



UNIVERSIDAD INTERNACIONAL DE LA RIOJA

ESCUELA SUPERIOR DE INGENIERÍA
Y TECNOLOGÍA

TESIS DOCTORAL

**Mejora en la enseñanza y estudio de la dinámica de métodos
iterativos**

Memoria presentada por

Íñigo Sarría Martínez de Mendivil
para optar al grado de Doctor
por la Universidad Internacional de La Rioja

Dirigida por los Doctores

Rubén González Crespo
y
Ángel Alberto Magreñán Ruiz
Madrid 2018



UNIVERSIDAD INTERNACIONAL DE LA RIOJA

ESCUELA SUPERIOR DE INGENIERÍA
Y TECNOLOGÍA

TESIS DOCTORAL

**Mejora en la enseñanza y estudio de la dinámica de métodos
iterativos**

Memoria presentada por

Íñigo Sarría Martínez de Mendivil
para optar al grado de Doctor
por la Universidad Internacional de La Rioja

Dirigida por los Doctores

Rubén González Crespo
y
Ángel Alberto Magreñán Ruiz
Madrid 2018

Agradecimientos

La realización tanto de la redacción de esta memoria como de la investigación que culmina en esta Tesis Doctoral se la debo a los Doctores Rubén González Crespo y Alberto Magreñán Ruiz a quienes agradezco enormemente que me empujasen a pensarla, trabajarla y redactarla. Por su inconmensurable esfuerzo y trabajo que culmina con este proyecto.

También al Grupo de Investigación MOMAIN - Modelación Matemática Aplicada a la Ingeniería sin cuyos integrantes nada de esto sería posible.

Agradezco también a todo el personal, tanto Docente como Administrativo de la Universidad Internacional de La Rioja por todas las facilidades que me han dado a la hora de poder realizar la investigación. Particularmente a la Escuela Superior de Ingeniería y Tecnología en la que tengo el placer de realizar mi labor docente.

Muy particularmente quiero acordarme del Doctor Fernando López Hernández, tutor durante todo el tiempo que ha durado el proyecto, y compañero en ESIT, por aportarme una visión clara y concisa del trabajo que tenía entre manos, y preocuparse de orientarme en los momentos más difíciles.

Tengo que agradecer todo el apoyo incondicional a la Doctora Elena Giménez de Ory, que me ha hecho caer en la secta del Latex, por todas las compilaciones y por todos esos conceptos previos.

También quiero agradecer a Mikel Perales Jarillo, compañero de fatigas durante estos 15 años dedicados a la Educación Superior, por todas esas vivencias pasadas y por las futuras que vendrán.

No puedo dejar de agradecer eternamente a mi mujer, Cristina Jiménez Hernández, compañera de Licenciatura, docente incansable por esos momentos de estudio cuando aún éramos jóvenes, por ese amor a las matemáticas y por todos esos momentos de no estar a su lado robados para finalizar esta Tesis. Quisiera darte siempre un poco más de lo que te pido, las cosas importantes son las que están detrás de la piel. Por esas dos maravillosas, y esperemos matemáticas y del Athletic Club, hijas que me ha dado, Uxue eta Haizea. Tan diferentes, tan iguales, tan empáticas, tan como nosotros.

Para finalizar, agradecer a mi aita, Don Javier Sarría del Álamo el amor por las cuentas que me ha inculcado, y a mi amatxu Doña Carmen Martínez de Mendibil Sanz, maestra de tantos y tantos niños que pasaron por casa, por dejarme jugar con su pizarra gigante, por adorar la docencia y a ambos por educarme tal y como soy. Por supuesto que a mis hermanos, Doña Inmaculada Sarría y Don Javier Sarría, porque me abrieron las puertas de la vida.

Madrid, España, 15 de junio de 2018

Prólogo

¿Por qué? Como profesor, en todo momento me pregunto el porqué de las cosas, por muy simples que parezcan a primera vista. Como Matemático, no puedo dejar de lanzar interrogantes al mundo. Uniendo ambos ejes de mi personalidad, obtengo la respuesta a la primera pregunta planteada. Porque me gusta enseñar matemáticas y que los estudiantes las descubran y las aprendan conmigo. Esa ha sido la motivación principal de este proyecto que ha durado una larga temporada y que finaliza con el depósito y defensa de la siguiente tesis doctoral.

Tras 15 años dedicado a la docencia en el ámbito de la Educación Superior, y embarcado ya desde casi los inicios en la Universidad Internacional de La Rioja, me planteo si verdaderamente los estudiantes de Matemáticas avanzadas están disfrutando de sus asignaturas tanto como yo lo haría. Los resultados de algunas de las materias observadas no son tan altos como se esperaban, por lo que decido investigar qué es lo que está ocurriendo con estas calificaciones. Además, coincide que en estos momentos de mi vida, la investigación me lleva justo por esas bifurcaciones matemáticas.

El contenido de esta Tesis Doctoral abarca la docencia e investigación de métodos iterativos para la resolución de ecuaciones no lineales, tan utilizados en matemática aplicada, especialmente, en Ingeniería.

Una vez introducimos la hipótesis de partida y los objetivos tanto principal como secundarios, realizamos un paseo por la Historia de las Matemáticas, la docencia de la matemática aplicada y los principales conceptos de dinámica.

En los siguientes dos capítulos, obtenemos condiciones de convergencia nuevas y menos estrictas que en anteriores estudios para el método de la Secante y estudiamos la dinámica de la familia. Reservamos los dos capítulos posteriores para familias eficientes o de orden alto ∞ como una familia libre de segundas derivadas. Realizamos el estudio dinámico de estas familias encontrando desde convergencias hasta comportamientos caóticos. En el noveno capítulo de esta Tesis, ponemos en funcionamiento una herramienta, programada en MATLAB, con un grupo de alumnos de la asignatura la asignatura “Sistemas dinámicos discretos y continuos”, del Máster de Ingeniería Matemática y Computación de la Universidad Internacional de La Rioja. Analizamos los trabajos que envían los estudiantes después de utilizar la herramienta y los comparamos con otros dos grupos de estudiantes que no dispusieron de ella.

Índice general

I	Justificación de la investigación	1
1.	Introducción	3
1.1.	Planteamiento del trabajo	4
1.2.	Hipótesis y objetivos	6
1.3.	Metodología seguida durante la investigación	7
1.4.	Organización de la tesis	8
1.4.1.	Introducción a la investigación	8
1.4.2.	Estado del arte y marco teórico	8
1.4.3.	Desarrollo la investigación	8
1.4.4.	Conclusiones y trabajo futuro	8
1.4.5.	Anexos	9
II	Estado del arte y marco teórico	11
2.	Breve historia de las matemáticas y su enseñanza	13
2.1.	Los orígenes de las Matemáticas. Antiguo Egipto y Mesopotamia .	14
2.2.	La Antigüedad Clásica: Grecia	17
2.2.1.	Periodo jónico	18
2.2.2.	Periodo ateniense	20
2.2.3.	El periodo alejandrino	23
2.3.	Oriente y Occidente: El Islam y la Europa Feudal	26
2.4.	El renacimiento: álgebra, métodos de cálculo y trigonometría . . .	29
2.5.	La revolución científica: geometría analítica, cálculo diferencial e integral.	36
2.6.	Los matemáticos en el siglo XVIII	38
2.7.	Los matemáticos en el siglo XIX	39
2.8.	Docencia de la matemática aplicada	40

3. Conceptos previos de matemática aplicada	43
3.1. Introducción a la dinámica compleja	44
3.1.1. Puntos fijos, puntos periódicos y puntos críticos	44
3.1.2. Caos	49
3.1.3. Conjugación topológica	50
3.1.4. Conjuntos de Julia y Fatou	52
3.2. Introducción a la iteración de funciones reales	60
3.3. Introducción operadores definidos en espacios de Banach	66
III Desarrollo de la investigación	75
4. Estudio de la convergencia y dinámica del método de la Secante I	79
4.1. Introducción	80
4.2. Sucesiones mayorantes para el método de la Secante	81
4.3. Convergencia semilocal del método de la Secante	88
4.4. Ejemplos numéricos	92
5. Estudio de la convergencia y dinámica del método de la Secante II	97
5.1. Introducción	98
5.2. Sucesiones mayorizantes para el método de la Secante	98
5.3. Convergencia semilocal del método de la Secante	103
5.4. Aplicaciones	106
6. Estudio de la convergencia y dinámica de un método de orden seis libre de segundas derivadas	109
6.1. Introducción	110
6.2. El método y el análisis de convergencia	111
6.3. Aplicaciones Numéricas	114
6.4. Estudio dinámico para un caso especial del método	115
6.4.1. Estudio de los puntos fijos	116
6.4.2. Estudio de los puntos críticos y planos de parámetros	117
6.5. Conclusiones	120
7. Estudio de la convergencia y dinámica de una familia altamente eficiente	131
7.1. Introducción	132

7.2.	Desarrollo del esquema de ecuaciones escalares	134
7.3.	Desarrollo del esquema para casos multidimensionales	137
7.4.	Estudio dinámico de un caso especial del método (7.3.1)	140
7.4.1.	Estudio de los puntos fijos	141
7.4.2.	Estudio de los puntos críticos y planos de parámetros	141
7.5.	Experimentos numéricos	146
7.5.1.	Resultados y discusiones	166
8.	Estudio de la convergencia y dinámica de una familia de métodos de orden alto	171
8.1.	Introducción	172
8.2.	Convergencia local	174
8.3.	Estudio dinámico de un caso especial del método (8.1.2)	181
8.3.1.	Estudio de los puntos fijos y su estabilidad	181
8.3.2.	Estudio de los puntos críticos y espacios de parámetros	182
8.4.	Aplicación	193
9.	Herramienta Pedagógica basada en el desarrollo de una aplicación informática para la mejora del aprendizaje en matemática avanzada	195
9.1.	Introducción	196
9.1.1.	Conceptos básicos de métodos iterativos	198
9.2.	Dificultades encontradas	203
9.3.	Trabajos	205
9.4.	Descripción herramienta	211
9.4.1.	GUI myfirst.m	212
9.4.2.	GUI DynamicTool	213
9.4.3.	GUI SelectFCP	213
9.4.4.	GUI FPandCP	214
9.4.5.	Ejemplo de uso	215
9.5.	Puesta en marcha de la herramienta	224
9.6.	Resultados	224
9.6.1.	Resultados Trabajo 1 sin NP	227
9.6.2.	Resultados Trabajo 1 con NP	234
9.6.3.	Resultados Trabajo 2 sin NP	241
9.6.4.	Resultados Trabajo 2 con NP	248
9.6.5.	Resultados Trabajo 3 sin NP	255
9.6.6.	Resultados Trabajo 3 con NP	262

IV Conclusiones	269
10.Conclusiones y trabajo futuro	271
V Anexos	277
Bibliografía	439

Índice de figuras

3.1. Cuencas de atracción asociadas a las raíces del polinomio $p(z) = z^3 - 1$ al aplicarle el método de Newton. En amarillo aparece la cuenca de $z = 1$, en azul la cuenca de $z = e^{2\pi i/3}$ y en rojo la cuenca de $z = e^{4\pi i/3}$	48
3.2. Diagrama resultante de una conjugación topológica.	51
3.3. Cuencas de atracción asociadas a las raíces del polinomio $p(z) = z^4 - 1$, al aplicarle el conocido método de Halley $\left(H_f(z) = z - \frac{f(z)}{f'(z)} \frac{2}{2 - L_f(z)}\right)$. En verde la cuenca de $z = -1$, en amarillo la de $z = 1$, en azul la cuenca de $z = i$ y en rojo la de $z = -i$	55
3.4. Cuencas de atracción asociadas a la función $p(z) = z^2 - 1$ que tiene una órbita periódica atractora de periodo 2, dicha órbita es $\{0, -1\}$. En amarillo aparece la cuenca de la órbita periódica, mientras que en rojo aparece la cuenca de infinito.	56
3.5. Cuencas de atracción asociadas a la función $p(z) = z^2 + 0.25$ que tiene en $z_0 = 0.5$ un punto fijo parabólico. En amarillo aparece la cuenca de $z_0 = 0.5$, mientras que en rojo aparece la cuenca de infinito.	57
3.6. En la parte izquierda aparecen los conjuntos de Julia universales asociados al método de Newton, al método de Halley y al método de super-Halley aplicados a un polinomio cuadrático con dos raíces diferentes. Notemos que son idénticos y coinciden con la circunferencia unidad. En la parte derecha aparece el conjunto de Julia universal asociado al método de Chebyshev aplicado a un polinomio cuadrático con dos raíces diferentes.	59
3.7. Aplicación del algoritmo de estudio gráfico de la dinámica con uno y dos pasos.	61
3.8. Aplicación del algoritmo de estudio gráfico de la dinámica con tres y cincuenta pasos.	61
3.9. Órbitas de $x_0 = 0.75$ bajo la función $f(x) = -x^3 + x$, donde puede verse que el 0 es un punto fijo con multiplicador $\mu = 1$, pero con comportamiento atractor.	62

3.10. Órbitas de $x_0 = 0.1$, bajo la función $f(x) = x^3 + x$, donde puede verse que el 0 es un punto fijo con multiplicador asociado $\mu = 1$, pero con comportamiento repulsor.	63
3.11. Diagrama de Feigenbaum que se obtiene al aplicar el método de dos pasos de Newton a la familia $f(x) = x^3 + \gamma x + 1$, donde se observa que para algunos valores de $\gamma \in (0.94, 1.01)$ se producen bifurcaciones.	65
3.12. Exponentes de Lyapunov asociados al punto $\frac{1}{3}$ en función del valor del parámetro $c \in [0, 4]$ para la función logística.	67
4.1. Proceso conocido como Ammonia.	95
6.1. Diagrama de bifurcación del punto fijo extraño de $G(z, \alpha)$	117
6.2. Plano de parámetros asociado al punto crítico libre $cr_5(\alpha)$	119
6.3. Plano de parámetros asociado al punto crítico libre $cr_6(\alpha)$	120
6.4. Cuencas de atracción asociadas al método con $\alpha = 13.75$	121
6.5. Cuencas de atracción asociadas al método con $\alpha = -30$	123
6.6. Cuencas de atracción asociadas al método con $\alpha = 35$	124
6.7. Cuencas de atracción asociadas al método con $\alpha = 20$	125
6.8. Cuencas de atracción asociadas al método con $\alpha = -15$	126
6.9. Cuencas de atracción asociadas al método con $\alpha = -50$	127
6.10. Cuencas de atracción asociadas al método con $\alpha = 0$	128
6.11. Cuencas de atracción asociadas al método con $\alpha = i$	129
7.1. Plano de parámetros asociado al punto crítico $cr_8(b_1)$	142
7.2. Detalle del plano de parámetros asociado al punto crítico $cr_1(b_1)$	143
7.3. Detalle del plano de parámetros asociado al punto crítico $cr_1(b_1)$	144
7.4. Detalle del plano de parámetros asociado al punto crítico $cr_8(b_1)$	145
7.5. Cuenca de atracción asociada al método con $b_1 = 3.25$	146
7.6. Cuenca de atracción asociada al método con $b_1 = 2.6$	147
7.7. Cuenca de atracción asociada al método con $b_1 = 3.375$	148
7.8. Cuenca de atracción asociada al método con $b_1 = -1.5 - i$	149
7.9. Cuenca de atracción asociada al método con $b_1 = -3$	150
7.10. Cuenca de atracción asociada al método con $b_1 = -3.6 - \frac{i}{10}$	151
7.11. Cuenca de atracción asociada al método con $b_1 = 1$	152
7.12. Cuenca de atracción asociada al método con $b_1 = i$	153
8.1. Diagrama de bifurcación de los puntos fijos.	182
8.2. Espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$	183

8.3. Detalle del espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$	184
8.4. Detalle del espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$	185
8.5. Cuencas de atracción asociadas al método con $\alpha = -10$	186
8.6. Cuencas de atracción asociadas al método con $\alpha = 4.25$	187
8.7. Cuencas de atracción asociadas al método con $\alpha = -2.5$	188
8.8. Cuencas de atracción asociadas al método con $\alpha = 11$	189
8.9. Cuencas de atracción asociadas al método con $\alpha = 10 - 13i$	190
8.10. Cuencas de atracción asociadas al método con $\alpha = 0.5$	191
8.11. Cuencas de atracción asociadas al método con $\alpha = -0.5i$	192
8.12. Obtenido de www.essentialchemicalindustry.org/chemicals/ammonia.html . Proceso del Amonio	193
9.1. Caracterización de las cuencas de atracción asociadas al método de Newton aplicado al polinomio $p(z) = z^2 - 1$	201
9.2. Caracterización de las cuencas de atracción asociadas al método de Newton aplicado al polinomio $p(z) = z^3 - 1$	201
9.3. Plano de parámetros en el que existen diferentes anomalías.	202
9.4. Primer Trabajo.	206
9.5. Segundo Trabajo.	208
9.6. Tercer Trabajo.	210
9.7. Primera pantalla de la aplicación.	212
9.8. GUI DynamicTool	214
9.9. GUI SelectFCP.	214
9.10. GUI FPandCP	215
9.11. Pantalla inicial. Introducimos el polinomio y seleccionamos el mé- todo de Newton.	216
9.12. Seleccionar el punto crítico libre para pintar el plano de parámetros.	217
9.13. Plano de parámetros.	218
9.14. Zoom del plano de parámetros.	219
9.15. Fijamos el valor del parámetro para pintar los planos dinámicos.	220
9.16. Plano dinámico.	221
9.17. Plano dinámico con una órbita dibujada.	222
9.18. Plano dinámico con varias órbitas dibujadas.	223
9.19. Descriptivos asociados al primer grupo.	227
9.20. Descriptivos asociados al primer grupo.	228
9.21. Descriptivos asociados al primer grupo.	229

9.22. Descriptivos asociados al segundo grupo.	230
9.23. Descriptivos asociados al segundo grupo.	231
9.24. Descriptivos asociados al segundo grupo.	232
9.25. Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el primer trabajo.	233
9.26. Descriptivos asociados al primer grupo.	234
9.27. Descriptivos asociados al primer grupo.	235
9.28. Descriptivos asociados al primer grupo.	236
9.29. Descriptivos asociados al segundo grupo.	237
9.30. Descriptivos asociados al segundo grupo.	238
9.31. Descriptivos asociados al segundo grupo.	239
9.32. Prueba T de Student para la comparación de medias de notas obtenidas en el primer trabajo.	240
9.33. Descriptivos asociados al primer grupo.	241
9.34. Descriptivos asociados al primer grupo.	242
9.35. Descriptivos asociados al primer grupo.	243
9.36. Descriptivos asociados al segundo grupo.	244
9.37. Descriptivos asociados al segundo grupo.	245
9.38. Descriptivos asociados al segundo grupo.	246
9.39. Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el segundo trabajo.	247
9.40. Descriptivos asociados al primer grupo.	248
9.41. Descriptivos asociados al primer grupo.	249
9.42. Descriptivos asociados al primer grupo.	250
9.43. Descriptivos asociados al segundo grupo.	251
9.44. Descriptivos asociados al segundo grupo.	252
9.45. Descriptivos asociados al segundo grupo.	253
9.46. Prueba T de Student para la comparación de medias de notas obtenidas en el segundo trabajo.	254
9.47. Descriptivos asociados al primer grupo.	255
9.48. Descriptivos asociados al primer grupo.	256
9.49. Descriptivos asociados al primer grupo.	257
9.50. Descriptivos asociados al segundo grupo.	258
9.51. Descriptivos asociados al segundo grupo.	259
9.52. Descriptivos asociados al segundo grupo.	260
9.53. Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el tercer trabajo.	261
9.54. Descriptivos asociados al primer grupo.	262

9.55. Descriptivos asociados al primer grupo.	263
9.56. Descriptivos asociados al primer grupo.	264
9.57. Descriptivos asociados al segundo grupo.	265
9.58. Descriptivos asociados al segundo grupo.	266
9.59. Descriptivos asociados al segundo grupo.	267
9.60. Prueba T de Student para la comparación de medias de notas obtenidas en el primer trabajo.	268

Índice de cuadros

7.1. Problemas de prueba	148
7.2. Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_1(x)$	154
7.3. Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_2(x)$	155
7.4. Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_3(x)$	156
7.5. Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_4(x)$	157
7.6. Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_5(x)$	158
7.7. Los valores de ψ_i y ϕ_i (en radianes) para el ejemplo 7.4.	159
7.8. Comportamiento de la convergencia de los diferentes métodos en el problema de la síntesis cinemática 7.4.	160
7.9. Comportamiento de la convergencia de los diferentes métodos en el problema de valores de contorno 7.5.	161
7.10. Comportamiento de la convergencia de los diferentes métodos en el ejemplo 7.6.	163
7.11. Comportamiento de la convergencia de los diferentes métodos en el problema bidimensional de Bratu 7.7.	164
7.12. Comportamiento de la convergencia de los diferentes métodos en el problema de Fisher 7.8.	165
7.13. (Abscisas y pesos de la fórmula de cuadratura de Gauss Legendre para $t = 8$)	167
7.14. Comportamiento de la convergencia de los diferentes métodos en el problema integral de Hammerstein 7.9.	168
7.15. Tiempo de CPU tomado por diferentes métodos para el ejemplo correspondiente.	169
9.1. Resultados del primer grupo	204
9.2. Resultados del primer grupo	205

9.3. Resultados del segundo grupo	225
9.4. Resultados del segundo grupo	226
9.5. Estadísticas del segundo grupo	226

Parte I

Justificación de la investigación

Capítulo 1

Introducción

Esta tesis surge de la demanda que ha surgido en los últimos años por resolver los problemas que provienen de otras ramas como pueden ser la ingeniería, la química o las ciencias en general. Los problemas que se estudian generalmente no tienen una forma de resolución sencilla, como pueden tener los polinomios de grados 2, 3 o incluso 4, ya que nos vamos a encontrar al modelizar con ecuaciones que utilicen polinomios de grados altos, expresiones no polinómicas, ecuaciones diferenciales ordinarias o ecuaciones en derivadas parciales, este tipo de ecuaciones necesitan de métodos mucho más sofisticados que la aplicación de una fórmula. Este tipo de problemas raramente tienen una resolución directa y es aquí donde entra la matemática aplicada, y más concretamente, los métodos iterativos. Estos métodos generan una sucesión de aproximaciones que, bajo ciertas condiciones, van a converger hacia la solución de nuestro problema. El número de métodos iterativos ha crecido exponencialmente en los últimos años debido a que numerosos autores de reconocido prestigio en el área han visto la necesidad de desarrollar métodos con diferentes características y así poder utilizar el más adecuado a cada problema concreto. Así un método puede garantizar la convergencia de forma rápida, en términos del número de iteraciones necesario, hacia la solución o puede ser un método que nos garantice que para cualquier punto que tomemos como valor inicial, proporcione una sucesión convergente a la solución del problema.

Dentro de las características deseables de un método iterativo, destacan la convergencia y el comportamiento dinámico, ya que cuanto mayor sea el dominio de convergencia y más estable sea el comportamiento dinámico del método, el método va a ser “mejor” para el problema que estemos tratando. Relacionado con la convergencia, existen multitud de resultados de investigación en los que se han conseguido mejoras para diferentes métodos iterativos, y entre ellos destacan 3 grandes bloques:

- Estudios sobre el método de Newton. Este método destaca por su simpleza y su buen comportamiento, ya que es un método óptimo, de orden 2 y con unas regiones de convergencia muy buenas.

- Estudios sobre métodos libres de derivadas. Dentro de estos métodos destaca el método de la secante, que sería la equivalencia al método de Newton, pero sin la necesidad de utilizar derivadas. El hecho de estudiar este tipo de métodos surge del hecho de que en multitud de problemas nos encontramos con que o bien la derivada es muy costosa de calcular, operacionalmente, o incluso esta derivada no se puede calcular, por lo que el método de Newton, y todos aquellos que utilizan derivadas en su definición no pueden ser usados.
- Estudios sobre métodos con órdenes de convergencia altos. La definición de estos métodos se centra en que el orden del método de Newton es cuadrático y por lo tanto, son necesarias muchas iteraciones para resolver ciertos problemas en los que la tolerancia al error es muy pequeña. Teniendo este hecho en cuenta, se diseñan métodos, que pueden usar derivadas o no dependiendo de las características que busquemos, en los que los órdenes de convergencia son altos y así garantizar que el número de iteraciones necesarias sea menor.

En este contexto, así como para garantizar la convergencia existen multitud de resultados de investigación, tales como artículos, libros, etc., así como libros de texto e incluso diferentes asignaturas en estudios que requieran de una base de matemática aplicada, pero sin embargo, el número de resultados de investigación, libros de texto y materias vinculadas con la dinámica es muy inferior al de convergencia. Este hecho es debido a que hasta hace pocos años y, gracias a la irrupción de los ordenadores, podemos estudiar la dinámica, algo que antes era impensable, como se mostrará en los capítulos posteriores. En el Máster de Ingeniería Matemática y Computación, sí que existe una asignatura dedicada a este tipo de estudios dinámicos y es esta asignatura la que nos va a servir como punto de partida de esta tesis.

1.1. Planteamiento del trabajo

Como hemos comentado anteriormente, el punto de partida de esta tesis es la asignatura “Sistemas dinámicos discretos y continuos”, del Máster de Ingeniería Matemática y Computación de la Universidad Internacional de La Rioja. En esta asignatura, dentro de los sistemas dinámicos discretos, que supone dos terceras partes del contenido de la materia, es donde hemos detectado diferentes problemas de conceptos que no se han llegado a interiorizar debido a la complejidad intrínseca de los mismos y se ha tenido que intervenir. La forma que se ha decidido intervenir es dándole un enfoque diferente a la asignatura y diseñando una herramienta, que ha sido programada en lenguaje MATLAB por un propio alumno del Máster como trabajo fin de máster, bajo la supervisión tanto del doctorando que defiende esta tesis así como de los dos directores de la misma, y diseñando los métodos necesarios, para garantizar que los alumnos obtienen una mejora en

términos de dominio de los contenidos y adquisición de competencias vinculadas al estudio dinámico.

El planteamiento del trabajo se ha realizado por lo tanto de la siguiente manera:

1. Modificación de la forma de impartición de la asignatura y diseño de la herramienta para tratar de paliar los problemas detectados en la misma y estudios de convergencia de diferentes métodos iterativos.

En la parte que está más vinculada con la impartición de la asignatura, los pasos que se han seguido, de forma cronológica son los siguientes:

- Observar el rendimiento de los alumnos en la asignatura.
- Detección de los problemas surgidos en la misma.
- Medición tanto de forma cualitativa como cuantitativa de los problemas surgidos.
- Diseño de la herramienta y modificación de la temporalización y la forma de impartir las clases.
- Puesta en marcha de la asignatura con los cambios.
- Medición de los resultados.
- Establecer conclusiones.

Tanto el estudio de esta parte, como su desarrollo, pueden verse en el capítulo 9 de esta tesis.

En la parte que está vinculada con la investigación más pura en matemáticas, se ha decidido no entrar a estudiar convergencia y dinámica del método de Newton, ya que en relación a la convergencia existen una infinidad de resultados, mientras que la dinámica del propio método es muy estable, salvo en contados casos, por lo que nos centraremos en los métodos que se detallan a continuación:

- Método de la secante. En este caso hemos realizado dos estudios diferentes vinculados al estudio de la secante.
- Método de orden 6 sin derivadas segundas.
- Familia de métodos altamente eficientes.
- Familia de métodos de orden alto.

Tanto el estudio de esta parte, como los resultados obtenidos, pueden verse en los capítulos 4-8 de esta tesis.

1.2. Hipótesis y objetivos

Como consecuencia de lo anterior, enunciamos la hipótesis de investigación de la manera siguiente:

La adquisición, de una manera eficaz y eficiente, de competencias complejas, dentro del área de los sistemas dinámicos continuos y discretos es factible. El desarrollo de una nueva metodología y el soporte otorgado por el uso de las tecnologías de la información, gracias al desarrollo de aplicativos orientados a diferentes métodos iterativos, permitirán que el objetivo de aprendizaje de los estudiantes sea alcanzado.

De esta hipótesis de partida se deriva el objetivo principal de esta tesis consistente en:

Diseñar una metodología basada en la aplicación de una herramienta para tratar de paliar los problemas detectados y aplicarla a diferentes métodos iterativos

Como objetivos parciales de la tesis se pueden resaltar:

- Objetivo 1.- Diseño y puesta en marcha de una modificación en temporalización como en impartición de las clases en la asignatura “Sistemas dinámicos continuos y discretos”.
- Objetivo 2.- Diseño de una herramienta, que será desarrollada en lenguaje MATLAB, y su aplicación en la asignatura “Sistemas dinámicos continuos y discretos”.
- Objetivo 3.-Mejorar los resultados tanto a nivel cualitativo como cuantitativo en la asignatura “Sistemas dinámicos continuos y discretos” utilizando la herramienta y la nueva forma de impartir la asignatura.
- Objetivo 4.- Encontrar condiciones de convergencia para el método de la secante más suaves que la ya existentes.
- Objetivo 5.- Estudiar el comportamiento dinámico de la secante para diferentes polinomios.

- Objetivo 6.- Encontrar condiciones de convergencia para diferentes métodos y familias de métodos de órdenes altos.
- Objetivo 7.- Estudiar el comportamiento dinámico de diferentes métodos y familias de métodos de órdenes altos.

Otro objetivo, en última instancia, será responder a la pregunta:

¿Se ha resuelto el problema del alto nivel de abstención en las entregas y se han mejorado los resultados obtenidos por los estudiantes?

1.3. Metodología seguida durante la investigación

La metodología seguida, desde el inicio de la investigación hasta la escritura de este documento está basada en el método científico y se puede dividir en cuatro fases:

Fase 1: Estudio del estado del arte.

Fase 2: Planteamiento de hipótesis y objetivos contrastado y verificado por el estado del arte.

Fase 3: Desarrollo de la investigación.

Fase 4: Diseño, implementación y aplicación de prototipos.

