler (3.1), algunos autores, como Diloné y Gutiérrez [62], Zhen-Shi-Ming [198], han considerado la anomalía media como punto de partida $E_0 = M$. Otros autores, como Danby-Burkardt [51] o Palacios [147], sugieren una elección de puntos de partida más sofisticada:

$$\begin{cases} E_0 = M + (\sqrt[3]{6M} - M)e^2 & \text{si } M < 0.1 \\ E_0 = M + 0.85e & \text{si } M \geq 0.1. \end{cases} \quad (3.7)$$

Con las elecciones de puntos de partida anteriores, no se garantiza siempre la convergencia

Cuadro 3.2: Valores de la anomalía media y de la excentricidad utilizados por M. Palacios [147].

Casos	M	e
Caso (1)	0.05	0.09
Caso (2)	0.05	0.95
Caso (3)	0.9	0.09
Caso (4)	1.5	0.5
Caso (5)	3.0	0.09
Caso (6)	3.0	0.9

Cuadro 3.3: Valores reales asociados a la anomalía media y a la excentricidad de diferentes asteroides y el planeta Mercurio [208].

Casos	M	e
Caso (7) Asteroide Amor	0.51π	0.4353621
Caso (8) Asteroide Apollo	0.088π	0.5600245
Caso (9) Asteroide Chiron	0.29π	0.3821080
Caso (10) Asteroide Damocles	0.29π	0.8662750
Caso (11) Asteroide Hidalgo	0.81π	0.6581069
Caso (12) Asteroide Hungaria	0.69π	0.0736162
Caso (13) Asteroide Pholus	0.11π	0.5717136
Caso (14) Planeta Mercurio	0.0067π	0.205635

del método de Newton (3.2) a una solución de la ecuación de Kepler (3.1). Por ello, Charles y Tatum [35], proponen tomar $E_0 = \pi$, ya que justifican que para ese valor se puede asegurar la convergencia del método de Newton (3.2) a una solución de la ecuación de Kepler (3.1) para cualquier elección de M y e .

En este capítulo mostramos varios experimentos numéricos para ilustrar el comportamiento dinámico de otros métodos iterativos, en particular los recogidos en el Cuadro 3.1, cuando se aplican a la ecuación de Kepler (3.1). En concreto, en el Cuadro 4.5 se muestra el comportamiento de dichos métodos cuando se toma como punto de partida $E_0 = M$. De forma análoga en los Cuadros 4.6 y 4.7 se analizan las situaciones $E_0 = \pi$ y E_0 dado por (3.7). En los tres cuadros se muestra el número total de iteraciones necesarias para alcanzar la solución para cada uno de los métodos iterativos recogidos en el Cuadro 3.1 y para los 14 pares de valores (e, M) dados en los Cuadros 3.3 y 4.5.

Junto al número total de iteraciones, T , aparece la media de las mismas, m . Notemos que el número de iteraciones no es un criterio muy razonable para comparar métodos con distinto orden de convergencia pues, evidentemente, los métodos con orden más elevado necesitarán

menor número de iteraciones para alcanzar la solución (en el caso de que haya convergencia en todos los casos). Por este motivo en este trabajo introducimos una media modificada, \bar{m} , calculada de la siguiente manera. Supongamos que para alcanzar una precisión prefijada con un método de orden q se necesitan j iteraciones. Entonces, para alcanzar la misma precisión con un método de orden p se necesitarían del orden de $j \frac{\log q}{\log p}$ iteraciones. Con este criterio, comparamos las medias obtenidas con la del método de Newton, modificando la media por $\frac{\log 3}{\log 2}$ para los métodos de orden 3 y por $\frac{\log 4}{\log 2} = 2$ para el método de Danby (orden 4).

En el caso de que un método iterativo no genere una sucesión convergente a una raíz de la ecuación de Kepler (3.1), lo indicaremos con el símbolo NC.

Lo primero que observamos (consúltese el Cuadro 4.5) es que no se obtiene una sucesión convergente en todos los casos. Así, se comprobó que el método de Chebyshev, bajo las condiciones del asteroide Damocles, presentó una divergencia oscilatoria.

Además, la media ponderada sitúa al método de Newton en mejor posición que el resto de los métodos considerados. En cuanto a los métodos de tercer orden, los mejor posicionados son los de Halley y super-Halley. La utilización de métodos de la familia (3.3) con valores de $\alpha > 1$ parece no aportar, en este sentido, ninguna ventaja. Del mismo modo, y atendiendo de nuevo a la media ponderada, el método de Danby, tampoco aporta ninguna mejora sustancial.

Los casos más conflictivos

En la sección anterior hemos analizado el comportamiento de diversos métodos iterativos para resolver la ecuación de Kepler, para distintos pares (e, M) . Como se ha podido observar, sólo en un caso, en concreto para el método de Chebyshev bajo las condiciones del asteroide Damocles, se presentó un comportamiento no convergente. Para todos los demás casos, los métodos considerados han convergido a una solución de (3.1).

Sin embargo, si se profundiza un poco más, se pueden detectar comportamientos no convergentes, incluso para el método de Newton. Así, tal y como indica Conway [48], también existen pares (e, M) para los cuales el método de Newton no converge a una solución de (3.1). Esto ocurre, por ejemplo, para $e = 0.992$, $M = 0.13\pi$. Por ese motivo, en esta sección particularizamos el estudio de los métodos iterativos para resolver la ecuación de Kepler aplicándolos a los pares (e, M) , en donde $e \in [0.980, 0.999]$ y $M \in [0.02\pi, 0.2\pi]$. Para cada uno de los métodos iterativos del Cuadro 3.1 realizaremos tres experimentos numéricos, tomando como puntos de partida $E_0 = M$, $E_0 = \pi$ y E_0 dado por las fórmulas (3.7). Dichos experimentos se muestran en los cuadros del Anexo II.

Los cuadros 3.4, 3.5 y 3.6 muestran de forma resumida la información obtenida a partir de los cuadros del Anexo II.

Cuadro 3.4: Media m del número de iteraciones de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.

E_0	Newton	Chebyshev	Halley	Super-Halley	Danby
$E_0 = M$	13.76	(*)	4.61	4.43	4.23
$E_0 = \pi$	5.4	4.04	4.01	3.83	3.63
$E_0 = (3.7)$	3.77	2.86	2.77	2.76	2.73

(*) Media no calculada debido al elevado número de situaciones de no convergencia.

Cuadro 3.5: Media ponderada \bar{m} del número de iteraciones de de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.

E_0	Newton	Chebyshev	Halley	Super-Halley	Danby
$E_0 = M$	20.7	(*)	8.93	7.02	8.46
$E_0 = \pi$	7.22	6.38	6.36	5.74	7.26
$E_0 = (3.7)$	5.04	4.54	4.39	4.37	5.46

(*) Media no calculada debido al elevado número de situaciones de no convergencia.

Cuadro 3.6: Porcentaje de fracasos de diversos métodos iterativos aplicados a la ecuación de Kepler, en función de los puntos de partida.

E_0	Newton	Chebyshev	Halley	Super-Halley	$\alpha = \frac{3}{2}$	$\alpha = 2$	Danby
$E_0 = M$	0	97.5	0	0	0	0.5	0
$E_0 = \pi$	0	0	0	0	0	0	0
$E_0 = (3.7)$	0	0	0	0	0	0	0

Conclusiones

A la vista de los cuadros 3.4, 3.5 y 3.6, obtenemos las siguientes conclusiones:

1. Los peores resultados se presentaron cuando los métodos iterativos fueron aplicados a partir de la condición inicial $E_0 = M$. Además de obtener el mayor porcentaje de fracasos (especialmente llamativo es el caso del método de Chebyshev, como se aprecia en el Cuadro 4.11), cuando lograron converger, obtuvieron las medias de iteraciones más altas, consúltese el Cuadro 3.4.
2. El porcentaje de fracasos del método de Chebyshev fue de un 97.5 por ciento, equivalentes a 195 casos de los 200 considerados, véase el Cuadro 3.6. Para dicho porcentaje, el método presentó una divergencia oscilatoria. Por ejemplo, para $e = 0.99$, $M = \pi/10$ y $x_0 = M$, las primeras iteraciones del método de Chebyshev aplicado a la ecuación de

Kepler dan lugar a la siguiente sucesión:

$$\{0.31416, -66.1291, -76.4247, 21154.9, 3.41 \times 10^{10}, -2.5 \times 10^{19}, -5.85 \times 10^{37}, \\ 9.46 \times 10^{75}, -1.8 \times 10^{150}, -3.06 \times 10^{299}, 1.95 \times 10^{602}, \dots\}$$

Como puede verse, se obtiene una sucesión que tiende a infinito en valor absoluto, pero con signos que pueden ser tanto positivos como negativos.

3. El método obtenido al sustituir a $\alpha = 2$ en (3.3), véase el Cuadro 4.23, no logró converger en todos los casos. En concreto su porcentaje de fracasos, como se aprecia en el Cuadro 3.6, fue un 0.5 por ciento, equivalente a un caso de los 200 considerados. El mismo presentó un ciclo atractor, y se obtuvo cuando $e = 0.996$ y $M = 0.04\pi$.

$$\{0.1256, 8.0357, 3.7986, 1.7757, -0.20402, \dots \\ 3.64608, 1.788161, -0.204782, 3.64608, 1.788161, -0.204782, \dots\}$$

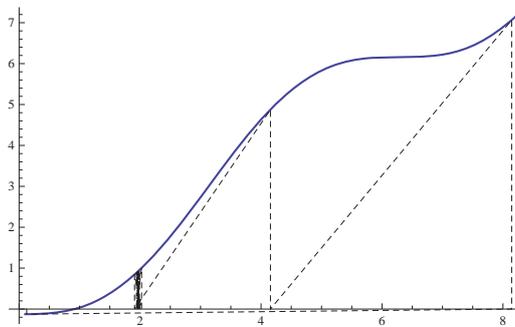


Figura 3.1: Ciclo atractor del método obtenido al sustituir $\alpha = 2$ en (3.3), para $e = 0.996$, $M = 0.04\pi$ y $x_0 = M$

4. Todos los demás métodos lograron su convergencia, pero en cada uno de ellos se presentaron situaciones para las cuales se necesitaron hasta 200 iteraciones para converger, consúltese el Cuadro 4.8, Cuadro 4.17, Cuadro 4.20 y el Cuadro 4.26. En concreto, el método de Newton, el Cuadro 3.4, obtuvo la media de las medias de las iteraciones más elevada $m_{Newton} = 23.73$. Se destaca que para muchos pares, véase nuevamente el Cuadro 4.8, se necesitaron más de 100 iteraciones para que dicho método lograra converger. Con respecto al método de Super-Halley, véase el Cuadro 4.17, para el par $(0.999, 0.2\pi)$ dicho método realizó más de 200 iteraciones para converger. En cuanto al método obtenido al sustituir a $\alpha = 3/2$ en (3.3), consúltese el Cuadro 4.20, se destaca el hecho de que para los pares $(0.996, 0.02\pi)$ y $(0.996, 0.04\pi)$, dicho método necesito más de 200 iteraciones para converger. El método de Danby, véase el Cuadro 4.26, fue el que obtuvo el mejor desempeño, aunque es conveniente destacar que para los pares (e, M) con $e \in [0.993, 0.998]$ y $M = 0.06\pi$, dicho método necesito entre 14 y 56 iteraciones para converger.

5. Bajo las condiciones iniciales dadas en (3.7), véase el Cuadro 4.10, Cuadro 4.13, Cuadro 4.16, Cuadro 4.19, Cuadro 4.22, Cuadro 4.25 y Cuadro 4.28, los métodos iterativos, aplicados a la ecuación de Kepler, experimentaron su mejor comportamiento de convergencia, esto así porque para tal condición inicial, los métodos iterativos obtuvieron la menor media m del número de de iteraciones, consúltese el Cuadro 3.4.
6. Asimismo, para la condición inicial $E_0 = M$, refiérase al Cuadro 4.5, Cuadro 4.8, Cuadro 4.11, Cuadro 4.14, Cuadro 4.17, Cuadro 4.20, Cuadro 4.23 y Cuadro 4.26, los métodos iterativos, al ser aplicados a la ecuación de Kepler, mostraron su peor comportamiento de convergencia, esto así porque para tal condición inicial, los métodos iterativos presentaron el mayor porcentaje de fracasos y cuando lograron converger, presentaron la mayor media m de iteraciones, consúltese el Cuadro 3.4 y el Cuadro 3.6.
7. Por otro lado, para la condición inicial $E_0 = M$, véase el Cuadro 3.4, se verificó que el método de Danby supera a los demás métodos, pues obtuvo la menor media m del número de iteraciones $m_{danby} = 5.6$. Asimismo, para las condiciones iniciales $E_0 = \pi$ y las dadas en (3.7), consúltese nuevamente el Cuadro 3.4, se comprueba que el método de super-Halley es el más efectivo, pues en ambos casos obtuvo la menor media m del número de iteraciones, $m_{super-Halley} = 4.23$ y $m_{super-Halley} = 3.18$ respectivamente.
8. Para algunos casos, consúltese el Cuadro 4.5, el Cuadro 4.11, el Cuadro 4.12 y el Cuadro 4.23, el comportamiento dinámico de la ecuación de Kepler presenta ciclos o comportamiento caótico. Estos comportamientos pueden ser motivos de posteriores estudios. Por ejemplo, en la sección 3.5 mostramos la existencia de ciclos atractores para el método de Newton aplicado a la ecuación de Kepler.

Para finalizar esta sección, a continuación presentamos un análisis gráfico de la convergencia de los métodos iterativos aplicados a la ecuación de Kepler. Para tal acción hemos seleccionado como punto de inicio a $x_0 = M$. Los gráficos están en blanco y negro. Se debe destacar que para el análisis de los gráficos asumimos que el blanco indica que se ha encontrado una raíz y el negro que no se ha encontrado raíz.

Visto así, en la Figura 3.2, podemos apreciar las gráficas generadas por el método de Newton aplicado a la ecuación de Kepler. Notemos que en la esquina superior izquierda de la Figura 3.2 de la izquierda, aparecen unos puntos negros. La Figura 3.2 de la derecha, es una ampliación que nos permite precisar más la resolución del gráfico. En dicha ampliación se observa una “nube” de puntos negros asociados a puntos de partida donde el método de Newton aplicado a la ecuación de Kepler no converge. Tal parece que dichos puntos están localizados dentro de una curva, la cual sería interesante caracterizar.

Al experimentar con el método de Halley, obtuvimos el gráfico que aparece a la izquierda de la Figura 3.3. Notemos que en la misma no aparecen manchas negras, por lo que parecería que el método de Halley, sin importar el punto de partida, siempre converge. En cuanto al

método de super-Halley, ver Figura 3.3 a la derecha, también presenta una “nube” de puntos negros que se encuentra delimitada por una curva. Notemos que la misma posee un radio de curvatura menor que la del método de Newton.

Repitiendo el experimento para el método de Chebyshev, se obtienen cosas interesantes. en la Figura 3.4 se verifica la existencia de bandas negras muy acentuadas. Observemos que en la curva del gráfico en donde la excentricidad está entre 1 y 0.93, ya no se presenta una “nube” de puntos negros, sino más bien una banda de escasos puntos blancos.

Los experimentos realizados para otros métodos considerados en esta sección, como por ejemplo para la familia con $\alpha = 3/2$ o $\alpha = 2$ aparece una situación similar a la obtenida para el método de Halley.

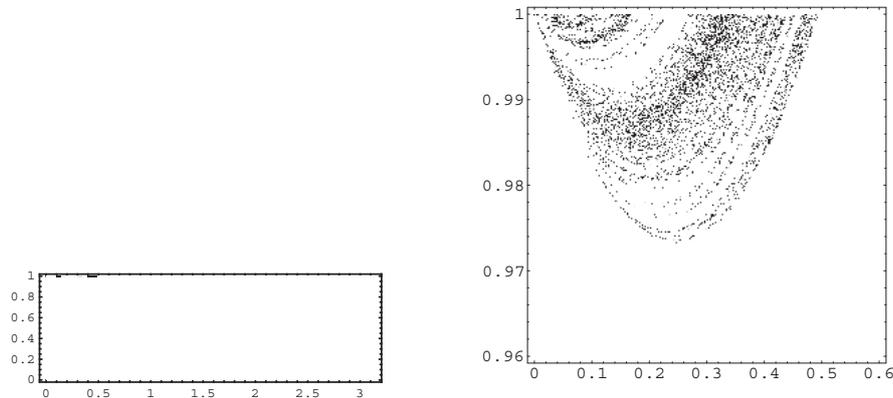


Figura 3.2: A la izquierda, gráfico de densidad del método de Newton aplicado a la ecuación de Kepler. A la derecha, gráfico de densidad ampliado del método de Newton aplicado a la ecuación de Kepler.

3.3. Métodos iterativos multipunto aplicados a la ecuación de Kepler

En esta sección se analizará el comportamiento de los métodos de la Secante, Yun-Petković y Bisección aplicado a la ecuación de Kepler. La finalidad de incluir el método de Bisección es tomarlo como referencia para compararlos con el método de la Secante y el método de Yun-Petković. Esto es así porque todos ellos inician sus iteraciones a partir de dos valores.

El método de Bisección nos sirve para encontrar una raíz de una función $f(x)$ continua en un intervalo $[a, b]$. Para encontrar la raíz, determinamos una sucesión de intervalos, cada uno la mitad del anterior, y cada uno conteniendo al menos una raíz de $f(x)$.

Su principal desventaja es que su convergencia es muy lenta, pues los cálculos aumentan considerablemente cuando se desea una buena aproximación de la raíz. Este inconveniente

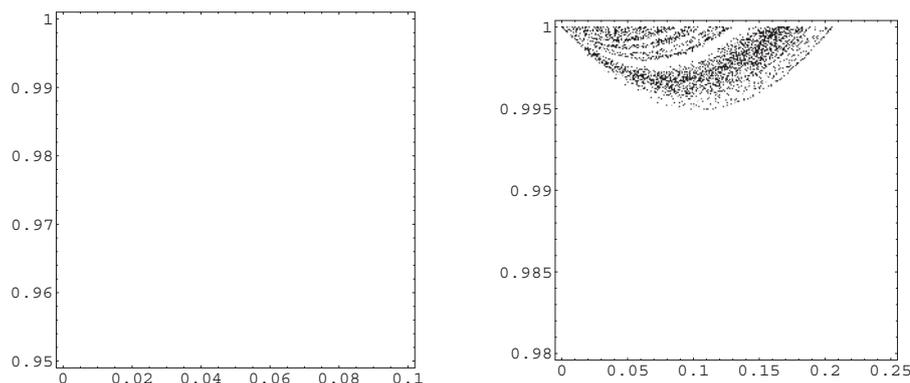


Figura 3.3: A la izquierda, gráfico de densidad ampliado del método de Halley aplicado a la ecuación de Kepler. A la derecha, gráfico de densidad ampliado del método super-Halley aplicado a la ecuación de Kepler.

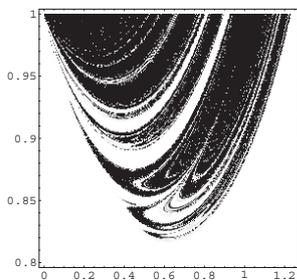


Figura 3.4: Gráfico de densidad ampliado del método de Chebyshev aplicado a la ecuación de Kepler.

hace que este método se aplique como un paso previo en la implementación de otros métodos iterativos de mayor orden de convergencia. Tiene la ventaja de que siempre converge a alguna de las raíces de la función $f(x)$ en $[a, b]$. En concreto, se busca una solución de una ecuación $f(x) = 0$, donde $f(x)$ es una función continua en un intervalo cerrado $[a, b]$, tal que $f(a)f(b) < 0$. En este caso, el teorema de Bolzano garantiza la existencia de soluciones de dicha ecuación. Además, si $f(x)$ es creciente (o decreciente en $[a, b]$), dicha solución es única.

Con la intención de describir este método, asumamos que

$$a_0 = a \quad \text{y} \quad b_0 = b,$$

y estimemos un punto inicial dado a partir del punto medio de $[a, b]$, es decir

$$x_0 = \frac{a_0 + b_0}{2}.$$

Si $f(x_0) = 0$, finaliza el proceso. En caso contrario, la raíz buscada debe estar en uno de los intervalos $[a_0, x_0]$, $[x_0, b_0]$. En concreto, la raíz estará en aquel intervalo en el que el

producto de los valores de $f(x)$ en sus extremos sea negativo. Por tanto, hallamos $f(a_0)f(x_0)$. Si $f(a_0)f(x_0) < 0$, ponemos

$$\begin{cases} a_1 = a_0, \\ b_1 = x_0. \end{cases}$$

mientras que si $f(a_0)f(x_0) > 0$, tomamos

$$\begin{cases} a_1 = x_0, \\ b_1 = b_0, \end{cases}$$

Entonces calculamos

$$x_1 = \frac{a_1 + b_1}{2}.$$

En general, dados a_n, b_n y x_n , si $f(a_n)f(x_n) < 0$, tomamos

$$\begin{cases} a_{n+1} = a_n, \\ b_{n+1} = x_n, \end{cases}$$

mientras que si $f(a_n)f(x_n) > 0$, elegimos a

$$\begin{cases} a_{n+1} = x_n, \\ b_{n+1} = b_n, \end{cases}$$

y determinamos la iteración $n + 1$ mediante la suceción

$$x_{n+1} = \frac{a_{n+1} + b_{n+1}}{2}, \tag{3.8}$$

la cual define la función de iteración del método de Bisección.

Cuadro 3.7: Función de iteración, orden de convergencia (p) e índice de eficiencia (δ) de los métodos iterativos de la Secante, Yun-Petković y Bisección.

Método	Función de iteración	p	δ
Secante	$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n)$	$\frac{1+\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$
Yun-Petković	$x_{n+1} = x_n - \left[\frac{2h_n}{f(b_n) - f(a_n)} \right] f(x_n)$	2	$\sqrt{2}$
Bisección	$x_{n+1} = \frac{a_{n+1} + b_{n+1}}{2}$	1	1

Notemos que en cada paso, la longitud del intervalo que encierra la raíz se reduce a la mitad, de forma que la convergencia del método de Bisección siempre está asegurada. Esto queda demostrado en el siguiente resultado. Parte de la demostración que presentamos a continuación aparece en [58] y [194].

Teorema 3.1. Sea $f(x)$ una función continua en un intervalo $[a, b]$ tal que $f(a)f(b) < 0$, entonces la sucesión definida por (3.8) converge a una raíz de $f(x) = 0$.

Demostración. La sucesión (a_n) es una sucesión no decreciente y acotada, mientras que la sucesión (b_n) es no creciente y acotada, por lo tanto, ambas son convergentes. Sean γ y β sus límites respectivos. Toda vez que

$$\lim_{n \rightarrow \infty} |a_n - b_n| = 0,$$

se sigue que $\gamma = \beta$. Además

$$a_n < x_n < b_n,$$

así que

$$\lim_{n \rightarrow \infty} x_n = \gamma.$$

Más aún, como $f(a_n)f(b_n) < 0$, para $n \geq 0$, se tiene que

$$0 \geq \lim_{n \rightarrow \infty} f(a_n)f(b_n) = \lim_{n \rightarrow \infty} f(a_n) \lim_{n \rightarrow \infty} f(b_n) = f(\gamma)^2,$$

luego

$$f(\gamma) = 0$$

y la sucesión generada por (3.8) converge a una raíz de $f(x)$.

□

A continuación presentamos una demostración alternativa del teorema anterior.

Demostración. Por el teorema de Bolzano existe, al menos, $\xi \in (a, b)$ tal que $F(\xi) = 0$. Suponiendo que se han separado las raíces de la ecuación y que ξ es la única raíz de

$$f(x) = 0 \quad \text{en } [a, b], \quad (3.9)$$

dividimos el intervalo $[a, b]$ por la mitad. Por lo que se pueden presentar dos casos

1. En el caso de que

$$f\left(\frac{a+b}{2}\right) = 0,$$

entonces

$$\xi = \frac{a+b}{2}.$$

2. Pero si se cumple que

$$f\left(\frac{a+b}{2}\right) \neq 0,$$

se elige el intervalo

$$\left[a, \frac{a+b}{2}\right] \quad \text{ó} \quad \left[\frac{a+b}{2}, b\right],$$

en cuyos extremos la función f toma signos opuestos. Esto se producirá porque al ser ξ la única raíz de f en $[a, b]$ sólo uno de estos intervalos tendrá la propiedad de que f cambie de signo en los extremos del mismo.

Denotando a éste intervalo por $[a_1, b_1]$, lo dividimos por la mitad y procedemos de forma análoga a la anterior.

Reiterando este proceso podemos obtener la raíz exacta de $f(x)$ o conseguir una sucesión de intervalos cerrados $\{[a_n, b_n]\}_{n=1}^{\infty}$, encajonados, es decir

$$a \leq a_1 \leq \dots \leq a_n \leq a_{n+1} \leq b_{n+1} \leq \dots \leq b_1 \leq b, \quad n \in \mathbb{N} \quad (3.10)$$

en los cuales

$$f(a_n)f(b_n) < 0, \quad (3.11)$$

y

$$b_n - a_n = \frac{b - a}{2^n}, \quad n \in \mathbb{N}. \quad (3.12)$$

La propiedad (3.10), indica que $\{a_n\}_{n=1}^{\infty}$ es una sucesión monótona creciente y acotada y que $\{b_n\}_{n=1}^{\infty}$ es una sucesión monótona decreciente y acotada. Por tanto existen

$$\eta_n \in [a, b], \quad n = 1, 2, \dots$$

tales que

$$\eta_1 = \lim_{n \rightarrow +\infty} a_n \leq \lim_{n \rightarrow +\infty} b_n = \eta_2. \quad (3.13)$$

Ahora, aplicando (3.12) se tiene que

$$\lim_{n \rightarrow +\infty} (b_n - a_n) = 0, \quad \text{luego } \eta_1 = \eta_2 = \eta,$$

por lo que la continuidad de f junto con la condición dada en (3.13) hacen que

$$(F(\eta))^2 \leq 0,$$

de donde $f(\eta) = 0$. Así que $\eta = \xi$ es la única raíz de la ecuación $f(x) = 0$.

□

Análisis numéricos

En esta sección compararemos las funciones de iteración presentadas en el Cuadro 3.7, cuando se aplican a la ecuación de Kepler. Para ello, continuaremos utilizando las 14 condiciones establecidas para los valores de los parámetros e y M dados en el Cuadro 3.2 y en el Cuadro 3.3.

En concreto, en el Cuadro 3.8 se muestra la media m del número de iteraciones necesarias para la convergencia de los métodos de la Secante, Yun-Petković y Bisección aplicados a la ecuación de Kepler, tomando como puntos de partida los extremos del intervalo $[0, \pi]$, mientras que en el Cuadro 3.9 representamos el porcentaje de fracasos que obtuvieron dichos métodos. Finalmente, en el Cuadro 3.10 indicamos el número de iteraciones que necesitó cada método para alcanzar la convergencia.

Cuadro 3.8: Media m del número de iteraciones necesarias para la convergencia de los métodos de la Secante, Yun-Petković y Bisección aplicados a la ecuación de Kepler, tomando como puntos de partida los extremos del intervalo $[0, \pi]$.

Intervalo	Secante	Yun-Petković	Bisección
$[0, \pi]$	5.07	3.43	14.43

Cuadro 3.9: Porcentaje de fracasos de los métodos de la Secante, Yun-Petković y Bisección aplicados a la ecuación de Kepler, tomando como puntos de partida los extremos del intervalo $[0, \pi]$.

Intervalo	Secante	Yun-Petković	Bisección
$[0, \pi]$	0	0	0

Cuadro 3.10: Número de iteraciones necesarias para alcanzar la solución de la ecuación Kepler, usando el método de la Secante, Yun-Petković y Bisección, y tomando como punto de partida los extremos del intervalo $[0, \pi]$.

Casos	Secante	Yun-Petković	Bisección
Caso (1)	4	3	14
Caso (2)	10	5	16
Caso (3)	5	3	16
Caso (4)	6	3	13
Caso (5)	3	3	10
Caso (6)	3	5	11
Caso (7)	5	2	16
Caso (8)	4	4	14
Caso (9)	5	3	15
Caso (10)	7	3	17
Caso (11)	5	4	13
Caso (12)	4	2	16
Caso (13)	6	4	13
Caso (14)	4	4	18
T	71	48	202
m	5.07	3.43	14.43

Conclusiones

Tomando en cuenta los datos representados en los cuadros 3.8, 3.9 y 3.10, obtenemos las siguientes conclusiones:

1. Los peores resultados de convergencia, refiérase al Cuadro 3.8 y al Cuadro 3.10, fueron obtenidos por el método de Bisección. Es importante destacar que para ciertas condiciones, véase el Cuadro 3.10, el método de la Secante y el método de Yun-Petković, en término de cantidad de iteraciones, tuvieron igual desempeño.
2. El porcentaje de fracasos de los métodos de la Secante, Yun-Petković y Bisección, al ser aplicados a la ecuación de Kepler bajo las condiciones dadas en los cuadros 3.2 y 3.3, fue nula. Esto es así, véase el Cuadro 3.9 y el Cuadro 3.10, porque para todos los casos establecidos, los métodos consiguieron converger.
3. Se debe destacar el buen comportamiento que tienen los métodos de la Secante, Yun-Petković y Bisección comparados con los métodos de un solo punto.