En la primera fase, se ha realizado un estudio detallado del estado del arte en cuanto a la historia de las matemáticas, la docencia de la matemática aplicada, en particular, en lo referente a dinámica compleja e iteración de funciones reales, así como operadores definidos en espacios de Banach. De igual manera se han estudiado los resultados previos obtenidos en la asignatura Sistemas Dinámicos del Máster en Ingeniería Informática y Computación de la Universidad Internacional de La Rioja. Durante la segunda fase, se ha planteado una hipótesis de partida y los objetivos necesarios para poder demostrar dicha hipótesis. Para ello, se ha buscado una hipótesis que no estuviese desarrollada e implementada en la actualidad. Tras el estudio del estado del arte que se realizó en la fase anterior. Como sugerencia de hipótesis preliminar se planteó el seguimiento del aprendizaje utilizando algún prototipo de aplicación a los estudiantes del Máster. Una vez establecida la hipótesis y los objetivos iniciales, la tercera fase ha constituido en el estudio de la convergencia y de la dinámica del método de la Secante, de un método de orden seis libre de segundas derivadas y de una familia altamente eficiente. Además, se ha puesto en marcha la herramienta para favorecer la adquisición de competencias por parte de los estudiantes. Tras el rediseño del orden de impartición de la asignatura, se ha realizado una prueba

empírico analítica de tipo cuasi experimental, ante la dificultad para controlar la formación de los grupos, con un grupo de control no equivalente y uno experimental. Esta metodología ha estado acompañada de una estricta planificación de trabajo durante el tiempo de la investigación y durante la misma se han realizado continuas revisiones y modificaciones del proyecto inicial, buscando la viabilidad de su desarrollo. Parte de los resultados parciales de la investigación se han ido publicando y presentando durante el periodo de la investigación a escrutinio público en diferentes congresos nacionales e internacionales, así como en revistas científicas.

1.4. Organización de la tesis

A continuación, vamos a dar un breve resumen de cada una de las partes en las que se ha organizado la presente tesis.

1.4.1. Introducción a la investigación

En la primer parte, hemos hecho una pequeña introducción a la investigación en la que se ha presentado el problema a investigar, las dos líneas de trabajo que se han llevado a cabo en esta tesis, la metodología utilizada y la organización de la tesis.

1.4.2. Estado del arte y marco teórico

En esta segunda parte, se han puesto de manifiesto un breve resumen de la historia fundamental de las matemáticas, así como de la matemática aplicada y su enseñanza a lo largo de los años. En esta parte también se introducen los conceptos fundamentales de convergencia en espacios de Banach y dinámica de diferentes métodos iterativos, tanto en la recta real como en el plano complejo.

1.4.3. Desarrollo la investigación

En esta parte, se han puesto de manifiesto todos los resultados obtenidos durante el desarrollo de la tesis, tanto a nivel educativo, como en el caso de la investigación en matemática pura. En cada uno de los capítulos que conforman esta parte se ponen de manifiesto las investigaciones llevadas a cabo durante esta tesis.

1.4.4. Conclusiones y trabajo futuro

En esta parte casi final, se ponen de manifiesto las conclusiones extraídas durante la fase de desarrollo del trabajo y el trabajo que ha quedado como

pendiente o futuro.

1.4.5. Anexos

En esta última parte, se adjuntan los diferentes méritos que se han obtenido durante el tiempo que ha durado el desarrollo de la tesis, así como los trabajos que han sido enviados y están en proceso de revisión.

Parte II

Estado del arte y marco teórico

Capítulo 2

Breve historia de las matemáticas y su enseñanza

Resumen

En este capítulo, vamos a hablar brevemente sobre la historia de la matemática en la que se verán los principales hitos que cambiaron la historia de esta importante rama milenaria. Comenzaremos viendo los orígenes de las Matemáticas centrándonos en el Antiguo Egipto y Mesopotamia. Seguimos con la Antigüedad Clásica, Grecia, tratando los tres periodos más importantes, el jónico, el ateniense y el periodo alejandrino. Avanzando en el tiempo, llegamos a la época del Islam y la Europa Feudal, donde destacamos a Leonardo de Pisa “Fibonacci”. Durante el Renacimiento se empiezan a utilizar símbolos para designar cantidades constantes o variables, operaciones o relaciones. Apareciendo los símbolos más, menos, igual, raíz cuadrada, los decimales y la raya para las fracciones. Los matemáticos se afanan en buscar fórmulas que les permitan obtener las raíces de ecuaciones de quinto grado o mayor en función de los coeficientes de la ecuación, pues ya sabían resolver los de primer, segundo y tercer grado.

Se alcanza la revolución científica, con el desarrollo de la geometría analítica, el cálculo diferencial y el cálculo integral. Terminamos destacando algunos de los matemáticos más importantes de los siglos XVIII y XIX. El último apartado del capítulo está dedicado a explicar la docencia de la matemática aplicada.

2.1. Los orígenes de las Matemáticas. Antiguo Egipto y Mesopotamia

Comenzamos situando el nacimiento del Homo sapiens, que tuvo lugar en el 200.000 A.C. junto a la formación de las primeras sociedades primitivas. Pero no fue hasta el 10.000 A.C. aproximadamente cuando se dieron los primeros conocimientos matemáticos y astronómicos. Nacieron los calendarios, los primeros elementos de los sistemas numéricos y la geometría.

Cabe recordar que en un principio el concepto de número era “mucho” o “poco” y que posteriormente utilizaban los miembros del cuerpo para poder contar. Comenzaban siempre con los dedos de la mano derecha para contar de 1 hasta 5, el número 6 era la muñeca, 7 el codo, 8 el hombro, 9 el pecho, 10 el hombro izquierdo, y así sucesivamente hasta alcanzar el número 33. También había hombres primitivos que entre 3 llegaban a contar hasta 1000 sólo con los dedos de las manos.

Nace aquí la capacidad de abstracción, lo que es la cantidad de cosas concretas frente a los números abstractos. Y con ello la construcción aditiva de la sucesión de números naturales, eligiendo un número como base de un sistema numérico, que fue diferente dependiendo de la civilización.

En el 6000 A.C. viene la Revolución Agrícola, de la subsistencia basada en la caza y la recolección se pasa a cultivar las plantas (agricultura) y a la cría de ganado (ganadería). ¿Cuáles fueron las consecuencias sociales, económicas y políticas de este cambio? Las colectividades nómadas se convirtieron en sedentarias, aparece la propiedad privada y se establece el Estado. Surgen con ello diferentes necesidades en estas nuevas Sociedades, como puede ser que hay que saber orientarse en el espacio y en el tiempo (astronomía y cartografía). Hay que establecer fechas de siembra y recolección (cronología y calendario). También se hace necesario predecir acontecimientos como por ejemplo las crecidas e inundaciones de los ríos. Al igual que determinar las dimensiones de los campos (geometría). Se realizan cálculos para poder regar los campos y hacer canales. La Sociedad se organiza en grupos, los artesanos manufacturan vasijas para almacenar alimentos y los mercaderes y los comerciantes tienen que llevar un control de las cuentas, con lo que utilizan la aritmética. Con todo lo anterior, vemos que las matemáticas de estas primeras sociedades agrarias eran elementales y de magnitudes constantes, pero matemáticas al fin y al cabo.

La civilización egipcia, en un principio estaba formada por dos sociedades, el alto y el bajo Egipto. Por medio del faraón Menes, que gobernaba el alto Egipto, se unen en una única sociedad bajo su mando. En esta sociedad existe la figura de los escribas, profesionales de la administración del estado pertenecientes a la clase dominante y que tienen bastante poder. Para desempeñar sus funciones necesitan tener conocimientos matemáticos como puede ser la medición de tierras, cálculo de impuestos y contribuciones, cálculos de la capacidad de los depósitos de provisiones, proyectos de obras arquitectónicas y de ingeniería.

Uno de estos escribas, Ahmes, en 1650 A.C. fue autor del más famoso papiro conocido como el Papiro de Rhind, que está copiado de unos textos escritos entre 1991 – 1786 A.C. (ver [154]). Exhumado en Tebas en 1855, mide 33 centímetros de ancho por 5,48 metros de longitud. Su título es: *Directrices para obtener un conocimiento de todas las cosas, inherentes a todo lo que existe, conocimiento de todos los secretos*. Contiene 85 problemas de aritmética, estereometría, geometría, cálculo de pirámides y problemas prácticos. También de esa época encontramos el Papiro de Moscú, de autoría anónima escrito en torno a 1850 A.C. Mide sólo 7 centímetros de ancho por 5,48 metros de largo. Actualmente en el Museo Estatal de Artes Plásticas Moscú o Museo Pushkin de Moscú, contiene 25 problemas relacionados con la vida cotidiana. Por ejemplo en geometría aparece la noción de ángulo, descomposición de figuras geométricas, determinación de áreas de figuras poligonales y el volumen del tronco de pirámide. Para hallar el área del círculo, utilizaban una aproximación de $\pi = 3$, posteriormente aproximarían π con 3.0976.

El tercer documento relevante de esa época es el rollo de cuero de las matemáticas egipcias. Con ancho de 25 centímetros y 5,18 metros de largo, se conserva en el British Museum de Londres desde 1864. Se consiguió terminar de desarrollarlo en el año 1927. Contiene una colección de 26 sumas en forma de fracciones unitarias y arroja mucha luz sobre el aspecto mecánico de las matemáticas egipcias.

Las matemáticas mesopotámicas eran notablemente superiores a las egipcias. Nos remontamos al año 5000 A.C. con los sumerios, pueblo pacífico estructurado en sociedad de clases. Con la aparición de las ciudades-estado tuvo lugar el primer sistema de escritura. Fueron conquistados por los acadios, pueblo guerrero, sobre 2500 A.C.. Fue Hammurabi quien creó el gran imperio de Babilonia alrededor del 1800 A.C.. La evolución de las matemáticas babilónicas se prolongó hasta el nacimiento de Cristo. Se transmitieron a los persas, fenicios y griegos. Cabe destacar que en Mesopotamia existía un importante sistema de regadío artificial, por lo cual tuvieron que acometer problemas relativos a obras hidráulicas, por ejemplo construcción de canales, pozos y diques así como medición de campos. También tuvieron una gran actividad comercial, lo que les llevó a alcanzar un extraordinario nivel de las técnicas de cálculo, rasgos de proceder de un modo genuinamente algebraico. Fueron capaces de calcular el rendimiento de los terrenos, el número de trabajadores necesarios por jornada de trabajo, presupuestos, etcétera.

Mediante estiletos o cuñas escribían en tablillas de arcilla que posteriormente dejaban secar. Esta escritura cuneiforme es una de las más antiguas de las que tenemos constancia. Contamos actualmente con diferentes textos matemáticos, aunque aún quedan miles por estudiar. Algunos de los problemas que se han encontrado en las tablillas tratan sobre sucesiones de números, relaciones geométricas, tablas de multiplicar, tablas de cuadrados y cubos y relaciones numéricas con potencias. Un dato reseñable es, que a diferencia de los egipcios que

trabajaban en base 10, los mesopotámicos lo hacían en base 60, este sistema sexagesimal tenía la ventaja de que era divisible por 12 números, mientras que el decimal lo es sólo por 4, pero tener que idear sesenta símbolos diferentes sería un gran inconveniente, que resolvieron con un sistema posicional, es decir, el mismo símbolo varía de significado según su posición. En un principio para separar las posiciones dejan un espacio en blanco, lo que llevaba muchas veces a confundir los valores finales, por lo que hacia el año 400 A.C. introdujeron símbolos para separar las cifras. Expresaban también con símbolos otras cifras como podía ser el menos uno o fracciones.

En Geometría, se piensa que era superior la egipcia a la mesopotámica, puesto que éstos utilizaban el número 3 para aproximar π mientras que en las tablillas de Susa daban un valor de

$$\pi = \frac{25}{8} = 3.125.$$

Vamos a recordar algunos de los resultados que se consiguieron en Mesopotamia, aunque cabe destacar que no se paraban a demostrarlos. De los materiales que se han encontrado observamos que sabían calcular el área de un triángulo y de un trapecio, los volúmenes de prismas rectos y cilindros y una buena aproximación al volumen de un tronco de cono. Además eran conocedores de bastantes propiedades de las figuras geométricas, como pueden ser:

- El ángulo inscrito en un semicírculo es recto,
- Los lados correspondientes de dos triángulos rectángulos semejantes son proporcionales
- La perpendicular trazada desde el vértice de un triángulo isósceles a su base, la divide en dos partes iguales.

El resultado más simple que demostraron los mesopotámicos es el que podemos encontrar en la tablilla *plimpton 322*, donde encontramos ternas pitagóricas.

Respecto a álgebra, no existían signos, todo era hablado, lo único que se usaba en los escritos era *lal* para la resta y *tab* indistintamente para la suma o el igual. Eran hábiles calculadoras y poseían una depurada técnica en la manipulación de ecuaciones así como eran capaces de resolver ecuaciones lineales, cuadráticas, cúbicas y hasta de cuarto grado con soluciones reales positivas. Incluso se ha encontrado un sistema de 10 ecuaciones con 10 incógnitas resuelto.

Sobre el año 1000 – 800 A.C. escriben los problemas con enunciados y resolución paso a paso. El problema práctico concreto, va quedando en segundo lugar, la matemática comienza a independizarse. Dentro de los problemas los que más dominan son los de distribución, por ejemplo, cálculo de los tributos según los rendimientos específicos de las tierras, cálculo del pago del transporte de ladrillos según la distancia recorrida, repartos de herencias, etcétera.

2.2. La Antigüedad Clásica: Grecia

Entre 3000 – 1500 A.C. en la llamada civilización del bronce se asientan diferentes civilizaciones a orillas del mar Egeo, desde el Asia Menor hasta Sicilia (Creta). Esta civilización vive su esplendor en los siglos XVI y XV A.C. con la construcción de grandes palacios (Knosso) y la expansión marítima y comercial de la civilización minoica. Simultáneamente en el continente comienza la civilización micénica, las tribus van bajando y van asentándose en la zona del Asia Menor, hacia Grecia. La cultura minoica, autóctona, mediterránea y con tez morena, se junta con la cultura micénica, que son invasores o pastores nómadas, de tez rubia y que procediendo del Danubio imponen el dialecto helénico. Todo ello da lugar a la cultura egea. Entre 1200 – 1100 A.C. los dorios llegan al Peloponeso, y al ya conocer éstos el hierro, se acaba la edad de bronce y se extiende la industria del hierro. Los eolios y los jónicos son expulsados del continente pero mantienen la cultura micénica en las islas y en el litoral del Asia Menor. Estas circunstancias geográficas ocasionan que la cultura que heredan los griegos sea original, diversificada y abierta a las civilizaciones.

A partir de Homero, en el siglo VIII A.C. podemos hablar ya de los griegos, que son curiosos, inteligentes, intuitivos, artistas y sensibles. Desde el siglo VI A.C. se preocuparon no solo de conocer el cómo, sino sobre todo, de saber el por qué. Hay que recordar que durante los siglos VIII y VII A.C. eran sociedades esclavistas, lo que propiciaba que un buen número de personas, al no tener que trabajar por tener esclavos, podían dedicarse al arte, a la cultura, a la filosofía y a la ciencia. En estos siglos, en las ciudades-estado del Asia Menor, se produjo el nacimiento del pensamiento matemático, se pasó de unas matemáticas empíricas y expuestas en forma de recetarios a unas matemáticas expuestas sistemáticamente bajo una forma lógico-deductivas. Sorprendentemente las fuentes de que se disponen para estudiar las matemáticas griegas son menos fiables que las que se tienen para estudiar las matemáticas egipcias y babilónicas, ya que la calidad del papiro era mala y sufrieron múltiples incendios, lo que hace que hayan llegado hasta nosotros muy pocos textos originales. Realmente conocemos los textos matemáticos griegos a través de traducciones y comentarios de autores posteriores. Normalmente del griego al árabe o al hebreo, de aquí al latín y de éste al italiano, francés o castellano.

Los griegos tuvieron diferentes sistemas de numeración. En el siglo VI A.C. contaban con el sistema ático, que era aditivo, en base diez y acrofónico, es decir, que se asignaba a cada número la primera letra de su significado. Era utilizado para trabajar pesos, medidas y sumas monetarias, bien en dracmas bien en talentos. Posteriormente utilizaron el sistema jónico, en el que daban valor numérico a su alfabeto en mayúsculas. Realmente, no utilizaban sus cifras para realizar sus operaciones aritméticas, si no ábacos. Las matemáticas griegas se dividen en tres periodos, a saber:

1. Periodo jónico, desde finales del siglo VII A.C. a la mitad del siglo V A.C.

El nombre viene dado por la similitud de las matemáticas que hacían los jónicos.

2. Periodo ateniense, entre el año 450 y el año 330 A.C. En este periodo, la ciudad de Atenas tiene una gran importancia.
3. Periodo helenístico, desde mediados del siglo IV A.C. hasta el siglo II A.C. también llamado periodo alejandrino, debido a que la ciudad donde se desarrollaron fue Alejandría.

2.2.1. Periodo jónico

Centrándonos en el periodo jónico, a partir del siglo VII A.C. las ciudades-estado jónicas eran importantes centros económicos, políticos y culturales. Los filósofos de la naturaleza fueron más allá de la simple descripción y recopilación de los fenómenos naturales, intentaron comprenderlos y explicarlos. Todo ello sin recurrir a la mitología y a la magia, es por esto que se considera que la Ciencia nació en Grecia.

Los conocimientos matemáticos formaban parte de la filosofía, del amor por la sabiduría. Provenían de Mesopotamia y Egipto y adquirieron una estructura lógica alcanzándose una clara distinción conceptual entre definición, hipótesis y demostración. La matemática demostrativa (ver [179]) es una de las máximas creaciones de la ciencia griega. De esta época destaca Tales de Mileto, aproximadamente 624-548 A.C., hombre polifacético puesto que fue comerciante, político de la federación de ciudades jónicas, astrónomo, ingeniero, filósofo y matemático. Se considera que por primera vez enunció explícitamente y a veces incluso demostró, lo que hace que sea el primer matemático famoso. Trabajó el cálculo de la altura de las pirámides, distancias y paralelismos, inició la organización racional de las matemáticas. Tales supone el punto de inflexión entre la etapa precientífica y el periodo caracterizado por el saber crítico, objetivo y científico.

Uno de los matemáticos más laureados del periodo jónico es Pitágoras (aproximadamente 580/569 a.C – 500/475 A.C.). Natural de Samos, era hijo de un cortador de piedras preciosas y tuvo que abandonar su ciudad natal por miedo a los Persas, trasladándose a Mileto. Posteriormente viajó a Fenicia y a Egipto donde aprendió matemáticas y astronomía. Pasó 12 años de su vida en Babilonia y se asentó en Crotona, la actual Italia. Allí fundó una sociedad secreta mística-religiosa muy próxima a la aristocracia esclavista. En esta secta, el conocimiento y la propiedad eran bienes comunes, aunque la gente se dividía en novicios (escuchar y callar) e iniciados (hablar y opinar). Les llamaban esotéricos o matemáticos, porque matema significa ciencia, sabiduría. Dividían la matemática en aritmética, geometría, música y astronomía. En la secta estaba la figura del Maestro, era invisible, sólo escuchaban su voz ya que hablaba tras una cortina. Estaba prohibido vestir con prendas de algodón, ir calzado, comer carne, pescado, habas o beber vino. Simulaban milagros, que no eran más que trucos de

magia para ganar adeptos. A diferencia de otras sectas, el acceso a los misterios, a lo trascendente, solo era posible a través de la profundización en el universo de los números, todo es número, la armonía es divina y consiste en relaciones numéricas, quien llega a comprender la armonía en términos numéricos se vuelve divino e inmortal. Por todo lo anterior, la investigación sobre los números pasó a ser prioritaria dentro de la secta.

Se cree que abrazaron estas ideas debido al mito órfico, de Orfeo, que trata del poder especial otorgado a la música. Se encargaban de encontrar vínculos entre las armonías musicales y numéricas. A partir de esta filosofía realizaron importantes descubrimientos matemáticos, como la clasificación de los números en pares e impares, la definición de números amigos (si cada uno es igual a la suma de los divisores propios del otro), de los números perfectos (son iguales a la suma de sus divisores, incluyendo el 1 y excluyendo el propio número), de los números figurados (números de puntos configurados geoméricamente). En geometría destaca el Teorema de Pitágoras y se cree que conocían los poliedros regulares o sólidos cósmicos. También utilizaban métodos geométricos para resolver problemas algebraicos, por ejemplo, el método de las proporciones o el método de la aplicación de las áreas para la resolución de ecuaciones cuadráticas. Del estudio de la música, dieron las definiciones de media aritmética, media geométrica y media armónica.

Hipaso de Metaponto, no más allá del año 420 A.C. descubrió los números irracionales. Este descubrimiento contribuyó al desmembramiento de la secta pitagórica, junto con otras causas de tipo político, puesto que se demostraba la existencia de números inconmesurables. Dos segmentos son inconmesurables cuando la longitud de cada uno de ellos no es a la vez múltiplo entero de la de un tercer segmento que se elija como unidad. El descubrimiento de la existencia de estos segmentos supuso el fracaso de la *Arithmetica Universalis* o sistema dogmático universal, en el que todos los acontecimientos del Universo son concebibles y expresables mediante relaciones entre números enteros. Existen dudas sobre con qué problema de las matemáticas se produjo dicho descubrimiento. Se piensa que pudo ser bien con el problema de determinar la diagonal del cuadrado de lado unidad o bien que tuviese que ver con el pentagrama o pentáculo mágico. Éste fue utilizado como emblema de la secta pitagórica, como símbolo de salud, belleza y amor y se le atribuyó un significado mágico hasta bien entrada la edad media. La principal secuela de carácter general de la labor de los pitagóricos es la exigencia de la demostración y la secuela de carácter coyuntural es el predominio de la geometría.

Destaca también en este periodo Demócrito de Abdera (aproximadamente 460 A.C. – 370 A.C.), más conocido como filósofo materialista que como matemático, viajó mucho por Egipto, Persia y Babilonia y seguía una teoría atómica materialista, donde todos los fenómenos han de ser explicados en términos de átomos infinitamente pequeños e infinitamente variados que se mueven en el espacio vacío. La base de sus concepciones matemáticas se encuentra en su doctrina física

del atomismo, por lo que los problemas matemáticos que le interesaban eran en cierta medida, de orden infinitesimal. Concebía un sólido como una suma de un número infinito de capas planas paralelas unas a otras infinitamente delgadas e infinitamente próximas. Logró expresar por primera vez los volúmenes de la pirámide y del cono, aunque sin demostración.

Se obtuvieron muchos resultados a raíz de los llamados los tres problemas clásicos de las matemáticas griegas, que fueron:

- Construir una figura de forma dada de área igual a otra figura dada.
- Dado que cualquier ángulo podía ser bisecado era natural plantearse la trisección.
- Dado que la diagonal de un cuadrado es el lado de un cuadrado de área doble a la del lado, resulta lógico plantearse el correspondiente para el cubo

Los filósofos y matemáticos griegos consideraban las rectas y los círculos como las curvas perfectas, a partir de las cuales todas las construcciones tenían que ser posibles. Se impusieron estas restricciones arbitrarias para que la geometría se conservara simple, armoniosa y estéticamente atractiva. Durante el siglo XIX se demostró que era imposible resolver estos problemas con dichas restricciones. Un matemático muy relacionado con el primero de los problemas fue Hipócrates de Quios (aproximadamente 470 – 410 A.C.). Comerciante arruinado fue el geómetra más célebre del siglo V A.C. y autor de los primeros *Elementos*, perdidos. En ellos argumentaba los teoremas distinguiendo entre hipótesis, tesis y demostración.

2.2.2. Periodo ateniense

Entre el 491 A.C. y el 429 A.C. durante el gobierno de Pericles, Atenas vive un periodo de esplendor inmerso en una democracia esclavista. La guerra del Peloponeso hace que este modelo torne a una aristocracia esclavista y en el año 338 A.C. Filipo de Macedonia conquista Atenas. Extendemos este periodo ateniense hasta la llegada del imperio de Alejandro Magno o Alejandro III de Macedonia.

En este periodo ateniense a los sofistas, maestros de la sabiduría, se les remuneraba por transmitir sus saberes, existía realmente un interés social por adquirir conocimientos en gramática, retórica, dialéctica, elocuencia, moral, astronomía, filosofía y geometría. Hacia el año 430 A.C., Antifón afirmaba que al ser todos los polígonos cuadrables, lo sería también el círculo, es decir, la cuadratura del círculo, ya que veía que dado un polígono inscrito en un círculo, siempre era posible inscribir otro de doble número de lados. Brisón en el año 450 A.C. también hizo una consideración parecida a la de Antifón pero con polígonos circunscritos, y conjeturaba que el área del círculo estará siempre comprendida entre la de un polígono inscrito y otro circunscrito. También se abordó en esta época otro de los problemas clásicos, la trisección del ángulo, por ejemplo

por parte de Hippias de Elis, aproximadamente en el año 460 A.C., que consistía en dividir un ángulo en tres partes iguales solamente con regla y compás. En 1837, Pierre Wantzel (1814-1848) demostró que la solución era imposible con las restricciones impuestas por los griegos. Teodoro de Cirene (aproximadamente 465-398 A.C.) conformó y consolidó el álgebra geométrica. Se cree que demostró que $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$, $\sqrt{7}$, $\sqrt{8}$, $\sqrt{10}$, $\sqrt{11}$, $\sqrt{12}$, $\sqrt{13}$, $\sqrt{14}$, $\sqrt{15}$ y $\sqrt{17}$ son números irracionales.

No fue hasta Arquitas de Tarento (aproximadamente 428 – 365 A.C.), discípulo de Fildas, a su vez discípulo de Pitágoras, cuando se estableció una clasificación de las disciplinas matemáticas llamadas Quadrivium matemático (ver [179]):

- Aritmética, números en reposo
- Geometría, magnitudes en reposo
- Música, números en movimiento
- Astronomía, magnitudes en movimiento

Posteriormente estableció el Trivium:

- Gramática
- Retórica
- Dialéctica

Formando conjuntamente las siete artes liberales. Se le relaciona con el tercer problema clásico, la duplicación del cubo, también llamado problema délico porque en la ciudad de Del, pidieron al Dios Apolo que acabase con una plaga y éste pidió al oráculo que duplicasen el altar que tenía forma de cubo. Podría enunciarse de la siguiente manera: Calcular la arista de un cubo cuyo volumen sea el doble del de un cubo dado. El principal problema era cómo hallar solo con regla y compás la raíz cúbica de dos. Llegamos a Platón (aproximadamente 427-347 A.C.) que fundó la Academia hacia el año 387 A.C. sirviendo de guía de las matemáticas del siglo IV, lo que hoy en día entenderíamos como Universidad. Se tiene noticia de que escribió 40 diálogos de los que no menos de 30 son auténticos y en 121 pasajes repartidos en 25 diálogos se refiere a las matemáticas. Pero veamos las principales ideas de Platón:

- “Las idealizaciones perfectas de los objetos físicos son la auténtica realidad.”
- “El mundo de las ideas y de las relaciones entre ellas es permanente, inmutable, sin edad, universal, es el mundo de las verdades absolutas.”
- “El mundo físico es una representación imperfecta del mundo ideal y está sujeto a la degradación.”

Según Platón: “Cada objeto concreto, una silla, un triángulo sobre el papel, es tan sólo una copia mala, defectuosa, de una idea objetiva.”

¿Qué papel desempeñaban las matemáticas en la propia concepción filosófica y en la construcción del mundo?

Los entes matemáticos venían a ser los intermediarios entre el mundo de las ideas y el mundo de las cosas.

Las Matemáticas eran el puente entre el mundo físico y el mundo ideal. A la entrada de la Academia de Platón estaba esculpido: que no entre nadie aquí que no sepa matemáticas. La escuela platónica contribuyó decisivamente a la estructura deductiva de las matemáticas.

Se interesó por la demostración y por la metodología del razonamiento.

Los platónicos fueron los primeros en sistematizar las reglas de la demostración rigurosa.

“El razonamiento deductivo es el único método de demostración en matemáticas.”

Esta exigencia ha caracterizado a las matemáticas desde entonces, distinguiéndola de los restantes campos de conocimiento e investigación. ¿Por qué insistieron tanto o los griegos en la demostración deductiva?

- Les gustaba el razonamiento y la especulación.
- Estaban interesados en alcanzar la verdad.
- Rechazaban los asuntos prácticos.
- La ocupación de los hombres libres en el comercio debería ser castigada como un crimen. (Platón).
- Ningún ciudadano debería practicar ningún arte mecánico. (Aristóteles) .

Por otra parte también decían que la observación, la inducción y las generalizaciones sólo podrían proporcionar un conocimiento probable. Como consecuencia, tuvo lugar un escaso desarrollo de la ciencia experimental y de la mecánica durante el periodo clásico griego.

Entre las aportaciones de la escuela platónica destacan la mejora en las definiciones, nuevos teoremas de geometría plana y que se interesaron por la geometría del espacio.

Entre los matemáticos que pertenecieron a la academia cabe destacar a Teeteto, que estableció una teoría sobre los irracionales, clasificándolos y demostrando identidades, además de demostrar que sólo existen cinco polígonos regulares.

Eudoxo de Cnido estableció la Teoría de las Proporciones. Dos cantidades son inconmensurables cuando no tienen una razón tal como la de un número natural a otro. ¿Cómo se pueden comparar entonces las razones entre cantidades inconmensurables? Cada vez se iban descubriendo más cantidades inconmensurables.

¿Cómo se podían extender las demostraciones geométricas que se habían hecho para longitudes, áreas y volúmenes conmensurables a los inconmensurables? Eudoxo introdujo la idea de magnitud continua. No eran números, sino entidades, como segmentos rectilíneos, ángulos, áreas, etcétera. Que podían variar de una manera continua.

A las magnitudes no se les asignaba ningún valor cuantitativo, sino que se definía una razón de magnitudes del mismo tipo y, a partir de ella, una proporción o igualdad de razones, que cubría los casos de razones conmensurables e inconmensurables.

No se utilizaba número alguno para expresar tales razones. Los conceptos de razón y proporción estaban ligados a la geometría.

Eudoxo evitó los números irracionales. Es decir, evitó dar valores numéricos a las longitudes de segmentos, tamaños de ángulos, etc., así como a las razones de magnitudes. También resolvió el problema de la comparación de figuras curvilíneas y rectilíneas mediante el método de exhaustión.

Menecmo (aproximadamente 380 A.C. – 320 A.C.) miembro también de la academia platónica, fue el descubridor de las secciones cónicas. Para introducirlas utilizó tres tipos de conos según sus ángulos en el vértice.