3.4. Análisis del método de Newton, según los puntos de inicio, aplicado a la ecuación de Kepler

La ecuación de Kepler tiene el privilegio de ser la primera ecuación no lineal que Newton, mediante un proceso más o menos iterativo, trató de resolver. Recordar que dicha ecuación tiene la forma

$$E = M + e \operatorname{sen}(E). \quad (3.14)$$

Para la solución de la ecuación de Kepler, se ha comprobado que la convergencia por iteración directa es muy lenta, por lo que es usual aplicar la función de iteración del método de Newton, que de hecho, vuelve a escribir la ecuación (3.14) en la forma

$$E_{n+1} = \frac{M - e \left(E_n \cos(E_n) - \operatorname{sen}(E_n) \right)}{1 - e \cos(E_n)}. \quad (3.15)$$

Charles y Tatum, [35], presentan un ejemplo, para el cual $e = 0.95$ y $M = 245$ grados, tomando como aproximación inicial a $E_0 = 215$ grados, en que la iteración directa de la ecuación de Kepler necesita de 68 iteraciones para alcanzar 10 cifras significativas, mientras que el método de Newton, para alcanzar el mismo resultado, tan sólo necesita 2 iteraciones.

Debido a la rápida convergencia del método de Newton, incluso en situaciones de grandes excentricidades para las cuales E y M son muy diferentes, se acostumbra a elegir a $E_0 = M$ como punto de inicio. Conway, [48], al estudiar el comportamiento de la ecuación de Kepler para $e > 0.99$, comprobó, después de aplicar más de 200 combinaciones de e y M , que al tomar a $E_0 = M$ como primera estimación, el método de Newton experimentó una rápida convergencia. Sin embargo, existen ciertos pares para los cuales el método de Newton, al tomar a $E_0 = M$, no convergen. Por ejemplo, si bien no hay dificultad alguna para $(e, M) = (0.991, 0.13\pi)$ o para $(e, M) = (0.993, 0.13\pi)$, el método de Newton, con $E_0 = M$, parece no converger para $(e, M) = (0.992, 0.13\pi)$

Conway, [48], mostró, que un procedimiento conocido como algoritmo de Laguerre, siempre converge tomando como primera iteración $E_0 = M$. Charles y Tatum, ([35]), demostraron que la función de iteración del método de Newton aplicada a la ecuación de Kepler, siempre converge tomando como primera iteración $E_0 = \pi$. Además, presentaron una modificación de la función de iteración de Newton, que puede converger a partir de $E_0 = M$, incluso en aquellos casos en los que se ha encontrado comportamiento caótico. Asimismo, Charles y Tatum, [35], realizaron un análisis gráfico, véase la Figura 3.5, en el que se presenta la función

$$f(E) = M - E + e \operatorname{sen}(E), \quad (3.16)$$

graficada para $M = 0.2$ y $e = 0.99$ para distintos intervalos de E .

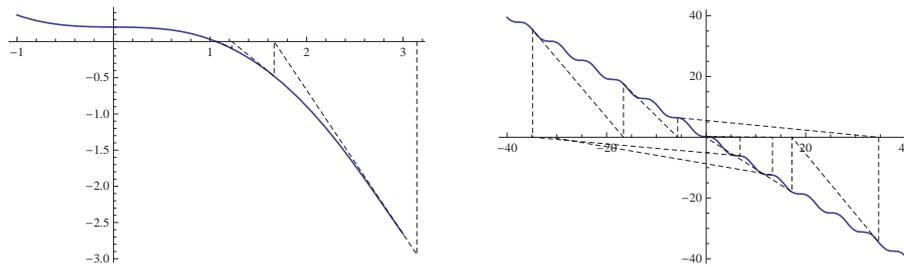


Figura 3.5: A la izquierda, gráfica de $f(E) = M - E + e \operatorname{sen}(E)$, para $e = 0.99$, $M = 0.2$ radianes y $E_0 = \pi$. A la derecha, gráfica de $f(E) = M - E + e \operatorname{sen}(E)$, para $e = 0.99$, $M = 0.2$ radianes y $E_0 = M$.

En la Figura 3.5, a la izquierda, la figura está graficada para valores pequeños de E , mientras que a la derecha, lo está para valores mucho más grandes. Concluyeron que independientemente de los valores M y de e en el rango $0 < e < 1$, la curva tiene las siguientes propiedades:

1. La intersección con el eje vertical, $E = 0$, es $f(0) = M$.
2. La curva de la función $f(E)$ es monótonamente decreciente.
3. En ninguna parte la curva es horizontal. Hay puntos de inflexión en $E = 2n\pi$, en donde n es un entero, positivo, negativo o cero. Sin embargo, estos puntos de inflexión no son horizontales. Para valores de e próximos a 1, la curva en el punto de inflexión está muy cerca de la horizontal, pero, sin embargo, no es del todo horizontal.
4. La solución a la ecuación de Kepler se encuentra donde el eje horizontal cruza la curva.
5. Cambiar el valor de M aumenta o disminuye la curva sin cambiar su forma. En consecuencia, la solución a la ecuación de Kepler para valores diferentes de M , se encuentra simplemente trazando una línea horizontal diferente a través de la curva. Un ejemplo de esto se presenta en la Figura 3.6 .

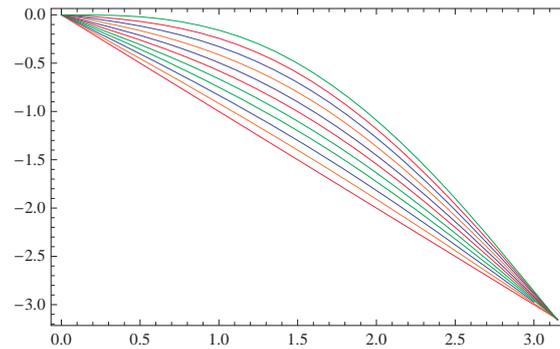


Figura 3.6: Gráfica de (3.16), para $M = 0$ y $e = j/10$ $j = 0, 1 \dots 10$.

En la Figura 3.5 a la izquierda, se muestra la gráfica de (3.16) para $E_0 = \pi$. Las iteraciones sucesivas se presentan en el Cuadro 3.11.

Cuadro 3.11: Valores obtenidos a partir de (3.15), en donde $M = 0.2$, $e = 0.99$ y $E_0 = \pi$.

E_n	E_{n+1}
0	π
1	1.663405391
2	1.225819545
3	1.094289037
4	1.067240193
5	1.066997414
6	1.066997365
7	1.066997365

Como se ilustra en la Figura 3.5, a la izquierda, una condición suficiente para la convergencia de (3.15) es que la curva sea convexa y sin inflexiones. Esta condición se satisface siempre por la ecuación de Kepler en la región comprendida entre la solución y π . Por lo tanto, siempre surge una rápida convergencia cuando la primera aproximación E_0 está en el rango $x^* \leq E_0 \leq \pi$, donde x^* es la solución de la ecuación de Kepler. Sin embargo, para valores $E_0 < x^*$, el límite inferior de este rango es desconocido. Por lo tanto, lo que tenemos como algo seguro es que tendremos convergencia cuando $E_0 = \pi$.

La Figura 3.5, a la derecha, muestra el intento fallido de resolver la ecuación $E = 0.2 + 0.99 \sin(E)$ utilizando $E_0 = M = 0.2$ como primera aproximación. Se debe destacar el hecho de que el valor inicial $E_0 = M$ siempre se encuentra donde la curva de $f(E)$ es cóncava hacia la línea de $f(E) = 0$, violando así la norma que garantiza la convergencia. El comportamiento errático de las iteraciones se muestran en el Cuadro 3.12.

En síntesis, la convergencia siempre se puede obtener si el valor inicial es elegido en el intervalo entre la raíz y π , esto es porque la curva de $f(E)$ es convexa hacia arriba, hacia

Cuadro 3.12: Valores obtenidos a partir de (3.15), en donde $M = 0.2$, $e = 0.99$ y $E_0 = M$.

E_n	E_{n+1}
0	0.2
1	6.814719036
2	-34.88627106
3	-16.60736508
4	-5.726152563
5	34.66910090
6	17.2386234
7	-0.100424545
8	13.32179834
9	-31.22883135
10	1127.740781
11	560.0144314
12	-1197.763463

el eje E , en una zona donde la raíz y el arco no incluye un punto de inflexión. De hecho, la derivación del método de Newton se basa en el supuesto de que se cumplan esos requisitos.

Por otro lado, Charles y Tatum, [35], al analizar las estimaciones iniciales de partida E_0 , en 3.15,

$$E_1 = \frac{M - e \left(E_0 \cos(E_0) - \operatorname{sen}(E_0) \right)}{1 - e \cos(E_0)}, \quad (3.17)$$

$$E_2 = \frac{M - e \left(E_1 \cos(E_1) - \operatorname{sen}(E_1) \right)}{1 - e \cos(E_1)}, \quad (3.18)$$

encontraron un caso especial de aparente falta de convergencia del método de Newton cuando $E_2 = E_0$. Establecieron que de imponerse la condición anterior, la iteración oscilará entre dos valores de E , ninguno de los cuales es la solución correcta. Por ejemplo, si $e = 0.9$, $M = 1$ y $E_0 = 5.32050617159410777$, se tiene que $E_2 = E_0$. Es decir, estamos ante la presencia de un ciclo de periodo 2:

$$\{5.32050617159410777, -5.093375863894274\}$$

Si denotamos N_f a la función de iteración del método de Newton, siendo $f(x)$ la ecuación de Kepler para los valores $e = 0.9$ y $M = 1$, tenemos que

$$N'_f(5.32050617159410777)N'_f(-5.093375863894274) = 207.053$$

por lo que resulta que el 2-ciclo es repulsor. Notemos que esta situación es computacionalmente inestable, y, a la menor desviación, las iteraciones dejan el ciclo y convergen a la solución correcta, $E = 1.86209$, como se muestra en la Tabla 3.13.

Cuadro 3.13: Valores obtenidos a partir de (3.15), en donde $M = 1$, $e = 0.9$ y $E_0 = 5.32$.

E_n	E_{n+1}
0	5.32051
1	-5.09344
2	5.32130
3	-5.10594
4	5.48711
5	-8.36212
6	-2.39797
7	-0.720473
8	2.76072
9	1.98373
10	1.86664
11	1.86209
12	1.86209

Por otro lado, podemos decir que la aparición de 2 ciclos es debido a la multitud de puntos de inflexión en la curva $y = f(x)$. Esto hace que la gráfica de $N_f(N_f(x))$ corte a la diagonal en un gran número de ocasiones, como se aprecia en la izquierda de la Figura 3.7. De hecho, el 2-ciclo encontrado por Charles y Tatum no es el único. Por ejemplo, hemos encontrado que también $\{0.268243, 7.60861\}$ y $\{0.539076, 4.59352\}$ son 2-ciclos, en ambos casos repulsivos.

Es más, también se puede comprobar gráficamente la existencia de muchos 3-ciclos (véase la parte derecha de la Figura 3.7). Como muestra, tenemos los tres ciclos

$$\{0.046938, 0.268243, 7.60861\}, \quad \{0.168312, 8.88436, 4.69565\}$$

ambos repulsivos. La existencia de 3-ciclos tiene una consecuencia interesante. Como la función de iteración $N_f(x)$ es continua (esto ocurre siempre que $0 < e < 1$), el teorema de Li y Yorke [127] garantiza la existencia de ciclos de todos los periodos.

En principio, todos los ciclos que hemos encontrado en este ejemplo son repulsivos. En la siguiente sección se analizamos la posible existencia de ciclos atractivos.

3.5. Ciclos atractivos para el método de Newton aplicado a la ecuación de Kepler

Sin ninguna duda, la ecuación de Kepler

$$E = M + e \operatorname{sen} E \tag{3.19}$$

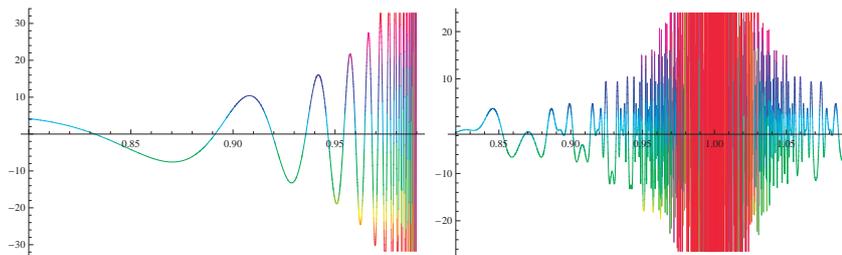


Figura 3.7: Gráficas de $N_f(N_f(x))$ y $N_f(N_f(N_f(x)))$ para $f(x)$ definida en (3.15), con $M = 1$ y $e = 0.9$.

es una de las ecuaciones más estudiadas en campos que van desde la mecánica celeste, el análisis numérico, pasando por sistemas dinámicos. A pesar de que durante de más de tres siglos ha sido profusamente considerada por muchos autores, véase a [49], la ecuación de Kepler continua siendo un tema de gran interés científico.

En el contexto orbital, la ecuación de Kepler (3.19), sólo tiene sentido para excentricidades $e \in [0, 1]$ y los valores de E y M en el intervalo $[0, 2\pi]$, o incluso, teniendo en cuenta la simetría, en el intervalo $[0, \pi]$. Sin embargo, también es común, véase a [35], [126], considerar la ecuación de Kepler como una ecuación puramente algebraica, sin ninguna interpretación orbital. En este acápite se considera la ecuación de Kepler como una función definida en la recta real. De hecho, consideramos la función

$$f(x) = x - a\pi - b \operatorname{sen} x, \quad x \in \mathbb{R}, \quad (3.20)$$

donde a y b están dados como números reales.

En muchos casos, véase [126], la extensión de la ecuación de Kepler (3.20) ha sido considerada como función test para comprobar, en la resolución de la misma, el comportamiento de diferentes métodos iterativos. Por ejemplo, el algoritmo de punto fijo dado por

$$x_{n+1} = F(x_n) = a\pi + b \operatorname{sen} x_n, \quad n \geq 0, \quad (3.21)$$

o el método de Newton definido por

$$x_{n+1} = N_f(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0, \quad (3.22)$$

son generalmente considerados para este propósito.

En esta sección centramos la atención en el método de Newton para resolver la ecuación (3.20). Es bien conocido que el método de Newton no siempre proporciona una solución a la ecuación de Kepler. En ese sentido, [35], [48] y [175], han considerado diferentes opciones de los parámetros a y b , así como del punto inicial x_0 .

Antes de continuar, es necesario incluir algunas notaciones y conceptos básicos sobre los sistemas dinámicos.

Definición 3.1 (Punto fijo). Sea $f(x)$ una función diferenciable definida en \mathbb{R} , decimos que ζ es un punto fijo de $f(x)$ si $f(\zeta) = \zeta$. Se dice que ζ será un punto fijo atractor, repulsor o indiferente si $|f'(\zeta)|$ es menor que, mayor que o igual a 1, respectivamente. Si $|f'(\zeta)| = 0$, entonces se dice que ζ es un punto fijo superatractor.

Definición 3.2 (Órbita de un punto por una función). La órbita de un determinado punto inicial $x_0 \in \mathbb{R}$, se define como el conjunto

$$\text{orb}(x_0) = \{x_0, f(x_0), f^2(x_0), \dots, f^n(x_0), \dots\}$$

donde f^n representa la n -ésima iteración de f .

Definición 3.3 (Ciclo de longitud n). Si x_0 es un punto fijo de f^n , pero no es un punto fijo de f^j , para $0 < j < n$, entonces el conjunto

$$\text{orb}(x_0) = \{x_0, f(x_0), f^2(x_0), \dots, f^n(x_0), \dots\}$$

se dice que es un ciclo de longitud n .

Nótese que un n -ciclo también se puede clasificar como *atractor*, *repulsor* o *indiferente*, si $|(f^n)'(x_0)|$ es menor que, mayor que o igual a 1, respectivamente. De la misma forma, un n -ciclo es superatractor si $|(f^n)'(x_0)| = 0$. Si escribimos $x_j = f^j(x_0)$ para $0 < j < n$, entonces, aplicando la Regla de la Cadena obtenemos que

$$(F^n)'(x_0) = \prod_{j=0}^{n-1} F'(F^j(x_0)) = \prod_{j=0}^{n-1} F'(x_j),$$

luego la derivada $(f^n)'$ tendrá el mismo valor en cualquier punto del n -ciclo.

La existencia de ciclos para el método de Newton aplicada a los polinomios ha sido estudiado por diferentes autores, véase a [22] o [170]. Estos autores probaron la existencia de 2-ciclos superatractores para el método de Newton aplicado a los polinomios $p_1(x) = x^3 - 2x + 2$ y $p_2(x) = 3x^5 - 10x^3 + 23x$. En concreto, los 2-ciclos superatractores son de la forma $\{0, 1\}$ y $\{1, 1\}$ respectivamente.

Como se señaló en [35], una de las razones del fracaso del método de Newton para resolver la ecuación de Kepler (3.20), es la existencia de ciclos atractores que obligan a las iteraciones a oscilar siempre entre dos valores, por lo que dicho método no será capaz de llegar a la solución. Para nuestra conveniencia en los cálculos siguientes, se introduce la notación

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}. \quad (3.23)$$

Un cálculo sencillo nos muestra que $N'_f(x) = L_f(x)$. Entonces, cuando $f(x)$ está definida por (3.20), los puntos críticos libres del método de Newton son precisamente las raíces de $f''(x) = b \operatorname{sen} x$, esto es, los puntos en la forma $x = k\pi$ con $k \in \mathbb{Z}$.

Caracterización de ciclos superatractores de periodo 2

En este apartado consideraremos la ecuación de Kepler como una función $f(x) = x - a\pi - b\sin x$ definida en la recta real, véase (3.20). Nuestro interés consiste en caracterizar los parámetros a y b para mostrar la existencia de ciclos superatractores de periodo 2 cuando el método de Newton (3.22) se aplica a (3.20). Los puntos críticos libres de $N_f(x)$, esto es, los puntos de inflexión de $f(x)$, son candidatos perfectos para formar parte de ciclos superatractores. Los siguientes resultados caracterizan la existencia de estos ciclos.

Teorema 3.2. *Si $b = 1/2$ y $a = j + k$, con $j, k \in \mathbb{Z}$, $j \neq k$, entonces el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 de la forma $\{2j\pi, 2k\pi\}$.*

Demostración. Notemos que las condiciones $N_f(2j\pi) = 2k\pi$ y $N_f(2k\pi) = 2j\pi$ llevan al siguiente sistema de ecuaciones lineales

$$\begin{cases} a + 2(k - j)b = 2k \\ a + 2(j - k)b = 2j. \end{cases}$$

Resolviendo este sistema para a y b obtenemos que $b = 1/2$ y $a = j + k$. En consecuencia, $\{2j\pi, 2k\pi\}$ es un 2-ciclo. Finalmente, tomando en cuenta que

$$(N_f^2)'(2j\pi) = N_f'(2j\pi)N_f'(2k\pi) = L_f(2j\pi)L_f(2k\pi) = 0,$$

deducimos que el 2-ciclo mencionado más arriba es superatractor. \square

Teorema 3.3. *Si $b = -1/2$ y $a = j + k + 1$, con $j, k \in \mathbb{Z}$, $j \neq k$, entonces el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 en la forma $\{(2j + 1)\pi, (2k + 1)\pi\}$.*

Demostración. La demostración sigue los pasos de la demostración del teorema anterior. Ahora las condiciones $N_f((2j + 1)\pi) = (2k + 1)\pi$ y $N_f((2k + 1)\pi) = (2j + 1)\pi$ llevan al siguiente sistema de ecuaciones lineales

$$\begin{cases} a + 2(j - k)b = 2k + 1 \\ a + 2(k - j)b = 2j + 1, \end{cases}$$

que tiene como única solución a $b = -1/2$ y $a = j + k + 1$. Como

$$(N_f^2)'((2j + 1)\pi) = N_f'((2j + 1)\pi)N_f'((2k + 1)\pi) = L_f((2j + 1)\pi)L_f((2k + 1)\pi) = 0,$$

concluimos que $\{(2j + 1)\pi, (2k + 1)\pi\}$ es un ciclo superatractor de periodo 2. \square

Por supuesto, los valores de a y b encontrados en los dos teoremas anteriores no son los únicos para los que la ecuación de Kepler presenta ciclos atractores de periodo 2. Es posible encontrar muchos otros. Como ejemplos, en los siguientes resultados se muestran algunos de ellos. Notemos que los teoremas 3.4–3.7 son una generalización directa de la dada en los teoremas 3.2–3.3.

Teorema 3.4. *Para $j, k \in \mathbb{Z}$, vamos a introducir las siguientes notaciones:*

$$\alpha = \frac{2}{(4k+1-4j)\pi}, \quad b = \frac{1}{1-\alpha}, \quad a = 2j - \frac{b}{\pi}.$$

Entonces, el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 en la forma $\{2j\pi, (4k+1)\pi/2\}$.

Demostración. En realidad las condiciones $N_f((2j)\pi) = (4k+1)\pi/2$ y $N_f((4k+1)\pi/2) = 2j\pi$ llevan al siguiente sistema de ecuaciones lineales

$$\begin{cases} 2a + (4k+1-4j)b &= 4k+1 \\ \pi a + b &= 2j\pi. \end{cases}$$

Resolviendo este sistema de ecuaciones en base a a y b , obtenemos los valores dados en el teorema. Como

$$(N_f^2)'(2j\pi) = N_f'(2j\pi)N_f'((4k+1)\pi/2) = L_f(2j\pi)L_f((4k+1)\pi/2) = 0,$$

concluimos que $\{2j\pi, (4k+1)\pi/2\}$ es un ciclo superatractor de periodo 2. \square

Teorema 3.5. *Para $j, k \in \mathbb{Z}$, introducimos las siguientes notaciones:*

$$\alpha = \frac{2}{(4k+3-4j)\pi}, \quad b = \frac{1}{1+\alpha}, \quad a = 2j + \frac{b}{\pi}.$$

Entonces, el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 en la forma $\{2j\pi, (4k+3)\pi/2\}$.

Demostración. Las condiciones $N_f(2j\pi) = (4k+3)\pi/2$ y $N_f((4k+3)\pi/2) = 2j\pi$ llevan al siguiente sistema de ecuaciones lineales

$$\begin{cases} 2a + (4k+3-4j)b &= 4k+3 \\ \pi a + b &= 2j\pi. \end{cases}$$

Resolviendo este sistema para a y b obtenemos que $b = \frac{1}{1+\alpha}$ y $a = 2j + \frac{b}{\pi}$. En consecuencia, $\{2j\pi, (4k+3)\pi/2\}$ es un 2-ciclo. Finalmente, tomando en cuenta que

$$(N_f^2)'(2j\pi) = N_f'(2j\pi)N_f'((4k+3)\pi/2) = L_f(2j\pi)L_f((4k+3)\pi/2) = 0,$$

deducimos que el 2-ciclo mencionado más arriba es superatractor. \square

Teorema 3.6. Para $j, k \in \mathbb{Z}$, definamos:

$$\alpha = \frac{2}{(4j+1-4k)\pi}, \quad b = \frac{1}{\alpha-1}, \quad a = 2j+1 - \frac{b}{\pi}.$$

Entonces, el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 de la forma $\{(2j+1)\pi, (4k+1)\pi/2\}$.

Demostración. En concreto, las condiciones $N_f((2j+1)\pi) = (4k+1)\pi/2$ y $N_f((4k+1)\pi/2) = (2j+1)\pi$ nos conducen al siguiente sistema de ecuaciones lineales

$$\begin{cases} 2a - (4j+1-4k)b = 4k+1 \\ \pi a - b = (2j+1)\pi. \end{cases}$$

Resolviendo este sistema de ecuaciones en a y b , obtenemos los valores dados en el teorema. Como

$$(N_f^2)'((2j+1)\pi) = N_f'((2j+1)\pi)N_f'((4k+1)\pi/2) = L_f((2j+1)\pi)L_f((4k+1)\pi/2) = 0,$$

concluimos que $\{(2j+1)\pi, (4k+1)\pi/2\}$ es un ciclo superatractor de periodo 2. \square

Teorema 3.7. Para $j, k \in \mathbb{Z}$, denotamos:

$$\alpha = \frac{2}{(4k+1-4j)\pi}, \quad b = \frac{1}{\alpha-1}, \quad a = 2j+1 + \frac{b}{\pi}.$$

Entonces, el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) tiene un ciclo superatractor de periodo 2 de la forma $\{(2j+1)\pi, (4k+3)\pi/2\}$.

Demostración. Ahora las condiciones $N_f((2j+1)\pi) = (4k+3)\pi/2$ y $N_f((4k+3)\pi/2) = (2j+1)\pi$ llevan al siguiente sistema de ecuaciones lineales

$$\begin{cases} 2a - (4k+1-4j)b = 4k+3 \\ \pi a + b = (2j+1)\pi, \end{cases}$$

que tiene como única solución a $b = \frac{1}{\alpha-1}$ y $a = 2j+1 + \frac{b}{\pi}$. Como

$$(N_f^2)'((2j+1)\pi) = N_f'((2j+1)\pi)N_f'((4k+3)\pi/2) = L_f((2j+1)\pi)L_f((4k+3)\pi/2) = 0,$$

concluimos que $\{(2j+1)\pi, (4k+3)\pi/2\}$ es un ciclo superatractor de periodo 2. \square

Para ilustrar numéricamente los resultados teóricos obtenidos en esta sección se muestran en el Cuadro 3.14 diferentes parámetros a y b dados por los teoremas 3.2–3.7, así como las órbitas de los diferentes puntos de la función de iteración del método de Newton (3.22) asociada a la ecuación de Kepler (3.20). En este cuadro presentamos los ciclos superatractores que se detectaron. También hemos generado algunos gráficos, véanse las Figuras 3.11–3.10,

Cuadro 3.14: Presencia de ciclos atractores para el método de Newton (3.22) aplicado a la ecuación de Kepler (3.20) para diferentes opciones de los parámetros a y b dados por los Teoremas 3.2–3.7.

j	0	0	3	3	-2	-3
k	4	4	2	2	0	-2
a	4	5	$\frac{9+18\pi}{2+3\pi}$	$\frac{6\pi-11}{\pi-2}$	$\frac{1-21\pi}{2+7\pi}$	$\frac{5-25\pi}{-2+5\pi}$
b	1/2	-1/2	$\frac{3\pi}{2+3\pi}$	$\frac{\pi}{\pi-2}$	$-\frac{7\pi}{2+7\pi}$	$\frac{5\pi}{2-5\pi}$
x_0	0.08	4	14.14	17.28	1.571	-15.708
Iteración	x_n	x_n	x_n	x_n	x_n	x_n
1	25.0528	21.9541	18.8385	18.8549	-9.4268	-23.5619
2	0.0799	9.4333	14.1385	17.2787	1.5705	-15.7079
3	25.053	21.9906	18.8443	18.8494	-9.4222	-23.5619
4	0.0793	9.4247	14.1374	17.2787	1.5703	-15.7079
5	25.0541	21.9911	18.8483	18.8495	-9.4207	-23.5619
6	0.0773	9.4247	14.1371	17.2787	1.5698	-15.7079
7	25.0581	21.9911	18.8494	18.8495	-9.4149	-23.5619
8	0.0697	9.4247	14.1371	17.2787	1.5649	-15.7079
9	25.0719	21.9911	18.8495	18.8495	-9.3664	-23.5619
10	0.0462	9.4247	14.1371	17.2787	1.3698	-15.7079
Ciclo de longitud 2	$\{0, 8\pi\}$	$\{\pi, 9\pi\}$	$\{6\pi, 9\pi/2\}$	$\{6\pi, 11\pi/2\}$	$\{-3\pi, \pi/2\}$	$\{-5\pi, -5\pi/2\}$

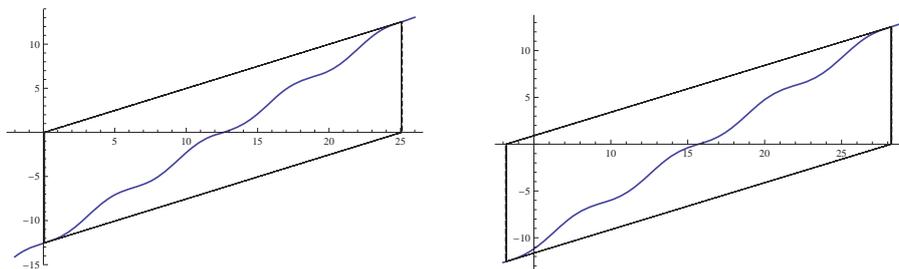


Figura 3.8: A la izquierda, un ciclo de periodo 2 de la forma $\{0, 8\pi\}$ para el método de Newton aplicado a $f(x) = x - 4\pi - 0.5 \sin x$. A la derecha, un ciclo de periodo 2 de la forma $\{\pi, 9\pi\}$ para el método de Newton aplicado a $f(x) = x - 5\pi + 0.5 \sin x$.

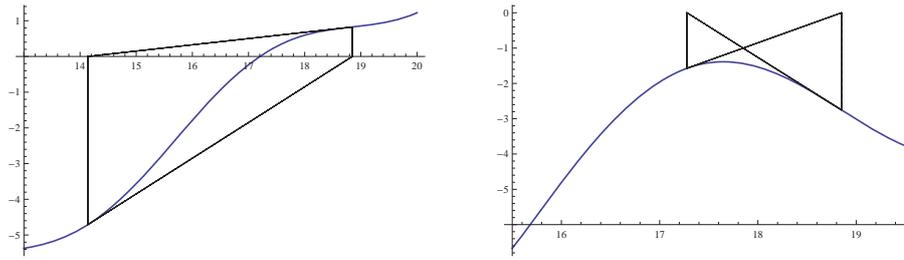


Figura 3.9: A la izquierda, un ciclo de periodo 2 en la forma $\{6\pi, 9\pi/2\}$, para el método de Newton aplicado a $f(x) = x - \left(\frac{9+18\pi}{2+3\pi}\right)\pi - \left(\frac{3\pi}{2+3\pi}\right)\text{sen } x$. A la derecha, un ciclo de periodo 2 de la forma $\{6\pi, 11\pi/2\}$, para el método de Newton aplicado a $f(x) = x - \left(\frac{6\pi-11}{\pi-2}\right)\pi - \left(\frac{\pi}{\pi-2}\right)\text{sen } x$.

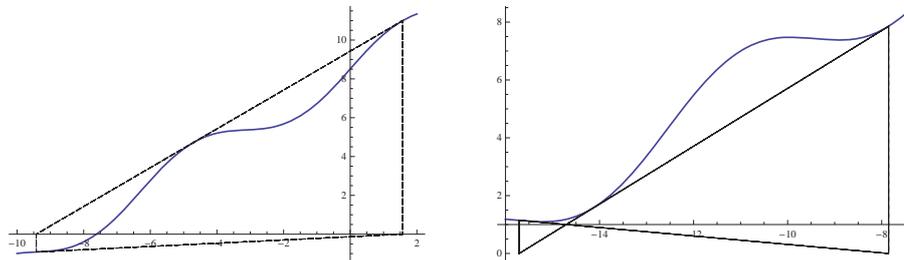


Figura 3.10: A la izquierda, un ciclo de periodo 2 de la forma $\{-3\pi, \pi/2\}$, para el método de Newton aplicado a $f(x) = x - \left(\frac{1-21\pi}{2+7\pi}\right)\pi + \left(\frac{7\pi}{2+7\pi}\right)\text{sen } x$. A la derecha, un ciclo de periodo 2 de la forma $\{-5\pi, -5\pi/2\}$, para el método de Newton aplicado a $f(x) = x - \left(\frac{5-25\pi}{5\pi-2}\right)\pi - \left(\frac{5\pi}{2-5\pi}\right)\text{sen } x$.

para mostrar el comportamiento de algunos puntos de partida y parámetros seleccionados. Los datos experimentales fueron obtenidos utilizando el sistema de álgebra computacional Mathematica *v8*.

Hemos de aclarar que, aunque en la segunda gráfica de la Figura 3.8 se ha obtenido un 2-ciclo de la forma $\{\pi, 9\pi\}$ para el método de Newton aplicado a $f(x) = x - 5\pi + 0.5 \text{sen } x$, esto no entra en contradicción con el resultado de Charles y Tatum analizado en la sección anterior. En efecto, el resultado de Charles y Tatum hace referencia a la ecuación de Kepler definida en el contexto orbital, es decir con valores de la anomalía media M entre 0 y 2π . En nuestro caso, $M = 5\pi$ está fuera de este intervalo.

Finalizamos esta sección con una aplicación interesante cuando la dinámica del método de Newton se considera con aritmética modular (mód 2π). Al trabajar con la aritmética usual en \mathbb{R} , es bien sabido que las raíces de una función $f(x)$ corresponden a los puntos fijos de la función de iteración del método de Newton $N_f(x)$ definida en (3.22). Además, las raíces simples corresponden a puntos fijos superatractores. Sin embargo, el uso de la aritmética modular (mód 2π) puede producir, para el método de Newton, la aparición de puntos fijos extraños, es decir, puntos fijos de $N_f(x)$ que no son raíces de $f(x)$. Este hecho es una simple

consecuencia de los Teoremas 3.2–3.7 vistos en la sección anterior.

Terminamos este acápite con algunos ejemplos de puntos fijos extraños para el método de Newton en la aritmética modular (mód 2π). Así definamos

$$\hat{x}_{n+1} = N_f(\hat{x}_n) \quad (\text{mód } 2\pi). \quad (3.24)$$

Entonces:

1. $\hat{x} = 0$, con $b = 1/2$ y $a = 1$, es un punto fijo extraño para el método de Newton (3.24) aplicado a la ecuación de Kepler (3.20). Notemos que existen puntos de partida próximos a $\hat{x} = 0$ cuyas órbitas convergen a $\hat{x} = 0$, pero $f(0) = \pi$. Por ejemplo, la órbita de $\hat{x}_0 = 0.3$ es

$$\{0.3, 6.023092, 0.198816, 6.164363, 0.043438, 6.277291, 0.000109, 6.283185, 3.5 \times 10^{-15}, 0, 0, \dots\}$$

2. $\hat{x} = \pi$ con $b = -1/2$ and $a = 2$, es un punto fijo extraño para el método de Newton (3.24) aplicado a la ecuación de Kepler (3.20). En este caso tenemos que la órbita de $\hat{x}_0 = 2.9$

$$\{2.9, 2.959726, 3.037684, 3.107514, 3.137933, 3.141551, 3.141593, \pi, \pi, \dots\}$$

converge a π , pero $f(\pi) = \pi$.

3.6. La derivada schwarziana y la ecuación de Kepler

En este acápite utilizaremos algunas técnicas de la teoría del caos, parecidas a la utilizadas por Devaney, [64] y compararemos los resultados con los obtenidos por Conway [48], Charles y Tatum, [35] y Stumpf, [175]. Debemos señalar que en los trabajos de [64] se investigan las características de la función iterativa del método de Newton, asociada a la ecuación de Kepler, mientras que en los trabajos de [35], se trata de explicar la naturaleza del problema a través del análisis de la ecuación de Kepler en sí misma.

Consideremos el método de Newton, cuya función de iteración aplicada a la ecuación de Kepler, en adelante definida por $N(E)$, está dada por la expresión

$$N(E) = \frac{M - eE \cos(E) + e \operatorname{sen}(E)}{1 - e \cos(E)}, \quad (3.25)$$

luego, encontrar las raíces de la ecuación de Kepler se reduce a encontrar los puntos fijos de (3.25).

En [156] y [106], se examina en profundidad el comportamiento caótico de la función de iteración del método de Newton aplicado a funciones diferentes a la de ecuación de Kepler. Una gran parte de las definiciones de la Teoría del Caos que se utilizan en este acápite se

abordan con más detalles en los citados documentos de referencia, algunas de las cuales se introducen a continuación.

Stumpf, [175] verificó que para valores pequeños de la excentricidad, el punto $E_0 = M$ es una buena condición inicial. Eso explica el porqué, en la práctica, dicha condición se utilice muy a menudo. Además comprobó que, cuando la excentricidad aumenta, la condición $E_0 = M$ se convierte en una mala condición inicial, pero que aún así, la ecuación de Kepler o la función de iteración, asumiendo que M pertenece a la cuenca de atracción de la función que se está aplicando, puede converger. También comprobó que algunas veces el método de Newton no converge o converge muy lentamente, destacando que la región problemática está conformada por grandes e y pequeñas M . Confirmó que las áreas problemáticas están separadas de las áreas no problemáticas por muy pequeños cambios en los valores de los parámetros. Conway, [48], así como Charles y Tatum, [35], ilustran este problema presentando cuadros típicos de las series de iteraciones necesarias para la convergencia en la región $0.980 < e < 0.999$ y $0.01 < M/\pi < 0.2$. Esta sensibilidad a las condiciones iniciales, no sólo se limita a cambios en la conjetura inicial. Los cambios en los otros parámetros, excentricidad o anomalía media, provocan diferencias igualmente dramáticas en las órbitas.

Al observar los puntos de inflexión, convexidad y concavidad de la ecuación de Kepler y el análisis de los resultados de las iteraciones utilizando el método de Newton, Charles y Tatum, [35], llegaron a la conclusión de que para $E_0 = \pi$, siempre habrá convergencia. ¿De alguna forma, se podría probar analíticamente esto aplicando la teoría del caos? La teoría del caos puede proporcionar algunas respuestas a través de una herramienta muy útil como es la *derivada schwarziana*.

En [64] encontramos que la derivada schwarziana de una función $f(x)$, viene dada por

$$S_f(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2} \left(\frac{f''(x)}{f'(x)} \right)^2. \quad (3.26)$$

En el texto anterior aparece un teorema que demuestra la relación que existe entre (3.26) y la cuenca de atracción de la función f . Recordemos que la cuenca de atracción de un punto fijo atractor de f está formada por los puntos x_0 cuyas órbitas por f , $f^n(x_0)$, son convergentes al punto fijo atractor. El mayor intervalo dentro de una cuenca de atracción que contiene al propio punto fijo atractor, se le llama *cuenca de atracción inmediata*.

Teorema 3.8. *Supongamos que $S_f(x) < 0$. Si x_0 es un punto periódico atractor de f , entonces, o la cuenca de atracción inmediata de x_0 se extiende desde $-\infty$ a $+\infty$, o existe un punto crítico de f , cuya órbita es atraída por la órbita de x_0 .*

Esto significa que si la derivada schwarziana de (3.25) fuese siempre negativa, entonces, la elección de un punto inicial igual al punto crítico de (3.25), podría asegurar la convergencia. Lo mismo ocurriría para el resto de métodos iterativos considerados en esta sección. En la Figuras 3.11, 3.12, 3.13 y 3.14, se representan las gráficas de la derivada schwarziana de

los métodos de Newton, Halley, Chebyshev, super Halley, Danby y los métodos obtenidos al sustituir $\alpha = 3/2$ y $\alpha = 2$ en (3.3). Los colores de los gráficos representan diferentes valores de e . Así, para $e = 0.1$ se asignó el color rojo, para $e = 0.2$ el color naranja, para $e = 0.3$ el color azul y para $e = 0.4$ el color verde.

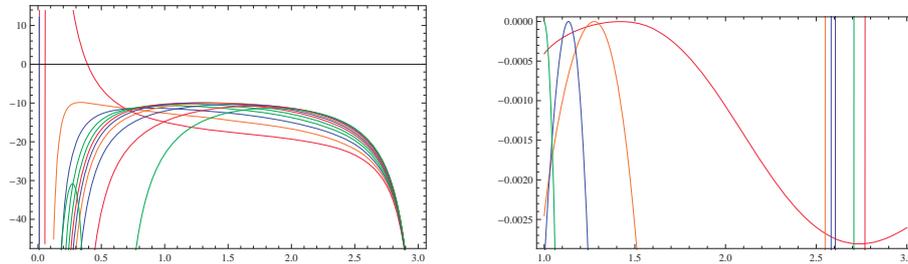


Figura 3.11: Curvas correspondientes a la derivada schwarziana de los métodos de Newton (izquierda) y Halley (derecha), aplicados a la ecuación de Kepler. Los colores de las curvas representan distintos valores de e con $M = 0.2$

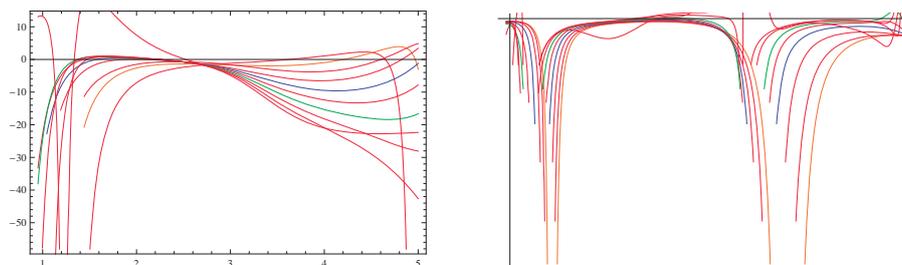


Figura 3.12: Curvas correspondientes a la derivada schwarziana de los métodos de Chebyshev (izquierda) y super-Halley (derecha), aplicados a la ecuación de Kepler. Los colores de las curvas representan distintos valores de e con $M = 0.2$

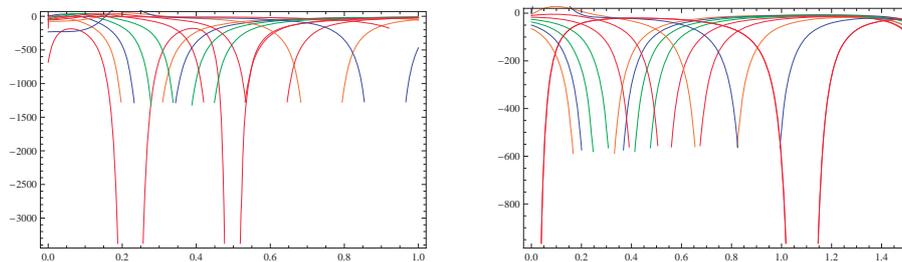


Figura 3.13: A la izquierda, derivada schwarziana del método de Danby. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$. Los colores de las curvas representan distintos valores de e con $M = 0.2$

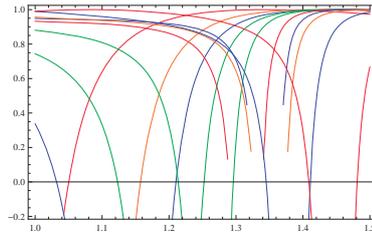


Figura 3.14: Curvas correspondientes a la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. Los colores de las curvas representan distintos valores de e con $M = 0.2$

Esta primera inspección gráfica no permite garantizar que la derivada schwarziana de los distintos métodos considerados sea siempre negativa. En lo que sigue profundizaremos un poco más en el análisis de la derivada schwarziana.

Notemos que los puntos críticos de la ecuación (3.25) aparecen cuando $N = 0$ o $N'' = 0$, lo cual significa que $E = n\pi$, es un punto crítico de (3.25) para cualquier entero n . Esto es consistente con los resultados de Charles y Tatum, [35], relacionados con los valores de la derivada schwarziana. A continuación presentamos las derivadas schwarzianas de los métodos de Newton, Halley, Chebyshev y super-Halley aplicadas a una función genérica $f(x)$.

$$\begin{aligned}
S_{N(x)} &= \frac{3f''(x)}{f(x)} + \frac{f^{iv}(x)}{f''(x)} - \frac{3f'(x)^2}{2f(x)^2} - \frac{3f'''(x)^2}{2f''(x)^2} - \frac{f'(x)f'''(x)}{f(x)f''(x)}. \\
S_{H(x)} &= 2[-8f'(x)^4 f'''(x)^2 + 4f'(x)^3 f'''(x)(6f''(x)^2 - f(x)f^{iv}(x)) + \\
&+ 6f(x)f''(x)^2(6f''(x)^3 + f(x)f'''(x)^2 - f(x)f''(x)f^{iv}(x)) - \\
&- f(x)f'(x)(4f'''(x)(15f''(x)^3 + 3f(x)f'''(x)^2 - 4f(x)f''(x)f^{iv}(x)) + 3f(x)f''(x)^2 f^v(x)) + \\
&+ f'(x)^2(-18f''(x)^4 + 24f(x)f''(x)f'''(x)^2 + 6f(x)f''(x)^2 f^{iv}(x) + \\
&+ f(x)^2(-3f^{iv}(x)^2 + 2f'''(x)f^v(x))]/f(x)^2(3f''(x)^2 - 2f'(x)f'''(x))^2]. \\
S_{C(x)} &= 1/((2f(x)^2 f'(x)^2)(-3f''(x)^2 + f'(x)f'''(x))^2(-3(12f(x)f''(x)^3 + 2f'(x)^3 f'''(x) - \\
&- 9f(x)f'(x)f''(x)f'''(x) + f'(x)^2(-6f''(x)^2 + f(x)f^{iv}(x)))^2 + \\
&+ 2(-3f''(x)^2 + f'(x)f'''(x))(60f(x)^2 f''(x)^4 - 2f'(x)^5 f'''(x) - 72f(x)^2 f'(x)f''(x)^2 f'''(x) + \\
&+ f'(x)^4(6f''(x)^2 - 4f(x)f^{iv}(x)) + 3f(x)f'(x)^2(-14f''(x)^3 + 3f(x)f'''(x)^2 + \\
&+ 4f(x)f''(x)f^{iv}(x) + f(x)f'(x)^3(34f''(x)f'''(x) - f(x)f^v(x))))). \\
S_{SH(x)} &= 1/(2f(x)^2(f(x)f''(x)^3 - f'(x)^3 f'''(x))^2)(-8f'(x)^8 f'''(x)^2 + 4f(x)f'(x)^5 f'''(x)(5f''(x)^3 + \\
&+ 2f'(x)f''(x)f'''(x) - f'(x)^2 f^{iv}(x)) + f(x)^4 f''(x)^4(-3f'''(x)^2 + 2f''(x)f'''(x)) + \\
&+ 2f(x)^3 f''(x)(15f''(x)^6 - 25f'(x)f''(x)^4 f'''(x) - 12f'(x)^3 f'''(x)^3 - 10f'(x)^2 f''(x)^3 + \\
&+ 8f'(x)^3 f''(x)f'''(x)f^{iv}(x) + f'(x)^2 f''(x)^2(30f'''(x)^2 - f'(x)f^v(x))) + \\
&+ f(x)^2 f'(x)^2(-15f''(x)^6 - 20f'(x)f''(x)^4 f'''(x) + 5f'(x)^2 f''(x)^2 f'''(x)^2 + \\
&+ 10f'(x)^2 f''(x)^3 f'''(x) + 2f'(x)^3 f'''(x)f^{iv}(x) + f(x)^2 f'(x)^2(-15f''(x)^6 - \\
&- 20f'(x)f''(x)^4 f'''(x) + 5f'(x)^2 f''(x)^2 f'''(x)^2 + 10f(x)^2 f''(x)^3 f'''(x) + \\
&+ 2f'(x)^3 f''(x)f'''(x)f^{iv}(x) + f'(x)^3(-6f'''(x)^3 - 3f'(x)f^{iv}(x)^2 + 2f'(x)f'''(x)f^v(x)))).
\end{aligned}$$

Asimismo, notemos que la derivada schwarziana de los métodos de Newton, Halley, Che-

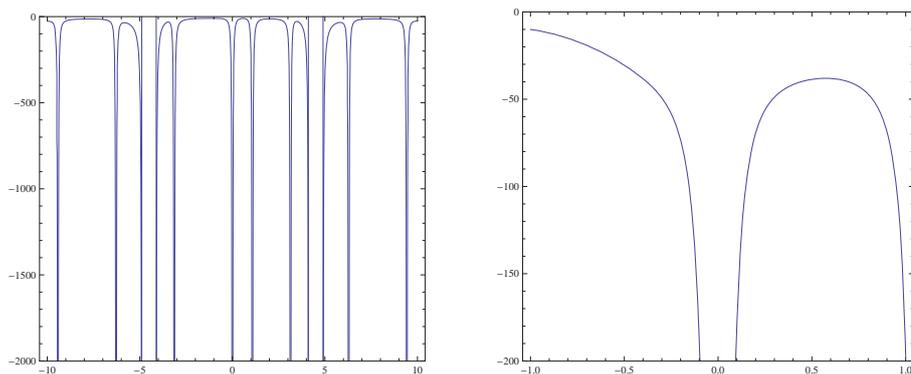


Figura 3.15: A la izquierda, la derivada schwarziana del método de Newton. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

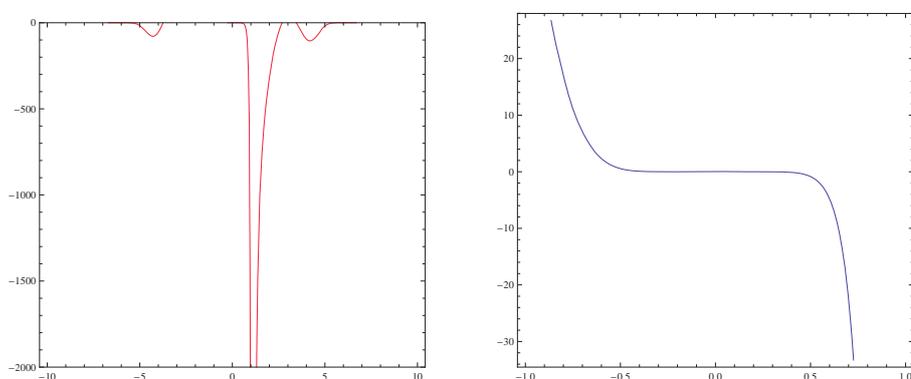


Figura 3.16: A la izquierda, la derivada schwarziana del método de Halley. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

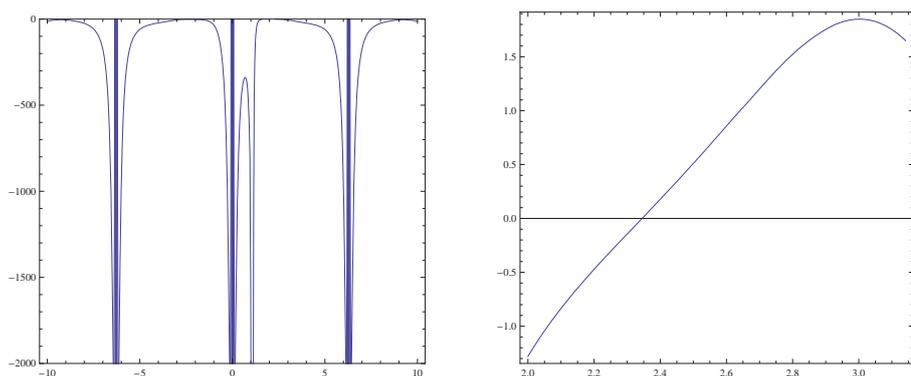


Figura 3.17: A la izquierda, la derivada schwarziana del método de Chebyshev. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

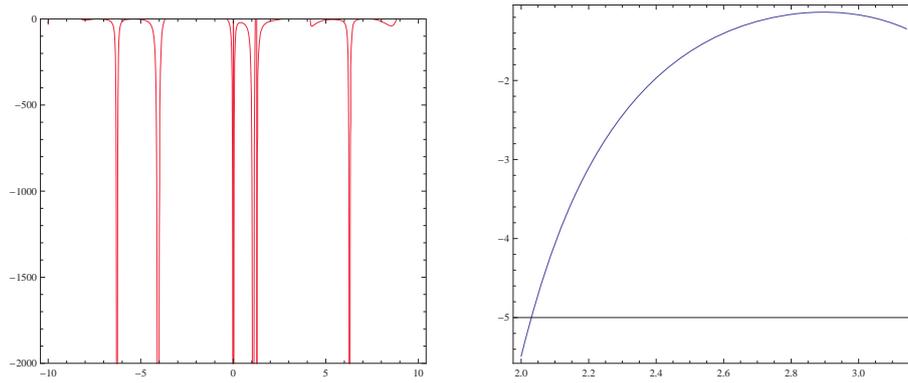


Figura 3.18: A la izquierda, la derivada schwarziana del método de super-Halley. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

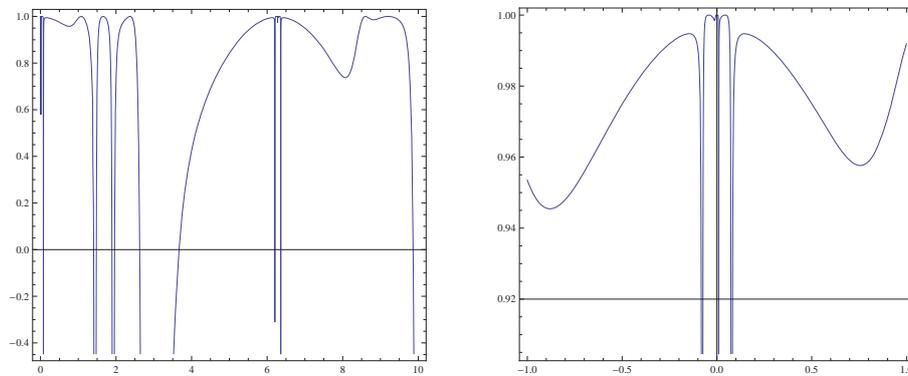


Figura 3.19: A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

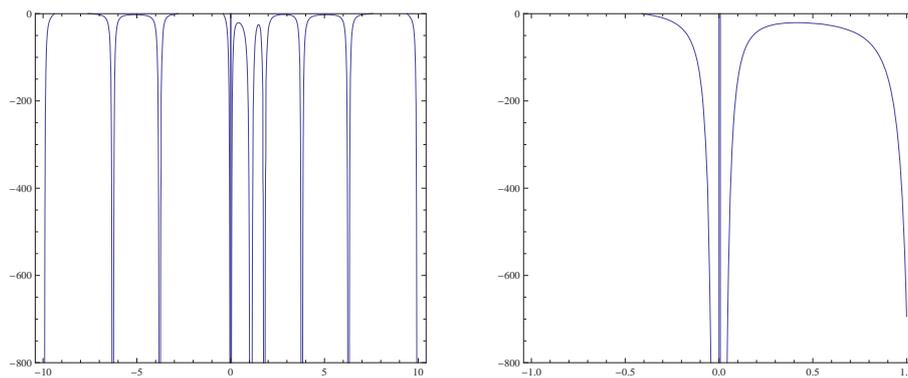


Figura 3.20: A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

byshev y super-Halley, aplicada a la ecuación de Kepler es

$$\begin{aligned}
S_N &= \frac{3b}{k} + \frac{d}{b} - \frac{3a^2}{2k^2} - \frac{3c^2}{2b^2} - \frac{ac}{kb}. \\
S_H &= 2[-8a^4c^2 + 4a^3c(6b^2 - kd) + 6kb^2(6b^3 + kc^2 - kbd) - \\
&\quad - ka(4c(15b^3 + 3kc^2 - 4kbd) + 3kb^2g) + \\
&\quad + a^2(-18b^4 + 24kbc^2 + 6kb^2d + k^2(-3d^2 + 2cg))]/k^2(3b^2 - 2ac)^2]. \\
S_C &= 1/((2k^2a^2)(-3b^2 + ac)^2)(-3(12kb^3 + 2a^3c - 9kabc + a^2(-6b^2 + kd))^2 + \\
&\quad + 2(-3b^2 + ac)(60k^2b^4 - 2a^5c - 72k^2ab^2c + a^4(6b^2 - 4kd) + \\
&\quad + 3ka^2(-14b^3 + 3kc^2 + 4kbd) + ka^3(34bc - kg))). \\
S_{SH} &= 1/(2k^2(kb^3 - a^3c)^2)(-8a^8c^2 + 4ka^5c(5b^3 + 2abc - a^2d) + k^4b^4(-3c^2 + 2bc) + \\
&\quad + 2k^3b(15b^6 - 25ab^4c - 12a^3c^3 - 10a^2b^3 + 8a^3bcd + a^2b^2(30c^2 - ag)) + \\
&\quad + k^2a^2(-15b^6 - 20ab^4c + 5a^2b^2c^2 + 10a^2b^3c + 2a^3cd + \\
&\quad + k^2a^2(-15b^6 - 20ab^4c + 5a^2b^2c^2 + 10a^2b^3c + 2a^3bcd + a^3(-6c^3 - 3ad^2 + 2acg))),
\end{aligned}$$

donde $k = E - M - e \operatorname{sen}(E)$, $a = k' = 1 - e \operatorname{cos}(E)$, $b = k'' = e \operatorname{sen}(E)$, $c = k''' = e \operatorname{cos}(E)$, $d = k^{iv} = -e \operatorname{sen}(E)$ y $g = k^v = -e \operatorname{cos}(E)$.