Aristóteles de Estagira, también discípulo de Platón, fue maestro de Alejandro Magno. Fundó su propia escuela, el Liceo y era principalmente filósofo y biólogo. Aunque estaba al corriente de las matemáticas de la época no hizo ninguna contribución importante. Se cree que fue autor de una biografía sobre Pitágoras y que promovió las matemáticas o más bien la filosofía de las matemáticas. Fue maestro de Eudemo de Rodas, el primer historiador de las matemáticas.

2.2.3. El periodo alejandrino

Periodo que cubre aproximadamente desde 336 A.C. desde la subida al trono de Alejandro Magno hasta el año 30 d.C. con la conquista de Egipto por parte de los romanos. Fue una época de importantes cambios, de tipo sociopolítico, nuevos adelantos en técnicas de producción, en la milicia, en los trabajos de construcción y en la extensión de la actividad comercial. La ciencia y la cultura griega siguieron ocupando un importante lugar en los sistemas políticos sucesivos y mantuvieron su influencia hasta bien entrada la época del Imperio Romano. Las matemáticas se liberaron, al igual que otras ciencias, de la filosofía.

La ciudad de Alejandría fue fundada en el año 331 A.C. como capital del imperio ptolomaico. Gran puerto comercial, fue el centro científico-cultural más importante y más duradero del mundo griego. Ptolomeo fundó dos Instituciones: el Museo y la Biblioteca. El Museo fue centro público de enseñanza y de investigación en el campo de las ciencias exactas y naturales. Allí acudían científicos y estudiantes de diversos lugares. Estaba dotado de aulas, material e instrumental científico, salas de trabajo, observatorio astronómico, jardín botánico y parque zoológico entre otras. Sus principales departamentos fueron los de matemáti-

cas, astronomía y medicina. Por el contrario la Biblioteca estaba dedicada a las Humanidades y contaba con 400000 papiros y numerosos copistas.

Entre los más famosos matemáticos de esta época encontramos a Euclides. Se conoce muy poco de él, todo lo que se sabe es por Proclo, ni siquiera podemos asegurar que nació en Megara. Es probable que estudiase con discípulos de Platón (ver [188]). Abandonó a Atenas para mejorar sus condiciones de trabajo, vivió y enseñó en Alejandría. Estuvo vinculado al Museo y a su biblioteca aunque es posible que no desempeñase ningún puesto oficial.

En el año 300 A.C. escribió el libro de las matemáticas de mayor éxito en todos los tiempos, *los elementos*. La clave del éxito era su habilidad expositiva, siendo un libro de texto, no pretendió ser original, únicamente se ven en él algunas demostraciones. Aunque se supone por la dificultad de algunas de las demostraciones que era un libro destinado a estudiantes avanzados y no a principiantes. Contiene casi la totalidad de las matemáticas de la época, compuesto por trece libros, 465 proposiciones, 93 problemas y 372 Teoremas.

En los elementos se trabaja el método axiomático, que hoy en día sigue siendo el método general empleado en las matemáticas. Consiste en establecer inicialmente aquellas propiedades o verdades que deben admitirse sin demostración para deducir a partir de ellas, sin otro recurso que la lógica, todo el conjunto de las proposiciones del sistema. Estas propiedades básicas son los llamados axiomas que Euclides designó con los nombres de “*postulados*” o “*nociones comunes*” (ver [188]).

Euclides no utilizaba las definiciones como argumentos deductivos, sino que las establecía para convencer al lector de que los axiomas son aplicables a tales conceptos. Basó su sistema en trece axiomas, cinco postulados y ocho nociones comunes. El libro I por ejemplo comienza con 23 definiciones y cinco postulados. El libro II consta de 14 proposiciones. Existe la creencia que los contenidos de los libros III y IV provienen en su mayor parte de Hipócrates de Quíos. El libro III contiene 37 proposiciones y definiciones relativas a la geometría de los círculos. El libro IV tiene 16 proposiciones y trata de la inscripción y circunscripción de polígonos regulares como triángulos, cuadrados, pentágonos y hexágonos a una circunferencia. Observar que no se hace alusión a los polígonos de 7, 9, 11, 13 lados, que no se pueden construir con regla y compás. El libro V se compone de 18 definiciones y 25 proposiciones sobre la teoría de las proporciones. Adquiere un nivel elevado comparado con los anteriores y se trata de una de las partes más sólidamente establecidas de toda la obra. En el libro VI se aplica la teoría de las proporciones a la geometría plana y se tratan las figuras semejantes. Compuesto por 4 definiciones y 33 proposiciones, entre las que se encuentran teoremas referidos a razones de triángulos semejantes, paralelogramos y otros polígonos. Los libros VII, VIII y IX tratan de las propiedades de los números enteros y de las razones entre números enteros. Son los únicos libros dedicados a la aritmética como tal. El libro VII consta de 32 definiciones y 39 proposiciones, el libro VIII contiene 27 proposiciones y trata las progresiones geométricas. El

libro IX comprende 36 proposiciones y en él se reflejan teoremas relativos a los números cuadrados, cúbicos, planos y sólidos y a las progresiones geométricas. El libro X es el más extenso, compuesto por 4 definiciones y 115 proposiciones sobre números irracionales mientras que el libro XI está formado por 28 definiciones y 39 proposiciones que trabajan cuerpo sólidos, inclinación de planos, planos paralelos, figuras sólidas semejantes, ángulo sólido, pirámide, prisma, cilindro, esfera, cono, octaedro, icosaedro, dodecaedro y el cubo. El libro XII está compuesto por 18 proposiciones sobre áreas y volúmenes y trabaja el método de exhaustión. En el libro XIII trata las propiedades de los cinco poliedros regulares y está compuesto por 18 proposiciones.

Además de los elementos, Euclides cuenta con otras obras como puede ser *De domena* en la que estudia qué partes de una figura están determinadas cuando otras han sido dadas previamente. Trabaja la descomposición de figuras y también estudiará cómo dividir según una relación dada el área de figuras limitadas por segmentos rectilíneos.

Una de las grandes figuras de la historia de la ciencia fue Arquímedes de Siracusa, autor de monografías sobre aritmética, geometría, astronomía, estática e hidroestática. Sus obras se caracterizaban por su estructura lógica, el rigor de sus demostraciones, la originalidad y profundidad de su pensamiento y por un enorme dominio del cálculo. Tuvo una extraordinaria influencia durante el Renacimiento y la Edad Moderna. Hijo de un astrónomo, de buena familia, estudió en Alejandría y dedicó toda su vida a la investigación científica. Plutarco dice que: “se olvidaba de comer y beber, a menudo se le debía arrastrar por la fuerza a los baños y ungüentos. Estaba totalmente cautivado por felicísimas sensaciones y realmente poseído por su musa matemática”. Se han conservado muchos de sus trabajos, no en dorio sino a través de traducciones al árabe y al latín. Sus obras más importantes son: sobre el equilibrio de los planos, sobre los cuerpos flotantes, sobre la esfera y el cilindro, sobre conoides y esferoides, cuadratura de la parábola y sobre el método de los teoremas deducibles mecánicamente, más conocido como *El Método*.

El Método de Arquímedes consistía en utilizar ideas procedentes de la mecánica para obtener teoremas matemáticos correctos. Además, construyó una especie de planetario movido hidráulicamente. Con Arquímedes las matemáticas griegas llegaron a su punto máximo. Perfeccionó y complementó las matemáticas ya existentes desvinculándolas casi totalmente de la filosofía.

Apolonio de Perga (aproximadamente 262-190 A.C.) fue otro de los más ilustres matemáticos de la época. Escribió cónicas, que consta de ocho libros en los que aportó nociones tan profundas a la geometría de las secciones cónicas, que las mismas no fueron obtenidas de nuevo hasta el siglo XVI ni superadas hasta el siglo XVII (ver [131]).

Desde la conquista de Egipto por los romanos en el año 30 D.C. hasta el final de la Antigüedad existieron otros matemáticos que realizaron grandes aportaciones como pueden ser Herón de Alejandría, Ptolomeo, Diofanto de Alejandría,

Pappus de Alejandría, Teón de Alejandría e Hipatia de Alejandría, primera mujer matemática conocida, que murió apedreada por impartir sus enseñanzas matemáticas.

2.3. Oriente y Occidente: El Islam y la Europa Feudal

El desmoronamiento de la parte occidental del Imperio Romano y la creciente influencia del cristianismo en todo el territorio dominado por Roma suponen un serio retroceso para las matemáticas. El paso de la sociedad esclavista al feudalismo no tiene consecuencias uniformes en el campo de las matemáticas. Mientras que en Europa las mismas decaen hasta niveles bajísimos, en Asia se mantienen e incluso se desarrollan apreciablemente en algunos lugares. Surgen nuevos centros de producción matemática como son China e India. Se vuelven a plantear los problemas prácticos que caracterizaron las matemáticas egipcias y mesopotámicas. Los conocimientos matemáticos se transmiten entre China y la India, de la India a Siria, desde Siria a Arabia, de Arabia al norte de África, del norte de África a Sicilia y a la Península Ibérica, y de aquí al resto de Europa gracias a las relaciones comerciales que se concretan en los viajes de las grandes caravanas.

Puede decirse que al período de la máxima decadencia de las matemáticas en Europa le corresponde la etapa de mayor esplendor de esta ciencia en la India. En efecto, surgen entre los siglos V y XV en la India, algunos matemáticos importantes que contribuyen al desarrollo del Álgebra, de la Aritmética y de la Geometría.

La máxima aportación de los matemáticos indios al saber universal es el desarrollo del sistema decimal posicional que, desde el siglo VII, se irradia desde la India, primero al mundo islámico y luego, sucesivamente, a todo el planeta. El sistema de numeración indoarábigo constituye uno de los grandes bienes del patrimonio cultural de la humanidad. Gracias a este sistema, se democratiza el conocimiento de los números y de las operaciones elementales hasta convertirse en un atributo mínimo de nivel de instrucción de cualquier persona. Los conocimientos matemáticos de los indios se transmiten a las capas intelectuales del vasto imperio creado siguiendo los impulsos derivados de la religión predicada por Mahoma, que era mercader en su juventud. Entre los siglos VII y IX dicho imperio se extiende desde la Península Ibérica hasta el Asia Central.

En un principio, en pleno proceso de expansión territorial del Islam, los conquistadores se ocupan poco de la ciencia en general y de las matemáticas en particular. Sin embargo, los musulmanes tienen dos preceptos que cumplir emanados del Corán que les obligaba a no separarse drásticamente de las preocupaciones científicas. Uno de ellos es el complicado sistema de reparto de herencias y el otro que todas las mezquitas (ver [197]) tienen que tener el mihrab orientado hacia

la Meca. Es decir, las herencias exigen hacer muchas cuentas (Aritmética) y los templos a saber orientarse (Astronomía y Geografía). A mediados del siglo VIII los grandes califas abásidas fomentan la cultura, comenzando a traducir al árabe algunas de las obras científicas de la antigüedad clásica. Por ejemplo, el califa al MAMUN (809-833) erige una “casa de la sabiduría” en Bagdad para promover e incrementar el volumen de las traducciones científicas. Se vierten al idioma de la que sería durante varios siglos la lengua de los intelectuales y sabios, entre otros libros, los Elementos de Euclides y el Almagesto de Ptolomeo.

No hay que pensar que los matemáticos de esta época son sólo de la península arábiga. En su mayor parte provienen de Siria, Persia, Uzbekistán, Mesopotamia, etc. Más todos ellos están unidos por una lengua común, el árabe.

Hoy en día todavía desconocemos los nombres de la gran mayoría de los matemáticos árabes de la Edad Media. Los orígenes y las causas de este desconocimiento hay que buscarlos en los procesos de asentamiento y consolidación de los estados absolutos en torno a la religión católica. La incomodidad derivada del hecho de tener que reconocer que los infieles tengan algo bueno, hace que se silencie que durante muchos siglos los árabes fueron más sabios, más cultos, más piadosos y más tolerantes que los cristianos. Además de salvar una parte importante del legado matemático de la antigüedad clásica, los árabes no sólo desarrollan saberes ya conocidos, sino que potencian decisivamente algunas nuevas ramas del árbol de las matemáticas. Por su relación con la astronomía, dedican muchos esfuerzos a la trigonometría, dando a conocer relaciones hoy básicas como, por ejemplo, los teoremas del seno y del coseno. También contribuyen decisivamente a la creación y el desarrollo del álgebra, palabra que proviene precisamente del árabe. Veamos los matemáticos más destacados de esta época.

Al-Khwarizmi (Bagdad, aproximadamente 780 - 850) en su tiempo fue tan popular como Euclides. Escribió sobre aritmética, haciendo difusión del sistema indoarábigo de numeración, sobre Algoritmia, dando un método descriptor de procesos que pueden tratarse mecánicamente por medio de máquinas y sobre operaciones ligadas a la resolución de ecuaciones de primer y segundo grado. Del término “Al-jabr” deriva la palabra álgebra que en árabe significa restaurar. Por lo tanto la operación de al-jabr (complementación o reemplazamiento) consiste en desembarazarse de los términos a sustraer en un miembro de la ecuación por la adición o sustracción de términos iguales en los dos miembros. La operación de al-muqabala (Puesta en posición, balanceamiento) Consiste en la reducción de términos iguales en los dos miembros más la reducción del coeficiente del término de segundo grado a la unidad. Al-khwarizmi y sus discípulos distinguieron seis tipos canónicos de ecuaciones de segundo grado.

Thabit ibn Qurra (Turquía 826 - Iraq 901) revisó la tercera versión árabe de los elementos de Euclides y dedicó su atención a los números amigos. El historiador matemático Hogendijk, le atribuye el descubrimiento de que los números 17.296 y 18.416 son amigos. Da una regla para determinar números perfectos y traduce las cónicas de Apolonio.

Al-Biruni (Uzbekistán 973 – Afganistán 1048) trabaja la suma de series, la combinatoria, la regla de tres, la resolución de ecuaciones algebraicas y las secciones cónicas, así como la resolución de triángulos esféricos.

Omar Khayyam (Irán 1048 – Irán 1131) fue director del Observatorio astronómico de Isfahán y abordó el estudio de las ecuaciones de tercer grado por procedimientos geométricos, utilizando las secciones cónicas. También propuso una reforma del calendario, pues quería añadir 5 horas, 49 minutos y 5,75 segundos al año de 365 días.

Al-Kashi (Irán 1380 – Samarkanda 1429) calcula el número π con 17 decimales exactos y establece un algoritmo para hallar las raíces de un polinomio de grado n , se trata de un caso especial del método dado por Ruffini siglos más tarde. Observó el eclipse de luna que tuvo lugar el día 2 de junio de 1406 y fue autor de un “Compendio de la Ciencia del Astronomía” (1410 – 1411).

En la Europa feudal, mientras las cortes musulmanas tienden hacia la decadencia, los reinos cristianos de Europa occidental se ejercitan en el arte de la guerra con objeto de intentar dominar el mundo. Sin embargo, también llega a Europa la necesidad de aumentar el nivel cultural, las nuevas ideas, y por supuesto las matemáticas, penetran por Sicilia y principalmente por la península ibérica. Empieza entonces un periodo de esplendor donde las matemáticas de Toledo, Madrid, Córdoba, Granada o Barcelona entre otras ciudades brillan con luz propia. Sirva para ilustrar esta situación la figura del judío barcelonés Abrahan Bar Hiyya (Savasorda 1065-70 – 1136) quien escribe en hebreo una “Geometría Práctica” que abrirá el camino a la utilización de esta lengua como vehículo de transmisión de conocimientos científicos a través de las traducciones del árabe al hebreo y de éste al latín. Con los libros de los sabios de Oriente llegan a la península ibérica el sistema de numeración indoarábigo y el álgebra de Al-Khwarizmi, transmitiéndose al resto de Europa gracias a la labor desempeñada por las Escuelas de Traductores. En efecto, en las Escuelas de Traductores de Toledo, Tarazona o de otros lugares, se vierten al latín las obras de los sabios griegos y árabes.

Gerardo de Cremona (ver [176]) es el traductor que más se distingue en estas tareas, aunque por supuesto hay muchos más, como Roberto de Chester, Platón de Tívoli, Juan de Sevilla o Gerberto de Aurillac, que fue elegido Papa en el año 999 con el nombre de Silvestre II.

En los siglos XII y XIII se crea en Europa una cierta infraestructura científica con la creación, bajo el auspicio de la iglesia católica, de las universidades de Bolonia, Oxford, París y Salamanca entre otras. En las mismas se explican las denominadas siete artes liberales divididas en dos grandes grupos. El Trivium, que comprende las tres disciplinas humanísticas: gramática, retórica y dialéctica, y el Quadrivium que se compone de las materias científicas: aritmética, geometría, música y astronomía. La difusión de las matemáticas griegas e islámicas a los países del occidente cristiano a partir del siglo XIII ha sido considerada por algunos autores como la etapa del primer renacimiento matemático.

Entre los matemáticos más famosos de la época, cabe destacar la figura de Leonardo de Pisa “Fibonacci”(1170 - 1250). Hijo de un mercader de Pisa, debido a los negocios del padre viaja a los países árabes, donde aprende las nuevas herramientas matemáticas que resultaban muy útiles para la buena marcha de los negocios familiares. A los 22 años escribe el libro titulado “Liber Abaci” o libro del ábaco. Se trata en realidad de una especie de manual de aritmética mercantil con muchos ejercicios sobre supuestos prácticos expuestos de forma retórica, es decir, sin fórmulas. Desde el punto de vista actual su lectura es muy difícil. No solamente con tiene problemas de la vida cotidiana, sino que también explica algunos elementos de la nueva metodología algorítmica, además de una defensa muy férrea del nuevo sistema de numeración indoarábigo. Puede afirmarse que el Liber Abaci contribuyó extraordinariamente a la difusión europea de los nuevos signos indios que representan los dígitos del 1 al 9, más un nuevo número, el 0, al que los árabes habían bautizado con la palabra de zephirum, del que proceden los conceptos y las palabras cero y cifra.

Las “justas” matemáticas eran competiciones entre matemáticos destinadas a la diversión de los Reyes y grandes Señores que eran sus mecenas. En estos torneos, los participantes recibían o bien preguntas del auditorio o bien de otro artista de este peculiar circo matemático. Fibonacci recopila algunos de los problemas propuestos en dichas justas en su obra Liber Quadratorum (1225) o libro de los números cuadrados, muchos de ellos recuerdan a los de Diofanto. Por todo lo anterior, Fibonacci es el matemático más destacado de la Europa Cristiana Medieval.

2.4. El renacimiento: álgebra, métodos de cálculo y trigonometría

En 1453 Constantinopla cae en poder de los turcos. Muchos autores se inclinan por considerar esta fecha como el inicio de un periodo que convulsionó la historia del mundo de una manera decisiva. El universo material, burguesía, y el espiritual, protestantismo, se transforman poderosamente. Las ciencias de la Naturaleza se colocan en el ojo del huracán ideológico, lo que supone que se les dedique una atención y estudio mucho mayor que las que se habían tenido hasta entonces. A este periodo se le conoce como El Renacimiento y al cambio de la situación que va a producirse en la Ciencia en consecuencia, se le conoce con el nombre de la Revolución Científica. La denominación “Renacimiento” proviene del hecho de la difusión de las obras de la antigüedad clásica, tanto de las que ya se conocían en Occidente, como de las que llegaron en los equipajes de los emigrados del desaparecido Imperio Bizantino. La lectura de estas obras con una nueva mentalidad contribuye a ampliar considerablemente los conocimientos que ya se tenían de la cultura y del pensamiento filosófico y científico griego y helénico, lo que les hizo renacer. Lo verdaderamente importante no es solamente

este renacer de las raíces paganas de la cultura europea, sino la tormenta de ideas que se desata y a la que contribuyen muchos factores. Aunque la batalla científica es muy importante (ver [70]), el frente principal de la confrontación se da en el terreno religioso. En efecto, el mundo establecido a lo largo de la Edad Media y santificado por la Iglesia Católica Romana era un mundo en el que “todo estaba atado y bien atado”. Desde que Tomás de Aquino cristianiza a Aristóteles se dispone de una teoría que aporta grandes dosis de racionalidad al cuerpo de doctrina religioso, basada en la fe. En este mundo establecido y cristianizado todo es fijo e inmutable. La Tierra inmóvil ocupa el centro del Universo, a su alrededor gira la luna, el sol y los planetas. La esfera del Universo la forman las estrellas fijas.

El Papa es el soberano espiritual y en cada estado europeo los reyes o príncipes son dueños absolutos de vidas y haciendas. Oficialmente nada queda por saber, todo está ya escrito en los libros. Sin embargo, las situaciones cuanto más rígidas y herméticas son, más endebles resultan. Basta con que se muevan algunas de las piezas que afectan al orden establecido para que todo el edificio se desmorone. Durante el renacimiento se remueven muchas de estas piezas, a causa de la corriente artística y cultural que viene de la Grecia Clásica. En 1492 se descubre un Nuevo Mundo. Si resulta que semejante continente desconocido hasta entonces está en nuestro planeta, surgen preguntas sobre lo que podrá encontrarse por ejemplo en el cielo. Desde mediados de siglo, se utiliza en Alemania una máquina que va a conmover los cimientos de la cerrazón y el hermetismo, la imprenta. Se pone en circulación a bajo coste el arma ideológica de mayor trascendencia en la historia de la humanidad hasta la aparición de la televisión, los libros. Las consecuencias de la tormenta de ideas, tienen una primera expresión en la división religiosa con la negativa de los protestantes a reconocer la autoridad espiritual del Papa. Esta situación conduce a las guerras de Religión y la defensa a sangre y fuego de las posiciones de las respectivas Iglesias. Científicos y matemáticos se ven inmersos en la crudeza de esta batalla ideológica. Alguno de ellos, como Giordano Bruno, son quemados vivos por la Inquisición, por razones de herejía científica. Hoy en día, nos resulta difícil entender el grado de apasionamiento que las controversias de carácter científico desataron, a lo largo del Renacimiento, en territorios tan aparentemente alejados del mundanal ruido de las contingencias políticas o religiosas como las matemáticas con la astronomía.

Pero no resulta difícil imaginar cómo al hilo de las disputas de tipo científico que ponían en entredicho aspectos ideológicos sacralizados por la Iglesia y asumidos como verdaderos, podría irse deshaciendo el ovillo de la injusticia subyacente en el ordenamiento social imperante. En efecto, si se demostraba que las Iglesias y los poderes fácticos defendían un sistema del mundo equivocado, se podrían poner en duda otros aspectos religiosos. Paradójicamente quienes llevaron a cabo la demostración de que las Iglesias estaban equivocadas en su interpretación de los hechos fueron hombres como Copérnico, Kepler o Newton que, en general, eran profundamente religiosos.

Las Matemáticas (ver [192]), también experimentan cambios decisivos en cantidad y calidad durante el Renacimiento. La creciente apreciación general de las matemáticas tiene mucho que ver con la relación cada vez más directa que mantienen las mismas con el desarrollo social. En el siglo XV, el orden social basado en la herencia, comienza a resquebrajarse gracias a la aparición de una nueva clase social, la burguesía. A partir de este siglo, las mercancías circulan más, el comercio se expande, crece el tráfico marítimo y mejora la construcción naval. La producción de productos manufacturados también crece y con ella las necesidades de carácter más o menos burocráticas. La sociedad precisa de hombres capaces de llevar la contabilidad, saber la capacidad de un tonel de vino o de un aljibe de agua, calcular un muro defensivo o diseñar una presa o un canal. El desarrollo de las matemáticas y la necesidad social de que existan más personas que sepan contar y medir tiene mucho que ver con esas cuestiones. Es durante esta época, cuando se produce un notable cambio en lo que se refiere a la consideración social de la gente que sabe matemáticas. Aun así, las universidades no atienden adecuadamente esta demanda social, por lo que se crean centros especiales de enseñanza para formar a los profesionales que precisan tener conocimientos científicos y matemáticos para el empeño de sus oficios. Es el caso de los pilotos de altura o de los artilleros. Las matemáticas que se enseñan en el siglo XV y en buena parte del siglo XVI se refieren fundamentalmente a los libros de geometría plana y aritmética de los elementos de Euclides, la exposición del sistema del mundo del *Almagesto* de Ptolomeo y de los contenidos matemáticos de Platón y Aristóteles. Las aplicaciones son todavía muy escasas, limitándose al cálculo con el ábaco y el *Computus*, cálculo de festividades móviles de carácter religioso. En matemáticas la influencia de la antigüedad clásica tiene como protagonista a Arquímedes cuyos resultados y métodos se contradicen a los de Aristóteles. Arquímedes es un valor en alza, mientras que Aristóteles es la autoridad a criticar. Esta tendencia se incrementa cuando en 1544 aparece impresa la primera edición de las obras completas de Arquímedes conocidas hasta entonces. Además de la influencia de Arquímedes hay que señalar que son las necesidades prácticas las que impulsan el florecimiento de las matemáticas.

Es en este periodo, cuando aparecen en Europa, los primeros profesionales seculares. Se trata de los llamados calculistas o “Maestros de Cálculo” que se ganan la vida haciendo cuentas. Michael Stifel (1487-1567), Adam Riese (1492-1559) o el mismísimo Kepler (1571-1630) son los maestros del cálculo más famosos de esta época.

Como consecuencia del aumento de la estructura administrativa de los estados, ejércitos y corporaciones y de la creciente actividad comercial, se publican un gran número de libros de aritmética escritos en las lenguas vernáculas que alcanzan una gran difusión. Contienen los elementos claves para entender las reglas del cálculo. La aritmética está relacionada con el álgebra, empieza a adquirir una nueva forma. Recordamos que los autores árabes y los cristianos medievales que les siguieron habían escrito el álgebra en forma retórica. Ya en el siglo XVI se empieza a utilizar símbolos para designar cantidades constantes o variables,

operaciones o relaciones. Al más importante de estos conceptos, la incógnita de los problemas, se le llama “la cosa”. Los maestros que efectúan cálculos utilizando símbolos o abreviaturas reciben el nombre de “Cosistas”. De esta época proceden los primeros intentos sistemáticos de escribir con formas abreviadas o simbólicas. El trabajo de los cosistas y calculistas italianos, alemanes u holandeses va preparando el terreno para la obra de François Viète, la “Vieta” que significa un punto de inflexión en el proceso de transformación de las matemáticas. Entre la obra de Luca Pacioli (1445-1514) y la Vieta se puede datar la aparición de los símbolos más, menos, igual, raíz cuadrada, decimales, la raya para fracción y algunos más. No todas las anotaciones surgen a la vez ni son automáticamente aceptadas por la comunidad científica. Cuando se utilizan estos símbolos los problemas se reducen a ecuaciones. Desde la Edad Media se sabía resolver ecuaciones de segundo grado y algunas de tercer grado. Pero así como había una fórmula para las ecuaciones de segundo grado que permitía hallar las raíces de la ecuación en función de sus coeficientes, no ocurría lo mismo para las ecuaciones de grado superior o igual al tercero. El siglo XVI va a ser el escenario temporal en el que se va a producir el hallazgo de la fórmula para resolver las ecuaciones de tercer grado.

De esta época podemos destacar a Tartaglia cuyo apodo viene por el hecho de ser tartamudo a causa de haber recibido un sablazo en la mandíbula durante la toma de Brescia en 1512. A pesar de ser de origen muy humilde, e hijo de viuda, consigue hacerse calculista público, gracias a su esfuerzo y a ser muy inteligente. Este título le permitía ganarse la vida haciendo cuentas para ingenieros, generales, comerciantes o para quien lo necesitara, siempre cobrando por ello. Después de vivir en Milán, Piacenza y Venecia, muere en esta última ciudad.

También cabe destacar a Cardano, que fue hijo natural de un jurisconsulto Milanés. Se graduó en medicina, aunque no tiene mucha fortuna en el ejercicio de esta profesión, debido en parte al rechazo corporativo por ser hijo ilegítimo y también por ser bastante pendenciero. A la edad de 60 años, su hijo es ejecutado por matar a su esposa. A los 70 años lo meten en la cárcel por deudas de juego y cuando sale es acusado de herejía. Algunos de sus biógrafos afirman que se deja morir de inanición para no sobrevivir al año 1576 que él mismo había pronosticado como el año de su muerte.

A comienzos del siglo XVI, A.M. Fior, un discípulo de Cardano y de S. del Ferro, compitió con Tartaglia en una de las justas matemáticas que se disputaban en aquella época. Le proponen 30 problemas cuyo planteamiento le lleva a Tartaglia a tener que resolver 30 ecuaciones de tercer grado que consigue resolver en el tiempo previsto. Ante semejante alarde, Cardano le pide a Tartaglia que le dé la fórmula que ha descubierto. Después de muchas vacilaciones y tras hacerle jurar que no se la contaría a nadie Tartaglia le da la fórmula. Sin embargo, Cardano lo público en el año 1545 en “Ars Magna” citándolo, así como a la del Ferro y atribuyéndoles el hallazgo. Tartaglia se enfada muchísimo y publica un año más tarde “Diferentes problemas e invenciones” en el que además de criticar a Cardano, se extiende en la historia de la resolución de la ecuación de tercer

grado y otros temas matemáticos.

A finales del siglo XVI los matemáticos saben resolver las ecuaciones de primer, segundo y tercer grado y haciendo cambios adecuados, bastantes de las de cuarto grado. Los algebristas se afanan en buscar fórmulas que les permitan obtener las raíces de ecuaciones de quinto grado o mayor en función de los coeficientes de la ecuación.

Albert Girard (1590 – 1633) se atreve a sostener en 1629 que una ecuación puede tener tantas raíces como indica su grado. A partir de este momento son muchos los que acometen la tarea de intentar hallar las raíces de una ecuación por medio de fórmulas. El final de esta historia lo van a protagonizar dos siglos más tarde Gauss, Abel y Galois.

Otro punto de interés surge para los matemáticos durante el renacimiento en relación con el sistema del mundo. N. Copérnico (1473 – 1543) (ver [49]) comprueba que cada vez resulta más complicado ajustar las observaciones astronómicas al rígido sistema de epiciclos y excéntricas del sistema Geocéntrico de Ptolomeo. En 1543 en su obra “De revolutionibus orbium coelestium” formula una idea radical: el centro del Universo no está ocupado por la Tierra inmóvil sino por el Sol. Aunque hoy en día sabemos que el modelo de Copérnico también era incorrecto, tuvo dos virtudes, simplificó los cálculos y dio a conocer otro modelo al santificado por la Iglesia y las Universidades.