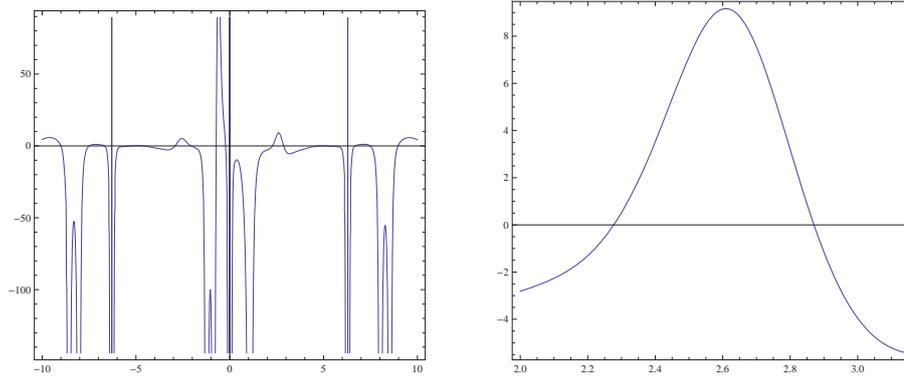


Figura 3.21: A la izquierda, la derivada schwarziana del método de Danby. A la derecha, ampliación de dicha derivada. En ambos casos $M = 0.2$ y $e = 0.997$.

El hecho de que los tres últimos términos de la derivada schwarziana del método de Newton aplicada a la ecuación de Kepler son siempre negativos, puede llevarnos a tener falsas esperanzas con respecto a que la condición $S_N(E) < 0$ fuese verdad. Entonces, a pesar de que la convergencia no está asegurada para algunas condiciones iniciales, tal como lo demuestran Conway, [48], Charles y Tatum, [35] y otros, esto estaría siempre asegurado si se tomase como punto inicial un punto crítico que sea igual a $n\pi$.

En síntesis, la derivada schwarziana proporciona una prueba de gran alcance para determinar si cualquier esquema particular de iteración será convergente, es decir, que la derivada

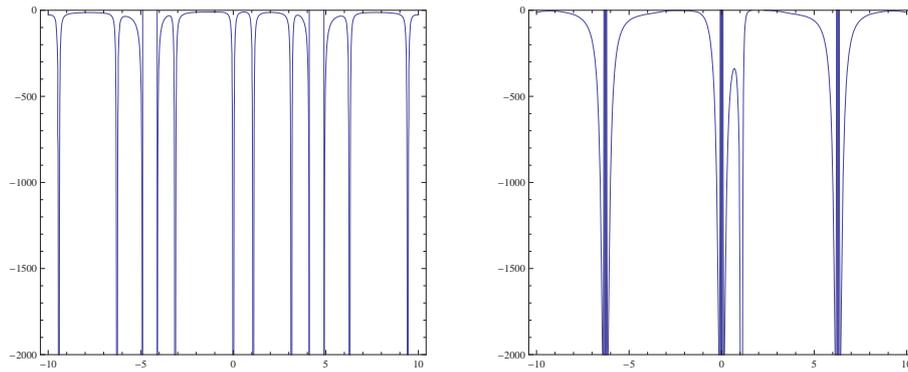


Figura 3.22: A la izquierda, la derivada schwarziana del método de Newton. A la derecha, la derivada schwarziana del método de Chebyshev. En ambos casos $M = 0.2$ y $e = 0.997$.

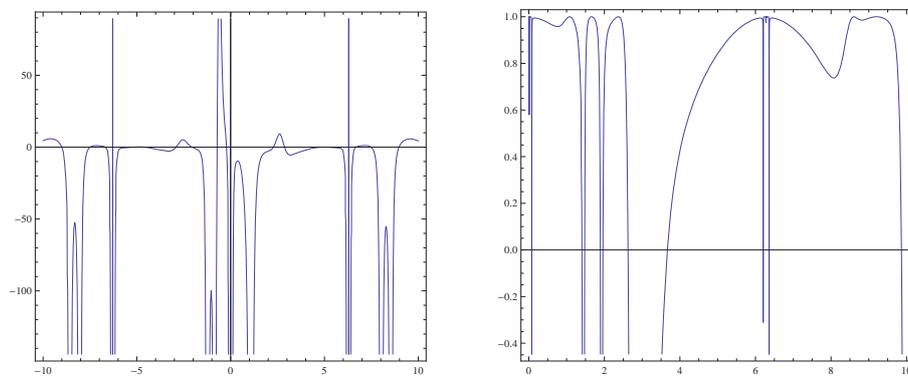


Figura 3.23: A la izquierda, derivada schwarziana del método de Danby. A la izquierda, la derivada schwarziana del método definido por (3.3) para $\alpha = 2$. En ambos casos $M = 0.2$ y $e = 0.997$

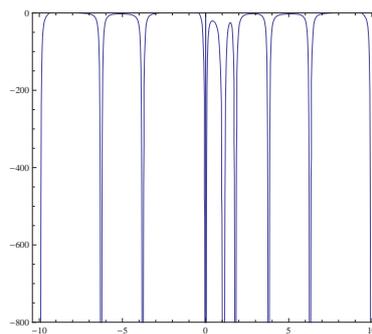


Figura 3.24: Derivada schwarziana del método definido por (3.3) para $\alpha = 3/2$, con $M = 0.2$ y $e = 0.997$

schwarziana prueba analíticamente que en una amplia gama de excentricidad y de anomalía media, la ecuación de Kepler, con una condición inicial $E_0 = \pi$, se puede resolver rápida-

mente mediante el uso del método de Newton. En el resto de métodos considerados, se ha comprobado que no se puede garantizar que la derivada schwarziana sea siempre negativa. En consecuencia, no se puede garantizar de esta manera que la elección de $E_0 = \pi$ como punto de partida garantice la convergencia a la solución del método iterativo considerado. No obstante, Charles y Tatum, [35], han demostrado que el método de Newton, cuando utiliza esta condición inicial, $E_0 = \pi$, siempre converge rápidamente, incluso en áreas donde el criterio de la derivada schwarziana falla. Como se ha puesto de manifiesto en la Sección 3.4 y, en concreto, en el Cuadro 4.6 también es cierto que la elección de $E_0 = \pi$ también produce una sucesión convergente para el resto de métodos iterativos considerados, con independencia de que no se satisfaga el criterio de la derivada schwarziana.

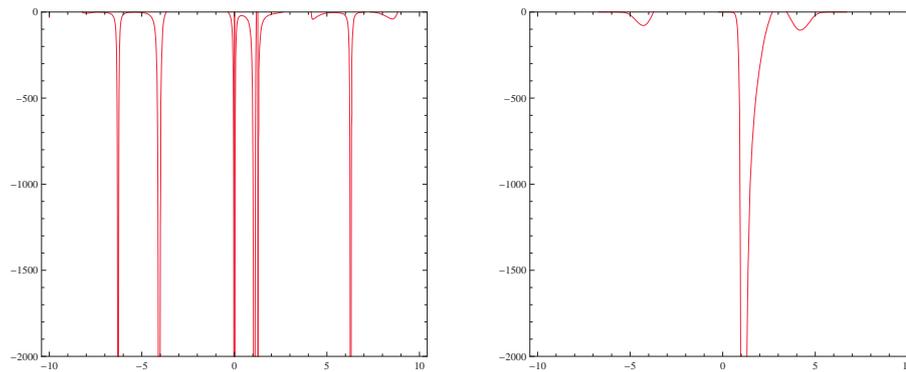


Figura 3.25: A la izquierda, la derivada schwarziana del método super-Halley. A la derecha, la derivada schwarziana del método de Halley. En ambos casos $M = 0.2$ y $e = 0.997$

3.7. Estudio dinámico de la ecuación de Kepler

Los problemas relacionados con la dinámica han centrado la atención de los científicos durante siglos. Los más conocidos hasta ahora están relacionados con la mecánica celeste, en concreto, con el estudio del movimiento de los cuerpos celestes pertenecientes al sistema solar. Son bien conocidos los intentos de Newton para comprender y modelar los movimientos de los planetas descritos en las famosas leyes de Kepler, estudios estos que sirvieron como base para el desarrollo del cálculo y el planteamiento de modelos de problemas dinámicos como ecuaciones diferenciales, teniendo especial interés los métodos de perturbaciones.

Hasta finales del siglo XIX, el análisis fue la herramienta preferida por los investigadores para estudiar los problemas dinámicos. Luego, a partir de los trabajos de Poincaré [149], quien demostró que los métodos de perturbaciones no siempre arrojaban resultados correctos, se fusionó el análisis con la geometría, surgiendo así lo que hoy en día se conoce como Sistemas Dinámicos. En [139] encontramos que Poincaré visualizaba un sistema dinámico como un campo de vectores, preocupándose principalmente por describir globalmente todas

las soluciones, mediante el llamado retrato de fase, y el efecto de pequeñas perturbaciones en las condiciones iniciales sobre la estabilidad del sistema. En la época de Poincaré, el estudio de la estabilidad de los sistemas dinámicos fue simultáneamente estudiado por Liapunov [128], quien con su tesis doctoral, sentó las bases de la teoría moderna de la estabilidad.

Diagramas de Feigenbaum

La ecuación logística se ha convertido en la ecuación por excelencia para verificar las características del caos. Pese a su aparente sencillez, constituye un buen ejemplo para mostrar el comportamiento de los sistemas caóticos. Se trata de una ecuación en diferencias que fue publicada por primera vez por Pierre Francois Verhulst en 1838, para explicar el crecimiento de una población perteneciente a la misma especie y que se reproduce en un entorno cerrado, sin ningún tipo de influencia externa.

La ecuación se puede escribir como $x_{n+1} = rx_n(1 - x_n)$, donde el parámetro r es una constante denominada parámetro de crecimiento, y la variable x_n puede verse como la fracción máxima de población que el ambiente puede soportar en un instante t_n .

Edward Albert Feigenbaum, nacido el 20 de enero de 1936 en Weehawken, Estados Unidos, es un científico de la computación, cuya área de especialidad es la inteligencia artificial. Uno de sus grandes aportes a las matemáticas lo constituyen los llamados Diagramas de Feigenbaum. En [?], encontramos que dicho diagrama aparece cuando se representan en los ejes, mediante la aplicación logística, los puntos a los que convergen la órbita del punto $1/2$, para los diferentes valores del parámetro r . La razón de utilizar el punto $1/2$ se basa en el hecho de que cuando existe un ciclo atractivo, el punto $1/2$, siempre está en la cuenca de atracción del ciclo.

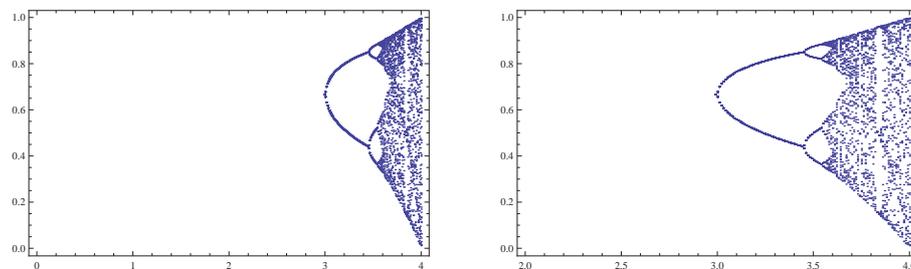


Figura 3.26: Diagrama de Feigenbaum para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para el intervalo $0.5 \leq r \leq 4$. A la derecha, para el intervalo $2 \leq r \leq 4$. Puede observarse la ruta del caos: sucesivos desdoblamientos de periodo que desembocan en un periodo infinito.

Si estudiamos experimentalmente la función logística, observaremos que para valores de $r < 3$, el sistema converge a un punto fijo estable.

Para $r < 1$ el punto fijo es el cero, para $r = 1$, el cero se convierte en un punto indiferente, pero atrae todos los puntos que están en el intervalo $(0, 1)$, mientras que para $r \geq 1$, las

órbitas de todos los puntos que están en $[0, 1]$ divergen a $-\infty$.

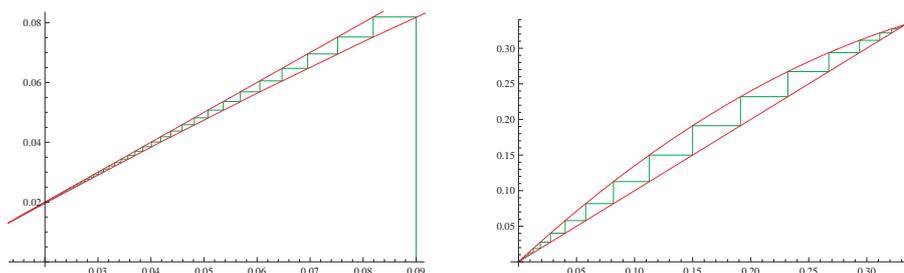


Figura 3.27: Diagrama de Telaraña para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para $r = 1$. A la derecha, para $r = 1.5$

Para $r > 1$ el punto cero se convierte en repulsivo, pero aparecen otros puntos atractivos, por ejemplo, para $r = 1.5$ las órbitas de cualquier punto de $(0, 1)$ tienden a $1/3$. Véase la Figura 3.27.

Para $r = 2$ en el intervalo $(0, 1)$, el punto 0.5 es superatractivo. Para $r = 2.5$ las órbitas de cualquier punto de $(0, 1)$ tienden al punto atractivo 0.6 . Refiérase a la Figura 3.28.

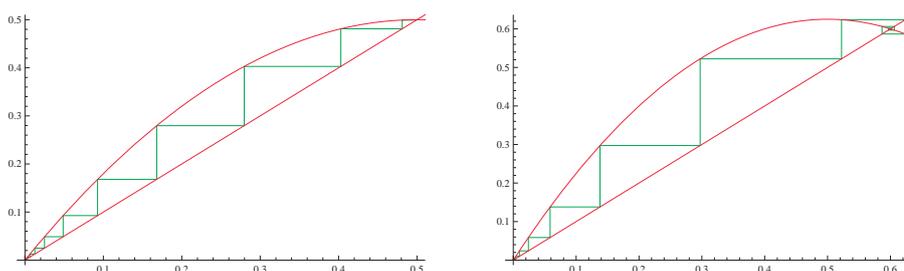


Figura 3.28: Diagrama de Telaraña para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$. A la izquierda para $r = 2$. A la derecha, para $r = 2.5$

Cuando $r = 3$, véase la Figura 3.29, el punto fijo se hace inestable, apareciendo una bifurcación, es decir, aparecen dos puntos periódicos de periodo 2.

En adelante abordaremos el estudio de la ecuación de Kepler desde un punto de vista puramente algebraico, tal y como lo hacen [35], [51], [126], [147] y [164]. Desde esta posición nos interesaremos por la ecuación algebraica y la dinámica de la función de iteración definida a partir de la misma ecuación de Kepler

$$E_{n+1} = M + e \operatorname{sen} E_n. \quad (3.27)$$

Debemos aclarar que desde el punto de vista de la mecánica celeste, el estudio de la ecuación de Kepler no tiene sentido para excentricidades que están fuera del intervalo $[0, 1)$.

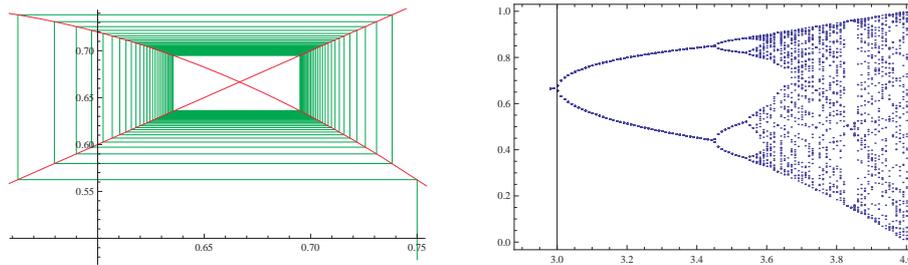


Figura 3.29: A la izquierda, diagrama de Telaraña, a la derecha, el Diagrama de Feigenbaum para la ecuación logística $x_{n+1} = rx_n(1 - x_n)$, con $r = 3$.

Así, sea la función $f(x)$ definida por

$$f(x) = M + e \operatorname{sen} x, \quad (3.28)$$

luego, la solución a la ecuación de Kepler es equivalente a encontrar los puntos fijos de (3.28). En [53] encontramos que la ecuación de Kepler tiene una solución única y real en el intervalo $[n\pi, (n+1)\pi]$, cuando $M \in [n\pi, (n+1)\pi]$ y $0 \leq e < 1$. Además, la ecuación (3.28) tiene solución si el valor absoluto de su primera derivada

$$f'(x) = e \cos x,$$

es menor que uno. Este es el caso cuando $|e| < 1$. La función (3.28) se puede considerar un desplazamiento de la función seno definida por

$$g(x) = \gamma \operatorname{sen} x, \quad (3.29)$$

la cual fue estudiada por Feigenbaum [73] en su famoso trabajo sobre el comportamiento universal de los sistemas no lineales. De acuerdo con [130], este comportamiento siempre es universal para la clase de funciones unimodales con derivada schwarziana negativa. Para el caso de (3.28), la derivada schwarziana viene dada por

$$S_K = -1 - \frac{3}{2} \tan^2 x, \quad (3.30)$$

la cual siempre es negativa, no importa los valores que tomen M o e . En lo que sigue, tal y como lo hacen en [126], analizaremos la influencia de la variación de los valores de M en (3.29), y el comportamiento de la misma cuando la excentricidad es compleja.

Bifurcaciones de la función de Kepler para excentricidades reales

En [126] encontramos que cuando $|e| \geq 1$, entonces, tomando en cuenta el valor absoluto de la derivada de (3.28), pueden aparecer uno o dos tipos de bifurcaciones:

- a) Si en el punto fijo se tiene que $f'(x) = -1$, aparecerá una bifurcación de duplicación del periodo o bifurcación de Pitchfork.
- b) Si en el punto fijo se tiene que $f'(x) = 1$, aparecerá una bifurcación tangencial.

En el primer caso, $f'(x) = -1$, el punto fijo perderá su propiedad de atractor, apareciendo en su lugar una órbita de ciclo 2. En el segundo caso, $f'(x) = 1$, el punto fijo también perderá su condición de atractor, pero aparecerán nuevos puntos fijos. En general, estos dos tipos básicos de bifurcaciones se pueden extender hasta la k -ésima iteración de la función (3.28), cuyos puntos fijos son ciclos atractores de periodo k . Es importante recordar que los ciclos de periodo impar sólo aparecen como consecuencia de una bifurcación tangencial.

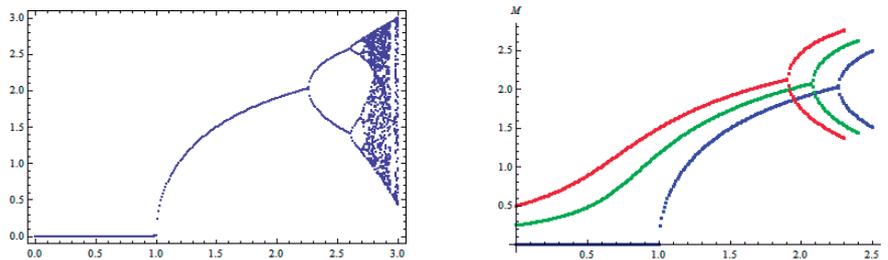


Figura 3.30: A la izquierda, cascada de bifurcación de duplicación del periodo para la función $f(x) = e \operatorname{sen}(x)$. A la derecha, desplazamiento de las primeras bifurcaciones de duplicación del periodo cuando M aumenta. Para $M = 0$, la curva es azul, para $M = 0.25$, la curva es verde y para $M = 0.5$, la curva es roja

De acuerdo con [126], si consideramos el caso en que $M = 0$, la primera bifurcación aparecerá cuando $e = 1$. Dicha bifurcación es tangencial, véase la parte izquierda de la Figura 3.30, y a medida que aumenta el valor de E , aparece una cascada de bifurcaciones de duplicación del periodo. Esto continúa hasta que e alcanza un valor crítico. Si e continúa aumentando, aparecen los ciclos de periodo impar. Notemos que el último ciclo en aparecer es el de periodo 3. Algo parecido ocurre cuando $M \neq 0$, con la novedad de que bajo esta condición, la bifurcación no es tangencial y la primera bifurcación de doble periodo aparece para valores pequeños de la excentricidad, véase la parte derecha de la Figura 3.30. En el caso extremo en que $M = \pi$, la primera bifurcación de duplicación de periodo aparece cuando $e = 1$. Para valores mayores que $M = \pi$ la primera bifurcación de duplicación de periodo ocurre de forma simétrica, como se deduce del siguiente resultado de Lanchares y Pérez-Barrón, [126], demostraron la siguiente proposición.

Proposición 3.9. *Sea $0 < \alpha < \pi$ y sea (e_α, E_α) los valores de e y E en la primera bifurcación de duplicación de periodo para $M = \pi + \alpha$, entonces $(e_\alpha, 2\pi - E_\alpha)$ son los valores de e y E en la primera bifurcación de duplicación periodo, para $M = \pi + \alpha$.*

Demostración. Se tiene que (e_α, E_α) son las soluciones del sistema no lineal

$$\pi - \alpha + e_\alpha \operatorname{sen} E_\alpha = E_\alpha. \tag{3.31}$$

$$e_\alpha \cos E_\alpha = -1. \tag{3.32}$$

Notemos también que la ecuación (3.32) no cambia si se sustituye E_α por $2\pi - E_\alpha$. Por otro lado, si añadimos la constante π a ambos lados de la ecuación (3.31), después de algunos cálculos obtenemos que

$$\pi + \alpha - e_\alpha \operatorname{sen} E_\alpha = 2\pi - E_\alpha. \tag{3.33}$$

Teniendo en cuenta que $E_\alpha = -\operatorname{sen}(2\pi - E_\alpha)$, concluimos que

$$\pi + \alpha + e_\alpha \operatorname{sen}(2\pi - E_\alpha) = 2\pi - E_\alpha. \tag{3.34}$$

Por lo tanto, $(e_\alpha, 2\pi - E_\alpha)$, para $M = \pi + \alpha$, es una solución del sistema de ecuaciones lineales presentado anteriormente. Luego, $(e_\alpha, 2\pi - E_\alpha)$, son los valores de la primera bifurcación de duplicación de periodo.

□

En [126] aparece una buena descripción gráfica, tanto de la simetría y el desplazamiento de la cascada de bifurcaciones de duplicación de periodo, como de M , las cuales se pueden obtener a partir de graficar las iteraciones de la función de Kepler (3.28), para el punto de partida inicial $E_0 = 0$.

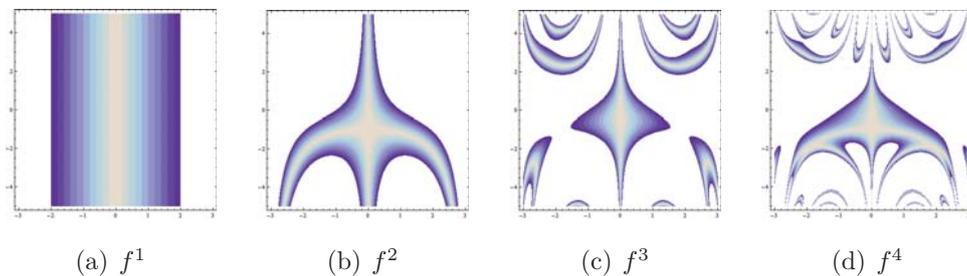


Figura 3.31: Iteración de de la función $E_{n+1} = M + e \operatorname{sen} E_n$ en el plano (M, e) para $E_0 = 0$. Se observan los contornos de las funciones $f^k(M, e; 0)$, donde $f(M, e; 0) = M + e \operatorname{sen} z$

Nótese que, al principio, todos los puntos en el plano (M, e) son del mismo color. Como la función (3.28) está anidada, el color de un punto cambia de acuerdo con el valor de $f^k(0)$. Para las primeras iteraciones, los colores se distribuyen de una manera suave, por lo que pequeños cambios en la elección del par (M, e) se convierten en pequeños cambios en el valor de $f^k(0)$. Sin embargo, si iteráramos un poco más la función (3.28), se presentan áreas de color mezcladas, mientras que las regiones de color uniforme se limitan, en su mayor parte, a la banda de $-1 < e < 1$, como se muestra en la Figura 3.31. El proceso de mezcla es una

consecuencia directa de la cascada de bifurcaciones producidas por la variación de M . Cabe destacar la simetría respecto a la línea $M = \pi$, como se señaló en la Proposición 3.9. Sin embargo, no existe simetría con respecto a la línea $e = 0$, excepto para la primera iteración, incluso en el caso en que aparece una función sinusoidal pura, ($M = 0$.)

3.7.1. La función de Kepler en el plano complejo

La función de Kepler (3.28) puede ser extendida a una función compleja si consideramos los parámetros y las variables en el plano complejo. De esta forma, obtendremos una familia de funciones paramétricas dadas por

$$f(M, e, z) = M + e \operatorname{sen} z, \quad e, M, z \in \mathbb{C}, \quad (3.35)$$

cuyos puntos fijos serían las soluciones de la ecuación de Kepler en su versión compleja. Es evidente que la sucesión generada por las iteraciones de (3.28), desde puntos iniciales reales, está acotada si los valores de M y e también son reales. Sin embargo, esto no sucede para funciones complejas, por lo que es suficiente que uno de los parámetros, M , e o z_0 sea complejo para que se pueda tener una sucesión no acotada.

Tomando esto en cuenta, Lanchares y Pérez-Barrón [126] generaron una familia de conjuntos fractales definidos de forma similar a los bien conocidos conjuntos de Mandelbrot

$$\mathcal{K}(M; z_0) = \{e \in \mathbb{C} \mid \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}. \quad (3.36)$$

También definieron una familia de conjuntos de Julia asociados a cada par (M, e)

$$\mathcal{J}(M, e) = \{z \in \mathbb{C} \mid \{f^n(M, e, z)\}_{n \in \mathbb{N}} \text{ acotada}\}. \quad (3.37)$$

Nosotros, además, hemos considerado una nueva familia de conjuntos fractales de tipo Mandelbrot a cada par (e, z_0)

$$\mathcal{I}(e, z_0) = \{M \in \mathbb{C} \mid \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}. \quad (3.38)$$

Se pueden obtener algunas propiedades de estos conjuntos a partir del hecho de que la función (3.35) es simétrica con respecto a su conjugada. En efecto, tomando en cuenta que $\operatorname{sen} \bar{z} = \overline{\operatorname{sen} z}$, obtenemos que

$$\overline{f(M, e, z)} = f(\bar{M}, \bar{e}, \bar{z}). \quad (3.39)$$

Como consecuencia de ello, el conjunto definido por (3.36) es simétrico con respecto al eje real, siempre y cuando M y z_0 son reales. De forma similar, el conjunto definido por (3.37) es también simétrico con respecto al eje real, si M y e son reales.

En el caso de la función seno puro ($M = 0$), Lanchares y Pérez-Barrón [126] obtuvieron dos simetrías adicionales

$$f^n(0, e; z) = (-1)^n f^n(0, -e; z), \quad f^n(0, e; z) = -f^n(0, e; -z), \quad (3.40)$$

por tanto, los conjuntos $\mathcal{K}(0; z_0)$ y $\mathcal{J}(0; e)$ son simétricos con respecto al origen.

En conclusión, Lanchares y Pérez-Barrón establecieron que como

$$f(M, e; z) = f(M, e; z + 2k\pi), k \in \mathbb{Z},$$

el conjunto de Julia (3.37) es, en cierta medida, periódico y extendido infinitamente a lo largo del eje real. Todas estas propiedades se pueden observar en el Anexo III. En el mismo mostramos las figuras asociadas a los conjuntos introducidos en el tercer capítulo (véanse las fórmulas 3.36, 3.37 y 3.38 respectivamente), así como los códigos de Mathematica empleados para su generación. Notemos que en los mismos se pueden cambiar los diversos parámetros para obtener figuras de mayor resolución y calidad, lo cual aumentaría considerablemente su peso.

Capítulo 4

Resultados de convergencia semilocal para el método de Newton aplicados a la ecuación de Kepler

En la sección 2.2 del Capítulo 2 establecimos algunas consideraciones históricas del primer resultado de convergencia semilocal, dado por Cauchy, y generalizado más tarde por Kantorovich.

En lo que sigue observaremos que los resultados de convergencia semilocal, además de garantizar la convergencia del método de Newton-Raphson a una solución de la ecuación considerada, proporcionan unos resultados sobre la existencia y la unicidad de solución.

Este último aspecto es el que vamos a explotar en esta sección. Vamos a considerar los dos principales teoremas sobre convergencia semilocal del método de Newton-Raphson, los de Kantorovich y Smale, junto con dos variantes de los mismos y vamos a comparar los resultados sobre existencia de unicidad que proporcionan cuando se aplican a una misma ecuación: la ecuación de Kepler.

En concreto, para

$$f(E) = E - e \operatorname{sen} E - M,$$

consideraremos en todos los casos el mismo punto de partida: $E_0 = M$.

Luego, aplicando distintos resultados obtendremos distintos dominios de convergencia de la forma

$$[M - t^*, M + t^*],$$

y de unicidad de solución

$$(M - t^{**}, M + t^{**}).$$

Tanto en la teoría de Kantorovich como en la de Smale (α -teoría), la idea de la demostración consiste en construir una función mayorizante de la cual obtener información sobre la convergencia del método de Newton-Raphson.

Así, sea $\phi(t)$ una función mayorizante obtenida por alguna de las teorías consideradas. Se comprobará que $\phi(t)$ tiene dos raíces: t^* y t^{**} que son las que proporcionan los dominios de existencia y unicidad buscados.

Aunque la idea que presentamos se aplica a la ecuación de Kepler, se podría seguir el mismo esquema para otros tipos de ecuaciones: Complejas, sistemas de ecuaciones, ecuaciones funcionales, etc.

En lo que sigue:

1. Vamos a construir la función mayorizante para la ecuación de Kepler siguiendo 4 teorías:
 - i) Kantorovich [109].
 - ii) Condiciones fuertes de Kantorovich [84].
 - iii) α -teoría de Smale [170]. (artículo de Zhen-Shi-Ming).
 - iv) α -teoría de Wang-Zhao [186]
2. Vamos a ver bajo qué condiciones la correspondiente función mayorizante tiene raíces reales. Esto nos va a dar un resultado que garantiza la existencia de solución para la ecuación de Kepler en términos de la excentricidad e .
3. Vamos a determinar los correspondientes dominios de existencia y unicidad en cada caso.

4.1. Las condiciones de convergencia del teorema de Kantorovich aplicadas a la ecuación de Kepler.

Para el estudio de la convergencia semilocal del método de Newton-Raphson normalmente se utilizan los resultados del teorema de Kantorovich y la α -teoría de Smale, pero aplicados a espacios de Banach y sistemas de ecuaciones. El estudio del teorema de Newton-Raphson en espacios de Banach es relativamente joven, de la segunda mitad del siglo XX.

Leyendo a [83], encontramos que fue a finales del 1940 cuando L. V. Kantorovich y L. B. Rall introdujeron el método de Newton-Raphson en estos tipos de espacios. Su planteamiento establece que si F es un operador suficientemente diferenciable definido entre dos espacios de Banach X y Y , a partir de un punto $x_0 \in X$ se construye la sucesión de Newton-Raphson

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n). \quad (4.1)$$

Teorema 4.1. Si $e \leq 0.4142$, entonces la ecuación de Kepler tiene solución. Dicha solución está localizada en el intervalo

$$[M - t^*, M + t^*]. \quad (4.2)$$

y es única en

$$(M - t^{**}, M + t^{**}), \quad (4.3)$$

donde

$$t^* = \frac{1 - \sqrt{1 - 2\delta}}{\delta}. \quad (4.4)$$

$$t^{**} = \frac{1 + \sqrt{1 - 2\delta}}{\delta}, \quad (4.5)$$

son las raíces del polinomio mayorizante

$$p(t) = \frac{1}{2}\delta t^2 - t + \delta, \quad \text{con } \delta = \frac{e}{1 - e}. \quad (4.6)$$

Además, el método de Newton-Raphson, definido por

$$E_{n+1} = E_n - \frac{f(E_n)}{f'(E_n)}, \quad n \geq 0, \quad (4.7)$$

iniciando en $E_0 = M$, converge a dicha solución.

Demostración. Tomemos $E_0 = M$ en el teorema de Kantorovich. Luego se deben cumplir las siguientes condiciones:

- i) $f'(E_0) \neq 0$.
- ii) $\left| \frac{f(E_0)}{f'(E_0)} \right| \leq \beta$.
- iii) $\left| \frac{f(E) - f(y)}{f'(E_0)} \right| \leq \gamma(E - y), \quad \forall E, y \in I$.
- iv) $h = \gamma\beta \leq \frac{1}{2}$.

Para ello, calcularemos los parámetros β y γ :

$$f'(E) = 1 - e \cos E \Rightarrow |f'(E_0)| \geq 1 - e. \quad (4.8)$$

$$f'(y) = 1 - e \cos y. \quad (4.9)$$

$$|f(E_0)| \leq |e \operatorname{sen} E_0| \leq e. \quad (4.10)$$

$$|f'(E) - f'(y)| = |(1 - e \cos E) - (1 - e \cos y)| = e|\cos E - \cos y| \leq e|E - y|. \quad (4.11)$$

Así

$$\left| \frac{f(E_0)}{f'(E_0)} \right| \leq \frac{e}{1-e} = \beta = \delta. \quad (4.12)$$

Además

$$\left| \frac{f(E) - f(y)}{f'(E_0)} \right| \leq \frac{e}{1-e} |E - y|, \quad \text{por lo que } \gamma = \delta. \quad (4.13)$$

Luego

$$h = \gamma\beta = \delta^2 = \left(\frac{e}{1-e} \right)^2 \leq \frac{1}{2}. \quad (4.14)$$

Y por tanto

$$2e^2 \leq (1-e)^2 \Leftrightarrow 2e^2 \leq 1 - 2e + e^2 \Leftrightarrow 0 \leq 1 - 2e - e^2 \Leftrightarrow 0 \leq 2 - (1 - 2e + e^2).$$

Luego

$$0 < 2 - (1 + e^2) \Leftrightarrow (1 + e)^2 \leq 2 \Leftrightarrow -\sqrt{2} \leq 1 + e \leq \sqrt{2}. \quad (4.15)$$

Tomando la parte positiva de (4.15) sigue

$$1 + e \leq \sqrt{2} \Leftrightarrow e \leq \sqrt{2} - 1 \approx 0.4142.$$

Notemos que la condición $e \leq 0.4142$. garantiza tres cosas:

1. La ecuación de Kepler tiene solución: E^* .
2. El polinomio mayorizante $p(t) = 1/2\delta t^2 - t + \delta$ tiene soluciones reales: t^* y t^{**} .
3. E^* está localizada en el intervalo $[M - t^*, M + t^*]$ y es la única solución en $(M - t^{**}, M + t^{**})$.

Por último, el método de Newton-Raphson definido por (4.7), iniciando en $E_0 = M$, converge a E^* . \square

4.2. Las condiciones fuertes de Kantorovich aplicadas a la ecuación de Kepler.

Como vimos, el teorema de Kantorovich presenta las siguientes condiciones:

- i) $x_0 \in I$.
- ii) $f'(x_0) \neq 0$.

$$\text{iii) } \left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta.$$

$$\text{iv) } \left| \frac{f'(x) - f'(y)}{f'(x_0)} \right| \leq \gamma |x - y|, \quad \forall x, y \in I.$$

$$\text{v) } h = \gamma\beta \leq 1/2.$$

En 1997, José Manuel Gutiérrez presentó un nuevo teorema de convergencia semilocal para el método de Newton-Raphson. En dicho teorema, Gutiérrez no toma en cuenta la condición v) de Kantorovich, sino que la sustituye por dos expresiones en que se aplica la derivada segunda de la función en x

$$1. \left| \frac{f''(x_0)}{f'(x_0)} \right| \leq b.$$

$$2. \left| \frac{f''(x) - f''(x_0)}{f'(x_0)} \right| \leq K |x - x_0|.$$

Además, aplica la condición *iii)* en la forma

$$\left| \frac{f(x_0)}{f'(x_0)} \right| \leq a.$$

Y el polinomio mayorizante

$$p(t) = a - t + \frac{b}{2}t^2 + \frac{K}{6}t^3. \quad (4.16)$$

En esta memoria, llamaremos a las condiciones anteriores condiciones fuertes de Kantorovich. Los detalles de la demostración de éste teorema se encuentran en [84]. En lo adelante veremos su versión en \mathbb{R} aplicado a la ecuación de Kepler.

Teorema 4.2. *Si $e \leq 0.3759$, entonces la ecuación de Kepler tiene solución. Dicha solución está localizada en el intervalo (4.2) y es única en (4.3) donde t^* y t^{**} son las raíces del polinomio mayorizante*

$$p(t) = \delta - t + \frac{1}{2}\delta t^2 + \frac{1}{6}\delta t^3, \quad \text{con } \delta = \frac{e}{1-e}. \quad (4.17)$$

Además, el método de Newton-Raphson, definido por (4.7), iniciando en $E_0 = M$, converge a dicha solución.

Demostración. Tomemos a $E_0 = M$ y tengamos en cuenta las condiciones fuertes de Kantorovich:

$$\text{i) } \left| \frac{f(E_0)}{f'(E_0)} \right| \leq a.$$

$$\text{ii) } \left| \frac{f''(E_0)}{f'(E_0)} \right| \leq b.$$

$$\text{iii) } \left| \frac{f''(E) - f''(E_0)}{f'(E_0)} \right| \leq K|E - E_0|.$$

Ahora calcularemos los parámetros a , b y K .

$$f'(E) = 1 - e \cos E \Rightarrow |f'(E_0)| = |1 - e \cos E| \geq 1 - e. \quad (4.18)$$

$$f''(E) = e \sen E \Rightarrow |f''(E_0)| = |e \sen E_0| \leq e. \quad (4.19)$$

$$|f''(E) - f''(E_0)| = e|\sen E - \sen E_0| \leq e|E - E_0|. \quad (4.20)$$

luego

$$\left| \frac{f(E_0)}{f'(E_0)} \right| \leq \frac{e}{1 - e} = a. \quad (4.21)$$

$$\left| \frac{f''(E_0)}{f'(E_0)} \right| \leq \frac{e}{1 - e} = b. \quad (4.22)$$

$$\left| \frac{f''(E) - f''(E_0)}{f'(E_0)} \right| \leq \frac{e}{1 - e}|E - E_0| \Rightarrow K = \frac{e}{1 - e}. \quad (4.23)$$

Así

$$a = b = K = \frac{e}{1 - e} = \delta. \quad (4.24)$$

Sustituyendo (4.24) en (4.16) se obtiene

$$p(t) = \delta - t + \frac{\delta}{2}t^2 + \frac{\delta}{6}t^3. \quad (4.25)$$

Analicemos a (4.25).

$$p'(t) = -1 + \delta t + \frac{1}{2}\delta t^2 \Rightarrow p'(0) = -1 < 0.$$

$$p''(t) = \delta + \delta t > 0, \quad \forall \delta > 0.$$

Entonces existe un único mínimo \hat{t} tal que $p'(\hat{t}) = 0$. Luego

$$p'(\hat{t}) = 0 \Leftrightarrow \hat{t} = \frac{-\delta \pm \sqrt{\delta^2 + 2\delta}}{\delta} = -1 \pm \sqrt{1 + \frac{2}{\delta}}.$$

Tomando la raíz positiva sigue

$$\hat{t} = -1 + \sqrt{1 + \frac{2}{\delta}}. \quad (4.26)$$

Notemos que haciendo

$$\Delta = 1/\delta, \quad \text{sigue} \quad (4.27)$$

$$\hat{t} = -1 + \sqrt{1 + 2\Delta}. \quad (4.28)$$

y

$$p(\hat{t}) = \frac{1}{\Delta} \left(1 - \Delta\hat{t} + \frac{1}{2}\hat{t}^2 + \frac{1}{6}\hat{t}^3 \right). \quad (4.29)$$

Apliquemos la condición de convergencia a (4.29). Sea $p(\hat{t}) < 0$, luego

$$\begin{aligned} p(\hat{t}) = \frac{1}{\Delta} \left(1 - \Delta\hat{t} + \frac{1}{2}\hat{t}^2 + \frac{1}{6}\hat{t}^3 \right) &= \frac{1}{3\Delta} (4 - \sqrt{1 + 2\Delta} + \Delta(3 - 2\sqrt{1 + 2\Delta})) < 0. \\ &= \frac{1}{\Delta} \left(\frac{4}{3} + \Delta - \frac{1}{3}\sqrt{1 + 2\Delta} - \frac{2}{3}\Delta\sqrt{1 + 2\Delta} \right) < 0. \end{aligned} \quad (4.30)$$

Recordar que $\Delta > 0$, luego (4.30) será cierto si y sólo si

$$\frac{4}{3} + \Delta - \frac{1}{3}\sqrt{1 + 2\Delta} - \frac{2}{3}\Delta\sqrt{1 + 2\Delta} < 0. \quad (4.31)$$

De (4.31) sigue

$$\begin{aligned} 4 + 3\Delta &< \sqrt{1 + 2\Delta} + \Delta\sqrt{1 + 2\Delta} = \sqrt{1 + 2\Delta}(1 + 2\Delta) = (1 + 2\Delta)^{\frac{3}{2}}. \\ (4 + 3\Delta)^2 &< (1 + 2\Delta)^3. \\ 16 + 24\Delta + 9\Delta^2 &< 1 + 6\Delta + 12\Delta^2 + 8\Delta^3. \end{aligned}$$

Luego

$$8\Delta^3 + 3\Delta^2 - 18\Delta - 15 > 0 \Leftrightarrow \Delta \approx 1.66032. \quad (4.32)$$

Sustituyendo (4.32) en (4.27) sigue

$$\delta < \frac{1}{\Delta}. \quad (4.33)$$

Sustituyendo (4.33) en (4.24)

$$\frac{e}{1 - e} < \frac{1}{\Delta} \Leftrightarrow e < \frac{1}{1 + \Delta} \approx 0.3759.$$

Notemos que la condición $e \leq 0.3759$, también nos garantiza tres cosas:

1. La ecuación de Kepler tiene solución: E^* .
2. El polinomio mayorizante $p(t) = \delta - t + \frac{1}{2}\delta t^2 + \frac{1}{6}\delta t^3$ tiene soluciones reales: t^* y t^{**} .
3. E^* está localizada en el intervalo $[M - t^*, M + t^*]$ y es la única solución en $(M - t^{**}, M + t^{**})$.

Por último, el método de Newton-Raphson definido por (4.7), iniciando en $E_0 = M$, converge a E^* .

□

4.3. Las condiciones de convergencia del teorema de la α -teoría de Smale aplicadas a la ecuación de Kepler.

En la segunda parte de la sección 4 del Capítulo 2 dimos las notas históricas del teorema de Smale, ahora demostraremos un teorema en el que se aplica la α -teoría de Smale a la ecuación de Kepler.

Teorema 4.3. *Si $e \leq 0.3594$, entonces la ecuación de Kepler tiene solución. Dicha solución está localizada en el intervalo (4.2) y es única en (4.3) donde t^* y t^{**} son las raíces de la función mayorizante*

$$\phi(t) = \delta - t + \frac{\delta^2 t^2}{1 - \delta t}, \quad \text{con } \delta = \frac{e}{1 - e}. \quad (4.34)$$

Además, el método de Newton-Raphson, definido por (4.7), iniciando en $E_0 = M$, converge a dicha solución.

Demostración. Tomemos a $E_0 = M$ en el teorema de la α -teoría de Smale. Luego se deben cumplir las siguientes condiciones:

- i) $f'(E_0) \neq 0$.
- ii) $\beta(E_0, f) = \left| \frac{f(E_0)}{f'(E_0)} \right|$.
- iii) $\frac{1}{k!} \left| \frac{f^{(k)}(E_0)}{f'(E_0)} \right| \leq \gamma^{k-1}, \quad k \geq 2$.
- iv) $\alpha = \beta\gamma \leq 3 - 2\sqrt{2}$.

Calculemos los parámetros β , γ y α .

$$\begin{aligned} f'(E) = 1 - e \cos E & \Rightarrow |f'(E_0)| \geq 1 - e. \\ f''(E) = e \operatorname{sen} E & \Rightarrow |f''(E_0)| \leq e. \\ f'''(E) = e \cos E & \Rightarrow |f'''(E_0)| \leq e. \\ & \vdots \\ & |f^k(E_0)| \leq e, \quad k \geq 2. \end{aligned}$$

Así

$$\beta(E_0, f) = \left| \frac{f(E_0)}{f'(E_0)} \right| \leq \frac{e}{1-e}. \quad (4.35)$$

$$\gamma(E_0, f) = \left| \frac{1}{k!} \frac{f^{(k)}(E_0)}{f'(E_0)} \right|^{\frac{1}{k-1}} \leq \left[\frac{e}{k!(1-e)} \right]^{\frac{1}{k-1}}, \quad k \geq 2. \quad (4.36)$$

$$\alpha = \beta\gamma \leq \left[\left(\frac{e}{1-e} \right)^k \frac{1}{k!} \right]^{\frac{1}{k-1}} \leq 3 - 2\sqrt{2}. \quad (4.37)$$

Consideremos a (4.37)

Sea

$$\alpha \leq \left[\left(\frac{e}{1-e} \right)^k \frac{1}{k!} \right]^{\frac{1}{k-1}} \leq \alpha_0 \leq 3 - 2\sqrt{2}. \quad (4.38)$$

Luego, de (4.38) sigue

$$\begin{aligned} \left(\frac{e}{1-e} \right)^k \frac{1}{k!} &\leq \alpha_0^{k-1}. \\ \left(\frac{e}{1-e} \right)^k &\leq \alpha_0^{k-1} k!. \\ \left(\frac{e}{1-e} \right) &\leq (\alpha_0^{k-1} k!)^{\frac{1}{k}}. \end{aligned} \quad (4.39)$$

Ahora, sea

$$\delta_k = (\alpha_0^{k-1} k!)^{\frac{1}{k}}, \quad k \geq 2. \quad (4.40)$$

Necesitamos conocer el mínimo valor de δ_k .

De (4.40) sigue

$$\delta_{k+1} = [\alpha_0^k (k+1)!]^{\frac{1}{k+1}}.$$

y

$$\delta_k^k = (\alpha_0^{k-1} k!), \quad k \geq 2.$$

Por tanto

$$\begin{aligned} \frac{\delta_{k+1}}{\delta_k} &= \frac{[\alpha_0^k (k+1)!]^{\frac{1}{k+1}}}{[\alpha_0^{k-1} k!]^{\frac{1}{k}}} = \left(\alpha_0 \right)^{\frac{k}{k+1} - \frac{k-1}{k}} \left(k! \right)^{\frac{1}{k+1} - \frac{1}{k}} (k+1)^{\frac{1}{k+1}}. \\ &= \left[\frac{\alpha_0 (k+1)^k}{k!} \right]^{\frac{1}{k(k+1)}}, \quad k \geq 2. \end{aligned} \quad (4.41)$$

Ahora utilizaremos a (4.41) para hallar el $\inf_{k \geq 2} \delta_k$. Lo haremos por inducción.

Si $k = 2$, luego

$$\frac{\delta_3}{\delta_2} = \left[\frac{9\alpha_0}{2} \right]^{\frac{1}{6}} = \left[\frac{9(3 - 2\sqrt{2})}{2} \right]^{\frac{1}{6}} = 0.772 \dots < 1.$$

Entonces $\delta_3 < \delta_2$.

Para $k = 3$ sigue

$$\frac{\delta_4}{\delta_3} = \left[\frac{64\alpha_0}{9} \right]^{\frac{1}{12}} = \left[\frac{64(3 - 2\sqrt{2})}{9} \right]^{\frac{1}{12}} = 1.83 \dots > 1.$$

Por tanto $\delta_4 > \delta_3$

Luego, si $k \geq 3$, entonces

$$\frac{\frac{\binom{k+2}{k+1}}{\binom{k+1}{k+1}!}}{\frac{\binom{k+1}{k+1}}{k!}} = \left(\frac{k+2}{k+1} \right)^{k+1} > 1.$$

Por tanto $\delta_{k+1} > \delta_k$, $\forall k \geq 3$. Luego

$$\inf_{k \geq 2} \delta_k = \delta_3 = [6\alpha_0^2]^{\frac{1}{3}} = 0.5610688.$$

Conocido el ínfimo de δ_k , inmediatamente de (4.39) podemos conocer el mínimo valor de e . Así

$$\frac{e}{1-e} \leq \delta_3 \Leftrightarrow e \leq (1-e)\delta_3 \Leftrightarrow e(1+\delta_3) \leq \delta_3 \Leftrightarrow e \leq \frac{\delta_3}{1+\delta_3} \approx 0.3594. \quad (4.42)$$

Notemos que la condición $e \leq 0.3594$ otra vez nos garantiza tres cosas:

1. La ecuación de Kepler tiene solución: E^* .
2. La función mayorizante $\phi(t) = \delta - t + \frac{\delta^2 t^2}{1-\delta t}$ tiene soluciones reales: t^* y t^{**} .
3. E^* está localizada en el intervalo $[M-t^*, M+t^*]$ y es la única solución en $(M-t^{**}, M+t^{**})$.

Por último, el método de Newton-Raphson definido por (4.7), iniciando en $E_0 = M$, converge a E^* .

□

4.4. Las condiciones de convergencia del teorema de la α -teoría de Wang-Zhao aplicadas a la ecuación de Kepler.

Como hemos visto, la α -teoría de Smale es una técnica muy conocida para estudiar la convergencia semilocal del método de Newton-Raphson, esto es, establece sobre un punto inicial x_0 y una función f ciertas condiciones que permiten al método de Newton-Raphson converger.

Recordar que las condiciones impuestas por la α -teoría son:

- i) $f'(x_0) \neq 0$.
- ii) $\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \beta$.
- iii) $\frac{1}{k!} \left| \frac{f^{(k)}(x_0)}{f'(x_0)} \right| \leq \gamma^{k-1}, \quad k \geq 2$.
- iv) $\alpha = \beta\gamma \leq 3 - 2\sqrt{2}$.

También recordar que que los resultados de la convergencia son probados a partir de de la siguiente función mayorante

$$\phi_S(t) = \beta - t + \sum_{k \geq 2} \gamma^{k-1} t^k = \beta - t + \frac{\gamma t^2}{1 - \gamma t}. \quad (4.43)$$

Wang y Zhao,[186], consideraron que la condición *iv)* es muy restrictiva y demostraron que la convergencia del método de Newton-Raphson se puede obtener, además de las dos primeras condiciones, a partir de la condición

$$\frac{1}{k!} \left| \frac{f^{(k)}(x_0)}{f'(x_0)} \right| \leq \gamma_k, \quad k \geq 2,$$

y la función mayorante

$$\phi_W(t) = \beta - t + \sum_{k \geq 2} \gamma_k t^k. \quad (4.44)$$

Es importante notar que en (4.43) se conoce la descomposición del sumatorio, pero en (4.44) no es así, pues se trata de una serie infinita.

Nótese también que $\gamma_k \leq \gamma^{k-1}$, para $k \geq 2$. Luego $\phi_W(t) \leq \phi_S(t)$, para $t \geq 0$. En consecuencia $\phi_S(t)$ mayorisa a $\phi_W(t)$, por tanto $\phi_S = 0$ implica que $\phi_W = 0$, pero no lo contrario.

El trabajo de Wang y Zhao tiene un gran significado teórico, pues los mismos pueden ser aplicados a situaciones donde la α -teoría de Smale falla. En [89] encontramos un ejemplo que ilustra éste caso.

A continuación presentamos un teorema en el que se aplica la α -teoría de Wang-Zhao a la ecuación de Kepler.

Teorema 4.4. *Si $e \leq 0.3678$, entonces la ecuación de Kepler tiene solución. Dicha solución está localizada en el intervalo (4.2) y es única en (4.3) donde t^* y t^{**} son las raíces de la función mayorizante*

$$\phi(t) = \delta \exp(t) - (1 + \delta)t, \quad \text{con } \delta = \frac{e}{1 - e}. \quad (4.45)$$

Además, el método de Newton-Raphson, definido por (4.7), iniciando en $E_0 = M$, converge a dicha solución.

Demostración. Tomemos a $E_0 = M$ en el teorema de la α -teoría de Wang-Zhao. Luego se deben cumplir las siguientes condiciones:

- i) $f'(E_0) \neq 0$.
- ii) $\beta(E_0, f) = \left| \frac{f(E_0)}{f'(E_0)} \right|$.
- iii) $\frac{1}{k!} \left| \frac{f^{(k)}(E_0)}{f'(E_0)} \right| \leq \gamma_k, \quad k \geq 2$.

Para ello calcularemos los parámetros β y γ .

$$\begin{aligned} f'(E) = 1 - e \cos E & \Rightarrow |f'(E_0)| \geq 1 - e. \\ f''(E) = e \sin E & \Rightarrow |f''(E_0)| \leq e. \\ f'''(E) = e \cos E & \Rightarrow |f'''(E_0)| \leq e. \\ & \vdots \\ & |f^{(k)}(E_0)| \leq e, \quad k \geq 2. \end{aligned}$$

Así

$$\beta(E_0, f) = \left| \frac{f(E_0)}{f'(E_0)} \right| \leq \frac{e}{1 - e}, \quad \text{por tanto } \beta = \delta. \quad (4.46)$$

$$\frac{1}{k!} \left| \frac{f^{(k)}(E_0)}{f'(E_0)} \right| \leq \frac{1}{k!} \left(\frac{e}{1 - e} \right) = \gamma_k, \quad \text{por lo que } \gamma_k = \frac{\delta}{k!}. \quad (4.47)$$

Sustituyendo (4.46) y (4.47) en (4.44) sigue

$$\begin{aligned} \delta - t + \sum_{k \geq 2} \gamma_k t^k &= \delta - t + \sum_{k \geq 2} \frac{1}{k!} \delta t^k. \\ &= \delta - t + \delta \sum_{k \geq 2} \frac{t^k}{k!}. \\ &= \delta - t + \delta [\exp(t) - 1 - t]. \\ &= \delta \exp(t) - (1 + \delta)t = p(t). \end{aligned} \quad (4.48)$$

Analícemos a (4.48)

$$\begin{aligned} p'(t) &= \delta[\exp(t)] - (1 + \delta) \Rightarrow p'(0) = \delta - 1 - \delta = -1 < 0. \\ p''(t) &= \delta[\exp(t)] \Rightarrow f''(t) > 0, \quad \forall t > 0. \end{aligned} \quad (4.49)$$

Luego, existe un único mínimo \hat{t} tal que $p'(\hat{t}) = 0$. Notemos que

$$p'(\hat{t}) = \delta[\exp(\hat{t})] - (1 + \beta) = 0.$$

Por tanto

$$\delta[\exp(\hat{t})] = 1 + \delta \Leftrightarrow \exp(\hat{t}) = \frac{1 + \delta}{\delta} \Leftrightarrow \exp(\hat{t}) = 1 + \frac{1}{\delta}.$$

Luego

$$\hat{t} = \ln\left(1 + \frac{1}{\delta}\right). \quad (4.50)$$

Ahora analicemos la condición de convergencia de $p(\hat{t})$. Sea $p(\hat{t}) < 0$, luego

$$p(\hat{t}) = \delta[\exp(\hat{t})] - (1 + \delta)\hat{t} < 0. \quad (4.51)$$

Sustituyendo (4.50) en (4.51) sigue

$$\begin{aligned} p(\hat{t}) &= \delta\left[\exp\left(\ln\left(1 + \frac{1}{\delta}\right)\right)\right] - (1 + \delta)\ln\left(1 + \frac{1}{\delta}\right) < 0. \\ &= \delta\left(1 + \frac{1}{\delta}\right) - (1 + \delta)\ln\left(1 + \frac{1}{\delta}\right) < 0. \\ &= (1 + \delta) - (1 + \delta)\ln\left(1 + \frac{1}{\delta}\right) < 0. \\ &= (1 + \delta)\left[1 - \ln\left(1 + \frac{1}{\delta}\right)\right] < 0. \end{aligned} \quad (4.52)$$

Recordar que $(1 + \delta) > 0$, luego (4.52) será cierto si y sólo si

$$1 - \ln\left(1 + \frac{1}{\delta}\right) < 0. \quad (4.53)$$

De (4.53) sigue

$$1 < \ln\left(1 + \frac{1}{\delta}\right) \Leftrightarrow \exp(1) < 1 + \frac{1}{\delta} \Leftrightarrow \exp(1) - 1 < \frac{1}{\delta}.$$

Luego

$$\delta = \frac{1}{\exp(1) - 1}. \quad (4.54)$$

Sustituyendo (4.46) en (4.54)

$$\frac{e}{1-e} < \frac{1}{\exp(1)-1} \Leftrightarrow e \exp(1) - e < 1 - e \Leftrightarrow e < \frac{1}{\exp(1)} \approx 0.3678.$$

La condición $e \leq 0.3678$. también nos garantiza tres cosas:

1. La ecuación de Kepler tiene solución: E^* .
2. La función mayorizante $\phi(t) = \delta \exp(t) - (1 + \delta)t$ tiene soluciones reales: t^* y t^{**} .
3. E^* está localizada en el intervalo $[M - t^*, M + t^*]$ y es la única solución en $(M - t^{**}, M + t^{**})$.