Las matemáticas juegan en este proceso un importante papel al suministrar el material para acometer la tarea que sería culminada por Isaac Newton (1643 – 1727). Los árabes habían desarrollado una rama de las matemáticas, de raíces indias, dedicada al estudio de los triángulos, la trigonometría. Dicho estudio resulta fundamental tanto para la elaboración e interpretación de horóscopos y otras cuestiones astrológicas como para la rigurosa composición de calendarios o para la resolución de otros problemas de carácter geográfico y topográfico. Resulta prácticamente imposible separar los trabajos astronómicos de los trigonométricos en este periodo de la Historia de la Ciencia. A lo largo de la Edad Media, numerosos hombres de ciencias árabes y europeos aportaron muchas observaciones y resultados trigonométricos que contribuyeron a la sustitución del sistema de cuerdas de Ptolomeo por el sistema de conceptos trigonométricos.

El matemático más importante del siglo XV fue Johanes Muller o Regiomontano (1436 - 1476) el cual sintetiza en su persona conocimientos científicos y humanísticos. A partir de su obra “Cinco libros sobre toda clase de triángulos” escrita en 1462 y 1463 pero publicada en 1533, la trigonometría se convierte en un cuerpo de doctrina cerrado y casi completo, mientras que a la astronomía todavía le quedará camino que recorrer. Es decir, en la primera mitad del siglo XVI se establecen casi de forma definitiva las estructuras teórico-prácticas de la trigonometría. Tras la publicación de dicha obra se empiezan a elaborar muchas tablas trigonométrica-astronómicas entre las que es justo destacar las de Copérnico, Vieta, Tycho Brahe y Kepler.

Con la teoría y la práctica trigonométrica se pueden obtener muchísimos resul-

tados. Sin embargo, la utilización de estas técnicas presenta un grave problema, los cálculos son muy costosos. Es necesario inventar un elemento que facilite los cálculos. El 1585 el ingenioso ingeniero flamenco Simón Stevin (1548 – 1620) se muestra partidario de utilizar en los cálculos números fraccionarios con parte entera y parte fraccionaria separadas pero seguidas. Es decir, clama por el empleo de los números decimales. El siguiente paso es de tipo operativo y computacional, los logaritmos. Los tres protagonistas de la invención de los logaritmos son el relojero suizo Joost Bürgi (1552 - 1632), el noble escocés John Napier (1550 – 1617) más conocido como Neper y el profesor inglés Henry Briggs (1561 -1631). A comienzos del siglo XVII realizan las primeras tablas de logaritmos. El cálculo logarítmico, que reduce los productos a sumas y las divisiones a restas, así como potencias y raíces a productos supone una extraordinaria ventaja a la hora de acometer los larguísimos cálculos inherentes a la trigonometría y a la astronomía. En su pretensión de mejorar las tablas existentes y facilitar su manejo, Bürgi calcula hacia 1600 unas “tablas de progresiones aritméticas y geométricas” que apenas tienen difusión, ya que se publicaron casi 20 años después de haber sido compuestas, cuando Europa Central andaba metida de nuevo en una guerra de religión, la “Guerra de los 30 años”. En cambio, sí llegan a difundirse los trabajos de Neper, sobre todo su obra “Mirifici logarithmorum canonis descriptio” (1614). En la misma aparece escrita por primera vez la palabra “logaritmo”, que literalmente significa “número para el cálculo”. La base del sistema utilizado por Neper es proporcional a $1/e$, por lo que a los logaritmos en base e se les llama también neperianos. Un año después de la publicación del trabajo de Neper, Briggs conviene con el noble escocés calcular una tabla de logaritmos en la que $\log(1) = 0$ y $\log(10) = 1$. Así aparecen los logaritmos decimales cuya primera tabla, con catorce cifras decimales, ve la luz en 1617. En el mismo siglo XVII la comunidad científica conoce y maneja dichas tablas. Puede decirse que con los instrumentos algorítmicos de cálculo, los decimales, logaritmos y las ecuaciones y las relaciones trigonométricas, los matemáticos renacentistas aportan el instrumental necesario para abordar los problemas fundamentales de la época. Tanto los relacionados con la navegación como con la artillería, tanto los que implican el conocimiento de la estructura y el funcionamiento del Universo, como los de la figura de la Tierra.

La sociedad se percató del hecho de que las matemáticas son, efectivamente, un instrumento imprescindible para poder resolver los problemas que le preocupan. Mas, no sólo hay que atender los problemas en los que están implicados los cálculos, sino que desde campos como la Arquitectura Civil y Militar, la Geografía o el Arte, surgen problemas que alimentan otra rama del árbol de las matemáticas, llamada Geometría. Así por ejemplo, en el caso de la pintura la perspectiva se va imponiendo como una forma de representar sobre el plano del lienzo escenas y figuras en tres dimensiones. Todo ello genera la necesidad de estudiar el espacio, su representación en el plano, y las propiedades y relaciones entre los entes geométricos que en él existan. En estas preocupaciones y en los problemas derivados de las mismas se encuentra el origen de dos importantes

ramas de la Geometría que, más adelante, van a tener un importante desarrollo, aunque hoy en día hayan quedado casi olvidadas. Nos estamos refiriendo a las geometrías descriptiva y proyectiva.

La vida en la Francia de finales del siglo XVI es bastante agitada. Ninguna persona puede sustraerse a las tensiones sociales que sobre todo en el terreno religioso adquieren tintes dramáticos. Como botón de muestra basta con mencionar la tristemente célebre “Noche de San Bartolomé” en agosto de 1572. En este ambiente vive el abogado François Viète o Vieta (1540 – 1603). No debe sorprender excesivamente la profesión de este innovador del álgebra, ya que en la profesión del matemático no está rigurosamente establecida en la época. La dedicación a las abstractas especulaciones algebraicas y geométricas entran más en el terreno de las aficiones que en el de la actividad profesional remunerada. Como abogado alcanza cierta notoriedad, llegando a defender asuntos de personas muy influyentes y a ser en 1571, letrado del Parlamento de París y Consejero de Enrique III para descifrar mensajes secretos. Vieta cultiva desde muy pronto su afición por las matemáticas. Su obra más importante no aparece hasta 1591 con el título “In artem analyticem Isagoge”, que significa “Introducción al método analítico”. Su gran mérito consiste en ser el iniciador del proceso de cambio desde el álgebra retórica hacia el álgebra simbólica en la que utilizando abreviaturas y símbolos se pueden describir las expresiones matemáticas por medio de fórmulas. Si no se hubiese dado este paso, puede decirse que tanto hoy en día, como desde hace varios siglos los matemáticos no podrían exponer prácticamente nada de su ciencia sino fuera por la utilización de notaciones que proceden del siglo XVI como los signos $+$, $-$ o la raíz. No debe entenderse que Vieta escribiera el álgebra como se hace ahora mismo, en los libros de texto. Lo que sí hace es aplicar sistemáticamente los criterios que defiende y además, se da cuenta de su importancia. Es decir, Vieta se distingue por utilizar siempre los mismos criterios. A las incógnitas, les aplica las vocales mayúsculas del alfabeto latino, a las cantidades conocidas, las consonantes en mayúsculas, signos $+$ y $-$ para la suma y la resta, la multiplicación la indica con la partícula “in”. Aunque el símbolo $=$ ya había sido utilizado por Robert Recorde (1510 – 1558) en 1557, Vieta utiliza las expresiones “aequibitur” o “aequale” para denotar igualdad. Con todo, Vieta no llegó a dar un compendio exhaustivo de notaciones simbólicas y todavía utilizaba expresiones del álgebra retórica. Al proceso de manejo y resolución de ecuaciones algebraicas Vieta le llama “Analysis exegetike”. Sólo publicó una parte de sus escritos, cosa normal entre los algebristas y geómetras de la época, debido a que tomaban las matemáticas como un hobby. Además del Isagoge su obra más trascendente es el “Canon Mathematicus”, obra muy del gusto de la época en la que aparecen tablas y más tablas.

2.5. La revolución científica: geometría analítica, cálculo diferencial e integral.

A comienzos del siglo XVII (ver [194]), el álgebra se está desarrollando gracias a la obra de Vieta y de la geometría, de la cual se sabe bastante gracias a las obras de la Antigüedad clásica. Pero ambas son disciplinas bastante separadas la una de la otra. Una de las grandes conquistas matemáticas de este siglo es la síntesis del álgebra y la geometría que da origen a una nueva rama de dicha ciencia, la geometría analítica. La crean dos matemáticos franceses entre los que se organiza una gran disputa sobre la autoría de tal descubrimiento. Ambos son verdaderos hombres de ciencia, de proyección universal, si bien como personas, por su origen, por su carácter, por su actividad profesional, por su forma de entender la ciencia, representan polos opuestos.

El primero de ellos, Pierre Fermat (1601 – 1665) fue hijo de un comerciante y estudió Derecho en Toulouse. Adquiere una sólida formación humanística, distinguiéndose por su dominio de las lenguas. No se interesó por publicar sus trabajos, de hecho, su primer trabajo no apareció hasta 14 años después de su muerte, publicado por su hijo. Sus obras completas empiezan a editarse a finales del siglo XIX. En 1629 escribe una “Introducción a los lugares planos y sólidos” considerado como el primer tratado de geometría analítica, es decir, aplicación del álgebra simbólica a la geometría. Los desarrollos de Fermat hunden sus raíces en las cónicas de Apolonio y en los Elementos de Euclides solo que cambiando las expresiones clásicas por las usuales del siglo XVII. Así lugares planos eran rectas y círculos, lugares sólidos eran cónicas, y lugares lineales las demás curvas planas.

En 1654 a través de la correspondencia que mantiene con Blaise Pascal (1623 – 1662) contribuye al desarrollo del cálculo de probabilidades y a la teoría de juegos. Con todo, lo que hace de Fermat uno de los matemáticos más famosos de todos los tiempos son sus aportaciones a la teoría de números, con el pequeño teorema de Fermat y la maravillosa historia escrita en el margen de la hoja, de que era conocedor de la solución, quedando así planteado uno de los grandes retos de la historia de las matemáticas. En 1993, Andrew Wiles, de la Universidad de Princeton, da a conocer una demostración de 200 páginas, que resulta contener un fallo, subsanado por él mismo en 1994.

El otro padre de la geometría analítica es René Descartes (1596 – 1650), que estudió en el Colegio de Jesuitas de La Fleche. Se licencia en Leyes en Poitiers y se alistó en diferentes ejércitos, lo que le permite viajar y conocer buena parte de Europa. Cuando deja la milicia, se instala durante veinte años en Holanda, y acomete la tarea de establecer un sistema del mundo que sustituya a todos los anteriores. Finalmente acepta la invitación de la reina de Suecia y se traslada a Estocolmo, donde muere.

Descartes es un autor polivalente y universal, que intenta dotar a todas las ciencias del rigor de las matemáticas. Su audaz visión unificadora le permite

hacer converger dos ramas de las matemáticas tan aparentemente alejadas como son el álgebra y la geometría, propiciando el nacimiento público de la geometría analítica. En 1637 publica su famoso “Discurso sobre el Método para conducir rectamente la razón y buscar la verdad en las Ciencias de la Naturaleza”, seguido de tres ensayos, “la Dióptrica, los Meteoros y la Geometría”. Otra de las grandes conquistas de las matemáticas del siglo XVII es el llamado “Cálculo Diferencial e Integral”

De esta época también tenemos que destacar a Isaac Newton (1643 – 1727) hijo póstumo de un arrendatario agrícola que recibió en vida toda clase de respetos y honores. Sus preocupaciones científicas se centraron en conocer el funcionamiento del mundo y en establecer los principios que rigen el desarrollo de los fenómenos físicos. En este sentido Newton, antes que cualquier otra cosa, es un filósofo de la Naturaleza que es como se llamaba a los físicos hasta el siglo XIX. Estudió en la Universidad de Cambridge, entrando en el Trinity College en 1661 con 18 años como subserver o fámulo, es decir, estudios gratis a cambio de realizar tareas domésticas. En 1663 un mecenas llamado H. Lucas dota una cátedra en la Universidad de Cambridge para que se dedique a la enseñanza de las ciencias. El primero en ocupar la cátedra “lucasiana” es I. Barrow hombre inteligente y culto. Newton asiste a sus clases y comienza su meteórica carrera, Scholar en 1664; Bachiller en 1665; Minor Fellow en 1667; Major Fellow y Master of Arts en 1668. En 1669, Barrow renuncia a su cátedra para que pueda ocuparla su discípulo Newton. El parón en su carrera en el año 1666 es debido a una terrible epidemia de peste que hace que Newton regrese a su aldea natal, donde parece ser que le vienen a la mente las ideas que le llevarían a la cima del olimpo científico: la Ley de la Gravitación Universal, la dinámica, la óptica, el binomio de Newton y el cálculo infinitesimal entre otros. Como científico de su tiempo, Newton sabe unir a su tremenda capacidad deductiva una gran afición hacia la construcción de instrumentos para la realización de sus experimentos. Sus investigaciones en el campo de la óptica le llevan a escribir un importante tratado, que no verá la luz hasta 30 años después de haberlo presentado a la Royal Society, y a construir un telescopio reflector, en 1671, por el que se le admite como miembro de dicha Sociedad. La obra cumbre de Newton es la titulada “Principios Matemáticos de la filosofía natural” (1687).

No podemos pasar sin hablar de Gottfried Wilhelm Leibniz (1666 – 1716), nacido en Leipzig en el seno de una familia burguesa que contaba con una gran biblioteca. De inteligencia prodigiosa, a los 10 años de edad ya dominaba el latín y el griego. La figura de Leibniz es impresionante, gran historiador, filósofo, organizador de instituciones científicas, hábil polemista y diplomático, viajero incansable y uno de los más grandes matemáticos de todos los tiempos. Durante 40 años reside en Hannover, al servicio del Príncipe de Hannover, desempeñando trabajos que no le entusiasman. Durante el siglo XVIII se producen dos sucesos sumamente necesarios para el desarrollo de las matemáticas, la creación de Academias y Sociedades Científicas y la aparición de las Revistas Científicas, en concreto del “Acta Eruditorum” (1682) en Leipzig. En esta publicación mensual

ven la luz los primeros y muy importantes trabajos de Leibniz sobre el “Calculus”. El primero de todos sobre el cálculo diferencial se publica en 1684 con el título: “Un nuevo método para máximos y mínimos y para tangentes que no cae en defecto para valores fraccionarios e irracionales y que a este respecto constituye un tipo de cálculo sin precedentes”. Contiene las primeras reglas para el cálculo diferencial y en él se resuelve la primera ecuación diferencial de la historia. En 1686 Leibniz publica las reglas del cálculo integral, al que denomina “Calculus Summatorius”. Para indicar la diferenciación utiliza la letra “d” y para la operación inversa la inicial gótica de la palabra suma “S”. La diferenciación aparece vinculada al problema del trazado de tangentes a curvas y la integración al cálculo de áreas. Ambos problemas se abordan de manera mucho más general a como habían sido atacados por los métodos particulares establecidos anteriormente a lo largo del siglo. Muy pronto se ve que lo que había nacido como una especulación de orden estrictamente matemático se puede aplicar a multitud de problemas prácticos. Antes de que termine el siglo, Leibniz inicia su esplendorosa marcha triunfal hacia el reconocimiento universal. Sus artículos se difunden con bastante rapidez llegando incluso a Inglaterra. Con la difusión de los mismos se inicia la gran polémica entre los matemáticos del continente y sus colegas ingleses por ver de quién es la autoría. Newton había desarrollado su método de las fluxiones en un texto titulado “Methodus fluxionum et serierum infinitarum” o método de las magnitudes fluentes y de las series infinitas, del que se dice que estaba terminado en 1671 pero que sin embargo no se publicó. Leibniz es acusado de plagio y comienza una campaña británica contra él. Hay que esperar hasta el siglo XIX para que las aguas vuelvan a su cauce y Leibniz sea puesto, ya sin recelos, en el lugar del que se había hecho acreedor.

2.6. Los matemáticos en el siglo XVIII

Destacamos a la familia Bernouilli (ver [224]), desde Jacob (1654 – 1705) que fue el primer Bernouilli y que se dedicó al cálculo, las probabilidades y la mecánica. Le siguió Johan (1667 – 1748) que trabajó en los campos del cálculo, ecuaciones diferenciales y mecánica. Daniel Bernouilli (1700 – 1782) estudió las ecuaciones diferenciales, la mecánica y la hidrodinámica. Leonard Euler (1707 – 1783) es el matemático más prolífico de la historia, sus obras completas constan de casi 90 volúmenes. Destacó por su memoria y por su poder de cálculo. Casi toda su obra se escribió en los últimos años de su vida, cuando ya estaba ciego y todo lo dictaba de memoria. Se le deben las siguientes notaciones, $f(x)$ como símbolo de función, e como base de los logaritmos naturales, (Σ) como símbolo de suma, i para la unidad imaginaria, y la conocida como fórmula de Euler.

Durante la revolución francesa fueron famosos muchos matemáticos, como Jean Le Rond D’Alembert (1717 – 1783), Joseph Louis Lagrange (1736 – 1813) que estudió el cálculo de variaciones y la mecánica celeste, el Marqués de Condorcet (1743 – 1794) que aplicó el análisis de probabilidades a la teoría de las

decisiones de carácter social, económico y político, Gaspard Monge (1746 – 1816) que destacó en geometría descriptiva y fue fundador de L'Ecole Polytechnique, Pierre Simon Laplace (1749 – 1827) estudioso de la mecánica celeste, la exposición del sistema del mundo y la teoría analítica de las probabilidades.

Ya en esta época destacaban varias mujeres en matemáticas, como la Marquesa de Chatelet (1706 – 1749) que realizó la versión francesa de la obra “Principios matemáticos de la filosofía natural” de Newton o María Gaetana de Agnesi (1718 – 1799) creadora de la “curva de Agnesi”.

2.7. Los matemáticos en el siglo XIX

Carl Friedrich Gauss (1777 – 1855) (ver [224]) fue uno de los matemáticos más importantes de la historia. En sus “Disquisitiones arithmetical” transformó la orientación y la estructura de la teoría de los números. No se prodigó excesivamente a la hora de publicar sus descubrimientos aun así, destacó por la búsqueda del rigor matemático y ejerció gran influencia en geometría, mecánica, óptica y astronomía.

Augustin Louis Cauchy (1789 – 1857) fue ingeniero y profesor en L'Ecole Polytechnique. En su Cours d'Analyse fundamentó el cálculo infinitesimal en el concepto esencial de límite, prácticamente en la misma forma en que hoy en día se sigue utilizando. Desarrolló el álgebra de límites, el concepto de sucesión de serie, la noción de convergencia e introdujo rigurosamente los conceptos de derivada e integral. También destacó en teoría de ecuaciones diferenciales ordinarias y en ecuaciones en derivadas parciales.

Niels Henrik Abel (1802 – 1829) demostró que las ecuaciones de quinto grado no tienen soluciones generales por radicales. Hasta entonces se conocían las soluciones de las ecuaciones de primer y segundo grado, las de tercer grado en el siglo XVI por Tartaglia y Cardano y las de cuarto orden mediante cambio de variable.

Evariste Galois (1811 – 1832) en sus intentos por hallar soluciones de las ecuaciones algebraicas de grado superior descubrió las relaciones entre ellas. A partir de esa idea, estableció la teoría de grupos abstractos en la que tiene el origen la teoría de las estructuras algebraicas en la que se fundamenta todo el álgebra moderna.

Sophie Germain (1776 – 1831) mandaba a la Escuela Politécnica sus trabajos firmados bajo el pseudónimo de Mr Antoine Leblanc, ya que no se admitían mujeres en su época.

Con el descubrimiento de las geometrías no euclídeas aparecen Lobachevski (1792 – 1856), Bolyai (1802 – 1860) y Weierstrass (1815 – 1897) profesor extraordinario a cuyas clases asistían 200 alumnos. El centro de producción matemática pasó de Gotinga a París y de ahí a Berlín. Entre las mujeres matemáticas destacamos a Sofia Kovalevskaya (1850 – 1891) que recibió clase de Weierstrass y fue

la primera mujer catedrática de matemáticas, en la Universidad de Estocolmo. Ada Lovelace (1815 – 1852), hija de Lord Byron dio los fundamentos matemáticos de la máquina de calcular. O Florence Nightingale que destacó en estadísticas sanitarias.

George Cantor (1845 – 1918) publicó un trabajo revolucionario sobre la teoría del infinito y la teoría de conjuntos en 1874 que levantó una gran polémica sobre la idea del infinito. Hoy en día esas ideas han sido aceptadas. El concepto del infinito es uno de los más fructíferos de las matemáticas. La exigencia de rigor supuso la aparición de contradicciones en el seno de las matemáticas, la crisis de los fundamentos, que para resolverla se adscriben tres escuelas. La escuela lógica con Frege (1848 – 1925), Peano (1858 – 1932) o Russell (1872 – 1970), la escuela intuicionista con Brouwer (1881 – 1966) y Weyl (1885 – 1955) y la escuela formalista con Hilbert (1862 – 1943) y Noether (1882 – 1935).

Se dice que ya en el siglo XX se han desarrollado del orden de 3000 subespecialidades o subdisciplinas matemáticas. Existen más de 1600 revistas especializadas que en su conjunto publican más de 200000 temas nuevos cada año.

2.8. Docencia de la matemática aplicada

Como ya hemos visto en los anteriores temas, desde la Antigüedad el ser humano se ha preocupado por la docencia de las matemáticas, que en sus inicios estaban ligadas a obras de ingeniería. Podemos establecer el inicio con Arquímedes, matemático e ingeniero. Posteriormente, ya en el Renacimiento, encontramos matemáticos preocupados por enseñar la matemática aplicada, como puede ser Tartaglia, cuyo primer libro original publicado fue *Nuova scientia* (1537). En él establece los principios de la balística y matematiza los conocimientos físicos en que se basa. Se dedicó aquí a estudiar el movimiento de un cuerpo lanzado con una resistencia nula por parte del aire. Leonardo da Vinci, pintor, inventor, escritor, músico y poeta, fue a su vez científico, arquitecto e ingeniero, más preocupado por cultivar sus múltiples talentos que por difundirlos, tuvo no obstante unos cuantos alumnos a los que supo transmitir las maquetas e instrumentos que diseñó. Realizando las primeras obras a gran escala de ingeniería. Durante la ilustración, diferentes científicos trabajaron la rama de la matemática aplicada a la ingeniería, como Leonhard Euler o D'Alembert. En España destacamos la Real Academia Militar de Matemáticas de Barcelona (1720), centro docente donde se formaban los oficiales que deseaban entrar en el Real Cuerpo de Ingenieros y recibían clases de matemáticas y de dibujo y fortificación. Los cursos que se impartían fueron exportados a otras Academias de Artillería, y se componía de ocho tratados: Aritmética, Geometría Elemental, Trigonometría y Geometría práctica, Fortificación, Artillería, Cosmografía, Estática y por último Arquitectura Civil. Destaca la docencia de la matemática aplicada, aun cuando no existían manuales de referencia en castellano, y las enseñanzas eran dictadas por los profesores a los estudiantes.

Desde hace unos pocos años, nos encontramos con el debate de cómo conseguir que las enseñanzas de matemáticas sean más multidisciplinares, de cómo formar a los jóvenes en lo que se conoce como Matemática aplicada. Pero la respuesta no es del todo obvia, y tendemos a crear asignaturas de modelización en las carreras de matemáticas. Mismo argumento nos encontramos en las Ingenierías, donde parece que las matemáticas que se enseñan son homogéneas y perfectamente uniformes, o que las matemáticas de las Escuelas de Ingeniería no albergan interés por sí solas y únicamente se estudian como vehículo para la consecución de otras competencias. En este último caso, queda aún mucho camino por recorrer para dotar a los ingenieros de una base científica, particularmente base matemática la cual complementa la formación de todos ellos.

Una de las disciplinas más relevantes de la matemática aplicada es la del Análisis Numérico ya que la llegada y el crecimiento de los ordenadores ha hecho que sea necesario traducir las Matemáticas a un lenguaje comprensible para la máquina, la cual posibilita la realización de cálculos cuyo número y complejidad eran hasta entonces inviables para los seres humanos. Esta disciplina, nace como ramificación del Análisis Matemático, es considerada hoy en día como la rama más potente y aplicable de las Matemáticas. Así pues, en los Congresos Internacionales de Matemáticas, de entre todas las secciones en que se dividen, por lo menos dos de ellas son directamente de Análisis Numérico, encontrando otras secciones que no son propiamente de Análisis Numérico pero en las que está íntimamente relacionado, lo que demuestra el vigor de esta disciplina.

En los años 70 y principios de los 80 el Análisis Numérico comenzaba a estudiarse en nuestros planes de estudios de matemáticas y poca gente en España dominaba esta disciplina considerada como una bifurcación del Análisis. Todavía tenía una importancia secundaria además de la dificultad de implementar estas herramientas teóricas en ordenadores que aún eran muy lentos y los resultados obtenidos no eran nada visuales y difíciles de interpretar. Pero la docencia de esta rama de las matemáticas aplicadas cambió radicalmente con la aparición de software como MatLab (<http://www.mathworks.com/>) y similares. Con ellos se podían efectuar cálculos complicados en ordenadores personales al alcance de todos los estudiantes con unas extraordinarias posibilidades a la hora de visualizar los resultados lo cual facilitaba su interpretación y permitía aproximar numéricamente el mundo real. En los años en los que apareció esta disciplina en nuestras Universidades, el Análisis Numérico se componía principalmente de las siguientes materias:

- Interpolación de funciones y fórmulas de cuadratura;
- Métodos iterativos para la resolución de sistemas lineales;
- Métodos de aproximación de Ecuaciones Diferenciales Ordinarias, más conocidas como EDOs;

- Métodos en Diferencias y Elementos finitos para Ecuaciones en Derivadas Parciales o EDPs.

La aparente simplicidad de los conceptos frente a otras materias como por ejemplo la Topología, hizo que no se tomase en serio el Análisis Numérico en un principio y que tardase en hacerse un hueco en el ámbito matemático internacional. Ahora ya no es así, de hecho, ha alcanzado no sólo a nivel teórico sino también en cuanto a complejidad y dimensión un desarrollo tan grande que podemos hablar de un nuevo paradigma científico fruto del mestizaje de la Informática con las Matemáticas. Vista una vez la dificultad de la docencia de la Matemática Aplicada a lo largo de la Historia, en particular a la rama del Análisis Numérico, no es extraño comprender que a los actuales alumnos de estas disciplinas les cueste entender la materia. De ahí que los docentes debemos buscar o fabricar herramientas muy visuales que haga que nuestros estudiantes puedan ver reflejados los resultados de una manera muy visual, como podemos comprobar en los siguientes capítulos de esta tesis.

Capítulo 3

Conceptos previos de matemática aplicada

Resumen

En este capítulo, veremos los conceptos fundamentales que vamos a utilizar a lo largo de toda la tesis, vinculados con la convergencia de métodos iterativos en espacios de Banach, la dinámica en la recta real así como la dinámica en el plano complejo.

Comenzamos definiendo qué es un punto fijo, un punto periódico y un punto crítico, así como n -ciclo, es decir, órbitas de puntos periódicos de periodo n . Definimos cuenca de atracción de un punto fijo e introducimos el concepto de caos o sistema dinámico discreto caótico. A lo largo de toda la tesis, siempre que trabajemos con dinámica compleja, utilizaremos la transformación de Möbius, y los conjuntos de Julia y de Fatou que también introducimos aquí. Para el estudio gráfico de la dinámica real, utilizaremos tanto los diagramas de Feigenbaum como los exponentes de Lyapunov, para trabajar con familias uniparamétricas de funciones de iteración, ya que el carácter de los puntos periódicos indiferentes puede depender del parámetro.

En el último punto de conceptos previos, introduciremos operadores definidos entre espacios de Banach y hablaremos de operadores invertibles y de la derivada de Fréchet, conceptos necesarios para entender varios de los capítulos de esta tesis. Enunciamos el Teorema del Valor Medio, tan necesario para la búsqueda de raíces mediante métodos iterativos. Para finalizar, enunciamos un teorema de gran utilidad, el Teorema de Taylor.

3.1. Introducción a la dinámica compleja

En este apartado introductorio presentaremos algunas de las definiciones y resultados relacionados con las funciones de variable compleja, restringidas a aplicaciones racionales. Sea $R : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ una función racional sobre la esfera de Riemann, $\bar{\mathbb{C}} = \mathbb{C} \cup \infty$, entonces $R(z) = \frac{P(z)}{Q(z)}$ con $P(z), Q(z)$ polinomios complejos sin factores comunes.

Nótese que en $\bar{\mathbb{C}}$, un polinomio de grado n tiene exactamente n raíces contadas con multiplicidad. Por esta razón, los polinomios en el plano complejo se van a distinguir según el número de raíces que tengan. Debido a que el objetivo principal de la presente tesis es el estudio del método de Newton amortiguado aplicado a ecuaciones polinómicas, haremos especial hincapié en ellos. Un estudio más extenso de estas y otras propiedades de la dinámica en el plano complejo puede verse en [40], [47], [55], [75], [91] y [175] entre otros.

3.1.1. Puntos fijos, puntos periódicos y puntos críticos

Definimos el grado de $R(z)$ como el máximo de los grados de $P(z)$ y $Q(z)$. De ahora en adelante supondremos que todas las aplicaciones que vamos a estudiar, salvo que se indique lo contrario, son de grado mayor o igual que 2.

En el estudio de la dinámica de aplicaciones racionales, uno de los problemas fundamentales consiste en estudiar el comportamiento de sucesiones $\{z_n\}$ generadas recursivamente a partir de un cierto valor inicial, $z_0 \in \bar{\mathbb{C}}$, donde cada

$z_k = R^k(z_0)$ y donde $R^k(z) = \overbrace{R \circ \dots \circ R}^k$ es la composición de $R(z)$ consigo misma k veces. Si dicha sucesión converge, significa que tiene límite y, por lo tanto, este límite será un punto fijo de $R(z)$.

Definición 3.1 *Decimos que un punto $z \in \bar{\mathbb{C}}$, es un punto fijo de $R(z)$ si verifica $R(z) = z$.*

Existen diferentes tipos de puntos fijos que se clasifican según sea su multiplicador, y éste viene definido como sigue.