Por último, el método de Newton-Raphson definido por (4.7), iniciando en $E_0 = M$, converge a E^* . □

4.5. Resultados sobre existencia y unicidad de soluciones

La aplicación a la ecuación de Kepler de las condiciones de convergencia semilocal de las técnicas vistas en la subsección anterior, nos proporciona una serie de resultados de existencia y unicidad de solución de la misma.

En concreto, si tomamos a

$$E_0 = M,$$

se garantiza la existencia de solución para la ecuación de Kepler.

Además, al comparar los dominios de existencia y unicidad de solución según las distintas técnicas, se verifica que bajo las condiciones de Cauchy-Kantorovich, véase el Cuadro 4.1, la ecuación de Kepler posee el mayor rango de solución ($e \leq 0.4142$).

La comparación también nos permite verificar que bajo las condiciones fuertes de Kantorovich, véase nuevamente el Cuadro 4.1, la ecuación de Kepler posee un rango de solución, ($e < 0.3759$), mayor que bajo las condiciones de la α -teoría de Wang-Zhao, ($e < 0.3678$), y de la α -teoría de Smale ($e \leq 0.3594$).

En el Cuadro 4.2 se muestran los valores de t^* y t^{**} para los intervalos de existencia de solución $[M - t^*, M + t^*]$ y de unicidad de solución $(M - t^{**}, M + t^{**})$ para las 4 teorías consideradas. En dicho cuadro

$$\delta = \frac{e}{1-e}$$

En los Cuadros 4.3 y 4.4 hemos calculado los valores de t^* y t^{**} para dos situaciones reales, correspondientes al planeta Mercurio ($M = 0.0067\pi$, $e = 0.205635$) y al asteroide Hungaria ($M = 0.69\pi$, $e = 0.0736162$). Mostramos los correspondientes intervalos de existencia y de

Cuadro 4.1: Función mayorizante para el método de Newton aplicado a la ecuación de Kepler para las diferentes teorías consideradas, junto con la condición en términos de la excentricidad para que dichas teorías puedan ser aplicadas.

Condición	Función mayorizante	Condición
Kantorovich	$p(t) = \frac{1}{2}\delta t^2 - t + \delta$	$e \leq 0.4142$
Kantorovich fuerte	$p_G(t) = \delta - t + \frac{1}{2}\delta t^2 + \frac{1}{6}\delta t^3$	$e \leq 0.3759$
α -teoría de Smale	$\phi(t) = \delta - t + \frac{\delta^2 t^2}{1 - \delta^2 t}$	$e \leq 0.3594$
α -teoría de Wang-Zhao	$\phi_{WZ}(t) = \delta \exp(t) - (1 + \delta)t$	$e \leq 0.3678$

Cuadro 4.2: Valores de t^* y t^{**} para las 4 teorías consideradas. En todos los casos, hemos denotado $\delta = \frac{e}{1-e}$.

Condición	t^*	t^{**}
Kantorovich	$\frac{1 - \sqrt{1 - 2\delta^2}}{\delta}$	$\frac{1 + \sqrt{1 - 2\delta^2}}{\delta}$
Kantorovich fuerte	Menor raíz positiva de $p_G(t)$	Mayor raíz positiva de $p_G(t)$
α -teoría de Smale	$\frac{(1 + \delta^3) - \sqrt{(1 + \delta^3)^2 - 8\delta^3}}{4\delta^2}$	$\frac{(1 + \delta^3) + \sqrt{(1 + \delta^3)^2 - 8\delta^3}}{4\delta^2}$
α -teoría de Wang-Zhao	Menor raíz positiva de $\phi_{WZ}(t)$	Mayor raíz positiva de $\phi_{WZ}(t)$

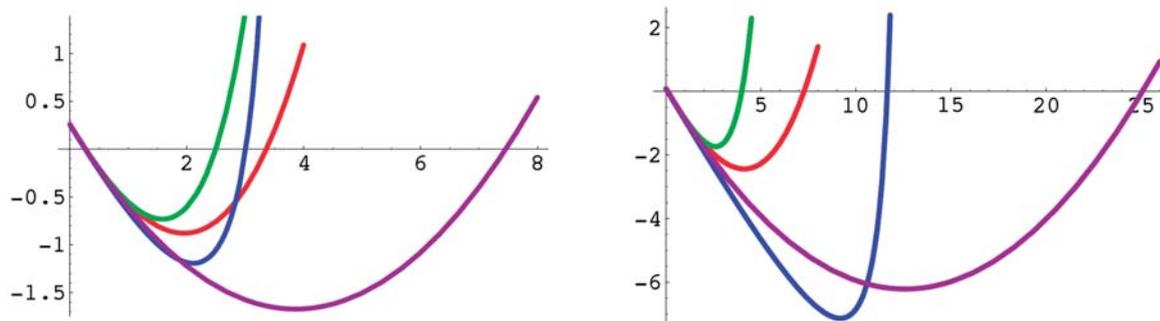
unicidad de solución para la ecuación de Kepler, junto con la solución exacta E^* . Hemos denotado los valores de t^* y t^{**} con un subíndice que hace referencia a la teoría con la que está relacionado: K para Kantorovich, G para fuertes de Kantorovich, S para Smale y WZ para Wang-Zhao.

Cuadro 4.3: Existencia de solución para dos casos reales.

Casos	t_K^*	t_G^*	t_S^*	t_{WZ}^*	E^*
Planeta Mercurio	0.269	0.268	0.267	0.269	0.0269
$[M - t^*, M + t^*]$	$[-0.247, 0.289]$	$[-0.247, 0.290]$	$[-0.242, 0.284]$	$[-0.248, 0.290]$	
Asteroide Hungaria	0.079	0.079	0.079	0.079	2.226
$[M - t^*, M + t^*]$	$[2.088, 2.247]$	$[2.087, 2.247]$	$[2.088, 2.247]$	$[2.087, 2.247]$	

Cuadro 4.4: Unicidad de solución para dos casos reales.

Casos	t_K^{**}	t_G^{**}	t_S^{**}	t_{WZ}^{**}	E^*
Planeta Mercurio	7.459	3.363	7.330	2.496	0.026
$(M - t^{**}, M + t^{**})$	$(-7.43, 7.48)$	$(-3.34, 3.38)$	$(-7.30, 7.35)$	$(-2.47, 2.51)$	
Asteroide Húngaria	25.094	7.271	12.659	3.993	2.226
$(M - t^{**}, M + t^{**})$	$(-22.92, 27.26)$	$(-5.10, 9.43)$	$(-10.42, 14.84)$	$(-1.82, 6.16)$	

Figura 4.1: Gráficas de las funciones mayorizantes $\phi_W(t)$, $\phi_G(t)$, $\phi_S(t)$ y $\phi_K(t)$ para el caso del planeta Mercurio (izquierda) y el asteroide Hungaria (derecha).

4.6. Líneas futuras de investigación y Conclusiones

Siguiendo lo presentado en el Capítulo 2, Métodos iterativos, quedan pendiente la construcción de los teoremas de convergencia semilocal para los métodos de Yun-Petković y de Danby. Esto nos permitirá conocer, con más profundidad, el comportamiento de los mismos, en términos de sus cotas del error y sus intervalos de existencia y unicidad de solución.

En concordancia con los resultados obtenidos en el Capítulo 3: Métodos iterativos y la ecuación de Kepler, pretendemos profundizar en lo que ocurre con el método de Chebyshev, cuando se aplica a la ecuación de Kepler a partir del punto inicial $E_0 = M$, así como extender el estudio de los métodos multipuntos aplicados a la ecuación de Kepler, a otros de mayor orden de convergencia. Tal sería el caso del método de Stirling. Otra línea abierta de investigación es extender la caracterización de 2-ciclos superatractores, no tan solo para el método de Newton, sino también a los demás métodos visto en esta memoria. Asimismo, caracterizar ciclos superatractores del orden $n > 2$. No menos importante es la caracterización de los intervalos en donde la derivada Schwarziana de los métodos vistos, sea negativa, pues esto nos permitirá identificar puntos de inicio óptimos, con los cuales se podría asegurar la convergencia de los mismos.

En cuanto a las teorías de convergencia semilocal aplicadas a la ecuación de Kepler, nos queda por profundizar en la aplicación de las mismas bajo otros puntos de partidas, en

especial cuando $E_0 = \pi$ o las condiciones iniciales dadas en (3.7).

En lo que concierne a los principales aportes de esta tesis, los mismos pueden ser enfocados a partir de tres aspectos:

1. Personales: conformada por la permanente crisis de conocimiento en que me ha hecho caer esta investigación. Gracias a ella, me he dado cuenta de los infinitos aspectos matemáticos que me faltan por aprender, aspectos que día tras día trataré de ir entendiendo y profundizando. Ya lo dijo Newton: “soy como un niño que se ha entretenido buscando en la arena algunos caracoles, mientras frente a mi se extiende el infinito mar de la sabiduría.”
2. Para mi país, la República Dominicana: en el sentido de que esta tesis es un proyecto de nación en el que la misma se convertirá en un libro de texto sobre Métodos numéricos. Con esta acción, el presente documento tendrá permanencia en el ambiente matemático dominicano, permitiendo que las futuras generaciones de doctores dominicanos, tengan una herramienta eficaz, con la cual realizar sus futuras investigaciones.
3. Matemáticos: en este aspecto, el principal aporte de esta memoria es la caracterización de los valores de la excentricidad para los cuales, las teorías de de convergencia semilocal de Kantorovich, Smale, Gutiérrez y Wang-Zhao, aseguran que la ecuación de Kepler tiene solución. Otro aporte significativo, además de los novedosos teoremas que se presentan, lo constituye la sistematización de todos los métodos que intervienen en la misma. Esto así, pues existen muy pocas investigaciones que presenten tantos métodos iterativos, con un esquema de desarrollo tan completo, en especial, lo correspondiente al análisis de convergencia.

4.7. Anexo I: Cuadros con experimentos numéricos de la sección 3.2.

En lo que sigue, se presentan una serie de cuadros que contienen los resultados de las iteraciones necesarias para que los métodos estudiados converjan. Para un mejor entendimiento de los mismos, damos a continuación el significado de las iniciales empleadas:

- T : Número total de iteraciones.
- m : Media aritmética del número de iteraciones que empleó el método para converger.
- \bar{m} : Media ponderada del número de iteraciones que empleó el método para converger.
- N/C: No convergencia del método en ese punto.

Cuadro 4.5: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos $E_0 = M$.

Casos	E_0	Newton	Chebyshev	Halley	super-Halley	$\alpha = \frac{3}{2}$	$\alpha = 2$	Danby
Caso (1)	0.05	2	2	2	2	2	2	2
Caso (2)	0.05	7	4	4	4	5	6	4
Caso (3)	0.9	3	2	2	2	2	2	2
Caso (4)	1.5	4	3	3	3	3	3	3
Caso (5)	3.0	4	3	2	2	2	2	2
Caso (6)	3.0	3	2	2	2	2	2	2
Caso (7)	0.51π	4	3	3	3	3	3	3
Caso (8)	0.88π	3	3	3	3	3	3	3
Caso (9)	0.29π	4	3	3	3	3	3	3
Caso (10)	0.29π	5	NC	4	4	4	4	4
Caso (11)	0.81π	4	3	3	3	3	3	3
Caso (12)	0.69π	3	3	2	2	2	2	2
Caso (13)	0.11π	4	3	3	3	3	3	3
Caso (14)	0.0067π	3	2	2	2	2	2	2
T		53	36*	38	38	39	40	40
m		3.79	2.77*	2.71	2.71	2.79	2.86	2.86
\bar{m}		3.79	4.39*	4.29	4.29	4.42	4.53	5.72

Nota*: Esta media se ha calculado únicamente con 13 experimentos.

4.8. Anexo II: Cuadros con experimentos numéricos de la sección 3.2.

Cuadro 4.6: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos $E_0 = \pi$.

Casos	E_0	Newton	Chebyshev	Halley	super-Halley	$\alpha = \frac{3}{2}$	$\alpha = 2$	Danby
Caso (1)	π	4	4	4	4	4	4	4
Caso (2)	π	8	6	5	5	17	17	5
Caso (3)	π	4	3	3	3	3	3	3
Caso (4)	π	5	4	3	3	3	3	3
Caso (5)	π	3	2	2	2	2	2	2
Caso (6)	π	3	2	3	3	3	2	2
Caso (7)	π	4	3	3	3	3	3	3
Caso (8)	π	3	3	3	3	3	3	2
Caso (9)	π	5	4	4	4	4	4	3
Caso (10)	π	5	4	4	4	4	4	3
Caso (11)	π	3	3	3	3	3	3	2
Caso (12)	π	4	3	3	3	3	3	2
Caso (13)	π	6	4	4	4	5	5	4
Caso (14)	π	5	4	4	4	4	4	3
T		62	49	48	48	61	60	41
m		4.43	3.5	3.43	3.43	4.36	4.29	2.93
\bar{m}		4.43	5.55	5.44	5.44	6.91	6.8	5.86

Tomando como punto inicial a $E_0 = M$, para el intervalo $[e, 0.20\pi]$, con $e \in [0.980, 0.999]$, el porcentaje de fracasos del método de Halley fue del 90 por ciento.

Cuadro 4.7: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando diversos procesos iterativos y tomando como punto de partida en todos los casos las condiciones dadas en (3.7).

Casos	E_0	Newton	Chebyshev	Halley	super-Halley	$\alpha = \frac{3}{2}$	$\alpha = 2$	Danby
Caso (1)	0.0550174	2	2	2	2	2	2	1
Caso (2)	0.6090380	4	3	3	3	3	3	3
Caso (3)	0.9765000	2	2	2	2	2	2	2
Caso (4)	1.9250000	3	2	2	2	2	2	2
Caso (5)	3.0765000	3	2	2	2	2	2	2
Caso (6)	3.7650000	4	3	3	3	3	3	3
Caso (7)	1.9722700	3	2	2	2	2	2	2
Caso (8)	3.2406200	3	3	3	3	3	3	2
Caso (9)	1.2358500	3	2	2	2	2	2	2
Caso (10)	1.6474000	4	3	3	2	3	3	3
Caso (11)	3.1040800	4	3	3	3	3	3	2
Caso (12)	2.2302700	3	2	2	2	2	2	2
Caso (13)	0.8315320	4	4	2	3	3	3	3
Caso (14)	0.0413741	3	2	2	2	2	2	2
T		45	35	33	33	34	34	31
m		3.21	2.5	2.36	2.36	2.43	2.43	2.21
\bar{m}		3.21	3.96	3.74	3.74	3.85	3.85	4.42

Cuadro 4.8: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Newton y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	8	25	11	9	10	9	7	6	6	6
0.981	8	11	15	7	9	12	7	6	6	6
0.982	8	12	22	45	21	50	7	6	6	6
0.983	8	12	14	16	13	13	7	5	6	6
0.984	8	21	21	32	11	12	7	5	6	6
0.985	8	108	13	195	85	9	8	6	6	6
0.986	8	10	15	14	17	8	8	6	6	6
0.987	7	39	20	18	12	19	8	6	6	6
0.988	15	53	28	213	10	10	9	6	6	6
0.989	17	204	13	13	22	11	9	6	6	6
0.990	13	101	12	53	52	14	31	6	6	6
0.991	18	12	138	444	10	11	8	7	6	6
0.992	59	8	13	17	10	15	74	7	6	6
0.993	99	8	8	11	18	14	10	7	6	6
0.994	8	12	8	31	29	92	25	7	6	6
0.995	11	12	8	7	11	63	9	7	5	6
0.996	11	8	10	28	199	12	7	8	5	6
0.997	15	24	11	10	275	15	9	8	6	6
0.998	8	28	162	28	19	17	16	8	6	6
0.999	8	8	8	7	16	36	10	8	6	6
T	345	716	550	1198	849	442	276	131	118	120
m	17.25	35.8	27.5	59.9	42.45	22.1	13.8	6.55	5.9	6.0
\bar{m}	27.34	56.74	43.58	94.94	67.28	35.03	21.87	10.38	9.354	9.51

Cuadro 4.9: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Newton y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	7	7	7	6	6	6	6	6	6	6
0.981	8	7	7	6	6	6	6	6	6	5
0.982	8	7	7	6	6	6	6	6	6	6
0.983	8	7	6	6	6	6	6	6	6	5
0.984	8	7	6	6	6	6	6	6	6	6
0.985	8	7	7	7	6	6	6	6	6	6
0.986	8	7	7	6	6	6	6	6	6	6
0.987	8	7	7	6	6	6	6	6	6	6
0.988	8	7	7	6	6	6	6	6	6	6
0.989	8	7	7	6	6	6	6	6	6	6
0.990	8	7	7	6	6	6	6	6	6	6
0.991	8	7	7	6	6	6	6	6	6	5
0.992	8	7	7	6	6	6	6	6	6	6
0.993	8	7	7	6	6	6	6	6	6	6
0.994	9	7	7	6	6	6	6	6	6	5
0.995	8	7	7	6	6	6	6	6	6	5
0.996	8	7	7	6	6	6	6	6	6	6
0.997	8	7	7	6	6	6	6	6	6	6
0.998	8	7	8	6	6	6	6	6	6	5
0.999	8	7	7	6	6	6	6	6	6	6
T	160	140	139	121	120	120	120	120	120	114
m	8.0	7.0	6.95	6.05	6.0	6.0	6.0	6.0	6.0	5.7
\bar{m}	12.68	11.09	11.02	9.59	9.51	9.51	9.51	9.51	9.51	9.03

Cuadro 4.12: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Chebyshev y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	4	5	5	5	4	4	5	4	4
0.981	6	5	5	5	5	4	4	5	5	5
0.982	5	4	5	5	5	5	4	4	5	4
0.983	6	5	5	5	5	4	4	4	4	4
0.984	5	5	5	5	5	4	4	4	4	5
0.985	5	5	5	6	5	4	5	4	5	4
0.986	5	5	5	5	5	4	4	4	4	5
0.987	5	5	5	5	5	4	5	4	4	5
0.988	5	5	5	5	5	4	4	5	4	4
0.989	7	5	5	5	5	4	4	4	4	4
0.990	5	5	5	5	5	4	4	4	4	5
0.991	5	5	5	5	5	4	5	4	4	5
0.992	5	5	5	6	5	4	4	4	5	4
0.993	5	5	6	5	5	4	5	4	5	5
0.994	5	5	5	5	5	4	5	4	4	4
0.995	5	5	5	5	5	4	4	4	4	5
0.996	5	5	5	5	5	4	4	4	4	4
0.997	5	5	5	5	5	4	4	4	4	5
0.998	5	5	5	5	5	4	4	4	4	4
0.999	5	5	5	5	5	5	4	4	4	5
T	104	98	101	102	100	82	85	83	85	90
m	5.2	4.9	5.05	5.1	4.0	4.1	4.25	4.15	4.25	4.5
\bar{m}	8.24	7.77	8.0	8.08	6.34	6.5	6.74	6.58	6.74	7.13

Cuadro 4.13: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Chebyshev y tomando como punto de partida las condiciones dadas en (3.7).

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	3	2	3	3	3	3	3	3	3
0.981	6	3	2	3	3	3	3	3	4	3
0.982	6	3	3	3	3	3	3	3	3	3
0.983	5	3	2	3	3	3	3	3	3	3
0.984	7	3	2	3	3	3	3	3	3	3
0.985	6	3	2	3	3	3	3	3	3	3
0.986	6	3	2	3	3	3	3	3	3	3
0.987	5	3	2	3	3	3	3	3	3	3
0.988	7	3	2	3	3	3	3	4	3	3
0.989	6	3	2	3	3	3	3	3	3	3
0.990	6	3	2	3	3	3	3	3	3	3
0.991	7	3	2	3	3	3	3	3	3	3
0.992	6	3	2	3	3	3	3	3	3	3
0.993	6	3	2	3	3	3	3	3	3	3
0.994	6	3	2	3	3	3	3	3	3	3
0.995	6	3	2	3	3	3	3	3	3	3
0.996	6	3	2	3	3	3	3	3	3	3
0.997	6	3	2	3	3	3	3	3	3	3
0.998	6	3	2	3	3	3	3	3	3	3
0.999	7	3	2	3	3	3	3	3	3	3
T	121	60	41	60	60	60	60	61	61	60
m	6.05	3.0	2.05	3.0	3.0	3.0	3.05	3.05	3.05	3.0
\bar{m}	9.59	4.76	3.25	4.76	4.76	4.76	4.83	4.83	4.83	4.76

Cuadro 4.14: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	4	5	5	5	6	5	5	5	4	4
0.981	4	5	5	5	5	6	5	5	5	5
0.982	4	5	5	5	5	5	5	5	5	4
0.983	4	5	5	5	5	5	5	5	4	4
0.984	5	6	5	5	5	5	5	5	4	4
0.985	4	5	6	5	5	5	5	5	5	4
0.986	4	5	5	5	5	5	5	6	5	4
0.987	4	5	5	5	5	5	5	5	4	4
0.988	4	5	5	5	6	5	5	5	4	5
0.989	5	5	5	5	5	5	5	5	5	5
0.990	5	5	5	5	5	5	5	5	5	4
0.991	5	6	5	5	5	5	5	5	5	5
0.992	5	5	5	6	5	5	5	6	5	5
0.993	5	5	5	6	5	5	5	5	5	5
0.994	5	6	6	6	5	5	5	5	5	4
0.995	5	6	5	5	5	5	5	5	5	4
0.996	5	6	5	5	6	5	5	5	5	4
0.997	6	7	5	5	5	5	5	6	5	5
0.998	6	6	6	5	6	5	5	5	5	5
0.999	6	6	6	5	5	6	5	6	5	5
T	95	109	104	103	104	102	100	104	95	85
m	4.75	5.45	5.2	5.15	5.2	5.1	5.0	5.2	4.75	4.25
\bar{m}	7.53	8.64	8.24	8.16	8.24	8.08	7.93	8.24	7.53	6.74

Cuadro 4.15: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	6	5	5	5	4	4	4	4	4	4
0.981	5	5	6	5	4	4	4	4	5	5
0.982	5	5	5	5	4	5	4	4	5	4
0.983	5	5	5	5	4	4	4	4	4	4
0.984	5	5	5	6	4	4	4	4	4	4
0.985	5	5	5	6	4	5	4	4	4	4
0.986	5	5	5	5	4	4	4	4	4	4
0.987	5	5	5	5	4	4	5	5	4	4
0.988	5	5	5	5	4	4	4	4	4	4
0.989	5	5	6	5	4	4	4	4	4	4
0.990	5	5	5	5	5	4	4	4	4	4
0.991	6	5	5	5	4	4	4	5	4	4
0.992	5	6	5	5	4	4	4	4	5	4
0.993	6	5	6	5	4	4	4	4	4	4
0.994	5	5	6	5	4	5	4	4	4	4
0.995	5	5	5	6	4	4	4	4	4	5
0.996	5	5	6	5	4	4	4	4	4	4
0.997	8	5	5	5	5	4	5	5	4	4
0.998	8	5	5	5	4	4	4	4	4	4
0.999	7	5	5	5	4	4	4	4	4	4
T	111	101	105	103	82	83	82	83	83	82
m	5.55	5.05	5.25	5.15	4.1	4.15	4.1	4.15	4.15	4.1
\bar{m}	8.8	8.0	8.32	8.16	6.5	6.58	6.5	6.58	6.58	6.5

Cuadro 4.16: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Halley y tomando como punto de partida las condiciones dadas en (3.7).

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	6	3	2	3	3	3	4	3	3	3
0.981	5	3	3	3	3	3	3	3	3	3
0.982	5	4	2	3	3	3	3	3	3	3
0.983	6	3	2	3	3	3	3	3	3	3
0.984	5	3	2	3	3	3	3	3	3	3
0.985	5	3	2	4	3	3	3	3	3	3
0.986	5	3	3	4	3	3	3	3	3	3
0.987	5	5	2	3	3	3	3	3	3	3
0.988	5	3	2	3	3	3	3	3	3	3
0.989	6	3	2	3	3	3	3	3	3	3
0.990	5	3	2	3	3	3	3	3	3	3
0.991	5	3	2	3	3	3	3	3	3	3
0.992	5	3	2	3	3	3	3	3	3	3
0.993	5	3	2	3	3	4	3	3	3	3
0.994	5	3	2	3	3	3	3	3	3	3
0.995	5	3	2	3	3	3	3	3	3	3
0.996	5	3	2	3	3	3	3	3	3	3
0.997	5	3	2	3	3	3	3	4	3	3
0.998	6	3	3	3	3	3	4	3	3	3
0.999	5	3	2	3	3	3	3	3	3	3
T	104	63	43	62	60	61	62	61	60	60
m	5.2	3.15	2.15	3.1	3.0	3.05	3.1	3.05	3.0	3.0
\bar{m}	8.24	4.99	3.41	4.91	4.75	4.83	4.91	4.83	4.75	4.75

Cuadro 4.17: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	5	5	5	5	4	5	4	4	4
0.981	5	7	5	5	5	4	4	4	4	4
0.982	5	5	5	5	5	5	4	4	4	4
0.983	5	6	5	5	5	4	4	4	4	4
0.984	5	5	5	5	5	5	4	4	4	5
0.985	5	5	5	5	5	5	4	4	4	4
0.986	5	5	5	5	5	5	4	4	4	4
0.987	4	6	5	5	5	5	4	4	5	4
0.988	5	6	6	5	5	5	4	4	5	4
0.989	6	6	6	5	5	5	4	4	4	4
0.990	5	6	6	5	5	5	4	4	4	4
0.991	6	6	6	5	5	5	5	4	4	4
0.992	6	6	6	5	5	6	5	4	4	4
0.993	6	6	6	5	5	5	5	4	4	4
0.994	6	5	6	5	5	5	5	4	4	4
0.995	6	6	5	5	5	5	5	4	4	4
0.996	7	10	5	5	5	5	5	4	4	4
0.997	11	14	5	5	5	5	5	4	4	5
0.998	8	30	5	5	5	5	5	4	4	5
0.999	223	6	9	5	5	5	5	4	4	4
T	334	151	111	100	100	98	90	80	82	83
m	16.7	7.55	5.55	5.0	5.0	4.9	4.5	4.0	4.1	4.15
\bar{m}	26.47	11.97	8.8	7.92	7.92	7.77	7.13	6.34	6.5	6.58

Cuadro 4.18: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	4	4	4	4	5	5	4	4	4
0.981	5	4	4	5	4	4	5	4	5	4
0.982	5	4	4	4	4	4	4	4	4	4
0.983	5	4	4	4	4	4	4	4	4	5
0.984	5	5	4	5	4	4	4	4	4	5
0.985	5	4	4	4	4	4	5	4	4	4
0.986	5	5	4	4	5	4	4	4	4	4
0.987	5	4	4	4	4	4	4	4	5	4
0.988	5	4	4	4	4	4	4	4	4	4
0.989	5	4	4	5	5	4	4	4	4	4
0.990	5	4	4	4	4	4	5	4	4	4
0.991	5	4	4	4	4	4	4	4	4	4
0.992	5	4	4	4	4	5	4	4	4	4
0.993	5	4	4	5	4	5	4	4	4	4
0.994	5	4	4	4	4	4	4	4	4	4
0.995	5	4	4	4	4	4	5	4	4	4
0.996	5	4	4	4	4	4	4	4	4	4
0.997	5	4	5	4	5	4	4	4	5	4
0.998	6	4	4	4	4	4	5	4	4	4
0.999	5	4	4	4	4	4	4	4	4	4
T	101	82	81	84	83	83	86	80	83	82
m	5.05	4.1	4.05	4.2	4.15	4.15	4.3	4.0	4.15	4.1
\bar{m}	8.0	6.5	6.42	6.66	6.58	6.58	6.82	6.34	6.58	6.5

Cuadro 4.19: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de super-Halley y tomando como punto de partida las condiciones dadas en (3.7).