Definición 3.2 *Si z es un punto fijo de $R(z)$, entonces llamaremos multiplicador o autovalor de z al valor $\mu = R'(z)$.*

Los puntos fijos de una aplicación racional son casos especiales de puntos periódicos, concretamente son puntos de periodo 1. Los puntos periódicos se definen de la siguiente manera.

Definición 3.3 *Decimos que un punto $z_0 \in \bar{\mathbb{C}}$, es un punto periódico de periodo p , si $R^p(z_0) = z_0$ y $R^n(z_0) \neq z_0$ para todo $n < p$.*

La órbita de un punto periódico z de periodo n está constituida por los n diferentes puntos

$$\mathcal{O}(z) = \{z, R(z), R^2(z), \dots, R^{n-1}(z)\}.$$

Definición 3.4 *La órbita de un punto periódico de periodo n recibe el nombre de n -ciclo.*

Otro tipo de puntos importantes, son aquéllos que no son periódicos, pero alguna de sus iteraciones sí lo es, éstos son los puntos eventualmente periódicos.

Definición 3.5 *Un punto z decimos que es eventualmente periódico si cumple que $R^k(z) = R^{k+p}(z)$ para algunos enteros positivos p y k .*

Los puntos periódicos también tienen asociado un multiplicador o autovalor que se define de la siguiente manera:

Definición 3.6 *Si z es un punto periódico de periodo p , entonces llamaremos multiplicador del p -ciclo al valor de la derivada $\mu = (R^p)'(z)$.*

Los puntos fijos y puntos periódicos se pueden clasificar según su carácter, y éste viene determinado por el autovalor asociado.

Definición 3.7 *Sea z_0 un punto fijo de $R(z)$ con multiplicador μ , entonces diremos que z_0 es:*

(i) *Superatractor, si $\mu = 0$.*

(ii) *Atractor, si $|\mu| < 1$.*

(iii) *Indiferente, si $|\mu| = 1$.*

(iv) *Repulsor, si $|\mu| > 1$.*

Los puntos periódicos de periodo p de una aplicación racional $R(z)$ se clasifican, de forma similar a los puntos fijos, en:

(i) Superatractores, si $\mu = 0$.

(ii) Atractores, si $|\mu| < 1$.

(iii) Indiferentes, si $|\mu| = 1$.

(iv) Repulsores, si $|\mu| > 1$.

Notemos que el autovalor μ asociado a un punto periódico de periodo p (o equivalentemente al p -ciclo), es de suma importancia. Los nombres que reciben los puntos según el valor que tenga su multiplicador son identificativos de lo que va a suceder en un entorno de dichos puntos. Así, por lo tanto, todo punto en un entorno de un punto fijo atractor se irá acercando a él a medida que iteremos y todo punto en un entorno de un punto fijo repulsor se alejará de él. El estudio de los puntos fijos indiferentes es mucho más complejo, ya que existen diversas opciones de lo que puede suceder y se corresponde con la situación frontera entre puntos atractores y repulsores. Probemos que esto es cierto, para ello supongamos que z_0 es un punto fijo (sería idéntico para puntos periódicos de cualquier periodo p) atractor de $R(z)$, es decir, $|R'(z_0)| < 1$. Entonces, para un z suficientemente cercano a z_0 existirá $\beta < 1$ tal que verifique

$$\frac{|R(z) - R(z_0)|}{|z - z_0|} < \beta < 1,$$

y, por ser punto fijo, se tiene que $R(z_0) = z_0$. Además, la desigualdad anterior resultaría

$$|R(z) - z_0| < \beta|z - z_0|,$$

por lo que, $R(z)$ está más cerca de z_0 que z . Si iteramos k veces tenemos que

$$|R(z) - z_0| < \beta^k|z - z_0|,$$

y, como $\beta < 1$, tenemos que

$$\lim_{n \rightarrow \infty} R^n(z) = z_0.$$

Consecuentemente, se puede concluir que todo z perteneciente a un entorno del punto fijo z_0 converge a z_0 bajo iteración de $R(z)$.

Si por el contrario, z_0 es un punto fijo (ídem para puntos periódicos de cualquier periodo p) repulsor de $R(z)$, es decir, $|R'(z_0)| > 1$, aplicando un procedimiento similar obtenemos que la desigualdad obtenida ahora sería

$$\frac{|R(z) - R(z_0)|}{|z - z_0|} > 1,$$

y, sin más que operar, $|R(z) - z_0| > |z - z_0|$, y por lo tanto, la iteración se aleja de z_0 .

Por último, centrando la atención en los puntos indiferentes, es decir, z_0 tal que verifica $|R'(z_0)| = 1$, el comportamiento de las órbitas de puntos en el entorno de estos puntos fijos indiferentes es mucho más complejo que en los casos anteriores, ya que representa la frontera entre ambos. El multiplicador asociado a puntos fijos indiferentes, como ya se ha dicho antes, es de la forma $|\mu| = 1 = e^{i\theta}$, donde $\theta \in [0, 2\pi)$ y el comportamiento dependerá de si θ es un número racional, es

decir, $\theta = \frac{p}{q}$, con p, q coprimos entre sí o si es irracional. Se puede profundizar en este tipo de puntos, como puede verse en [51], [71] y [208] entre otros. Se distinguen los siguientes tipos de puntos fijos indiferentes:

- Parabólicos, si existe $q \in \mathbb{Z}$ tal que $\mu^q = 1$.
- Neutrales o indiferentes, si $|\mu| = 1$ y $\mu^q \neq 1$ para todos los $q \in \mathbb{Z}$. En este caso existen dos posibilidades:
 - Si es linealizable, punto de Siegel.
 - Si no es linealizable, punto de Cremer.

El carácter del ∞ como punto fijo necesita de un estudio especial.

Definición 3.8 *El punto ∞ es un punto fijo de $R(z)$ si y sólo si $z = 0$ es un punto fijo de la función*

$$F : z \rightarrow \frac{1}{R(\frac{1}{z})}.$$

Si ∞ es un punto fijo de $R(z)$, entonces su multiplicador asociado es $\mu = F'(0)$. En particular tenemos que ∞ es un superatractor si y sólo si $F'(0) = 0$.

Además de los distintos tipos de puntos fijos que existen, también es importante el concepto de cuenca de atracción.

Definición 3.9 *Llamaremos cuenca de atracción de un punto fijo ω de $R(z)$ al conjunto de puntos del plano cuyas iteraciones por $R(z)$ convergen hacia el punto fijo, es decir,*

$$\{z \in \bar{\mathbb{C}} \mid \lim_{n \rightarrow \infty} R^n(z) = \omega\}.$$

En la Figura 3.1 vemos las cuencas de atracción asociadas a las raíces del polinomio $p(z) = z^3 - 1$ al aplicarle el método de Newton, es decir,

$$R(z) = z - \frac{z^3 - 1}{3z^2} = \frac{2z^3 + 1}{3z^2}.$$

Debido a que las cuencas de atracción pueden tener infinitas componentes se define el siguiente concepto.

Definición 3.10 *La cuenca de atracción inmediata de un punto fijo ω de $R(z)$ es la componente conexa que contiene el punto fijo.*

Además, dentro del estudio de la dinámica de aplicaciones racionales van a jugar un papel importante los puntos críticos, ya que en cierta medida, pueden caracterizar el comportamiento de las distintas órbitas.

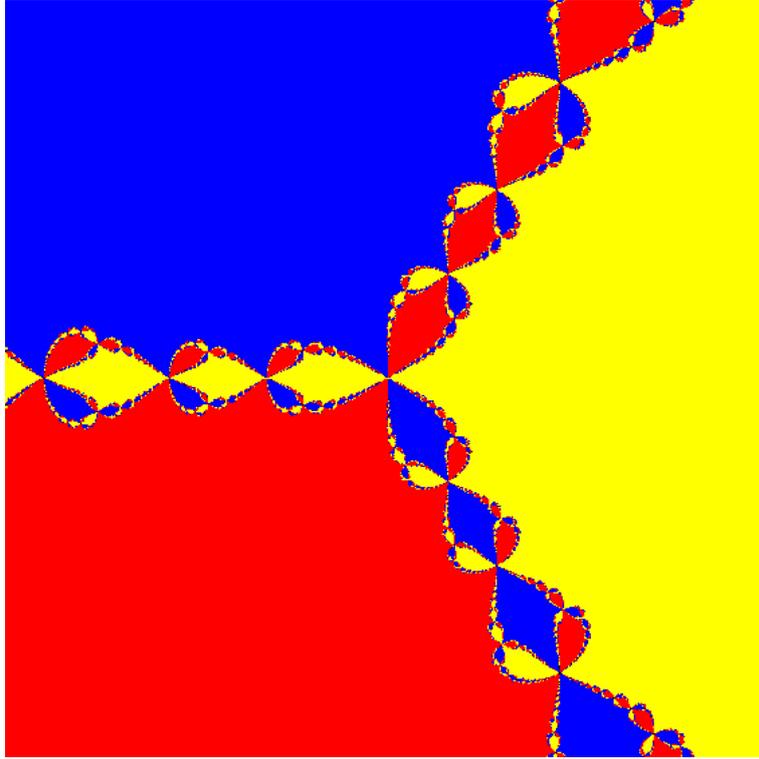


Figura 3.1: Cuencas de atracción asociadas a las raíces del polinomio $p(z) = z^3 - 1$ al aplicarle el método de Newton. En amarillo aparece la cuenca de $z = 1$, en azul la cuenca de $z = e^{2\pi i/3}$ y en rojo la cuenca de $z = e^{4\pi i/3}$.

Definición 3.11 Sea $R(z)$ una función racional de grado d . Un punto $w \in \bar{\mathbb{C}}$, para el cual la cardinalidad de $R^{-1}(w)$ es menor que d , es llamado un valor crítico de $R(z)$. Un punto $z \in R^{-1}(w)$ que es una raíz de $R(z) - w$, con multiplicidad mayor que 1, es llamado un punto crítico de $R(z)$. En particular, z es un punto crítico de una función holomorfa $p(z)$ si $p'(z) = 0$.

Por ejemplo, la función $R(z) = \frac{z}{(z-1)^2}$ tiene un punto crítico en $z = -1$, para el que se anula $R'(z) = \frac{z+1}{(1-z)^3} = 0$. El valor crítico asociado a $z = -1$ es $-\frac{1}{4}$. Además, si $\omega = \infty$, $R^{-1}(\omega) = 1$, por lo que $\omega = \infty$ es otro valor crítico, en este caso, asociado al punto crítico $z = 1$. Los puntos críticos juegan un papel fundamental en el estudio de la dinámica de una función $p(z)$. De hecho, son puntos alrededor de los cuales la función $p(z)$ no es un homeomorfismo local. Es decir, si p_1 es un punto crítico de $p(z)$, ninguna rama de la función inversa de $p(z)$ está definida en ningún entorno del valor crítico $p(p_1)$.

Estos dos conceptos suelen confundirse a menudo y es muy importante saber distinguir entre punto crítico y valor crítico. Por otro lado, se tiene la siguiente nota que relaciona los puntos críticos con los ciclos atractores.

Nota 3.12 Si el máximo de los grados del denominador y el numerador de una

función racional es d , entonces la función tendrá como mucho $2d - 2$ puntos críticos y, por lo tanto, $2d - 2$ ciclos atractores.

Consecuentemente, la existencia de órbitas o ciclos atractores interfieren en la búsqueda de raíces de $p(z) = 0$, al aplicar un método iterativo.

3.1.2. Caos

En esta sección introducimos el concepto de caos. Para ello, nos apoyaremos en la definición dada por Devaney (véase [75]), que ha sido modificada con el paso del tiempo, primero por Banks et al. en 1992 [38] y, posteriormente por Touhey en 1997 [214].

La aparición de caos en el método de Newton ha sido estudiada por diferentes autores, como podemos ver en el libro de Gutiérrez y Plaza *Estudio dinámico del método de Newton para resolver ecuaciones no lineales* (véase [109]) y también en el método de Newton amortiguado para polinomios cuadráticos, en concreto, para el caso del polinomio $p(z) = z^2 + 1$ (véase [45]).

Para poder definir y entender la noción de caos, es necesario definir varios conceptos previos, que aparecen a continuación.

Definición 3.13 *Un sistema dinámico discreto es un par (X, f) donde X es un espacio métrico y f es una función definida de X en X , es decir, $f : X \rightarrow X$.*

Una vez que tenemos definido lo que es un sistema dinámico discreto pasaremos a ver cuándo se dice que es topológicamente transitivo.

Definición 3.14 *Se dice que un sistema dinámico discreto (X, f) , es topológicamente transitivo si dados dos subconjuntos abiertos cualesquiera U y V de X , existe un entero $n \in \mathbb{N}$ tal que se verifica que $f^n(U) \cap V \neq \emptyset$.*

También es necesaria la definición de subconjunto denso en otro.

Definición 3.15 *Se dice que un subconjunto Y de un espacio métrico X es denso en X , si para cualquier subconjunto abierto U de X , existe siempre un punto de Y en U .*

Por último, necesitamos definir el concepto de sistema dinámico discreto sensible con respecto a las condiciones iniciales.

Definición 3.16 *Un sistema dinámico discreto (X, f) se dice que es sensible a las condiciones iniciales si existe un $\delta > 0$ tal que, para todo $x \in X$ y para todo $\epsilon > 0$, existen $y \in X$ y $n \in \mathbb{N}$ tales que se cumple que la distancia entre x e y es menor que ϵ ($d(x, y) < \epsilon$) y la distancia entre $f^n(x)$ y $f^n(y)$ es mayor que δ ($d(f^n(x), f^n(y)) > \delta$).*

Este último fenómeno es conocido como «efecto mariposa». Toda vez que ya tenemos definidos los conceptos anteriores, ya podemos ver la definición de caos dada por Devaney en 1992.

Definición 3.17 *Se dice que un sistema dinámico discreto (X, f) es caótico si cumple todas y cada una de las siguientes condiciones:*

- (i) *Es topológicamente transitivo.*
- (ii) *El conjunto de puntos periódicos de f es denso en X .*
- (iii) *El sistema es sensible con respecto a las condiciones iniciales.*

Diversos autores han ido refinando la definición de sistema caótico (en [38] se demuestra que la sensibilidad con respecto a condiciones iniciales es una propiedad redundante) y en 1997 Touhey [214], da una caracterización diferente del concepto de sistema caótico sin utilizar directamente las tres condiciones de la definición de Devaney, como se observa a continuación.

Teorema 3.18 *Sea (X, f) un sistema dinámico discreto tal que para dos conjuntos abiertos no vacíos cualesquiera $U, V \subseteq X$, existe un punto periódico $p \in U$ tal que para $f^n(p) \in V$ para algún $n \in \mathbb{N}$, entonces el sistema dinámico es caótico en el sentido de la definición de Devaney.*

Esta caracterización establece que un sistema dinámico discreto es caótico si y sólo si para cualesquiera dos conjuntos abiertos U, V de X , existe una órbita periódica que visita a ambos conjuntos.

Por último, en [109] se pueden ver ejemplos de sistemas dinámicos discretos caóticos como la función «diente de sierra», la función «tienda de campaña» o la función logística entre otras.

3.1.3. Conjugación topológica

Un concepto fundamental para comprender el comportamiento de las aplicaciones racionales, desde el punto de vista dinámico, es el de conjugación. Para introducir este concepto hacemos uso de las transformaciones de Möbius, que son aplicaciones racionales de grado 1.

Definición 3.19 *Llamamos transformación de Möbius de parámetros $a, b, c, d \in \mathbb{C}$ a la función τ de la siguiente forma:*

$$\tau(z) = \frac{az + b}{cz + d}, \forall z \in \bar{\mathbb{C}}$$

con a, b, c, d tales que $ad - bc \neq 0$.

Las conjugaciones topológicas son muy utilizadas en el estudio de funciones racionales que dependen de parámetros, ya que en ocasiones, pueden reducir el número de parámetros involucrados en el estudio.

Definición 3.20 Sean $R_1, R_2 : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ dos funciones racionales. Decimos que son conjugadas si y sólo si existe una transformación de Möbius $\varphi : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$, tal que $\varphi \circ R_1 \circ \varphi^{-1}(z) = R_2(z)$ para todo z .

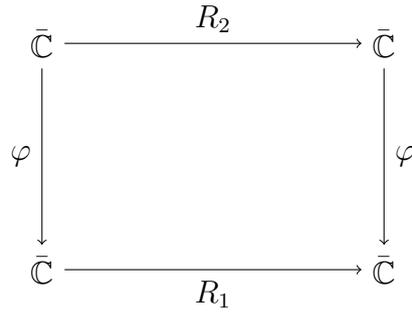


Figura 3.2: Diagrama resultante de una conjugación topológica.

La conjugación es una relación de equivalencia, y las clases de equivalencia se denominan clases de conjugación de las aplicaciones racionales. Para un estudio más profundo sobre conjugaciones se pueden consultar los textos [40], [91] y [109], entre otros. En el siguiente teorema se pueden observar de manera resumida las propiedades elementales de las conjugaciones topológicas.

Teorema 3.21 Sean $R_1 : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ y $R_2 : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ dos funciones, y sea $\varphi : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ una conjugación topológica entre $R_1(z)$ y $R_2(z)$, como la que aparece en la Figura 3.2. Entonces,

- (i) $\varphi^{-1} : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$, es también una conjugación topológica entre $R_2(z)$ y $R_1(z)$.
- (ii) $\varphi \circ R_1^n(z) = R_2^n \circ \varphi(z)$, para todo $n \in \mathbb{N}$.
- (iii) p es un punto periódico de $R_1(z)$ si y sólo si $\varphi(p)$ es un punto periódico de $R_2(z)$. Además, p y $\varphi(p)$ tienen periodos iguales.
- (iv) Si p es un punto periódico de $R_1(z)$ y $\varphi'(z)$ no se anula en la órbita de p , entonces p y $\varphi(p)$ tienen el mismo carácter (atractor, repulsor, indiferente).
- (v) Si p es un punto periódico de $R_1(z)$ con cuenca de atracción $\mathcal{B}(p)$, entonces la cuenca de atracción de $\varphi(p)$ es $\varphi(\mathcal{B}(p))$.
- (vi) Los puntos periódicos de $R_1(z)$ son densos en $\bar{\mathbb{C}}$, si y sólo si, los puntos periódicos de $R_2(z)$ son densos en $\bar{\mathbb{C}}$.

(vii) $R_1(z)$ es caótica sobre $\bar{\mathbb{C}}$, si y sólo si, $R_2(z)$ es caótica sobre $\bar{\mathbb{C}}$.

Obsérvese que, en ocasiones, el problema de descubrir el comportamiento tanto cualitativo como cuantitativo de una aplicación racional puede simplificarse transfiriendo este problema en otro más sencillo, en términos de su conjugado. Veámoslo con un ejemplo, supongamos que queremos estudiar el método de Newton aplicado a polinomios con dos raíces diferentes, es decir, sea $p(z) = (z - a)(z - b)$, entonces

$$N_p(z) = z - \frac{p(z)}{p'(z)} = z - \frac{(z - a)(z - b)}{2z - a - b}.$$

El estudio de $N_p(z)$ depende de las dos raíces a y b , sin embargo, si utilizamos la transformación de Möbius

$$M(z) = \frac{z - a}{z - b}$$

y componiendo, de la siguiente manera

$$S_p(z) = M \circ N_p \circ M^{-1}(z) = z^2,$$

se obtiene que la aplicación a estudiar es mucho más sencilla que la primitiva, pues es independiente de las raíces. Este hecho ya fue puesto de manifiesto por Cayley en 1879 y es clave en el estudio de procesos iterativos aplicados para resolver ecuaciones cuadráticas complejas, como se verá en el Capítulo ??.

3.1.4. Conjuntos de Julia y Fatou

Esta sección está dedicada a las propiedades de los conjuntos de Julia (véase [124]) y de Fatou (véase [93]), denominados así en honor a sus descubridores, Gaston Julia y Pierre Fatou, quienes mantuvieron una dura batalla por *le grand prix des sciences mathématiques* de 1918 (para más información, puede consultarse el apartado dedicado a su enfrentamiento en [91]). Ambos conjuntos tienen comportamientos complementarios, es más, la definición de ambos se corresponde con el complementario el uno del otro. En términos coloquiales y de una manera superficial puede definirse el conjunto de Fatou como los puntos cuyas órbitas bajo un método iterativo tienen un comportamiento estable, mientras que el conjunto de Julia se correspondería con los puntos cuyas órbitas tienen un comportamiento inestable.

Diversos autores se han sentido atraídos por el estudio de los conjuntos de Julia y Fatou, ya que son esenciales en el estudio de la dinámica compleja (véanse [91], [161] ó [174], entre otros). La distinción entre ambos conjuntos está ligada al concepto de normalidad que se define de la siguiente manera.

Definición 3.22 *Dados dos espacios métricos (X, d) e (Y, ρ) y una familia τ de aplicaciones de X en Y , decimos que τ es equicontinua en x_0 si para cada $\varepsilon > 0$*

existe un $\delta > 0$ tal que para todo $x \in X$ y para todo $f \in \tau$, se cumple que si $d(x_0, x) < \delta$, entonces $\rho(f(x_0), f(x)) < \varepsilon$. Decimos que τ es equicontinua en un subespacio Ω , si es equicontinua para cada punto $x \in \Omega \subseteq X$.

Definición 3.23 Suponemos que $R(z)$ es una aplicación racional no constante. Se denomina conjunto de Fatou asociado a $R(z)$, y se denota por $\mathcal{F}(R)$, al mayor subconjunto abierto de $\bar{\mathbb{C}}$ en el que la familia

$$\tau = \{R, R \circ R, \dots, R \circ \dots \circ R, \dots\}$$

es equicontinua con respecto a una de las métricas en $\bar{\mathbb{C}}$.

El complementario del conjunto de Fatou es el conjunto en el que la dinámica es más complicada de entender.

Definición 3.24 Suponemos que $R(z)$ es una aplicación racional no constante. El complementario del conjunto de Fatou, denotado por $\mathcal{J}(R) = \bar{\mathbb{C}} - \mathcal{F}(R)$, se denomina conjunto de Julia asociado a $R(z)$.

Las propiedades de los conjuntos de Julia y Fatou han sido estudiadas por numerosos autores en multitud de obras, un ejemplo de ello son el libro de Fagella y Jarque (véase [91]) o el artículo de Peitgen, Saupe y Haeseler (véase [174]). Entre los resultados sobre las propiedades más importantes cabe destacar el siguiente listado.

Sea $R(z)$ una aplicación racional y $\mathcal{J}(R)$ el conjunto de Julia asociado a $R(z)$. Entonces:

- (i) $\mathcal{J}(R) \neq \emptyset$ y $\mathcal{J}(R)$ es denso en sí mismo (es perfecto).
- (ii) $\mathcal{J}(R) = \mathcal{J}(R^n)$, para todo $n \in \mathbb{N}$.
- (iii) $R(\mathcal{J}(R)) = R^{-1}(\mathcal{J}(R)) = \mathcal{J}(R)$, es decir, $\mathcal{J}(R)$ es completamente invariante.
- (iv) Si $z_0 \in \mathcal{J}(R)$, entonces la clausura de $\{z \in \bar{\mathbb{C}} \mid R^n(z) = z_0\} = \mathcal{J}(R)$.
- (v) Si γ es un ciclo atractor de $R(z)$, entonces la cuenca de atracción de γ está contenida en $\mathcal{F}(R)$ y, además, su frontera está en $\mathcal{J}(R)$.
- (vi) Si $\mathcal{J}(R)$ tiene un interior no vacío, entonces $\mathcal{J}(R) = \bar{\mathbb{C}}$.
- (vii) $R(z)$ restringida a su conjunto de Julia es sensible respecto a condiciones iniciales.
- (viii) $\mathcal{J}(R)$ es autosimilar.
- (ix) Para todo punto $z \in \mathcal{J}(R)$ el conjunto de preimágenes de z es denso en $\mathcal{J}(R)$.

- (x) $R(z)$ es topológicamente transitiva en $\mathcal{J}(R)$.
- (xi) Si P_R es el conjunto de los puntos periódicos repulsores de $R(z)$, entonces $P_R \subset \mathcal{J}(R)$ y además $\bar{P}_R = \mathcal{J}(R)$.
- (xii) Si $R(z)$ es no constante y sea $\tau(z)$ una transformación de Möbius. Definimos una aplicación racional nueva $S(z) = \tau \circ R \circ \tau^{-1}(z)$. Entonces $\mathcal{F}(S) = \tau(\mathcal{F}(R))$ y $\mathcal{J}(S) = \tau(\mathcal{J}(R))$.

Como hemos visto con anterioridad, pueden existir infinitas componentes conexas de una cuenca de atracción, cada una de ellas pertenece al conjunto de Fatou, recordemos que el conjunto de Fatou es abierto. Para ver una clasificación detallada de las diferentes componentes que puede presentar el conjunto de Fatou, se pueden consultar las siguientes referencias [90], [91], [174] o [213].

Decimos que un conjunto X es errante si la intersección $R^n X \cap R^m X$ es vacía para toda $n > m \geq 0$. La posibilidad de aparición de dominios errantes en aplicaciones racionales fue descartada por Sullivan en 1985, como indica el siguiente resultado.

Teorema 3.25 *El conjunto de Fatou de una aplicación racional no tiene componentes errantes. Es decir, cada componente del conjunto de Fatou es eventualmente periódica.*

Sin embargo, I. N. Baker en [35] mostró que los conjuntos de Fatou de algunas funciones enteras en $\bar{\mathbb{C}}$ tienen dominios errantes.

De forma esquemática se pueden encontrar las siguientes componentes de Fatou:

1. Cuencas de atracción de puntos fijos atractores. Las órbitas de puntos cercanos a los puntos fijos convergen hacia dichos puntos. Un ejemplo puede verse en la Figura 3.3, en la que se observan las órbitas de los 4 puntos fijos atractores, con colores diferentes.
2. Cuencas de atracción de ciclos periódicos. Las órbitas convergen a una órbita periódica atractora. Por ejemplo, la función $p(z) = z^2 - 1$ tiene una órbita $\{0, -1\}$ de periodo 2 que es atractora. Dicha órbita está rodeada de una cuenca inmediata de atracción que se corresponde con las dos componentes de color amarillo que la contienen, véase la Figura 3.4. La unión de todas las preimágenes de éstas, forma la totalidad de la cuenca dibujada en amarillo. Su complementario, en rojo, es la cuenca de atracción del infinito. En negro aparece el conjunto de Julia.
3. Cuencas de atracción parabólicas. En términos de comportamiento son muy similares a las anteriores, pero con la diferencia que los puntos fijos o periódicos se encuentran en la frontera de la cuenca y no en su interior. En

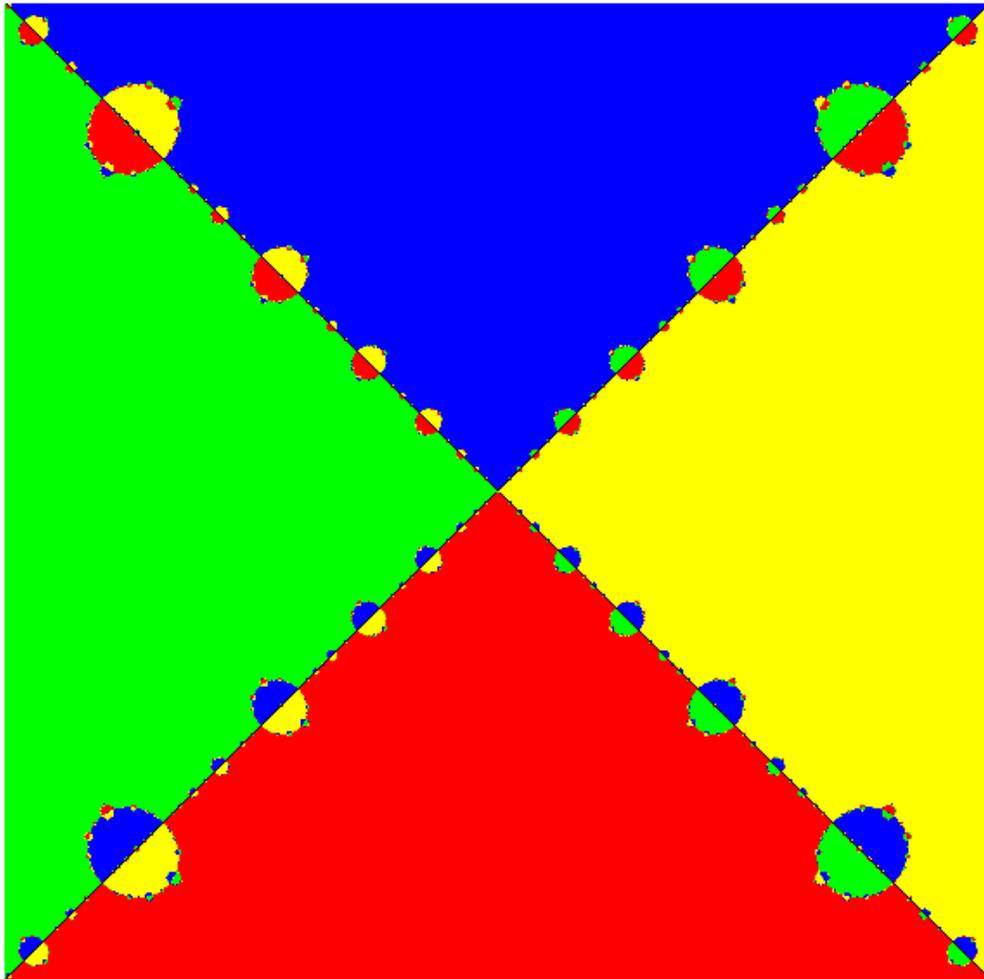


Figura 3.3: Cuencas de atracción asociadas a las raíces del polinomio $p(z) = z^4 - 1$, al aplicarle el conocido método de Halley $\left(H_f(z) = z - \frac{f(z)}{f'(z)} \frac{2}{2-L_f(z)}\right)$. En verde la cuenca de $z = -1$, en amarillo la de $z = 1$, en azul la cuenca de $z = i$ y en rojo la de $z = -i$.

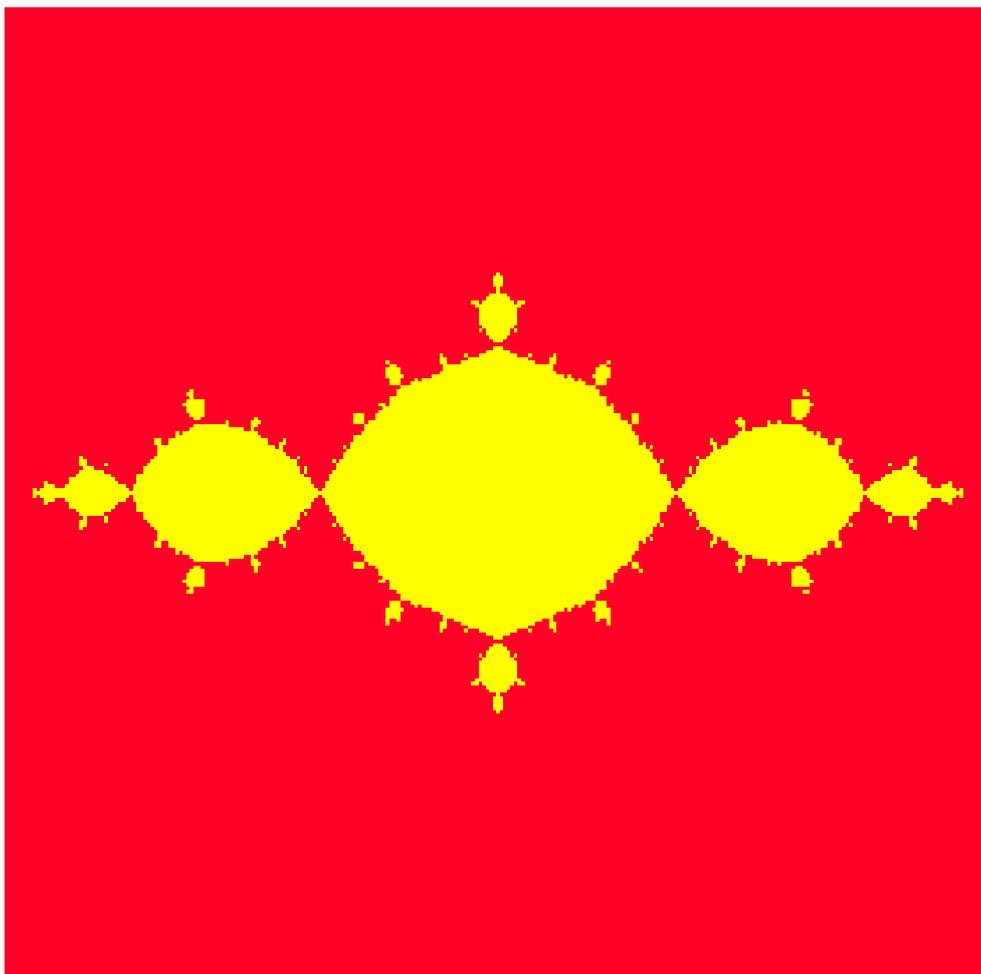


Figura 3.4: Cuencas de atracción asociadas a la función $p(z) = z^2 - 1$ que tiene una órbita periódica atractora de periodo 2, dicha órbita es $\{0, -1\}$. En amarillo aparece la cuenca de la órbita periódica, mientras que en rojo aparece la cuenca de infinito.

la Figura 3.5 podemos ver las cuencas de atracción asociadas al polinomio $p(z) = z^2 + 0.25$, el cual tiene a $z_0 = 0.5$ como punto fijo con multiplicador asociado $\mu = 1$. En dicha figura se observa en amarillo su cuenca de atracción parabólica, en este caso formada por una sólo componente conexa. Todos los puntos de este color tienen órbitas que convergen a z_0 , situado en la frontera. Además, en rojo se aprecia la cuenca de atracción del infinito y en negro el conjunto de Julia.

Además, aunque en el desarrollo de esta tesis no aparecen, pueden encontrarse las siguientes componentes de Fatou:

- Discos de rotación o de Siegel. Se conoce como disco de rotación fijo de $P(z)$, a un abierto U conforme equivalente al disco unidad, que contiene

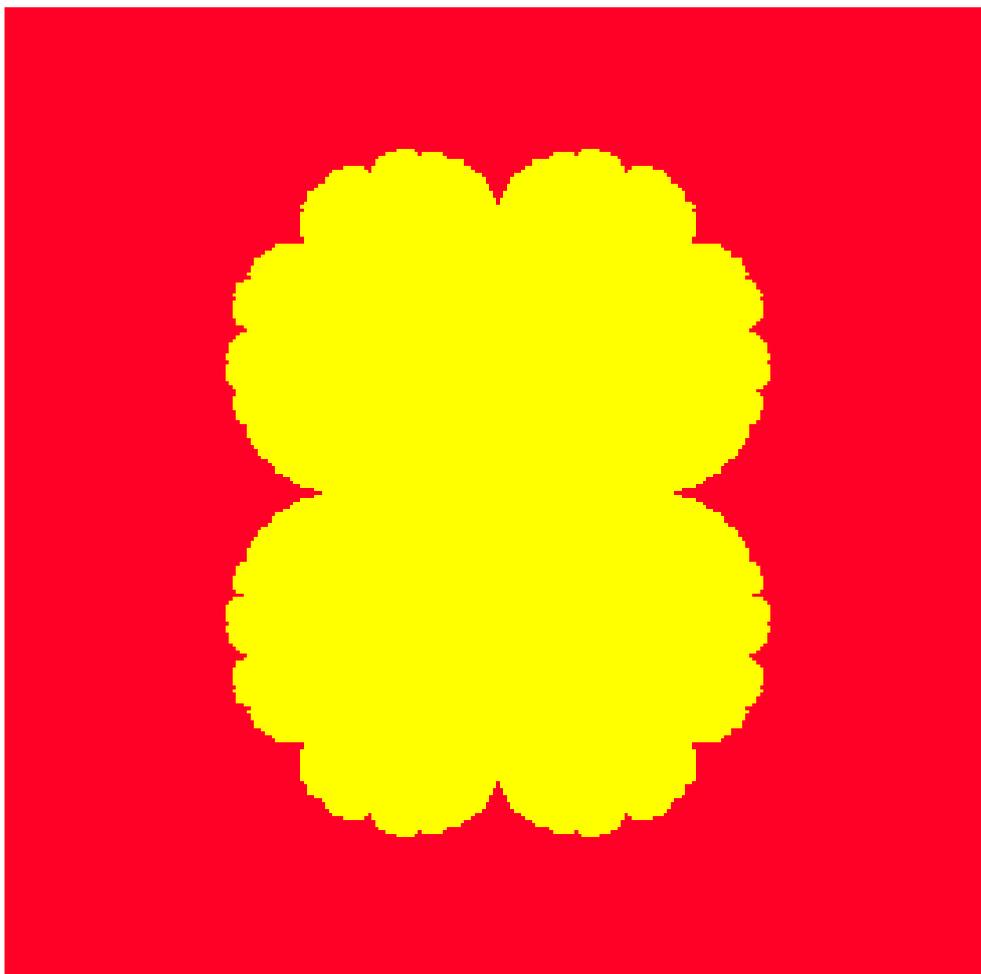


Figura 3.5: Cuencas de atracción asociadas a la función $p(z) = z^2 + 0.25$ que tiene en $z_0 = 0.5$ un punto fijo parabólico. En amarillo aparece la cuenca de $z_0 = 0.5$, mientras que en rojo aparece la cuenca de infinito.

un punto fijo z_0 en su interior, y tal que $P|_U$ es conformemente conjugada a una rotación rígida $R_\theta(z) = e^{2\pi i\theta}$ en el disco unidad, donde $\theta \in \mathbb{R} \setminus \mathbb{Q}$. Esto quiere decir que existe un homeomorfismo conforme $h : U \rightarrow D$ tal que $h \circ P(z) = R_\theta \circ h(z)$. Ésto significa que las órbitas densas en U viven sobre curvas invariantes y, sobre ellas, se comportan como una rotación irracional, es decir, se acumulan de forma densa sobre toda la curva.

- Anillos de rotación o de Herman. El comportamiento es muy similar al anterior, con la única salvedad de que en lugar de ser conformemente conjugado a un disco de rotación, lo es a un anillo de rotación. Los anillos de rotación se diferencian de los discos de rotación en que no están asociados a ninguna órbita periódica.

Como consecuencia del principio de módulo máximo, se tiene que todo

anillo de rotación requiere de un polo de la función, por eso nunca encontraremos anillos de rotación en los planos de fase de polinomios o funciones enteras.

- Dominios parabólicos en el ∞ o de Baker. Un dominio parabólico infinito fijo es un abierto U del plano tal que todas sus órbitas convergen hacia el punto del infinito, que debe ser una singularidad esencial. Por definición, sólo existen en funciones trascendentes.
- Dominios errantes. Un dominio errante es un abierto U tal que $F^n(U) \cap F^m(U) = \emptyset$ para toda $n, m \in \mathbb{N}$. Es decir, los dominios errantes no son componentes ni periódicos ni eventualmente periódicos.

Una vez que ya tenemos definidos los conjuntos de Julia y Fatou, nos preguntamos qué tipo de puntos pertenecen a cada uno de ellos.

Nota 3.26 *Los puntos fijos:*

- *Superatractores, atractores y linealizables (Siegel) pertenecen al conjunto de Fatou.*
- *Repulsores, parabólicos y no linealizables (Cremer) pertenecen al conjunto de Julia.*

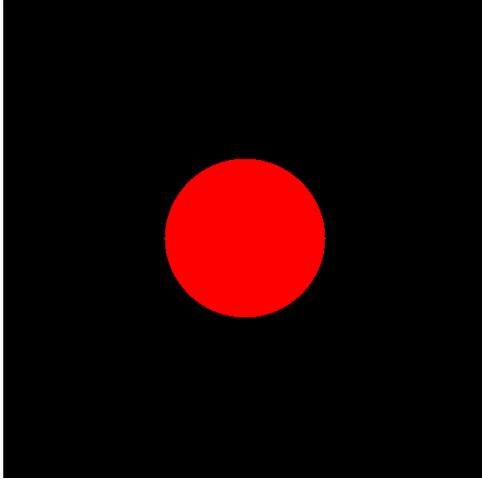
El concepto de conjunto de Julia universal introducido por Kneils en [132], nos permitirá reducir el estudio de la dinámica de un método iterativo como buscador de soluciones aplicado a polinomios, al estudio de la dinámica de polinomios particulares. Por lo tanto, este concepto va a ser de gran utilidad y se define de la siguiente manera.

Definición 3.27 *Decimos que un algoritmo iterativo de un punto, para encontrar raíces $p \rightarrow T_p$, tiene un conjunto de Julia universal (para polinomios de un cierto grado d) si existe una aplicación racional $R(z)$ tal que para cada polinomio $p(z)$ de grado d , el conjunto de Julia asociado $\mathcal{J}(T_p)$ es conjugado vía una transformación de Möbius a $\mathcal{J}(R)$, es decir, existe una transformación de Möbius $\tau(z)$ tal que $T_p(z) = \tau \circ R \circ \tau^{-1}(z)$.*

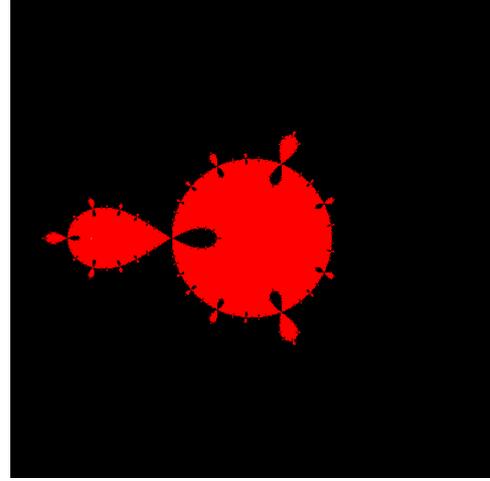
Para ilustrar este concepto, consideramos la bien conocida familia de métodos iterativos de Chebyshev-Halley (véase [16] o [69]) para resolver ecuaciones no lineales

$$z_{n+1} = z_n - \left(1 + \frac{1}{2} \frac{L_p(z_n)}{1 - \lambda L_p(z_n)}\right) \frac{p(z_n)}{p'(z_n)}, \quad L_p(z_n) = \frac{p(z_n)p''(z_n)}{(p'(z_n))^2}, \quad (3.1.1)$$

donde $\lambda \in \mathbb{R}$. La mayoría de los métodos de tercer orden más famosos están incluidos en esta familia. Por ejemplo para $\lambda = 0$ se obtiene el método de Chebyshev,



Newton, Halley y super-Halley



Chebyshev

Figura 3.6: En la parte izquierda aparecen los conjuntos de Julia universales asociados al método de Newton, al método de Halley y al método de super-Halley aplicados a un polinomio cuadrático con dos raíces diferentes. Notemos que son idénticos y coinciden con la circunferencia unidad. En la parte derecha aparece el conjunto de Julia universal asociado al método de Chebyshev aplicado a un polinomio cuadrático con dos raíces diferentes.

para $\lambda = 0.5$ obtenemos el método de Halley, para $\lambda = 1$ se obtiene el método de super-Halley y si $\lambda \rightarrow \pm\infty$ obtenemos el método de Newton.

Si consideramos polinomios cuadráticos de la forma $p(z) = (z - a)(z - b)$ con $a, b \in \mathbb{C}$, $a \neq b$. Sea $R_\lambda(z)$ la función racional obtenida al aplicar la familia anterior a estos polinomios y sea $M(z)$ la transformación de Möbius definida por

$$M(z) = \frac{z - a}{z - b}.$$

Podemos definir la aplicación racional $S_\lambda(z) = M \circ R_\lambda \circ M^{-1}(z)$ como

$$S_\lambda(z) = z^3 \frac{z + 2(1 - \lambda)}{2(1 - \lambda)z + 1}. \quad (3.1.2)$$

Así, tenemos para Newton $S_\infty(z) = z^2$, para super-Halley $S_1(z) = z^4$, para Halley $S_{1/2}(z) = z^3$ y, por último, para Chebyshev $S_0(z) = z^3 \frac{z+2}{2z+1}$. Los gráficos asociados pueden verse en la Figura 3.6.

Existen multitud de resultados relacionados con los conjuntos de Julia y Fatou, en esta tesis se presentan sólo algunos (véase [109] para un listado más detallado).

Teorema 3.28 *Si z_r es un punto de un ciclo repulsor, entonces el conjunto de Julia, $\mathcal{J}(R)$, está formado por la clausura de las preimágenes de z_r , es decir,*

$$\mathcal{J}(R) = \text{clausura}\{z \in \bar{\mathbb{C}} : R^n(z) = z_r, n \in \mathbb{N}\}.$$

Uno de los resultados más importantes es el siguiente teorema.

Teorema 3.29 (*Teorema Fundamental de Fatou y Julia*) *Los ciclos repulsores son densos en $\mathcal{J}(R)$, es decir,*

$$\mathcal{J}(R) = \text{clausura}\{z \in \bar{\mathbb{C}} : z \text{ pertenece a un ciclo repulsor de } R(z)\}.$$

En particular, existe una cantidad infinita de ciclos repulsores y cada $z \in \mathcal{J}(R)$ es obtenido como límite de puntos en ciclos repulsores.

Además, otro concepto íntimamente ligado con el estudio de la dinámica es el siguiente.

Definición 3.30 *Llamaremos conjunto postcrítico de una aplicación racional $R(z)$ a la clausura de la unión de las órbitas de todos los puntos críticos de $R(z)$.*

La importancia del conjunto postcrítico en el estudio de la dinámica de funciones racionales se resume en la siguiente proposición que aparece en [55] y que es consecuencia de los trabajos de Fatou.

Proposición 3.31 *El conjunto postcrítico de una función racional $R(z)$ contiene, los ciclos atractores de $R(z)$, los ciclos indiferentes que pertenecen al conjunto de Julia y la frontera de cada disco de Siegel y anillo de Herman.*

3.2. Introducción a la iteración de funciones reales

Además de todos los conceptos introducidos en la sección de nociones básicas de iteración compleja, en el caso real aparecen diferentes conceptos que deben ser comentados, como por ejemplo, la aparición de asíntotas, los diagramas de Feigenbaum o los exponentes de Lyapunov entre otros.

Una de las herramientas en la que nos vamos a apoyar es el estudio gráfico de la dinámica. Para llevar a cabo este estudio necesitamos un gráfico de la función de iteración junto con la identidad. Partimos de un punto inicial, x_0 , y se traza una línea vertical hasta que interseca con la función de iteración. En el punto de intersección se traza una línea recta horizontal hasta que interseca la identidad, de nuevo en ese punto trazamos otra recta vertical hasta que interseca la función. La primera coordenada de este último punto se corresponde con la primera iteración del método. Repitiendo el proceso desde este nuevo punto se van obteniendo las siguientes iteraciones. En la Figura 3.7 pueden verse los dos primeros pasos de este procedimiento, mientras que en la Figura 3.8 se observa el procedimiento, con tres y cincuenta pasos, respectivamente.

Por otro lado, resulta de gran utilidad el homeomorfismo $G : \mathbb{R} \rightarrow (0, 1)$ que nos permite obtener compactificaciones de métodos al intervalo $(0, 1)$ y que viene definido por

$$G(x) = \frac{1}{\pi} \arctan(x) + \frac{1}{2}. \quad (3.2.1)$$

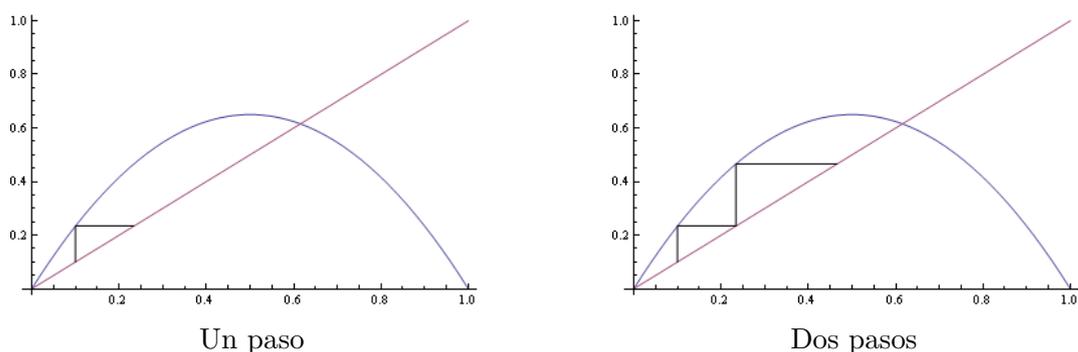


Figura 3.7: Aplicación del algoritmo de estudio gráfico de la dinámica con uno y dos pasos.

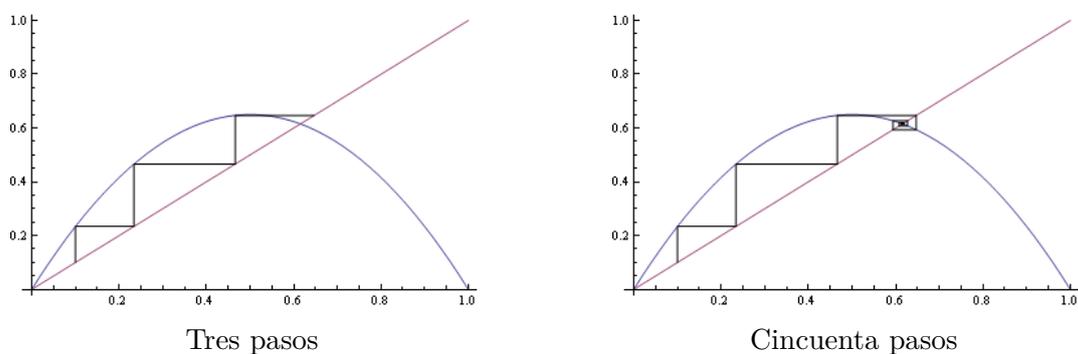


Figura 3.8: Aplicación del algoritmo de estudio gráfico de la dinámica con tres y cincuenta pasos.

Esta función tiene como inversa

$$G^{-1}(x) = \tan\left(\pi x - \frac{\pi}{2}\right). \quad (3.2.2)$$

Sea $\tau : \mathbb{R} \rightarrow \mathbb{R}$ un polinomio de grado mayor que 1, la compactificación que vamos a construir es la siguiente

$$\tilde{\tau}(x) = G \circ \tau \circ G^{-1}(x). \quad (3.2.3)$$

Además, introduciendo las siguientes notaciones

$$G(-\infty) = \lim_{x \rightarrow -\infty} G(x) = 0,$$

y

$$G(\infty) = \lim_{x \rightarrow \infty} G(x) = 1,$$

y teniendo en cuenta que $\tau(x)$ es un polinomio de grado mayor que 1, se puede extender la compactificación a una aplicación definida en $[0, 1]$. Cabe destacar que al utilizar esta compactificación aparecen dos puntos fijos adicionales que

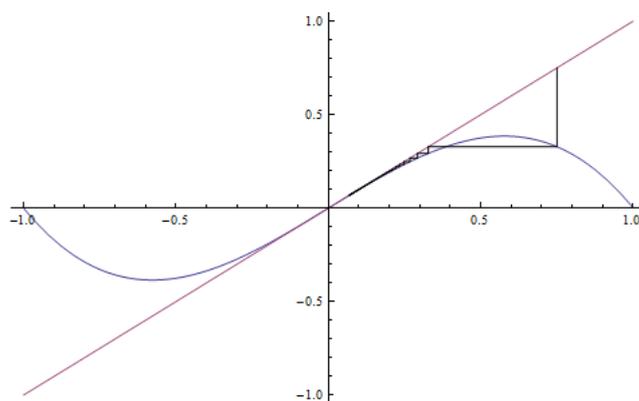


Figura 3.9: Órbitas de $x_0 = 0.75$ bajo la función $f(x) = -x^3 + x$, donde puede verse que el 0 es un punto fijo con multiplicador $\mu = 1$, pero con comportamiento atractor.

son $x = 0$ y $x = 1$. Como estos puntos fijos extraños son repulsores, no van a interferir en la dinámica del método.

Con respecto a los puntos fijos, también hay variaciones con respecto a la dinámica compleja, como demuestran los siguientes dos teoremas.

Teorema 3.32 *Sea $f : I \rightarrow \mathbb{R}$, donde I es un intervalo y $f \in \mathcal{C}^1(I)$. Sea $x^*i \in I$ un punto fijo indiferente de $f(x)$, es decir, el módulo de su multiplicador asociado $|f'(x^*i)| = 1$, tal que $|f'(x)|$ presenta un máximo local en x^*i , entonces x^*i es un punto fijo atractor.*

La demostración a este teorema puede verse en [101].

Ejemplo 3.33 ([101]) *El 0 es un punto fijo indiferente atractor de la función $f(x) = -x^3 + x$, como puede verse en la Figura 3.9. Notemos que $|f'(x)| = |-3x^2 + 1|$ presenta un máximo local en $x = 0$.*

Teorema 3.34 *Sea $f : I \rightarrow \mathbb{R}$, donde I es un intervalo y $f \in \mathcal{C}^1(I)$. Sea $x^*i \in I$ un punto fijo indiferente de $f(x)$, es decir, el módulo de su multiplicador asociado $|f'(x^*i)| = 1$, tal que $|f'(x)|$ presenta un mínimo local estricto en x^*i , entonces x^*i es un punto fijo repulsor.*

La demostración a este teorema se puede encontrar también en [101].

Ejemplo 3.35 ([101]) *El 0 es un punto fijo indiferente repulsor de la función $f(x) = x^3 + x$ como puede verse en la Figura 3.10. En este caso, $|f'(x)| = 3x^2 + 1$ tiene un mínimo absoluto en $x = 0$.*

En el estudio de familias uniparamétricas de funciones de iteración, el carácter de los puntos periódicos puede depender del parámetro. En concreto, los puntos

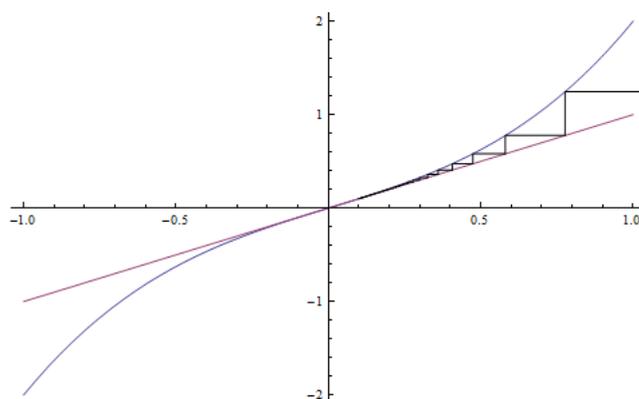


Figura 3.10: Órbitas de $x_0 = 0.1$, bajo la función $f(x) = x^3 + x$, donde puede verse que el 0 es un punto fijo con multiplicador asociado $\mu = 1$, pero con comportamiento repulsor.

periódicos indiferentes marcan el paso de un carácter atractor a repulsor o viceversa. Cuando se produce un cambio cualitativo en el comportamiento de los puntos periódicos, se dice que se ha producido una bifurcación. Los casos más comunes de bifurcación son los siguientes.

Bifurcación tangente. Cuando un punto fijo indiferente se desdobla en dos puntos fijos, uno atractor y otro repulsor, o cuando dos puntos fijos, uno repulsor y otro atractor colapsan en un único punto fijo indiferente.

Ejemplo 3.36 La familia cuadrática $f_c(x) = x^2 + c$ presenta una bifurcación tangente en $c = 0.25$. Los puntos fijos de $f_c(x) = x$ son

$$x = \frac{1 \pm \sqrt{1 - 4c}}{2}.$$

Por otro lado, como la derivada es $f'_c(x) = 2x$, se tiene que:

- Si $c < 0.25$, $f_c(x)$ tiene dos puntos fijos cuyos multiplicadores asociados son $1 \pm \sqrt{1 - 4c}$ y, por lo tanto, uno es atractor y el otro es repulsor.
- Si $c = 0.25$, $f_c(x)$ tiene un único punto fijo indiferente en $x = 1$.
- Si $c > 0.25$, $f_c(x)$ no tiene puntos fijos.

Bifurcación transcítica. Cuando dos puntos fijos intercambian su carácter.

Ejemplo 3.37 La familia logística $f_c(x) = cx(1 - x)$, con $c > 0$, presenta una bifurcación transcítica en $c = 1$. Los puntos fijos de $f_c(x)$, son $x_1 = 0$ y $x_2 = 1 - \frac{1}{c}$. Además, como la derivada es $f'_c(x) = c(1 - 2x)$, se tiene que:

- Si $c < 1$, x_1 es atractor y x_2 repulsor.

- Si $c = 1$, $x = 0$ es el único punto fijo indiferente.
- Si $c > 1$, x_1 es repulsor y x_2 atractor.

Bifurcación horca. Cuando un punto fijo atractor se convierte en repulsor a la vez que aparecen dos puntos fijos atractores nuevos.

Ejemplo 3.38 La familia de funciones $f_c(x) = c(-x^3 + x)$, con $c > 0$, presenta una bifurcación horca en $c = 1$. Los puntos fijos $f_c(x) = x$, son $x_1 = 0$, $x_2 = \sqrt{\frac{c-1}{c}}$ y $x_3 = -\sqrt{\frac{c-1}{c}}$. Ahora bien, como la derivada es $f'_c(x) = c(-3x^2 + 1)$, se tiene que:

- Si $c = 1$, $f_c(x)$ tiene un único punto fijo indiferente en $x = 0$.
- Si $c > 1$, x_1 es repulsor, x_2 atractor y x_3 también es atractor.

Para estudiar la aparición de bifurcaciones, utilizaremos el conocido diagrama de Feigenbaum o de bifurcaciones que viene definido en [101] de la siguiente forma:

Definición 3.39 Un diagrama de bifurcación de un sistema dinámico, es una estratificación de su espacio de parámetros inducida por la equivalencia topológica, junto con los retratos de fase representativos de cada estrato. Las soluciones estables suelen representarse mediante líneas continuas, mientras que las soluciones inestables se representan con líneas punteadas.

Una ejemplo de este tipo de diagramas puede verse en la Figura 3.11. A la hora de interpretar el diagrama, hay que tener en cuenta que muchas veces el rango de bifurcaciones es tan pequeño que, dependiendo de la escala con la que se trabaje, no se llega a apreciar, lo cual no significa que no haya bifurcaciones. Sin embargo, los exponentes de Lyapunov constituyen una buena herramienta que permitirá obviar el problema con las escalas.

Como se ha visto anteriormente, las órbitas cercanas a un n -ciclo se acercan o se alejan del ciclo, bien sean atractores o repulsores y además, el ritmo viene regulado por la derivada de $f^n(x)$ en cualquier punto del ciclo. Para un ciclo $\{x_1, x_2, \dots, x_n\}$ tras n iteraciones sobre un punto cercano la distancia de éste al ciclo se habrá multiplicado por

$$|(f^n)'(x_1)| = \dots = |(f^n)'(x_n)| = |f'(x_1)||f'(x_2)| \dots |f'(x_n)|$$

y la variación media tras cada iteración será

$$\sqrt[n]{|f'(x_1)||f'(x_2)| \dots |f'(x_n)|}.$$

Este razonamiento también es válido para órbitas sin ciclos.