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	3	2	3	3	3	4	3	3	3
0.981	5	3	2	3	3	3	3	3	3	3
0.982	5	3	2	3	3	3	3	3	3	3
0.983	5	3	2	3	3	3	3	3	3	3
0.984	6	3	2	3	3	3	3	3	4	3
0.985	6	3	2	3	3	3	3	3	3	3
0.986	5	3	2	3	3	3	3	3	3	3
0.987	6	3	2	3	3	3	3	3	4	3
0.988	5	3	2	3	3	3	3	3	3	3
0.989	5	3	2	3	4	3	3	3	3	3
0.990	5	3	3	3	3	3	3	3	3	3
0.991	5	3	2	3	3	3	3	3	3	3
0.992	5	3	2	3	3	3	3	3	3	3
0.993	5	3	2	3	3	3	3	3	3	3
0.994	5	3	2	3	3	3	3	3	3	3
0.995	5	3	2	3	3	3	3	3	3	3
0.996	5	3	2	3	3	3	3	3	3	3
0.997	5	3	2	3	3	4	3	4	3	3
0.998	5	2	2	3	3	3	4	3	3	3
0.999	5	2	3	3	3	3	3	3	3	4
T	103	58	42	60	61	61	62	61	62	61
m	5.15	2.9	2.1	3.0	3.05	3.05	3.1	3.05	3.1	3.05
\bar{m}	8.16	4.6	3.33	4.75	4.83	4.83	4.91	4.83	4.91	4.83

Cuadro 4.20: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = \frac{3}{2}$ y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	4	6	6	6	5	5	5	5	5	5
0.981	5	6	6	6	5	5	5	5	4	5
0.982	5	6	6	6	5	6	5	5	6	5
0.983	6	6	6	6	6	5	5	5	5	6
0.984	6	6	6	6	6	5	5	5	5	6
0.985	6	6	7	6	5	6	5	5	6	5
0.986	7	6	6	6	6	5	5	6	6	5
0.987	7	6	6	6	5	5	5	5	6	5
0.988	7	6	6	6	5	6	5	5	5	6
0.989	6	5	6	6	5	5	5	5	5	5
0.990	6	6	5	6	5	5	5	6	5	6
0.991	6	36	14	5	5	5	5	5	5	5
0.992	7	33	21	5	5	5	5	5	6	5
0.993	5	40	20	5	5	6	5	5	5	5
0.994	8	23	14	5	5	6	5	5	5	5
0.995	21	34	34	6	5	5	5	5	5	6
0.996	219	244	26	11	5	6	5	6	5	6
0.997	16	7	55	10	5	5	5	5	5	6
0.998	7	34	57	11	5	5	5	5	5	5
0.999	9	14	8	16	5	5	5	6	5	5
T	363	530	315	140	103	106	100	104	104	107
m	18.15	26.5	15.75	7.0	5.15	5.3	5.0	5.2	5.2	5.35
\bar{m}	28.77	42.0	24.96	11.09	8.16	8.4	7.92	8.24	8.24	8.48

Cuadro 4.21: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = \frac{3}{2}$ y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	6	5	5	5	5	4	4	4	5	4
0.981	6	5	5	5	5	4	5	4	5	4
0.982	6	5	5	5	4	5	4	4	4	4
0.983	6	5	5	5	4	5	4	4	4	4
0.984	6	5	5	5	4	5	4	5	4	5
0.985	6	5	5	5	4	4	4	4	5	4
0.986	6	5	5	5	4	4	5	5	4	4
0.987	6	5	5	5	5	4	4	4	5	4
0.988	6	5	5	6	4	4	4	4	4	5
0.989	6	5	6	5	4	4	4	5	5	4
0.990	6	5	5	5	4	4	4	4	4	5
0.991	6	5	5	6	4	4	4	4	4	4
0.992	6	5	5	5	4	5	4	4	4	4
0.993	6	6	5	5	4	5	4	4	4	4
0.994	6	5	5	5	4	4	4	4	4	4
0.995	6	5	5	5	4	4	4	4	4	4
0.996	6	6	5	5	4	5	5	4	4	4
0.997	6	5	5	5	5	4	4	5	4	4
0.998	6	5	5	5	4	4	4	5	4	5
0.999	6	5	4	5	5	4	4	5	4	5
T	120	102	100	102	85	86	83	86	85	85
m	6.0	5.1	5.0	5.1	4.25	4.3	4.15	4.3	4.25	4.25
\bar{m}	9.51	8.08	7.92	8.08	6.74	6.82	6.58	6.82	6.74	6.74

Cuadro 4.23: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = 2$ y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	12	13	7	6	6	6	6	5	5	5
0.981	6	9	7	6	6	6	6	5	4	5
0.982	9	8	6	6	6	6	6	5	5	5
0.983	7	8	6	6	6	6	6	5	5	5
0.984	17	9	6	6	6	6	6	5	5	5
0.985	90	12	5	6	6	6	6	5	5	5
0.986	11	5	6	5	6	6	6	5	5	5
0.987	44	5	8	5	6	6	6	5	5	5
0.988	98	10	7	5	6	6	6	5	5	5
0.989	53	26	19	6	6	6	6	5	5	5
0.990	35	226	14	8	5	6	6	5	5	5
0.991	93	11	118	11	5	6	6	5	5	5
0.992	232	26	7	8	5	6	6	5	5	5
0.993	6	25	16	12	5	5	6	5	5	5
0.994	127	85	56	12	5	6	6	5	5	5
0.995	90	21	31	44	6	6	6	5	5	5
0.996	32	NC	17	13	7	6	6	5	5	5
0.997	268	215	29	7	11	5	6	5	5	5
0.998	20	209	14	17	9	5	6	5	5	5
0.999	46	80	7	17	120	5	6	5	5	5
T	1296	1003*	386	206	238	117	120	100	99	100
m	64.8	52.8*	19.3	10.3	11.9	5.85	6.0	5.0	4.95	5.0
\bar{m}	102.71	83.69*	31.06	16.32	18.86	9.27	9.51	7.92	7.85	7.92

Nota*: Esta media se ha calculado únicamente con 19 experimentos.

Cuadro 4.24: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método para $\alpha = 2$ y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	13	8	6	5	5	5	4	4	4	4
0.981	13	16	6	5	5	5	4	4	6	4
0.982	43	8	6	5	5	5	4	4	4	4
0.983	13	7	6	5	5	5	4	4	4	4
0.984	17	9	6	5	5	5	4	4	4	4
0.985	69	24	6	5	5	5	4	4	4	4
0.986	89	8	6	5	5	5	4	4	4	4
0.987	30	9	6	5	5	5	4	4	4	4
0.988	167	22	6	5	5	5	4	4	4	4
0.989	29	31	6	5	5	5	4	4	4	4
0.990	73	14	6	5	5	5	4	4	4	4
0.991	15	16	6	5	5	5	4	4	4	4
0.992	24	26	6	5	5	5	4	4	4	4
0.993	25	17	6	5	5	6	4	4	4	4
0.994	21	101	6	5	5	5	4	4	4	4
0.995	80	114	6	5	5	5	4	4	4	4
0.996	158	14	6	5	5	5	5	4	4	4
0.997	12	13	6	5	5	4	4	4	4	4
0.998	31	59	6	5	5	4	4	4	4	4
0.999	75	52	6	5	5	5	4	4	4	4
T	997	568	120	100	100	99	81	80	82	80
m	49.85	28.4	6.0	5.0	5.0	4.95	4.05	4.0	4.1	4.0
\bar{m}	79.01	45.01	9.51	7.92	7.92	7.85	6.92	6.34	6.5	6.34

Cuadro 4.26: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida $E_0 = M$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	4	5	5	5	5	5	5	5	5	4
0.981	5	5	5	5	5	5	5	5	4	4
0.982	4	5	5	5	5	5	5	5	4	4
0.983	4	5	5	5	5	5	5	5	4	4
0.984	5	5	5	5	4	5	5	5	4	4
0.985	4	5	5	5	5	4	5	5	5	4
0.986	4	5	5	5	5	5	5	5	5	4
0.987	5	5	5	5	5	5	5	5	5	4
0.988	5	5	5	6	5	5	5	5	5	4
0.989	5	5	7	5	5	5	5	5	5	4
0.990	5	5	5	5	5	5	5	5	5	4
0.991	5	5	5	5	5	5	5	5	5	4
0.992	5	5	7	5	5	5	5	5	5	4
0.993	5	5	16	5	5	5	5	5	5	4
0.994	5	5	56	6	5	5	5	5	5	4
0.995	5	5	31	5	5	5	6	5	5	4
0.996	5	6	17	6	6	5	5	5	5	4
0.997	6	6	30	5	5	5	5	5	5	4
0.998	6	6	14	5	5	5	5	5	5	4
0.999	7	6	5	5	5	5	5	5	5	4
T	99	104	238	103	100	99	101	100	96	80
m	4.95	5.2	11.9	5.15	5.0	4.95	5.05	5.0	4.8	4.0
\bar{m}	9.9	10.4	23.8	10.3	10.0	9.9	10.1	10.0	9.6	8.0

Cuadro 4.27: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida $E_0 = \pi$.

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	4	4	4	4	4	4	4	4	4
0.981	5	5	4	4	4	4	4	4	4	4
0.982	5	4	4	6	4	5	4	4	4	4
0.983	5	4	4	5	5	4	4	4	4	4
0.984	6	4	4	5	4	4	4	4	4	4
0.985	6	4	4	5	4	4	4	4	4	4
0.986	5	4	4	5	4	4	4	4	4	4
0.987	5	5	4	4	4	4	4	4	4	4
0.988	5	4	4	5	4	4	4	4	4	4
0.989	5	4	4	5	4	4	4	4	4	4
0.990	5	4	4	4	4	4	4	4	4	4
0.991	5	4	6	4	4	4	4	4	4	4
0.992	5	5	6	4	4	4	4	4	4	4
0.993	5	5	6	5	4	4	4	4	4	4
0.994	5	5	6	4	5	4	4	4	4	4
0.995	6	5	6	4	4	4	4	4	4	4
0.996	6	5	6	4	6	4	4	4	4	4
0.997	6	5	6	4	4	4	4	4	4	4
0.998	7	4	6	4	4	4	5	4	4	4
0.999	5	5	4	4	4	4	4	4	4	3
T	107	89	96	89	84	81	81	80	80	79
m	5.35	4.45	4.8	4.45	4.2	4.05	4.05	4.0	4.0	3.95
\bar{m}	10.7	8.9	9.6	8.9	8.4	8.1	8.1	8.0	8.0	7.9

Cuadro 4.28: Número de iteraciones necesarias para alcanzar la solución de la ecuación de Kepler usando el método de Danby y tomando como punto de partida las condiciones dadas en (3.7).

$e \backslash M$	0.02π	0.04π	0.06π	0.08π	0.10π	0.12π	0.14π	0.16π	0.18π	0.20π
0.980	5	3	3	3	3	3	3	3	3	3
0.981	7	3	3	3	3	3	3	3	3	3
0.982	6	3	4	3	3	3	3	3	3	3
0.983	5	3	4	3	4	3	3	3	3	3
0.984	5	3	3	3	3	4	3	3	3	3
0.985	5	3	3	3	3	3	3	3	3	3
0.986	5	3	3	3	3	3	3	3	3	3
0.987	5	3	3	3	3	3	3	3	3	3
0.988	5	3	3	3	4	3	3	3	3	3
0.989	5	3	3	3	3	3	3	3	3	3
0.990	5	3	3	3	3	3	3	3	3	3
0.991	5	3	3	3	3	3	3	3	3	3
0.992	5	3	3	3	3	3	3	3	3	3
0.993	6	3	3	3	3	3	3	3	3	3
0.994	5	3	3	3	3	3	3	3	3	3
0.995	5	3	3	3	3	3	3	3	3	3
0.996	5	3	3	3	3	3	3	3	3	3
0.997	5	3	3	3	3	3	3	3	3	3
0.998	5	3	4	3	3	3	3	3	3	3
0.999	5	3	3	3	3	3	3	3	3	3
T	104	60	63	60	62	61	60	60	60	60
m	5.2	3.0	3.15	3.0	3.1	3.05	3.0	3.0	3.0	3.0
\bar{m}	10.4	6.0	6.3	6.0	6.2	6.1	6.0	6.0	6.0	6.0

Anexo III: Galería de gráficos para el capítulo 3

En este anexo mostramos las figuras asociadas a los conjuntos introducidos en el tercer capítulo (véanse las fórmulas (3.36), (3.37) y (3.38) respectivamente), así como los códigos de Mathematica empleados para su generación.

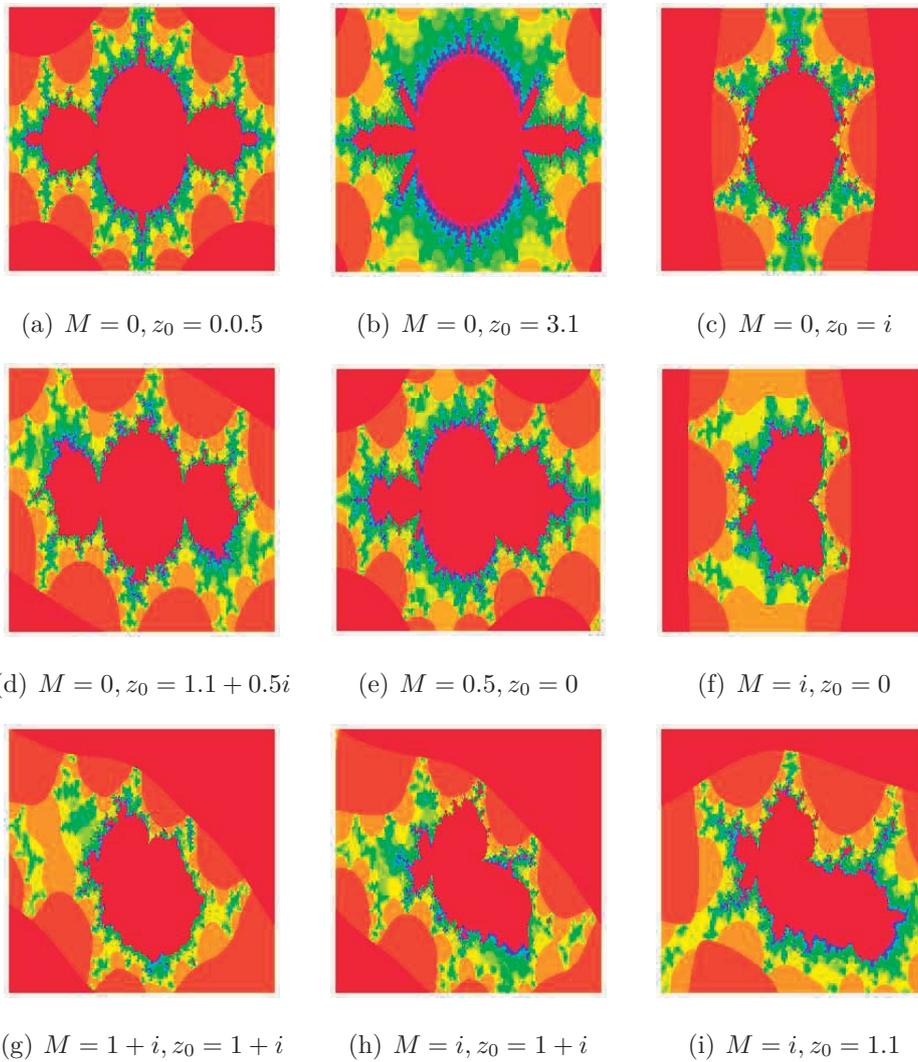


Figura 4.2: Conjuntos del tipo $\mathcal{K}(M; z_0) = \{e \in \mathbb{C} \mid \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}$, para distintos valores de M y z_0 .

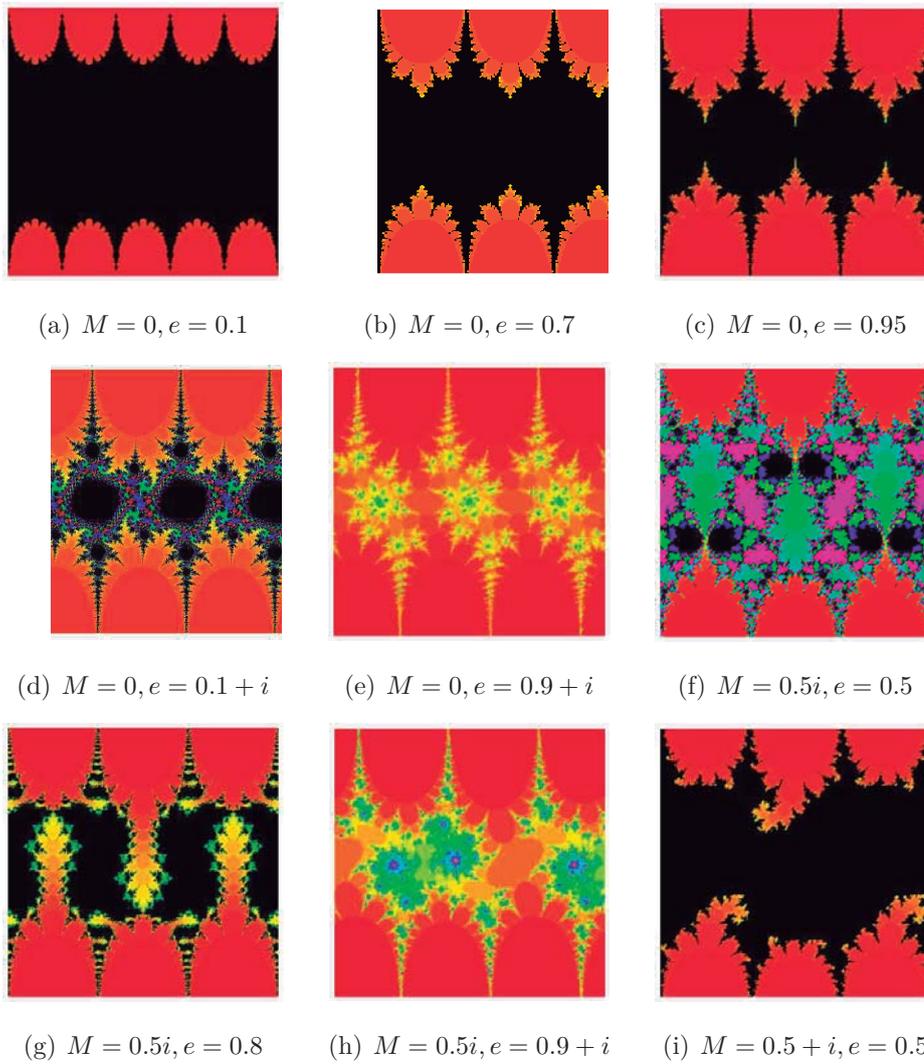


Figura 4.3: Conjuntos del tipo $\mathcal{J}(M, e) = \{z \in \mathbb{C} \mid \{f^n(M, e, z)\}_{n \in \mathbb{N}} \text{ acotada}\}$ para distintos valores de M y e .

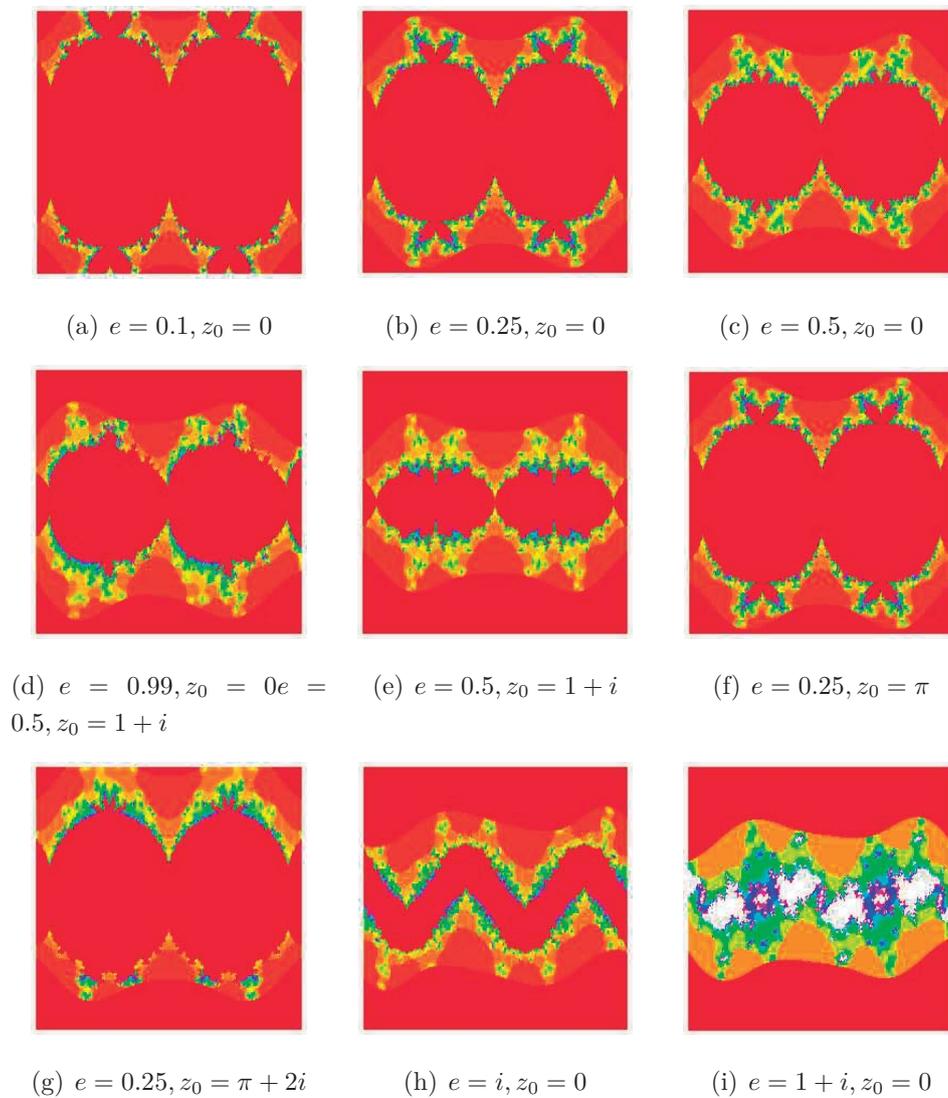


Figura 4.4: Conjuntos del tipo $\mathcal{I}(e, z_0) = \{M \in \mathbb{C} \mid \{f^n(M, e, z_0)\}_{n \in \mathbb{N}} \text{ acotada}\}$ para distintos valores de e y z_0 .

Cuadro 4.29: Código de Mathematica para generar fractales tipo Mandelbrot con la ecuación de Kepler. Está desarrollado para el caso particular en que $M = 0.01$, $e = 0.9$ y nivel de resolución de 0.1

```
KeplerFunction =  
  Compile[{{z, _Complex}},  
    Module[{w, k = 0}, w = 0.9 + 0.*I;  
      While[Abs[w] <= 100. && k <= 100, w = z*Sin[w] + 0.01 + 0*I;  
        k = k + 1]; k];  
  
datosKepler =  
  Table[KeplerFunction[x + y I], {y, -2., 2., .10}, {x, -Pi, Pi, .10}];  
  
ListDensityPlot[datosKepler, Mesh -> False, Frame -> True,  
  ColorFunction -> "BlueGreenYellow", ImageSize -> {156, 156}]
```

Cuadro 4.30: Código de Mathematica para generar fractales tipo Julia con la ecuación de Kepler. Está desarrollado para el caso particular en que $M = 0.01$, $e = 0.9$ y nivel de resolución de 40 iteraciones o pasos

```
KeplerIter = Compile[{{z, _Complex}},
Module[{cnt, nz=N[z]},
For[cnt=0,
(Abs[nz] <100) && (cnt<100), cnt++,
nz = anom+exc*Sin[nz]];
N[cnt]]];

anom=0.01; exc=0.9;KeplerFrac[z_, {{remin_, remax_}, {immin_, immax_}}, steps_] :=
{{{remin, remax}, {immin, immax}},
Table[KeplerIter[x+I*y],
{y, immin, immax, (immax-immin)/steps},
{x, remin, remax, (remax-remin)/steps}]];

FracColor = (If[#==1, Hue[1,1,0], Hue[2 #]]&);
FracPlot[{{remin_, remax_}, {immin_, immax_}}, matrix_,
colorfn_] :=
ListDensityPlot[matrix,
Mesh -> False, Frame -> True,
ColorFunction -> "BlueGreenYellow",
PlotRange -> {0, 100},
AspectRatio -> (immax-immin)/(remax - remin)]

pasos=40;FracPlot[KeplerFrac[0,
{{-1.5 Pi,1.5 Pi},{-1.5 Pi,1.5 Pi}},pasos],FracColor]
```


Bibliografía

- [1] A. Abad: *Astrodinámica*, Ed. Bubok Publishing S.L. Ed. Prensas Universitarias de Zaragoza España, 2012.
- [2] A. Abad, J. A. Docobo y A. Elipe: *Curso de Astronomía*, Ed. Prensas Universitarias de Zaragoza. España, 2002.
- [3] J. C. Adams: *On Newton's solution of Kepler's problem*, Monthly Notices of Royal Astronomical Society, **Vol. 43**, (1882), 43–49.
- [4] C. Altman: *Iterative methods of higher order*, Bull. Acad. Pol. Sci., Serie Sci. Math., Ast. et Phys., **Vol. 9**, (1961), 62–68.
- [5] S. Amat, S. Busquier, J. M. Gutiérrez: *Geometric constructions of iterative functions to solve nonlinear equations*, Journal of Computational and Applied Mathematics, **Vol. 157**, (2003), 197–205.
- [6] S. Amat, S. Busquier, J. M. Gutiérrez, M.A. Hernández: *On the global convergence of Chebyshev's iterative method*, Journal of Computational and Applied Mathematics, **Vol. 220**, (2008), 17–21.
- [7] S. Amat, S. Busquier, D. El Kebir, J. Molina: *A fast Chebyshev's method for quadratic equations*, Applied Mathematics and Computation, **Vol. 148**, (2004), 461–474.
- [8] S. Amat, S. Busquier, S. Plaza: *Review of some iterative root-finding methods from a dynamical point of view*, Scientia, Series A: Math. Sciences, **Vol. 10**, (2004), 3–35.
- [9] S. Amat, M.A. Hernández, N. Romero: *A modified Chebyshev's iterative method with at least sixth order of convergence*, Applied Mathematics and Computation, **Vol. 206**, (2008), 164–174.
- [10] I. K. Argyros: *Newton methods*, Nova Science Publisher, Inc. New York, 2004.
- [11] I. K. Argyros: *The convergence of a Halley-Chebysheff-type method under Newton-Kantorovich hypotheses*, Appl. Math. Lett. **Vol. 6**, (1993), 71–74.

- [12] I. K. Argyros: *A new convergence theorem for Steffensen's method on Banach spaces and applications*, Electronic Journal: Southwest Journal of Pure and Applied Mathematics, **Vol. 01**, (1997), 23–29.
- [13] I. K. Argyros: *The super-Halley method using divided differences*, Applied Mathematics Letters, **Vol. 10**, (1997), 91–95.
- [14] I. K. Argyros and D. Chen: *Results on Chebyshev method in Banach space*, Proyecciones, **Vol. 12**, (1993), 119–128.
- [15] I. K. Argyros and H. Ren: *Local convergence of efficient Secant-type methods for solving nonlinear equations*, Applied Mathematics and computation, **Vol. 218**, (2012), 7655–7664.
- [16] I. K. Argyros and H. Ren: *Local convergence of a Secant type methods for solving least squares problems*, Applied Mathematics and computation, **Vol. 217**, (2010), 3816–3824.
- [17] A. Aubanell. A. Benseny y A. Delshams: *Útiles básicos de cálculo numérico*, Editorial Labor, S. A. Barcelona, España, 1993.
- [18] M. Awais, A. Rafiq and F. Zafar: *Modified efficient variant of super-halley method*, Appl. Math; and Comp. **Vol. 189**, (2007), 2004–2010.
- [19] D. K. Babajee: *Analysis of higher order variants of Newton's method and their application to differential and integral equations and in ocean acidification*, Thesis, University of Mauritius, 2010.
- [20] D. K. Babajee and M. Z. Dauhoo: *An analysis of the properties of the variants of Newton's method with third order convergence*, Applied Mathematics and Computation, **Vol. 183**, (2006), 659–684.
- [21] M. Balazs and G. Goldner: *On Existence of Divided Differences in Linear Spaces*, Rev. Anal. Numer. Theorie Approximation, **Vol. 2**, (1973), 3–6.
- [22] B. Barna: *Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischen Gleichungen. II*, Publ. Math. Debrecen, **Vol. 4**, (1956), 384–397.
- [23] R. Behl, V. Kanwar and K. K. Sharma: *Another simple way of deriving several iterative functions to solve nonlinear equations*, Journal of Applied Mathematics, Received 25 June 2012; Revised 20 September 2012; Accepted 4 October 2012.
- [24] A. Ben-Israel: *Newton's Method with modified functions*, Contemporary Mathematics, **Vol. 204**, (1997), 39–50.