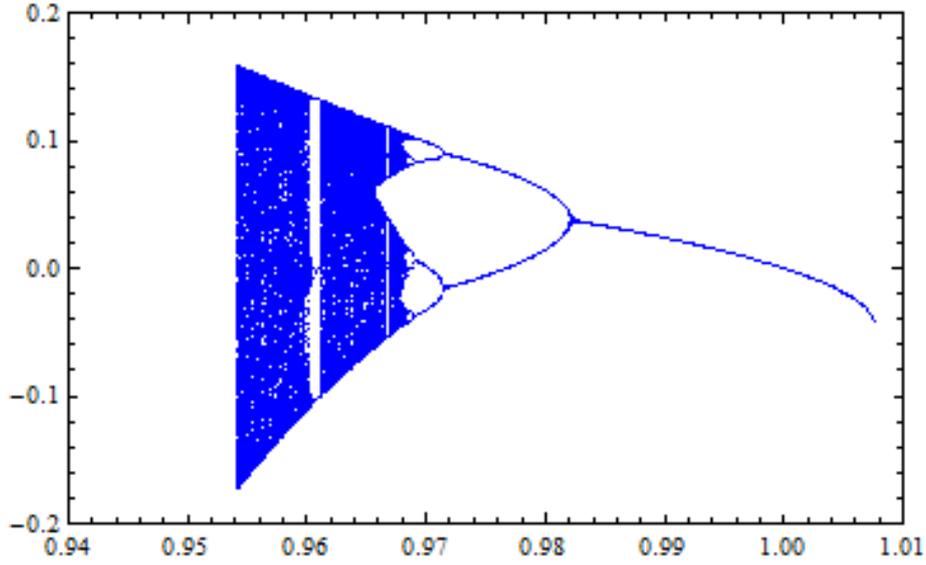


Figura 3.11: Diagrama de Feigenbaum que se obtiene al aplicar el método de dos pasos de Newton a la familia $f(x) = x^3 + \gamma x + 1$, donde se observa que para algunos valores de $\gamma \in (0.94, 1.01)$ se producen bifurcaciones.

Definición 3.40 Sea la órbita $\{x_1, x_2, \dots, x_n\}$, entonces llamaremos número de Lyapunov de la órbita al valor

$$L(x_1) = \lim_{n \rightarrow \infty} \sqrt[n]{|f'(x_1)| |f'(x_2)| \cdots |f'(x_n)|}.$$

Notemos además, que este valor es el mismo para todos los elementos de la órbita.

El número de Lyapunov mide la contracción o expansión local asintótica en cada iteración en la proximidad de una órbita.

Nota 3.41 Si una órbita $\{x_1, x_2, \dots, x_n\}$ es asintótica a una órbita periódica $\{y_1, y_2, \dots, y_n\}$, es decir, $\lim_{n \rightarrow \infty} |x_n - y_n| = 0$, entonces $L(x_1) = L(y_1)$, siempre que ambos estén definidos.

Una vez definido el número de Lyapunov, otro concepto importante es el de exponente de Lyapunov que se define de la siguiente manera.

Definición 3.42 Sea la órbita $\{x_1, x_2, \dots, x_n\}$, entonces llamaremos exponente de Lyapunov al valor

$$h(x_1) = \lim_{n \rightarrow \infty} \frac{1}{n} (\log |f'(x_1)| + \log |f'(x_2)| + \cdots + \log |f'(x_n)|) = \log(L(x_1)).$$

Notemos además, que este valor es el mismo para todos los elementos de la órbita.

Si el exponente de Lyapunov de x_1 es negativo, las órbitas de puntos cercanos a x_1 serán atraídas por la de x_1 , mientras que si es positivo tenderán a separarse.

Nota 3.43 *Notar que $h(x)$ no está definida si $f'(x)$ se anula en algún punto de la órbita.*

Una manera de relacionar el exponente de Lyapunov con el concepto de caos es el siguiente (véase [45], [101]).

Nota 3.44 *Se tiene que una órbita $\{x_1, x_2, \dots, x_n\}$ es caótica si*

$$h(x_1) > 0, \text{ o lo que es lo mismo, } L(x_1) > 1.$$

Ejemplo 3.45 *La familia logística, es la formada por las aplicaciones*

$$f_c : [0, 1] \rightarrow [0, 1],$$

de la forma $f_c(x) = cx(1 - x)$, con $c \in \mathbb{R}$. En la Figura 3.12 pueden verse los exponentes de Lyapunov asociados a la iteración del punto $x_0 = \frac{1}{3}$ en función del parámetro c . Se observa, de forma muy resumida, que:

- *Para $c \in (0, 3]$, las iteraciones de x_0 convergen hacia una de las raíces.*
- *Para $c \in (3, 1 + \sqrt{6}]$, las iteraciones de x_0 convergen hacia un 2-ciclo.*
- *Para $c \in (1 + \sqrt{6}, c_\infty]$, donde $c_\infty \approx 3.5699$, las iteraciones de x_0 convergen hacia un 4-ciclo.*
- *Para $c_\infty < c \leq 4$, las iteraciones son, en su mayoría, caóticas.*
- *Para $c \geq 4$, las iteraciones de x_0 divergen.*

3.3. Introducción operadores definidos en espacios de Banach

En el último capítulo de esta memoria se estudia el método de Newton amortiguado para operadores definidos en espacios de Banach. Por este motivo, introducimos a continuación algunos de los conceptos básicos que facilitarán la lectura de los capítulos siguientes. Para un estudio más detallado sobre espacios de Banach se puede consultar cualquiera de los numerosos tratados dedicados al tema que hay en la bibliografía matemática. Como fuente bibliográfica pueden consultarse las tesis de Gutiérrez [108] y, más reciente la de González [102], además de los libros de Argyros [16] y [17]. Otros textos muy recomendables, ya que

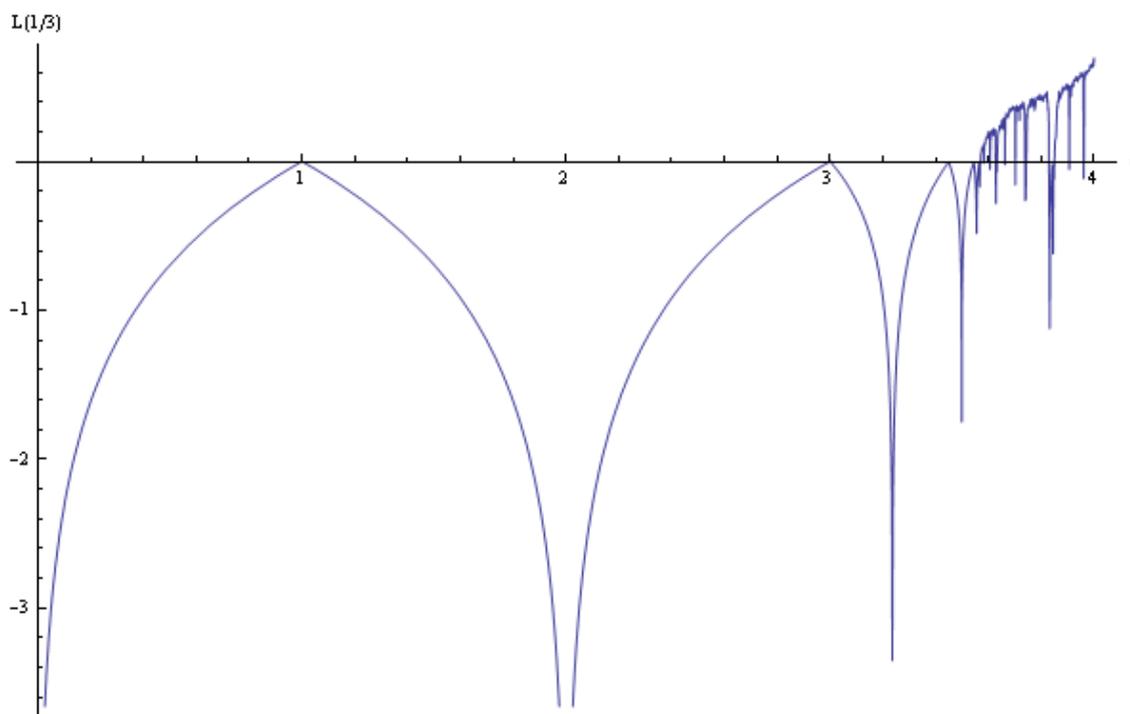


Figura 3.12: Exponentes de Lyapunov asociados al punto $\frac{1}{3}$ en función del valor del parámetro $c \in [0, 4]$ para la función logística.

introducen los espacios de Banach con un objetivo similar al nuestro son los de Kantorovich [128] y Rall [190].

Sea $F : X \rightarrow Y$ un operador definido entre dos espacios de Banach X e Y . Se pretende estudiar la resolución de una ecuación $F(x) = 0$, bajo ciertas condiciones sobre el operador F . Al conjunto de los operadores lineales y acotados entre los espacios X e Y lo denotaremos a partir de ahora $\mathcal{L}(X, Y)$. Si $X = Y$ denotaremos simplemente $\mathcal{L}(X)$. Sea $L \in \mathcal{L}(X, Y)$, denotamos $\mathcal{D}(L)$ a su dominio y $\mathcal{R}(L)$ a su rango. Si existe un operador L^{-1} que aplica $\mathcal{R}(L)$ en $\mathcal{D}(L)$, de manera que

$$L^{-1}Lx = x, \quad \text{para todo } x \in \mathcal{D}(L),$$

$$LL^{-1}y = y, \quad \text{para todo } y \in \mathcal{R}(L),$$

diremos que L es *invertible* y que L^{-1} es el *inverso* de L . Si L^{-1} existe, entonces también es un operador lineal.

A continuación citamos unos resultados conocidos sobre inversión de operadores. Obsérvese que estos resultados están dados para operadores de un espacio X en sí mismo.

Lema 3.46 (Banach) *Sea $L \in \mathcal{L}(X)$ verificando que*

$$\|L\| \leq k < 1,$$

entonces el operador $I - L$ tiene inverso continuo y, además,

$$\|(I - L)^{-1}\| \leq \frac{1}{1 - k}.$$

Unas ligeras variantes del lema anterior son los siguientes resultados.

Lema 3.47 Sea $L \in \mathcal{L}(X)$. Existe L^{-1} si y sólo si existe un operador invertible $M \in \mathcal{L}(X)$ tal que

$$\|I - ML\| < 1.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - ML)^n M$$

y

$$\|L^{-1}\| \leq \frac{\|M\|}{1 - \|I - ML\|}.$$

Lema 3.48 Sea $L \in \mathcal{L}(X)$. Existe L^{-1} si y sólo si existe un operador invertible $M \in \mathcal{L}(X)$ tal que

$$\|M - L\| < \frac{1}{\|M^{-1}\|}.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - M^{-1}L)^n M^{-1}$$

y

$$\|L^{-1}\| \leq \frac{\|M^{-1}\|}{1 - \|I - M^{-1}L\|} \leq \frac{\|M^{-1}\|}{1 - \|M^{-1}\|\|M - L\|}.$$

Presentamos ahora algunos conceptos y resultados básicos del cálculo diferencial e integral en espacios de Banach.

Definición 3.49 Dado $x_0 \in X$, si existe un operador $L \in \mathcal{L}(X, Y)$ de manera que, para todo $x \in X$,

$$\lim_{h \rightarrow 0} \frac{F(x_0 + hx) - F(x_0)}{h} = L(x), \quad (1.1)$$

entonces F se dice diferenciable Gâteaux (o diferenciable débilmente) en x_0 . En esta situación, el operador lineal L es la derivada de Gâteaux (o la diferencial de Gâteaux) de F en x_0 , y se denota $L = dF(x_0)$.

Definición 3.50 Si el límite de la ecuación (1.1) es uniforme en el conjunto $\{x \in X; \|x\| = 1\}$, entonces F se dice diferenciable Fréchet, o simplemente diferenciable en x_0 . En este caso, el operador lineal L se llama derivada de Fréchet (o diferencial de Fréchet) de F en x_0 , y se denota $L = F'(x_0)$.

Equivalentemente, el concepto de diferenciabilidad se puede expresar de la siguiente forma. Si dado $x_0 \in X$, existe un operador lineal y continuo $L = F'(x_0)$ de manera que

$$\lim_{\|v\| \rightarrow 0} \frac{\|F(x_0 + v) - F(x_0) - Lv\|}{\|v\|} = 0,$$

entonces F es diferenciable Fréchet en x_0 .

Obsérvese que si F es un operador diferenciable Fréchet en x_0 , entonces también es diferenciable Gâteaux en x_0 . Además, $F'(x_0) = dF(x_0)$. Salvo que se indique lo contrario, cuando nos referimos a un operador diferenciable, lo entenderemos en el sentido de Fréchet.

Es bien conocido que las propiedades de la diferencial en espacios de Banach son análogas a las de la derivada en el caso escalar, salvo el Teorema del Valor Medio. El siguiente resultado generaliza el Teorema del Valor Medio conocido para funciones escalares.

Teorema 3.51 (del Valor Medio) Sean X e Y dos espacios de Banach. Sea $F : X \rightarrow Y$ un operador diferenciable en un conjunto convexo $\Omega \subseteq X$. Entonces, si $x_0, x_1 \in \Omega$, se tiene

$$\|F(x_1) - F(x_0)\| \leq \sup_{0 < \theta < 1} \|F'(x_0 + \theta(x_1 - x_0))\| \|x_1 - x_0\|.$$

Corolario 3.1 Con la notación anterior se tiene que

$$\begin{aligned} & \|F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0)\| \\ & \leq \sup_{0 < \theta < 1} \|F'(x_0 + \theta(x_1 - x_0)) - F'(x_0)\| \|x_1 - x_0\|. \end{aligned}$$

A continuación comentamos algunos aspectos sobre el cálculo integral en espacios de Banach.

Definición 3.52 Sea F definida en un intervalo real $[a, b]$ y con valores en un espacio de Banach Y . Entonces podemos definir su integral como el límite de las siguientes sumas

$$\sum_{k=0}^{n-1} F(\tau_k)(t_{k+1} - t_k),$$

donde $a = t_0 < t_1 < \dots < t_n = b$ y $\tau_k \in [t_k, t_{k+1}]$, cuando $\max_k \{t_{k+1} - t_k\} \rightarrow 0$. Si el límite dirigido anterior existe, lo llamaremos integral de F y lo denotaremos

$$\int_a^b F(t) dt.$$

Evidentemente, si la integral existe, es un elemento de Y . Una condición suficiente para que exista dicha integral es que la función F sea continua en el intervalo $[a, b]$.

Las propiedades de esta integral se deducen de las conocidas para la integral de Riemann en el caso real. De entre ellas, destacamos las tres siguientes por su posterior utilidad.

(i) Si consideramos un operador lineal y acotado $L \in \mathcal{L}(Y, Z)$, entonces

$$\int_a^b L(F(t)) dt = L \left(\int_a^b F(t) dt \right).$$

(ii) Si $F(t) = \phi(t)y_0$, donde y_0 es un elemento fijo de Y y $\phi(t)$ es una función real integrable, entonces

$$\int_a^b F(t) dt = y_0 \int_a^b \phi(t) dt.$$

(iii) $\left\| \int_a^b F(t) dt \right\| \leq \int_a^b \|F(t)\| dt.$

Definición 3.53 Supongamos ahora que T es un operador definido en un segmento $[x_0, x_1] \subseteq X$ y con valores en el espacio $\mathcal{L}(X, Y)$. En este caso, se define su integral

$$\int_{x_0}^{x_1} T(x) dx = \int_0^1 T(x_0 + t(x_1 - x_0))(x_1 - x_0) dt.$$

Los siguientes resultados sobre cálculo integral en espacios de Banach pueden consultarse en los libros de Kantorovich [128] o Rall [190].

Teorema 3.54 Sea F un operador de X en Y que tiene derivada continua en el segmento $[x_0, x_1] \subseteq X$. Entonces

$$\int_{x_0}^{x_1} F'(x) dx = F(x_1) - F(x_0).$$

Lema 3.55 Si se verifica

$$\left\| F(x_0 + \tau(x_1 - x_0)) \right\| \leq \phi(t_0 + \tau(t_1 - t_0)), \quad \tau \in [0, 1]$$

y

$$\|x_1 - x_0\| \leq t_1 - t_0,$$

entonces

$$\left\| \int_{x_0}^{x_1} F(x) dx \right\| \leq \int_{t_0}^{t_1} \phi(t) dt.$$

Como variante del lema anterior, tenemos el siguiente resultado.

Corolario 3.2 *Si se cumple la desigualdad*

$$\|F(x)\| \leq \phi(t)$$

para x y t tales que

$$\|x - x_0\| \leq t - t_0,$$

entonces

$$\left\| \int_{x_0}^{x_1} F(x) dx \right\| \leq \int_{t_0}^{t_1} \phi(t) dt,$$

donde x_1 es un elemento que cumple $\|x_1 - x_0\| \leq t_1 - t_0$.

Si $F : X \rightarrow Y$ es un operador dos veces diferenciable en un punto $x_0 \in X$, entonces $F''(x_0) \in \mathcal{B}(X^2, Y)$, donde $\mathcal{B}(X^2, Y)$ es el conjunto de los operadores bilineales acotados de $X^2 = X \times X$ en Y , con la norma

$$\|B\| = \sup_{\|x_1\|, \|x_2\| \leq 1} \|B(x_1, x_2)\|, \quad B \in \mathcal{B}(X^2, Y).$$

Es conocido, (véase [56]), que existe una isometría entre el espacio $\mathcal{B}(X^2, Y)$ y $\mathcal{L}(X, \mathcal{L}(X, Y))$. A partir de ahora, y teniendo en cuenta esta isometría, identificaremos los elementos de ambos espacios. También se sabe, véase [190], que $F''(x_0)$ es un operador bilineal simétrico, es decir,

$$F''(x_0)x_1x_2 = F''(x_0)x_2x_1 \quad \text{para todo } x_1, x_2 \in X.$$

Para un operador bilineal de $\mathbb{R}^2 \times \mathbb{R}^2$ en \mathbb{R}^2 empleamos la notación dada por una matriz de bloques [16] y [17]

$$B = \begin{pmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{pmatrix}. \quad (1.2)$$

Entonces, si $x = (x_1, x_2)$, $y = (y_1, y_2)$, se tiene que

$$\begin{aligned} B(x, y) &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \begin{pmatrix} b_1^{11}x_1 + b_1^{21}x_2 & b_1^{12}x_1 + b_1^{22}x_2 \\ b_2^{11}x_1 + b_2^{21}x_2 & b_2^{12}x_1 + b_2^{22}x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} b_1^{11}x_1y_1 + b_1^{21}x_2y_1 + b_1^{12}x_1y_2 + b_1^{22}x_2y_2 \\ b_2^{11}x_1y_1 + b_2^{21}x_2y_1 + b_2^{12}x_1y_2 + b_2^{22}x_2y_2 \end{pmatrix}.$$

En \mathbb{R}^2 consideramos la norma del máximo

$$\|(x_1, x_2)\| = \text{máx} \{|x_1|, |x_2|\},$$

y en $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ la norma dada por

$$\|L\| = \text{máx} \{|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|\},$$

con

$$L = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Entonces se comprueba fácilmente [189] que la norma del operador bilineal B definido por (1.2) viene dada por

$$\|B\| = \sup_{\|x\|=1} \text{máx}_i \left| \sum_{j=1}^2 \sum_{k=1}^2 b_i^{jk} x_k \right|.$$

Además, se tiene la siguiente estimación de la norma anterior

$$\|B\| \leq \text{máx} \{|b_1^{11}| + |b_1^{12}| + |b_1^{21}| + |b_1^{22}|, |b_2^{11}| + |b_2^{12}| + |b_2^{21}| + |b_2^{22}|\}.$$

El resultado de componer una aplicación bilineal B y una lineal L es una aplicación bilineal LB . Con la notación anterior, tenemos que la matriz asociada a LB es

$$LB = \begin{pmatrix} a_{11}b_1^{11} + a_{12}b_2^{11} & a_{11}b_1^{12} + a_{12}b_2^{12} \\ a_{11}b_1^{21} + a_{12}b_2^{21} & a_{11}b_1^{22} + a_{12}b_2^{22} \\ a_{21}b_1^{11} + a_{22}b_2^{11} & a_{21}b_1^{12} + a_{22}b_2^{12} \\ a_{21}b_1^{21} + a_{22}b_2^{21} & a_{21}b_1^{22} + a_{22}b_2^{22} \end{pmatrix}.$$

Si F es un operador de \mathbb{R}^2 en \mathbb{R}^2 que a un par $(x, y) \in \mathbb{R}^2$ le hace corresponder $(f(x, y), g(x, y)) \in \mathbb{R}^2$, tenemos que la derivada segunda de F en un punto (x_0, y_0) es el operador de $\mathbb{R}^2 \times \mathbb{R}^2$ en \mathbb{R}^2 dado por la matriz de la forma (1.2)

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \\ g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix},$$

donde las derivadas parciales anteriores están evaluadas en el punto (x_0, y_0) .

Finalizamos esta pequeña introducción a los espacios de Banach con un resultado que será de gran utilidad para nosotros, como es el Teorema de Taylor.

Teorema 3.56 (Taylor) *Supongamos que F es un operador n -veces diferenciable en la bola $B(x_0, r)$, $r > 0$, y que $F^{(n)}$ es integrable en el segmento $[x_0, x_1]$ con $x_1 \in B(x_0, r)$. Entonces*

$$F(x_1) = F(x_0) + \sum_{k=1}^{n-1} \frac{1}{k!} F^{(k)}(x_0)(x_1 - x_0)^k + R_n(x_0, x_1)$$

donde

$$\begin{aligned} R_n(x_0, x_1) &= \frac{1}{(n-1)!} \int_{x_0}^{x_1} F^{(n)}(x)(x_1 - x)^{n-1} dx \\ &= \frac{1}{(n-1)!} \int_0^1 F^{(n)}(x_0 + t(x_1 - x_0))(x_1 - x_0)^n (1-t)^{n-1} dt. \end{aligned}$$

Parte III

Desarrollo de la investigación

En esta parte se presenta la investigación llevada a cabo en la que destacan dos partes bien diferenciadas: por un lado, se muestra la investigación vinculada con la pedagogía de los métodos iterativos, es decir, una investigación más aplicada a la educación y posteriormente se presentan los resultados de investigación en matemática aplicada, es decir, de matemática más pura. Todos los capítulos de esta parte están basados en artículos, publicados, aceptados o en proceso de revisión en diferentes revistas indexadas en diferentes bases de datos como Journal Citation Report o Scopus. Se ha tratado de hacer esta sección lo más autocontenida posible y es por ello, por lo que van a aparecer, algunos de los conceptos ya presentados en secciones anteriores, pero de esta manera se consigue que cada capítulo dentro de esta parte pueda ser leído de manera independiente.

En concreto, en esta parte aparecen:

- Estudio del método de la secante I
- Estudio del método de la secante II
- Estudio de un método iterativo de orden 6 libre de derivadas segundas.
- Estudio de una familia de métodos iterativos altamente eficiente.
- Estudio de una familia de métodos iterativos de alto orden.
- Aplicación de una herramienta pedagógica para la enseñanza de la matemática aplicada

Capítulo 4

Estudio de la convergencia y dinámica del método de la Secante I

Resumen

En este capítulo vamos a realizar un nuevo análisis de convergencia semilocal para los métodos de la Secante a la hora de aproximar una solución local única de una ecuación no lineal en las condiciones de un espacio de Banach. Aunque ya existen estudios sobre la convergencia semilocal, vamos aquí a obtener criterios de convergencia más débiles y estimaciones más precisas sobre las distancias $\|x_n - x_{n-1}\|$, $\|x_n - x^*\|$, donde x^* es solución local de la ecuación no lineal y x_n y x_{n-1} son dos aproximaciones consecutivas obtenidas mediante el método de la Secante. Además obtendremos una información al menos tan precisa sobre la ubicación de la solución x^* como la que podemos obtener por ejemplo mediante el método de Newton. Estos criterios de convergencia son más suaves que los que ya se conocen en estudios previos y tienen el mismo desgaste computacional en cuanto a uso de recursos físicos a utilizar sobre los parámetros involucrados, e incluso los límites de error son más precisos que en anteriores estudios. Pero no solo nos vamos a centrar en el estudio teórico, también vamos a presentar un ejemplo numérico para ilustrar estos resultados obtenidos en el estudio de la convergencia semilocal. En dicho ejemplo vamos a poder comprobar que estamos en condiciones de garantizar la convergencia del método de la Secante debido a que cumple las condiciones obtenidas en este estudio, pero no cumple las de los estudios anteriores, demostrando así la importancia de nuestro trabajo en cuanto a que se exigen condiciones menos difíciles de alcanzar. Posteriormente lo aplicamos a un problema concreto real utilizando la ecuación que describe la fracción de nitrógeno-hidrógeno que se convierte en amoníaco, llamada conversión fraccionaria, para unos valores fijados de presión y temperatura. Este artículo está basado en el artículo [19] y que puede verse en la sección de Anexos.

4.1. Introducción

En este estudio nos ocupamos del problema de aproximar una solución única local x^* de la ecuación no lineal

$$F(x) = 0, \quad (4.1.1)$$

donde, F es un operador Fréchet-diferenciable definido sobre un subconjunto no vacío \mathbb{D} de un espacio de Banach \mathbb{X} con valores en un espacio de Banach \mathbb{Y} . Muchos problemas de las Ciencias Aplicadas se pueden expresar como (4.1.1) usando modelos matemáticos [20, 23, 59, 89, 141, 151, 171, 185, 209, 215]. Las soluciones de estas ecuaciones solo pueden ser encontradas de forma exacta en algunos casos especiales. Razón por la que la mayoría de métodos para resolver estas ecuaciones son iterativos. El análisis de la convergencia de los métodos iterativos está normalmente dividido en dos categorías: análisis de convergencia local y análisis de convergencia semilocal. En el análisis de convergencia semilocal se obtienen criterios de convergencia a partir de la información en torno a un punto inicial, mientras que en el análisis local se encuentran estimaciones de los radios de bolas de convergencia a partir de la información en torno a una solución. Si $\mathbb{X} = \mathbb{Y}$ y $Q(x) = F(x) + x$, entonces la solución x^* de la ecuación (4.1.1) es muy importante en la teoría del punto fijo.

Estudiamos la convergencia del método de la Secante

$$x_{n+1} = x_n - \mathcal{A}_n^{-1}F(x_n), \quad \mathcal{A}_n = \delta F(x_n, x_{n-1}) \quad \text{para cada } n = 1, 2, \dots, \quad (4.1.2)$$

donde x_{-1}, x_0 son los puntos iniciales. Aquí $\mathcal{A}_n \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $x, y \in \mathbb{D}$ es una aproximación del derivado de Fréchet de F y $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ representa el espacio de operadores lineales acotados de \mathcal{X} a \mathcal{Y} . Existen un gran número de criterios de suficiencia de convergencia para los métodos de la secante (4.1.2) bajo condiciones tipo Lipschitz (4.1.2) (ver [5, 10, 11, 13, 16, 20, 21, 23, 31, 57, 59, 74, 79, 89, 104, 128, 151, 171, 185, 186, 193, 200, 215] y sus referencias). Por lo tanto, es importante estudiar la convergencia del método de la Secante de forma unificada. Es interesante darse cuenta de que aunque usamos sucesiones mayorantes muy generales para $\{x_n\}$ nuestra técnica lleva, en el caso de convergencia semilocal a: criterios de suficiencia de la convergencia más débiles; estimaciones más precisas sobre las distancias $\|x_n - x_{n-1}\|$, $\|x_n - x^*\|$ y una información al menos tan precisa sobre la ubicación de la solución x^* en muchos casos especiales interesantes, como el método de Newton o el método de la Secante.

El resto del artículo se organiza de la siguiente manera. En la sección 4.2 estudiamos la convergencia de las sucesiones mayorantes $\{x_n\}$. La sección 4.3 contiene el análisis de la convergencia semilocal para $\{x_n\}$. Los ejemplos numéricos se dan en la sección final 4.4. Además, en el caso semilocal proporcionamos un ejemplo que involucra una ecuación integral no lineal de tipo Chandrasekhar [59] y otro en el que aparece el proceso conocido como *ammonia*.

4.2. Sucesiones mayorantes para el método de la Secante

En esta Sección, primero estudiaremos algunas sucesiones escalares que están relacionadas con el método de la Secante.

Sean los parámetros $c \geq 0$, $\nu \geq 0$, $k > 0$, $k_0 > 0$, $k_1 > 0$ y $k_2 \geq 0$. Definimos la sucesión escalar $\{\alpha_n\}$ por

$$\begin{cases} \alpha_{-1} = 0, \alpha_0 = c, \alpha_1 = c + \nu, \\ \alpha_{n+2} = \alpha_{n+1} + \frac{k_1(\alpha_{n+1} - \alpha_n) + k_2(\alpha_n - \alpha_{n-1})(\alpha_{n+1} - \alpha_n)}{1 - [k_0(\alpha_{n+1} - c) + k\alpha_n]}, \forall n = 0, 1, 2, \dots \end{cases} \quad (4.2.1)$$

En algunos casos especiales de la sucesión $\{\alpha_n\}$ varios autores han utilizado sucesiones mayorantes para el método de la Secante. Por ejemplo:

- **Caso 1** (Método de la Secante): $k_0 = k$ y $k_1 = k_2$ ha sido estudiado en [5, 13, 21, 23, 31, 74, 89, 171, 185, 200] y para $k_0 = k$, $k_1 = k_2$ y $k_0 \leq k_1$ en [16].
- **Caso 2** (Método de Newton): $k_2 = 0$, $k = 0$, $c = 0$ y $k_0 = k_1$ ha sido estudiado en [10, 16, 23, 57, 74, 79, 104, 128, 151, 171, 185, 186, 193, 215] y para $k_0 \leq k_1$ en [10, 11, 13].