- [25] E. Bodewig: *On types of convergence and on the behavior of approximations in the neighborhood of a multiple root of an equation*, Quart. Appl. Math. **Vol. 7**, (1949), 325–333.
- [26] G. H. Brown Jr: *On Halley's variation of Newton's method*, Amer. Math. Monthly, **Vol. 84**, (1977), 726–728.
- [27] R. Burden and D. Faires: *Análisis Numérico*, Grupo Editorial Iberoamérica. España, 1985.
- [28] S. Busquier: *Métodos iterativos de alto orden y aplicaciones*, Tesis Doctoral, Universidad Politécnica de Cartagena, 2003. .
- [29] R. Campbell: *Les intégrales eulériennes et leurs applications*, Dunod, 1966.
- [30] V. Candela and A. Marquina: *Recurrence Relations for Rational Cubic Methods II: The Chebyshev Method*, Computing **Vol. 45**, (1990), 355–367.
- [31] B. Carnahan and H. A. Luther: *Applied Numerical Method*, John Wiley and Sons. New York, 1969.
- [32] A. L. Cauchy: *Lecons sur le Calcul Differentiel*, Buré frères. París, 1829.
- [33] O. Ciaurri: *Instantáneas Diferenciales. Métodos Elementales de Resolución de Ecuaciones Diferenciales Ordinarias, Estudio del Problema de Cauchy y Teoría de Ecuaciones y Sistemas Lineales*, Servicio de Publicaciones de la Universidad de la Rioja. Logroño, España, 2013.
- [34] S. C. Chapra and R. P. Canale: *Numerical methods for engineers*, McGraw-Hill. EEUU, 1998.
- [35] E. D. Charles y J. B. Tatum: *The convergence of Newton-Raphson iteration with Kepler's equation*, Celestial Mechanics and Dynamical Astronomy, **Vol. 69**, (1998), 357–372.
- [36] J. Chavarriga, I. A. García, J. Giné: *Manual de métodos numéricos*, Ediciones de la Universitat De Lleida. España, 1999.
- [37] F. Chicharro, A. Cordero, J. M. Gutiérrez and J. R. Torregrosa: *Complex dynamics of derivative-free methods for nonlinear equations*, Applied Mathematics and Computation, **Vol. 219**, (2013), 7023–7035.
- [38] P. L. Chebyshev: *Complete collected works*, Izdatel'stvo Akad. Nauk SSR. Moscow, 1946–1951.

- [39] J. Chen and W. Li: *An improved exponential regula falsi method with quadratic convergence of both diameter and point for solving nonlinear equations*, Applied Numerical Mathematics, **Vol. 57**, (2007), 80–88.
- [40] D. Chen, I. K. Argyros and Q. S. Quian: *A note on the Halley's method in Banach spaces* Appl. Math. and Comp. **Vol. 58**, (1993), 215–224.
- [41] D. Chen, I. K. Argyros and Q. S. Quian: *A local convergence theorem for the super-Halley's method in a Banach spaces*, Appl. Math. Lett. **Vol. 7**, (1994), 49–52.
- [42] C. Chun: *Certain improvement of Chebyshev-Halley methods with accelerated fourth-order convergence*, Appl. Math. Comput. **Vol. 189**, (2005), 1559–1568.
- [43] C. Chun: *Some variants of Chebyshev-Halley methods free from second derivative*, Appl. Math. Comput. **Vol. 191**, (2007), 193–198.
- [44] C. Chun: *Some second-derivative-free variants of Chebyshev-Halley methods*, Appl. Math. Comput. **Vol. 191**, (2007), 410–414.
- [45] C. Chun and Y. Ham: *Some second-derivative-free variants of super-Halley method with fourth-order convergence*, Appl. Math. Comput. **Vol. 195**, (2008), 537–541.
- [46] A. M. Cohen: *Análisis Numérico*, Editora Reverté, España, 1977.
- [47] D. Conte y C. De Boor: *Análisis Numérico elemental*, McGraw-Hill. Inc. USA, 1972.
- [48] B. A. Conway: *An improved algorithm due to Laguerre for the solution of Kepler's equation*, Celest. Mech. **Vol. 39**, (1986), 199–211.
- [49] P. Colwell: *Solving Kepler's Equations. Over three Centuries*, Willmann-Bell, Inc, Richmond. Virginia, 1993.
- [50] A. Cordero, J. L. Hueso, E. Martínez and J. R. Torregrosa: *Steffensen type methods for solving nonlinear equations*, Journal of Computational and Applied Mathematics, **Vol. 36**, (2012), 3058–3064.
- [51] J. M. Danby and T. M. Burkardt: *The solution of Kepler equation I*, Celestial Mechanics, **Vol. 31**, (1983), 95–107.
- [52] J. M. Danby and T. M. Burkardt: *The solution of Kepler's equation III*, Celestial Mechanics, **Vol. 31**, (1987), 303–312.
- [53] J. M. A. Danby: *Fundamentals of Celestial Mechanics*, 2nd. edition. Willmann-Bell, Inc. Richmond, EEUU, 1988.

- [54] M. Davies and B. Dawson: *On the global convergence of Halley's iteration formula*, Numer. Math. **Vol. 24**, (1975), 133–135.
- [55] J. P. Dedieu and M. H. Kim: *Newton's method for analytic system of equations with constant rank derivatives*, Journal of Complexity, **Vol. 18**, (2002), 187–209.
- [56] J. J. De Orús, M. A. Català y J. Núñez de Murga: *Astronomía esférica y mecánica celeste*, Edicions Universitat Barcelona, (2007).
- [57] N. L. De la Caille: *Sur les aelaements de la theorie du soleil, Premier Memoire*, Memoires de l'Academie Royale des Sciences, **Vol. A2**, (1750), 166–178.
- [58] J. M. Díaz y F. Benítez: *Introducción a los métodos numéricos para la resolución de ecuaciones*, Servicio de publicaciones de la Universidad de Cádiz. España, 1998
- [59] P. Díez: *A note on the convergence of the Secant method for simple and multiple roots*, Applied Mathematics Letters, **Vol. 16**, (2003), 1211–1215.
- [60] M. A. Diloné: *Una introducción al método de Newton-Raphson para resolver ecuaciones no lineales*, Editora ServiCentro. Nagua, Rep. Dominicana, 2009.
- [61] M. A. Diloné, M. García-Olivo and J. M. Gutiérrez: *A note on the semilocal convergence of Chebyshev's method*, aceptado en Bulletin of the Australian Mathematical Society, (2012).
- [62] M. A. Diloné y J. M. Gutiérrez: *A comparative study of different semilocal convergence results applied to Kepler's equation*, Computational Mathematics: Theory, Methods and Applications. Nova Science Publisher. EEUU, 2011.
- [63] M. A. Diloné y J. M. Gutiérrez: *On the existence of attractive cycles for Newton's method applied to Kepler's equation*, aceptado en Journal of Advanced Research in Applied Mathematics, 2013.
- [64] R. L. Devaney: *A First Course in Chaotic Dynamical Systems, Theory and Experiment*, Addison-Wesley. New York, 1992.
- [65] G. C. Donovan, A. R. Miller and T. J. Moreland: *Pathological functions for Newton's method*, American Mathematical Monthly, **Vol. 100**, (1993), 53–58.
- [66] H. Epstein: *Nuevas pruebas de la existencia de las funciones Feigenbaum*, Inst. Hautes Études Sco., Informe No. IHES, (1985), 55–85.
- [67] J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández y M. A. Salanova: *El método de Halley: Posiblemente, el método más redescubierto del mundo*, Margarita Mathematica

en memoria de José Javier (Chicho) Guadalupe, (L. Español y J. L. Varona, editores), Servicio de Publicaciones, Universidad de la Rioja. Logroño, España, 2001.

- [68] J. A. Ezquerro: *A Modification of the Chebyshev Method*, IMA Journal of Numerical Analysis, **Vol. 17**, (1996), 511–525
- [69] J. A. Ezquerro and M. A. Hernández: *A Modification of the Super-Halley Method under mild differentiability conditions*, Journal of Computational and Applied Mathematics, **Vol. 114**, (2000), 405–409
- [70] J. A. Ezquerro and M. A. Hernández: *A Generalization of the Kantorovich Type Assumptions for Halley's Method*, International Journal of Computer Mathematics, **Vol. 84**, (2007), 1771–1779.
- [71] J. A. Ezquerro: *Construcción de procesos iterativos mediante aceleración del método de Newton*, Tesis Doctoral, Servicio de Publicaciones de la Universidad de la Rioja. Logroño, España, 1996.
- [72] J. A. Ezquerro, J. M. Gutiérrez and M. A. Hernández: *A Construction Procedure of Iterative Methods with Cubical Convergence*, Applied Mathematics and Computation, **Vol. 85**, (1997), 181–199.
- [73] M. J. Feigenbaum: *Universal behavior in nonlinear systems*, Physica D, **Vol. 7**, (1983), 16–39.
- [74] J. S. Frame: *A variation of Newton's method*, American Mathematical Monthly, **Vol. 51**, (1944), 36–38.
- [75] C. E. Froberg: *Introduction to numerical analysis*, Addison-Wesley, Second edition. EE UU, 1970.
- [76] W. Gander: *On Halley's iteration method*, Amer. Math. Monthly, **Vol. 92**, (1985), 131–134.
- [77] F. García y A. Nevot: *Métodos numéricos en forma de ejercicios resueltos*, UPCU Departamento Publicaciones. Madrid, 1997.
- [78] M. García-Olivo y J. M. Gutiérrez, *La obra "El cálculo de las raíces de una ecuación" de P. L. Chebyshev*, Editorial Académica Española, Alemania, 2012.
- [79] W. Gautschi: *Numerical Analysis. An introduction*, Birkhauser. Boston, USA, 1997.
- [80] S. B. Gaythorpe: *Jeremiah Horrocks and his New Theory of the Moon*, Journal of the British Astronomical Association, **Vol.67**, (1957), 134–144.

- [81] J. Gerlach: *Accelerated convergence in Newton's Method*, SIAM Review, **Vol. 36**, (1994), 272–276.
- [82] W. J. Gilbert: *Generalizations of Newton's method*, Fractals, **Vol. 9**, (2001), 251–262.
- [83] J. M. Gutiérrez: *El método de Newton en espacios de Banach*, Tesis doctoral, Universidad de La Rioja. Logroño, España, 1995.
- [84] J. M. Gutiérrez: *A new semilocal convergence theorem for Newton's method*, Journal of Computational and Applied Mathematics, **Vol. 79**, (1997), 131–145.
- [85] J. M. Gutiérrez y M. A. Hernández: *Recurrence relations for the Super-Halley's method*, Computers Math. Applic. **Vol. 36**, (1998), 1–8.
- [86] J. M. Gutiérrez y M. A. Hernández: *Newton's method under weak Kantorovich conditions*, Journal of numerical analysis, **Vol. 20**, (2000), 521–532.
- [87] J. M. Gutiérrez y M. A. Hernández: *An acceleration of Newton's method: Super-Halley method*, Applied Mathematics and Computation, **Vol. 117**, (2001), 223–239.
- [88] J. M. Gutiérrez y M. A. Hernández: *An acceleration of Newton's method: Super-Halley method*, Applied Mathematics and Computation, **Vol. 117**, (2001), 222–239.
- [89] J. M. Gutiérrez, M. A. Hernández y M. Amparo Salanova: *α -theory for nonlinear Fredholm integral equations*, Grazer Mathematische Berrichte, **Vol. 346**, (2004), 187–196.
- [90] M. Grau-Sánchez and J. L. Díaz-Barrero: *An improvement of the Euler-Chebyshev iterative method*, J. Math. Anal. Appl. **Vol. 135**, (2006), 1–7.
- [91] M. Grau and J. M. Gutiérrez: *Some family of zero-finder methods derive from Obreshkov's techniques*, Comunicación particular.
- [92] M. Grau and J. M. Gutiérrez: *Zero-finder methods derived from Obreshkov's techniques*, Applied Mathematics and Computation, **Vol. 215**, (2009), 2992–3001.
- [93] M. Grau, M. Noguera y J. M. Gutiérrez: *On some computational ordeers of convergence*, Applied Mathematical Letters, **Vol. 23**, (2009), 472–478.
- [94] M. Grau, J. M. Peris and J. M. Gutiérrez: *Accelerated iterative methods for finding solutions of a system of nonlinear equations*, Applied Mathematics and Computation, **Vol. 190**, (2007), 1815–1823.
- [95] E. Hansen and M. Patrick: *A family of root finding methods*, Numer. Math. **Vol. 27**, (1977), 257–269.

- [96] P. Henrici: *Elements of numerical analysis*, John Wiley and Sons, Inc. New York, 1964.
- [97] M.A. Hernández: *Newton-Raphson's method and convexity*, Zb. Rad. Prirod.-Mat. Fak. Ser. Mat. **Vol. 22**, (1992), 159–166.
- [98] M. A. Hernández y M. A. Salanova: *La convexidad en la resolución de ecuaciones escalares no lineales*, Servicio de Publicaciones, Universidad de la Rioja. Logroño, España, 1996.
- [99] M.A. Hernández and N. Romero: *Accelerated convergence in Newton's method for approximating square roots*, Journal of Computational and Applied Mathematics, **Vol. 177**, (2005), 225–229.
- [100] M.A. Hernández y M. J. Rubio: *Un estudio unificado de la convergencia semilocal de métodos tipo Newton de dos puntos en espacios de Banach*, XX Congreso de ecuaciones diferenciales y aplicaciones. Sevilla, España, (2007), 1–8.
- [101] M. A. Hernández y M. J. Rubio: *Semilocal Convergence of the Secant Method under Mild Convergence Conditions of Differentiability*, Comput. Math. Appl. **Vol. 44**, (2004), 277–285.
- [102] M. A. Hernández y M. A. Salanova: *A Family of Chebyshev Type Methods in Banach Spaces*, International Journal of Computer Mathematics, **Vol. 61**, (1996), 1–2.
- [103] M. A. Hernández y M. A. Salanova: *An acceleration procedure of Regula Falsi method*, Tamkang Journal of Mathematics, **Vol.28**, (1997), 67–77.
- [104] H. H. H. Homeier: *A modified Newton method with cubic convergence: the multivariate case*, Journal of Computational and Applied Mathematics, **Vol. 169**, (2004), 161–169.
- [105] A. S. Householder: *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill. New York, 1970.
- [106] M. Hurley and C. Martin: *Newton's algorithm and chaotic dynamical systems*, SIAM J. Math. Anal. **Vol. 15**, (1984), 238–252.
- [107] J. A. Infante y J. M. Rey: *Métodos numéricos. Teoría, problemas y prácticas con MATLAB*, Ediciones Pirámide, tercera edición. Madrid, 2007.
- [108] A. N. Johnson and B. Neta: *High order nonlinear solver*, Journal of Computational Methods in Sciences and Engineering, **Vol. 8**, (2008), 245–250.
- [109] L. V. Kantorovich: *On Newton's method for functional equations*, Dokl. Akad. Nauk. **Vol. 59**, (1948), 1237–1240.

- [110] L. V. Kantorovich: *The majorant principle and Newton's method*, Dokl. Akad. Nauk. **Vol. 76**, (1951), 17–20.
- [111] L. V. Kantorovich and G. P. Akilov: *Functional Analysis*, Pergamon. Oxford, U.K., 1982.
- [112] V. Kanwar, S. Singh and S. Bakshi: *Simple geometric constructions of quadratically and cubically convergent iterative functions to solve nonlinear equations*, Numerical Algorithms, **Vol. 4**, (2008), 95–107.
- [113] A. Kharab and R. B. Guen: *An introduction to numerical methods. A MATLAB approach*, Chapman and Hall/ CRC. Florida, EE UU, 2002.
- [114] C. Chun and Y-I. Kim: *Several New Third-Order Iterative Methods for Solving Nonlinear Equations*, Acta Applicandae Mathematicae, **Vol. 109**, (2010), 1053–1063.
- [115] M. Kim: *On approximate zeros and rootfinding algorithms for a complex polynomial*, Mathematics of Computation, **Vol 51**, (1988), 707–719.
- [116] M. Kim: *Computational Complexity of the Euler Type Algorithms for the Roots of Complex Polynomials*, Thesis, City University of New York, 1985.
- [117] D. Kincaid: *Análisis Numérico*, Addison-Wesley Iberoamericana. España, 1991.
- [118] N. Kjurkchiev: *Initial approximations in Euler–Chebyshev's method*, Journal of Computational and Applied Mathematics, **Vol. 58**, (1995), 233–236.
- [119] E. Kobal: *Notice concerned with the calculation of roots of numerical equations*, Monatsh. Math. **Vol. 02**, (2002), 331–332.
- [120] J. Kou: *On Chebyshev-Halley with sixth-order convergence for solving nonlinear equations*, Appl. Math. Comput. **Vol. 190**, (2007), 126–131.
- [121] J. Kou and Y. Li: *A family of modified super-Halley methods with fourth-order convergence*, Appl. Math. Comput. **Vol. 186**, (2007), 366–370.
- [122] J. Kou, Y. Li and X. Wang: *A variant of super-Halley method with accelerate fourth-order convergence*, Appl. Math. Comput. **Vol 186**, (2007), 535–539.
- [123] K. Kneisl: *Julia sets for the super-Newton method, Cauchy's method and Halley's method*, Chaos, **Vol. 11**, (2001), 359–370.
- [124] R. J. Knill: *A modified Babylonian algorithm*, Amer. Math. Monthly, **Vol. 99**, (1997), 734–737.

- [125] J.D. Lambert: *Computational Methods in Ordinary Differential Equations*, John Wiley & Sons. Chichester, EE UU, 1976.
- [126] V. Lanchares and I. L. Pérez-Barrón: *The dynamics of Kepler equation*, Analytic and numerical techniques in orbital dynamics, **Vol. 22**, (2003), 75–82.
- [127] T. Li and J. Yorke: *Period three implies chaos*, Amer. Math. Monthly, **Vol. 82**, (1975), 985–993.
- [128] A. Liapunov: *Probleme Général de la stabilité du mouvement*, Ann. Fac. Sci. Univ. Toulouse, **Vol. 9**, (1907).
- [129] R. Lozada: *Análisis Matemático*, Ediciones Pirámide. Madrid, 1978.
- [130] M. Lyubich: *The quadratic family as a qualitatively solvable model of chaos*, Notices of the American Mathematical Society, **Vol. 47**, (2000), 1042–1052.
- [131] J. H. Mathews and K. D. Fink: *Métodos Numéricos con MATLAB*, Prentice Hall Iberia. Madrid, 2000.
- [132] A. Melman: *Geometry and convergence of Euler's and Halley's methods*, SIAM Rev. **Vol. 39**, (1997), 728–735.
- [133] V. Muto: *Curso de métodos numéricos*, Servicio editorial, Universidad del País Vasco. España, 1998.
- [134] F. R. Moulton: *An Introduction to Celestial Mechanics*, 2nd. revised edition. Dover. New York, 1970.
- [135] S. Nakamura: *Análisis numérico y visualización gráfica con MATLAB*, Prentice-Hill Inc, A Simon y Schuster Company. México, 1997.
- [136] B. Neta: *Several new methods for solving equations*, Inter. J. Computer. Math. **Vol. 23**, (1988), 265–282.
- [137] C. Neuhauser: *Matemática para ciencias*, Pearson Educación, S. A. Madrid, 2004.
- [138] A. Newmaier: *Introduction to numerical analysis*, Cambridge University Press. UK, 2001.
- [139] E. A. Lacomba: *Los sistemas dinámicos, ¿qué son y para qué sirven?*, Miscelánea Matemática, **Vol. 32**, (2000), 39–50.
- [140] N. Obreshkov: *Neue Quadraturformeln*, Abh. Preuss. Akad. Wissenschaft. Math. Nat. Kl. **Vol. 4**, (1940) 1–20.

- [141] J. M. Ortega and W. C. Rheinbolt: *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM. Philadelphia, EE UU, 2000.
- [142] M. A. Ostrowski: *Solution of Equations and Systems of Equations*, Academic Press. New York, 1966.
- [143] M. A. Ostrowski: *Solution of Equations in Euclidean and Banach Spaces*, Academic Press. New York, 1976.
- [144] A. Y. Ozban: *Some new variants of Newton's method*, Applied Mathematics Letters, **Vol. 17**, (2004), 667–682.
- [145] J. R. Pastor, P. Callejas y C. Trejo: *Análisis Matemático*, Editora Kapeluz. Argentina, 1969.
- [146] M. Pal: *Numerical analysis for scientists and engineers*, Alpha Science International Ltd. U. K, 2007.
- [147] M. Palacios: *Kepler equation and accelerated Newton method*, Journal of Computational and Applied Mathematics, **Vol.1** , (2002), 335–346.
- [148] M. S. Petković, L. D. Petković y D. Herceg: *On Schröder's families of root-finding methods*, J. Comput. Appl. Math. **Vol. 233**, (2010), 1755–1762.
- [149] H. Poincaré *Les méthodes nouvelles de la mécanique céleste*, Gauthier-Villars. Paris, 1892.
- [150] A. Quarteroni y F. Saleri: *Cálculo científico con MATLAB y Octave*, Springer-Verlag. Roma, 2006.
- [151] V. Ramírez, D. Barrera, M. Pasadas y P. González: *Cálculo numérico con Mathematica*, Editora Ariel. Barcelona, España, 2001.
- [152] A. Ralston: *Introducción al análisis numérico*, Editora Limusa. México, 1978.
- [153] F. J. Rodríguez: *Cálculo y Métodos Numéricos. Teoría, Algoritmos y Problemas resueltos*, Universidad Pontificia de Comillas, Colección Ingeniería. Madrid, España, 2003.
- [154] N. Romero: *Familias paramétricas de procesos iterativos de alto orden de convergencia*, Tesis doctoral, Universidad de La Rioja, Servicio de publicaciones, Madrid, España, 2006.
- [155] M. J. Rubio: *Procesos iterativos definidos mediante diferencias divididas*, Tesis doctoral, Universidad de La Rioja, Servicio de publicaciones, Madrid, España, 2006.

- [156] D. G. Saari and J. B. Urenko: *Newton's method, circle maps, and chaotic motion*, Amer. Math. Monthly, **Vol. 91**, (1984), 3–17.
- [157] R. A. Salehov: *The method of tangent hyperbolas*, Dokl.Akad.Nauk SSSR, **Vol. 4**, (1963), 482–485.
- [158] G.S. Salehov: *On the convergence of the process of tangent hyperbolas*, Dokl.Akad.Nauk SSSR, **Vol. 82**, (1952), 525–528.
- [159] G.S. Salehov and M. A. Mertvetsova: *On the convergence of some iterative processes*, Izv. Kazansk. Phil. ANSSSR. Ser. Phiz-Matem. i Techn. **Vol. 5**, (1954), 77–108.
- [160] T. R. Scavo and J. B. Thoo: *On the geometry of Halley's method*, American Mathematical Monthly, **Vol. 102**, (2002), 417–426.
- [161] F. Scheid: *Análisis Numérico*, McGrau-Hill. México, 1998.
- [162] E. Schroder: *Uber unendlich viele Algorithmen zur Au osung der Gleichungen*, Math. Ann. **Vol. 2**, (1870), 317–365.
- [163] R. E. Scraton: *Métodos numéricos básicos*, McGrau-Hill. México, 1987.
- [164] M. A. Sharaf and M.A. Banajh: *Error analysis for stationary one point iteration formulae of arbitrary order of convergence for solving universal Kepler's equation*, Applied Mathematical and Computation, **Vol. 155**, (2004), 699–707.
- [165] J. R. Sharma: *A family of third-order methods to solve nonlinear equations by quadratic curves approximation*, Applied Mathematics and Computation, **Vol. 184**, (2007), 210–215.
- [166] W. T. Shaw: *Complex analysis with Mathematica*, Cambridge University Press. United Kingdom, 2006.
- [167] M. Shub y S. Smale: *Computation complexity: O the geometry of polynomials and a theory of cost, I*, Ann. Sci. Éc. Norm. Supér. **Vol. 18**, (1985), 107–142.
- [168] S. Smale: *The fundamental theory for solving equations*, The international congress of mathematicians. Providence, USA, 1987.
- [169] S. Smale: *Newton's method estimates from data at one point*, The Merging of Disciplines: New Directions in Pure Applied and Computational Mathematics, (1986) 185–196.
- [170] S. Smale: *On the efficiency of algorithms of analysis*, Bull. AMS. **Vol. 13**, (1985), 87–121.

- [171] Smart, W. M: *Text Book of Spherical Astronomy*, Cambridge University Press. London, 1987.
- [172] J. F. Steffensen: *Factorial moments and discontinuous-frequency functions*, Scandinavian Actuarial Journal, **Vol. 1923**, (1923), 73–89.
- [173] G. W. Steward: *Aftersnotes on numerical analysis*, SIAM. Philadelphia, EE UU, 1996.
- [174] K. Stumpff: *On the application of Lie-series to problems of celestial mechanics*, National Aeronautics and Space Administration. Springfield, EE UU, 1968.
- [175] L. Stumpf: *Chaotic behaviour in the newton iterative function associated with kepler's equation*, Celestial Mechanics and Dynamical Astronomy, **Vol. 74**, (1999), 95–109.
- [176] C. O. Suárez-Alemán: *Aceptación en España de los criterios rigurosos del análisis matemático durante los siglos XIX y XX*, Tesis Doctoral, Universidad de Cádiz, 2007.
- [177] E. Suli and D. Mayers: *An introduction to numerical analysis*, Cambridge University Press. UK, 2003.
- [178] M. Tabor: *El caos y la integrabilidad en Dinámica no lineal: una introducción*, Wiley. New York, 1989.
- [179] B. Taylor: *Methodus Incrementorum Directa et Inversa*. Londres, 1715.
- [180] J. F. Traub: *Iterative methods for the solution of equations*, Prentice -Hall. N.J, EE UU, 1964.
- [181] S. Ulm: *Ob oboscennych Razdelenych Raznostjach*, I, II, Eesti NSV Tead. Akad. Toimised. Mat. **Vol. 16**, (1967), 146–156.
- [182] N. Ujeviec: *An iterative method for solving nonlinear equations*, Journal of Computational and Applied Mathematics, **Vol. 2**, (2006), 1–9.
- [183] O. Vaarmann: *High order iterative methods for decomposition-coordination problems*, Technological and economic development of economy, **Vol. XII**, (2006), 56–61.
- [184] J. M. Vallejo: *Tratado Elemental de Matemáticas*. Madrid, 1813. Segunda edición corregida y aumentada de 1832.
- [185] J. L. Varona: *Representación gráfica de fractales mediante un programa de cálculo numérico simbólico*, La Gaceta de la RSME, **Vol. 6**, (2003), 214–230.
- [186] D. Wang y F. Zhao: *The Theory of Smale's Point Estimation and its Applications*, Journal of Computational and Applied Mathematics, **Vol. 60**, (1995), 253–269.

- [187] X. H. Wang: *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal. **Vol. 20**, (2000), 123–134.
- [188] S. Weerakoon and T. G. I. Fernando: *A variant of Newton's method with accelerated third-order convergence*, Applied Mathematics Letters, **Vol. 13**, (2000), 87–93.
- [189] D. T. Whiteside: *The mathematical papers of Isaac Newton*, Cambridge University Press, **Vol. 06**, (1974), 308–323.
- [190] S. Wolfram: *The Mathematica Book*, Cambridge University Press. U. K, 1996.
- [191] J. C. Yakoubsohn: *Finding Zeros of Analytic Functions: α Theory for Secant Type Methods*, Journal of Complexity, **Vol. 15**, (1999), 239–281.
- [192] T. Yamamoto: *Historical development in convergence analysis for Newton's and Newton-like methods*, Journal of Computational and Applied Mathematics, **Vol. 124**, (2000), 1–23.
- [193] T. Yamamoto: *On the method of tangent hyperbolas in Banach spaces*, J. Comput. Appl. Math. **Vol. 21**, (1988), 75–86
- [194] D. M. Young. and R. G. Todd: *A Survey of Numerical Mathematics*, Dover Publications, Inc. New York, 1973.
- [195] T. J. Ypma: *Historical Developments of Newton-Raphson Method*, SIAM Review, **Vol. 37**, (1965), 531–551.
- [196] B. I. Yun: *A non-iterative method for solving non-linear equations*, Appl. Math. Comput. **Vol. 198**, (2008), 691–699.
- [197] B. I. Yun and M. S. Petković: *A quadratically convergent iterative method for nonlinear equations*, J. Korean Math. Soc. **Vol. 48**, (2011), 487–497.
- [198] Zhen Shi-Ming: *Point estimates for Halley's iteration*, Acta Mathematicae Applicatae Sinica, **Vol. 14**, (1991), 381–383.
- [199] S. Zheng: *The convergence of Halley's method and its optimal error bounds*, J. Hangzhou University, **Vol. 9**, (1982), 285–289.
- [200] S. Zhen and D. Robbie: *A note on the convergence of Halley's method for solving operator equations*, J. Austral. Math. Soc. Ser. B. **Vol. 37**, (1995), 16–25.
- [201] <http://www4.ncsu.edu/~njrose/Special/Bios/Danby-Obit.html>
- [202] <http://www.research.microsoft.com/~minka/papers/newton.html>

- [203] <http://www.math.heldermann.de/JAA/JAA09/jaa0919.html>
- [204] <http://numbers.computation.free.fr/Constants/Algorithms/Newton.html>
- [205] http://en.wikipedia.org/wiki/Halley%27s_method
- [206] <http://mathworld.wolfram.com/HalleysMethod.html>
- [207] http://math.fullerton.edu/mathews/n2003/halleysmethod/HalleysMethodMod_ink_3.html
- [208] <http://asteroidi.uai.it>