En el presente capítulo estudiaremos la convergencia de la sucesión $\{\alpha_n\}$ simplificándola primero. De hecho, el propósito de las siguientes transformaciones es estudiar la sucesión (4.2.1) después de usar secuencias más fáciles de estudiar definidas por (4.2.3), (4.2.6) y (4.2.8). Sea

$$\lambda = \frac{k_2}{k_1}, \quad L_0 = \frac{1}{1 + k_0 c} \quad \text{y} \quad L = \frac{k_1}{1 + k_0 c}. \quad (4.2.2)$$

Usando (4.2.1) y (4.2.2), podemos escribir la sucesión $\{\alpha_n\}$ como

$$\begin{cases} \alpha_{-1} = 0, \alpha_0 = c, \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \frac{L(\alpha_{n+1} - \alpha_n + \lambda(\alpha_n - \alpha_{n-1}))(\alpha_{n+1} - \alpha_n)}{1 - L_0(k_0\alpha_{n+1} + k\alpha_n)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (4.2.3)$$

Tengamos en cuenta que

$$L = k_1 L_0 \quad (4.2.4)$$

y

$$\beta_n = L_0 \alpha_n. \quad (4.2.5)$$

Entonces, podemos definir la sucesión $\{\beta_n\}$ como

$$\begin{cases} \beta_{-1} = 0, \beta_0 = L_0 c, \beta_1 = L_0(c + \nu) \\ \beta_{n+2} = \beta_{n+1} + \frac{k_1(\beta_{n+1} - \beta_n + \lambda(\beta_n - \beta_{n-1}))(\beta_{n+1} - \beta_n)}{1 - (k_0\beta_{n+1} + k\beta_n)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (4.2.6)$$

Además, tomando

$$\gamma_n = \frac{1}{k_0 + k} - \beta_n \text{ for each } n = 0, 1, 2, \dots \quad (4.2.7)$$

Entonces, la sucesión $\{\gamma_n\}$ está definida por

$$\begin{cases} \gamma_{-1} = \frac{1}{k_0 + k}, \gamma_0 = \frac{1}{k_0 + k} - L_0 c, \gamma_1 = \frac{1}{k_0 + k} - L_0(c + \nu) \\ \gamma_{n+2} = \gamma_{n+1} - \frac{k_1(\gamma_{n+1} - \gamma_n + \lambda(\gamma_n - \gamma_{n-1}))(\gamma_{n+1} - \gamma_n)}{k_0\gamma_{n+1} + k\gamma_n}, \forall n = 0, 1, 2, \dots \end{cases} \quad (4.2.8)$$

Finalmente, sea

$$\delta_n = 1 - \frac{\gamma_n}{\gamma_{n-1}} \quad \forall n = 0, 1, 2, \dots \quad (4.2.9)$$

Entonces, definimos la sucesión $\{\delta_n\}$ como

$$\begin{cases} \delta_0 = 1 - \frac{\gamma_0}{\gamma_{-1}}, \delta_1 = 1 - \frac{\gamma_1}{\gamma_0} \\ \delta_{n+2} = \frac{k_1\delta_{n+1}(\lambda\delta_n + (1 - \delta_n)\delta_{n+1})}{(1 - \delta_n)(1 - \delta_{n+1})(k_0(1 - \delta_{n+1}) + k)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (4.2.10)$$

Para el estudio de la convergencia de la sucesión $\{\alpha_n\}$ es conveniente definir el polinomio p como

$$p(t) = k_0 t^3 - (k_1 + 3k_0 + k)t^2 + (2k + 3k_0 + k_1(\lambda + 1))t - (k_0 + k). \quad (4.2.11)$$

Tenemos que $p(0) = -(k_0 + k) < 0$ y $p(1) = k_1\lambda$ para $\lambda > 0$. Del teorema del valor intermedio se sigue que p tiene raíces en $(0, 1)$. Denotemos por δ a la raíz más pequeña. Si $\lambda = 0$, entonces $p(t) = (t - 1)(k_0 t^2 - (k_1 + k + 2k_0)t +$

$k_0 + k$). Por consiguiente, podemos elegir la raíz más pequeña de p dada por $\frac{2k_0+k_1+k-\sqrt{(2k_0+k_1+k)^2-4k_0(k_0+k)}}{2k_0} \in (0, 1)$ que es δ en este caso.

Notemos también que

$$p(t) \leq 0 \forall t \in (-\infty, \delta]. \quad (4.2.12)$$

A continuación, estudiamos la convergencia de estas sucesiones a partir de $\{\delta_n\}$.

Lema 4.1 Sean $\delta_1 > 0$, $\delta_2 > 0$ y $k_1 > 0$ los parámetros dados. Supongamos que

$$0 < \delta_2 \leq \delta_1 \leq \delta, \quad (4.2.13)$$

donde δ es la definida en (4.2.11). Sea $\{\delta_n\}$ la sucesión escalar definida en (4.2.10). Entonces, las siguientes afirmaciones son verdaderas:

(A₁) Si

$$\delta_1 = \delta_2 \quad (4.2.14)$$

entonces,

$$\delta_n = \delta, \forall n = 1, 2, 3, \dots \quad (4.2.15)$$

(A₂) Si

$$\delta_2 < \delta_1 < \delta \quad (4.2.16)$$

entonces, la sucesión $\{\delta_n\}$ es decreciente y converge a 0.

Demostración. Es claro de (4.2.10) y $\delta_2 \leq \delta_1$ que $\delta_3 > 0$. Debemos demostrar que

$$\delta_3 \leq \delta_2. \quad (4.2.17)$$

En vista de (4.2.10) para $n = 1$, basta con demostrar que

$$p_1(\delta_2) = k_0(1 - \delta_1)\delta_2^2 - (1 - \delta_1)(2k_0 + k + k_1)\delta_2 - (k_0 + k + \lambda k_1)\delta_1 + k_0 + k. \quad (4.2.18)$$

El discriminante Δ del polinomio cuadrático p_1 viene dado por

$$\Delta = (1 - \delta_1) \left[(1 - \delta_1)(2k_0 + k + k_1)^2 + 4k_0((k_0 + k + \lambda k_1)\delta_1 + k_0 + k) \right] > 0. \quad (4.2.19)$$

Por lo tanto, p_1 tiene dos raíces diferentes δ_s y δ_l con $\delta_s < \delta_l$. Entonces, (4.2.18) será verdadero, si

$$\delta_2 \leq \delta_s, \quad (4.2.20)$$

o si $\delta_2 \leq \delta_1$, si $\delta_1 \leq \delta_s$ o si $p(\delta_1) \leq 0$ lo cual es verdadero por (4.2.12), y $\delta_1 \leq \delta$. Por lo tanto, hemos demostrado (4.2.17). Por lo tanto, la relación

$$0 < \delta_{k+1} < \delta_k, \quad (4.2.21)$$

se cumple para $k = 2$. Entonces, debemos demostrar que

$$0 < \delta_{k+2} < \delta_{k+1}. \quad (4.2.22)$$

Es claro de (4.2.10), $\delta_k < 1$ y $\delta_{k+1} < 1$ que $\delta_{k+2} > 0$. Entonces, en vista de (4.2.10) el lado derecho de (4.2.22) es verdad, si

$$\frac{k_1 \delta_{k+1} [\lambda \delta_k + (1 - \delta_k) \delta_{k+1}]}{(1 - \delta_k)(1 - \delta_{k+1}) [k + k_0(1 - \delta_{k+1})]} \leq \delta_{k+1} \quad (4.2.23)$$

o

$$p(\delta_k) \leq 0, \quad (4.2.24)$$

lo cual es verdad por (4.2.12) ya que $\delta_k \leq \delta_1 \leq \delta$. Hemos completado la inducción para (4.2.21). Si $\delta_1 = \delta_2 = \delta$, entonces se sigue de (4.2.10) para $n = 1$ que $\delta_3 = \delta$ y $\delta_n = \delta$ para $n = 4, 5, \dots$, lo cual demuestra (4.2.15). Si $\delta_2 < \delta_1$, la sucesión $\{\delta_n\}$ es decreciente, delimitado por debajo por 0 y, como tal, converge a su único límite inferior más grande denotado por γ . Tenemos de (4.2.10) que

$$\gamma = \frac{k_1 \gamma [\lambda \gamma + (1 - \gamma) \gamma]}{(1 - \gamma)^2 [k + k_0(1 - \gamma)]} \Rightarrow \gamma = \delta \text{ or } \gamma = 0. \quad (4.2.25)$$

Pero $\gamma \leq \delta_1 \leq \delta$. Por lo tanto, concluimos que $\gamma = 0$. □

A continuación, presentamos tres resultados de la convergencia de las sucesiones $\{\alpha_n\}$, $\{\beta_n\}$ y $\{\gamma_n\}$ bajo condiciones que no son del todo las mismas que las del Lema 4.1.

Lema 4.2 *Supongamos que se satisface la hipótesis (4.2.16). Entonces, la sucesión $\{\gamma_n\}$ es decreciente y convergente y las sucesiones $\{\alpha_n\}$ y $\{\beta_n\}$ son crecientes y convergentes.*

Demostración. Usando (4.2.2) y (4.2.9) obtenemos que

$$\gamma_n = (1 - \delta_n) \gamma_{n-1} = \dots = (1 - \delta_n) \dots (1 - \delta_1) \gamma_0 = (1 - \delta_n) \dots (1 - \delta_1) \left(\frac{1}{k_0 + k} - L_0 c \right) > 0$$

De la ecuación precedente deducimos que $\gamma_n > 0$ para cada $n = 1, 2, \dots$ y

$$\gamma_n < \gamma_{n-1} \forall n = 1, 2, \dots,$$

desde $\delta_n < 1$. Por lo tanto, la sucesión $\{\gamma_n\}$ converge a su único límite inferior más grande denotado por γ^* . También tenemos que $\beta_n = \frac{1}{k_0 + k} - \gamma_n < \frac{1}{k_0 + k}$. En consecuencia, la sucesión $\{\beta_n\}$ es creciente, delimitada superiormente por $\frac{1}{k_0 + k}$ y como tal, converge a su único límite superior mínimo denotado por β^* . Entonces, a vista de (4.2.5) la sucesión $\{\alpha_n\}$ es también creciente, delimitada por superiormente por $\frac{L_0^{-1}}{k_0 + k}$ y, por lo tanto, también converge a su único límite superior mínimo denotado por α^* . □

Lema 4.3 *Supongamos que se cumplen (4.2.13) y (4.2.14). Entonces, las siguientes afirmaciones son verdaderas para cada $n = 1, 2, \dots$*

$$\delta_n = \delta$$

$$\gamma_n = (1 - \delta)^n \gamma_0, \quad \gamma^* = \lim_{n \rightarrow \infty} \gamma_n = 0,$$

$$\beta_n = \frac{1}{k_0 + k} - (1 - \delta)^n \gamma_0, \quad \beta^* = \lim_{n \rightarrow \infty} \beta_n = \frac{1}{k_0 + k}$$

y

$$\alpha_n = \frac{1}{L_0} \left[\frac{1}{k_0 + k} - (1 - \delta)^n \gamma_0 \right], \quad \alpha^* = \lim_{n \rightarrow \infty} \alpha_n = \frac{1}{L_0(k_0 + k)}$$

Corolario 4.4 *Supongamos ciertas las hipótesis del Lema 4.1 y del Lema 4.2. entonces, la sucesión $\{\alpha_n\}$ definida en (4.2.1) es no decreciente y converge a*

$$\alpha^* = \beta^*(1 + k_0 c).$$

A continuación, presentamos los límites inferior y superior en el punto límite α^* .

Lema 4.5 *Supongamos que se satisface la condición (4.2.16). Entonces, la siguiente afirmación es cierta*

$$b_1^1 \leq \alpha^* \leq b_2^1, \quad (4.2.26)$$

donde

$$b_1^1 = \frac{1 + k_0 c}{1 - k c} \left[\frac{1}{k_0 + k} - \exp \left(-2 \left(\frac{\delta_1}{2 - \delta_1} + \frac{\delta_2}{2 - \delta_2} \right) \right) \right],$$

$$b_2^1 = \frac{1 + k_0 c}{1 - k c} \left[\frac{1}{k_0 + k} - \exp(\delta^*) \right], \quad (4.2.27)$$

$$\delta^* = - \left[\frac{1}{1 - \delta_1} \left(\delta_1 + \frac{\delta_2}{1 - r} \right) + \ln \left(\frac{(k_0 + k)(1 - k_0 c)}{1 - k c} \right) \right]$$

y

$$r = k_1 \frac{\lambda \delta_1 + \delta_2 (1 - \delta_1)}{(1 - \delta_1)(1 - \delta_2)(k + k_0(1 - \delta_2))}.$$

Demostración. Usando (4.2.16) tenemos que $0 < \delta_3 < \delta_2 < \delta_1$. Supongamos que $0 < \delta_{k+1} < \delta_k < \dots < \delta_1$. Entonces, se deduce de las hipótesis de inducción que

$$\delta_{k+2} = \delta_{k+1} k_1 \frac{\delta_k + \delta_{k+1}(1 - \delta_k)}{(1 - \delta_k)(1 - \delta_{k+1})(2 - \delta_{k+1})} < r \delta_{k+1} < r^2 \delta_k \leq \dots \leq r^{k-1} \delta_3 \leq r^k \delta_2.$$

Tenemos que

$$\gamma^* = \lim_{n \rightarrow \infty} \gamma_n = \prod_{i=1}^{\infty} (1 - \delta_n) \gamma_0.$$

Es equivalente a

$$\ln\left(\frac{1}{\gamma^*}\right) = \sum_{n=1}^{\infty} \ln\left(\frac{1}{1-\delta_n}\right) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right),$$

recordando que $\gamma_0 = \frac{1-kc}{(k_0+k)(1+k_0c)}$. Usaremos los siguientes límites para $\ln t$, $t > 1$:

$$2\left(\frac{t-1}{t+1}\right) \leq \ln t \leq \frac{t^2-1}{2t}.$$

Primero, encontraremos un límite superior para $\ln(1/\gamma^*)$. Tenemos que

$$\begin{aligned} \ln(1/\gamma^*) &\leq \sum_{n=1}^{\infty} \frac{\delta_n(2-\delta_n)}{2(1-\delta_n)} + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\ &\leq \frac{1}{1-\delta_1} \sum_{n=1}^{\infty} \delta_n + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\ &\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2 + \delta_3 + \dots) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\ &\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2 + r\delta_2 + \dots + r^n\delta_2 + \dots) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\ &\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2(r+r^2+\dots+r^n+\dots)) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\ &\leq \frac{1}{1-\delta_1} \left(\delta_1 + \frac{\delta_2}{1-r}\right) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) = -\delta^*. \end{aligned}$$

Como $\beta^* = 1/(k_0+k) - \gamma^*$ y $\alpha^* = L_0^{-1}\beta^*$, obtenemos el límite superior de 4.2.24. Además, para obtener el límite inferior para $\ln(1/\gamma^*)$, tenemos que

$$\ln(1/\gamma^*) \geq 2 \sum_{n=1}^{\infty} \frac{\delta_n}{2-\delta_n} > 2 \left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2} \right),$$

lo que implica el límite inferior por 4.2.24. \square

A partir de ahora, denotaremos por (C^1) a las hipótesis del Lema 4.1 y del Lema 4.2.

Nota 4.6 (a) *Vamos a introducir algo de notación necesaria*

$$c^N = \alpha_{N-1} - \alpha_{N-2}, \quad \nu^N = \alpha_N - \alpha_{N-1}$$

para algún entero $N \geq 1$. Notemos que $c^1 = \alpha_0 - \alpha_{-1} = c$ y $\nu^1 = \alpha_1 - \alpha_0 = \nu$. Los resultados de los Lemas precedentes pueden suavizarse de la siguiente manera. Consideremos el criterio de convergencia (C_*^N) para $N > 1$: (C^1) con c, ν reemplazado por c^N, ν^N , respectivamente

$$\alpha_{-1} < \alpha_0 < \alpha_1 < \dots < \alpha_N < \alpha_{N+1},$$

$$k_0(\alpha_{N+1} - c^N) + k\alpha_N < 1.$$

Entonces, los resultados precedentes se mantienen con $c, \nu, \delta_1, \delta_2, b_1^1, b_2^1$ reemplazados, respectivamente por $c^N, \nu^N, \delta_N, \delta_{N+1}, b_1^N, b_2^N$.

(b) Notemos que si

$$k_0(\alpha_{N+1} - c^N) + k\alpha_N < 1 \text{ se mantienen para cada } n = 0, 1, 2, \dots, \quad (4.2.28)$$

entonces, de (4.2.1) se sigue que la sucesión $\{\alpha_n\}$ es creciente, con límite superior $\frac{1+k_0c}{k_0+k}$ y como tal, converge a su único límite superior mínimo α^* . El criterio (4.2.28) es el criterio más débil de todos los criterios de convergencia anteriores para la sucesión $\{\alpha_n\}$. Obviamente todos los criterios anteriores implican (4.2.28). Finalmente, definimos el criterio para $N \geq 1$

$$(I^N) = \begin{cases} (C_*^N) \\ (4.2.28) \text{ si } (C_*^N) \text{ falla.} \end{cases} \quad (4.2.29)$$

Lema 4.7 Supongamos ciertas las condiciones de (4.2.16). Entonces, las siguientes afirmaciones son verdaderas

$$b_1^1 \leq \alpha^* \leq b_2^1, \quad (4.2.30)$$

donde

$$\begin{aligned} b_1^1 &= \frac{1+k_0c}{1-kc} \left[\frac{1}{k_0+k} - \exp\left(-2\left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2}\right)\right) \right], \\ b_2^1 &= \frac{1+k_0c}{1-kc} \left[\frac{1}{k_0+k} - \exp(\delta^*) \right], \\ \delta^* &= - \left[\frac{1}{1-\delta_1} \left(\delta_1 + \frac{\delta_2}{1-r} \right) + \ln \left(\frac{(k_0+k)(1+k_0c)}{1-kc} \right) \right] \end{aligned} \quad (4.2.31)$$

y

$$r = k_1 \frac{\lambda\delta_1 + \delta_2(1-\delta_1)}{(1-\delta_1)(1-\delta_2)(k+k_0(1-\delta_2))}.$$

Demostración. Usando (4.2.16) tenemos que $0 < \delta_3 < \delta_2 < \delta_1$. Asumiendo que $0 < \delta_{k+1} < \delta_k < \dots < \delta_1$. Entonces, se deduce de las hipótesis de inducción que

$$\delta_{k+2} = \delta_{k+1} k_1 \frac{\delta_k + \delta_{k+1}(1-\delta_k)}{(1-\delta_k)(1-\delta_{k+1})(2-\delta_{k+1})} < r\delta_{k+1} < r^2\delta_k \leq \dots \leq r^{k-1}\delta_3 \leq r^k\delta_2.$$

Tenemos

$$\gamma^* = \lim_{n \rightarrow \infty} \gamma_n = \prod_{i=1}^{\infty} (1-\delta_n)\gamma_0.$$

Que es equivalente a

$$\ln\left(\frac{1}{\gamma^*}\right) = \sum_{n=1}^{\infty} \ln\left(\frac{1}{1-\delta_n}\right) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right),$$

recordando que $\gamma_0 = \frac{1-kc}{(k_0+k)(1+k_0c)}$. Usaremos los siguientes límites para $\ln t$, $t > 1$:

$$2\left(\frac{t-1}{t+1}\right) \leq \ln t \leq \frac{t^2-1}{2t}.$$

Primero, encontraremos un límite superior para $\ln(1/\gamma^*)$. Tenemos que

$$\begin{aligned}
\ln(1/\gamma^*) &\leq \sum_{n=1}^{\infty} \frac{\delta_n(2-\delta_n)}{2(1-\delta_n)} + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\
&\leq \frac{1}{1-\delta_1} \sum_{n=1}^{\infty} \delta_n + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\
&\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2 + \delta_3 + \dots) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\
&\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2 + r\delta_2 + \dots + r^n\delta_2 + \dots) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\
&\leq \frac{1}{1-\delta_1} (\delta_1 + \delta_2(r+r^2+\dots+r^n+\dots)) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) \\
&\leq \frac{1}{1-\delta_1} \left(\delta_1 + \frac{\delta_2}{1-r}\right) + \ln\left(\frac{(k_0+k)(1+k_0c)}{1-kc}\right) = -\delta^*.
\end{aligned}$$

Como $\beta^* = 1/(k_0+k) - \gamma^*$ y $\alpha^* = L_0^{-1}\beta^*$, obtenemos el límite superior de (4.2.30). Además, para obtener el límite inferior para $\ln(1/\gamma^*)$, tenemos que

$$\ln(1/\gamma^*) \geq 2 \sum_{n=1}^{\infty} \frac{\delta_n}{2-\delta_n} > 2 \left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2} \right),$$

lo cual implica el límite inferior de (4.2.30). \square

4.3. Convergencia semilocal del método de la Secante

En esta sección, primero presentamos la convergencia semilocal del método de la Secante utilizando $\{\alpha_n\}$ (definida en (4.2.1)) como una sucesión mayorante. Sea $U(x, R)$ que representa una bola abierta centrada en $x \in X$ con radio $R > 0$. Denotemos su clausura por $\bar{U}(x, R)$. Estudiaremos el método de la Secante para las ternas $(\mathcal{F}, x_{-1}, x_0)$ pertenecientes a la clase $\mathcal{K} = \mathcal{K}(\nu, c, k, k_0, k_1, k_2)$ definidas de la siguiente manera.

Definición 4.8 Sean ν, c, k, k_0, k_1, k_2 constantes que satisfacen las hipótesis (I^N) para algún entero fijo $N \geq 1$. Una terna $(\mathcal{F}, x_{-1}, x_0)$ pertenece a la clase $\mathcal{K} = \mathcal{K}(\nu, c, k, k_0, k_1, k_2)$ si:

(\mathbb{D}_1) \mathcal{F} es un operador no lineal definido en un subconjunto convexo D de un espacio de Banach \mathcal{X} con valores en un espacio de Banach \mathcal{Y} .

(\mathbb{D}_2) x_{-1} y x_0 son dos puntos pertenecientes al interior D^0 de D que satisfacen la inecuación

$$\|x_0 - x_{-1}\| \leq c,$$

para alguna constante $c \geq 0$.

(\mathbb{D}_3) \mathcal{F} es Fréchet-diferenciable en D^0 y existe un operador $\delta\mathcal{F} : \mathbb{D}^0 \times D^0 \rightarrow \mathbb{L}(X, Y)$ tal que $\delta\mathcal{F}(x, y)(x - y) = F(x) - F(y)$ para cada $x \neq y$, $\delta\mathcal{F}(x, x) = F'(x)$, $x \in D^0$, $\mathcal{A}^{-1} = \delta\mathcal{F}(x_0, x_{-1})^{-1} \in \mathcal{L}(Y, X)$ para todos $x, y, z \in D$ entonces, la siguiente afirmación

$$\|\mathcal{A}^{-1}\mathcal{F}(x_0)\| \leq \nu,$$

$$\|\mathcal{A}^{-1}(\delta\mathcal{F}(x, y) - \delta\mathcal{F}(z, w))\| \leq k_1\|x - z\| + k_2\|y - w\|$$

y

$$\|\mathcal{A}^{-1}(\delta\mathcal{F}(x, y) - \mathcal{A})\| \leq k_0\|x - x_0\| + k\|y - x_{-1}\|$$

para algunas constantes $k > 0$, $k_0 > 0$, $k_1 > 0$, $k_2 \geq 0$ y $\nu \geq 0$.

(\mathcal{D}_4)

$$\bar{U}(x_0, \alpha^* - c) \subseteq D,$$

donde α^* es la dada en el Lema 4.3.

A continuación, presentamos el resultado de convergencia semilocal para el método de la Secante.

Teorema 4.9 Si $(\mathcal{F}, x_{-1}, x_0) \in \mathcal{K}(\nu, c, k, k_0, k_1, k_2)$ entonces, la sucesión $\{x_n\}$ ($n \geq -1$) generada por el método de la Secante está bien definida, permanece en $\bar{U}(x_0, \alpha_0^*)$ para cada $n = 0, 1, 2, \dots$ y converge a la solución única $x^* \in \bar{U}(x_0, \alpha^* - c)$ de (4.1.1). Además, las siguientes afirmaciones son ciertas para cada $n = 0, 1, 2, \dots$

$$\|x_n - x_{n-1}\| \leq \alpha_n - \alpha_{n-1} \quad (4.3.1)$$

y

$$\|x^* - x_n\| \leq \alpha^* - \alpha_n, \quad (4.3.2)$$

donde la sucesión $\{\alpha_n\}$ ($n \geq 0$) es la dada en (4.2.1). Además, si existe R tal que

$$\bar{U}(x_0, R) \subseteq D, \quad R \geq \alpha^* - c \text{ y } k_0(\alpha^* - \alpha_0) + k(R + c) < 1, \quad (4.3.3)$$

entonces, la solución x^* es única en $\bar{U}(x_0, R)$.

Primero demostraremos (4.3.1). Si (4.3.1) es cierto para todos $n \leq m$ y si $\{x_n\}$ ($n \geq 0$) está bien definida para $n = 0, 1, 2, \dots, m$, entonces

$$\|x_n - x_0\| \leq \alpha_n - \alpha_0 < \alpha^* - \alpha_0, \quad n \leq m. \quad (4.3.4)$$

Y (4.1.2) está bien definida para $n = m + 1$. Para $n = -1$ y $n = 0$, (4.3.1) se reduce a $\|x_{-1} - x_0\| \leq c$ y $\|x_0 - x_{-1}\| \leq \nu$. Supongamos cierto (4.3.1) para $n = -1, 0, 1, \dots, m$ ($m \geq 0$). A continuación, demostramos que $\mathcal{A}m + 1 =$

$\delta\mathcal{F}(x_{m+1}, x_m)$ es inversible para $x_m, x_{m+1} \in \overline{U}(x_0, \alpha^* - c)$. Por (\mathbb{D}_2) y (\mathbb{D}_3) , tenemos que

$$\begin{aligned} \|I - \mathcal{A}^{-1}\mathcal{A}_{m+1}\| &= \|\mathcal{A}^{-1}(\mathcal{A}_{m+1} - \mathcal{A})\| \\ &\leq k_0\|x_{m+1} - x_0\| + k\|x_m - x_{m-1}\| \\ &\leq k_0\|x_{m+1} - x_0\| + k(\|x_m - x_0\| + \|x_0 - x_{-1}\|) \\ &\leq k_0(\alpha_{m+1} - c) + k(\alpha_{m+1} - c + c) < 1. \end{aligned} \quad (4.3.5)$$

Usando el Lema de Banach para operadores inversibles y (4.3.5), deducimos que \mathcal{A}_{m+1} es inversible y

$$\|\mathcal{A}_{m+1}^{-1}\mathcal{A}\| \leq (1 - k_0(\alpha_{m+1} - c) + k\alpha_m)^{-1}. \quad (4.3.6)$$

Por (4.3.6) y

$$\mathcal{F}(x_{m+1}) = \mathcal{F}(x_{m+1}) - \mathcal{F}(x_m) - \mathcal{A}_m(x_{m+1} - x_m) \quad (4.3.7)$$

obtenemos a su vez las siguientes estimaciones

$$\begin{aligned} \|\mathcal{A}^{-1}\mathcal{F}(x_{m+1})\| &= \|\mathcal{A}^{-1}(\delta\mathcal{F}(x_{m+1}, x_k) - \mathcal{A}_m)(x_{m+1} - x_m)\| \\ &\leq (k_1\|x_{m+1} - x_m\| + k_2\|x_m - x_{m-1}\|)\|x_{m+1} - x_m\| \\ &\leq (k_1(\alpha_{m+1} - \alpha_m) + k_2(\alpha_m - \alpha_{m-1}))(\alpha_{m+1} - \alpha_m) \end{aligned} \quad (4.3.8)$$

y

$$\begin{aligned} \|x_{m+2} - x_{m+1}\| &= \|\mathcal{A}_{m+1}^{-1}\mathcal{F}(x_{m+1})\| \\ &\leq \|\mathcal{A}_{m+1}^{-1}\mathcal{A}\| \|\mathcal{A}^{-1}\mathcal{F}(x_{m+1})\| \\ &\leq \frac{(k_1(\alpha_{m+1} - \alpha_m) + k_2(\alpha_m - \alpha_{m-1}))(\alpha_{m+1} - \alpha_m)}{1 - (k_0(\alpha_{m+1} - c) + kc)} \\ &\leq \alpha_{m+2} - \alpha_{m+1}. \end{aligned}$$

Completando así la inducción para (4.3.1). Se sigue de (4.3.1) y del Lema 4.1 que $\{x_n\}$ ($n \geq -1$) es una secuencia completa en un espacio de Banach \mathbb{X} y como tal converge a algún $x^* \in \overline{U}(x_0, \alpha^* - c)$ ya que $\overline{U}(x_0, \alpha^* - c)$ es un conjunto cerrado. Tomando $k \rightarrow \infty$ en (4.3.8), obtenemos $\mathcal{F}(x^*) = 0$. Además, la estimación (4.3.2) se sigue de (4.3.1) mediante el uso de técnicas estándar de sucesiones mayorizantes [16, 23, 128, 151, 215]. Finalmente, para demostrar la unicidad en $\overline{U}(x_0, R)$, sea $y^* \in \overline{U}(x_0, R)$ una solución (4.1.1). Tomemos

$$\mathcal{T} = \delta F(x^*, y^*)$$

Usando (\mathcal{D}_3) y (4.3.3) obtenemos a su vez que

$$\begin{aligned} \|\mathcal{A}^{-1}(\mathcal{A} - \mathcal{T})\| &= k_0\|x^* - x_0\| + k\|y^* - x_{-1}\| \\ &\leq k_0(\alpha^* - \alpha_0) + k(R + c). \end{aligned} \quad (4.3.9)$$

De (4.3.9) y del Lema de Banach para operadores inversibles se tiene que \mathcal{T}^{-1} existe. Usando la identidad:

$$\mathcal{F}(x^*) - \mathcal{F}'(y^*) = \mathcal{T}(x^* - y^*), \quad (4.3.10)$$

deducimos que $x^* = y^*$. \(\square\)

Nota 4.10 *Se sigue de la demostración del Teorema 4.9 que la sucesión $\{r_n\}$ definida por*

$$\begin{cases} r_{-1} = 0, r_0 = c, r_1 = c + \nu \\ r_2 = r_1 + \frac{(k_0(r_1-r_0)+k(r_0-r_{-1}))(r_1-r_0)}{1-(k_0(r_1-c)+kc)} \\ r_{n+2} = r_{n+1} + \frac{(k_1(r_{n+1}-r_n)+k_2(r_n-r_{n-1}))(r_{n+1}-r_n)}{1-(k_0(r_{n+1}-c)+kr_n)} \end{cases} \quad (4.3.11)$$

es una sucesión mayorante más precisa para $\{x_n\}$. Obviamente, la sucesión $\{r_n\}$ también converge bajo las hipótesis (I^N) .

Un simple argumento inductivo muestra que si $k_0 < k_1$ o $k < k_2$ para cada $n = 2, 3, \dots$

$$r_n < \alpha_n \quad (4.3.12)$$

$$r_{n+1} - r_n < \alpha_{n+1} - \alpha_n \quad (4.3.13)$$

y

$$r^* = \lim_{n \rightarrow \infty} r_n \leq \alpha^* = \lim_{n \rightarrow \infty} \alpha_n. \quad (4.3.14)$$

Tengamos en cuenta también que la sucesión $\{r_n\}$ puede converger bajo hipótesis incluso más débiles. El criterio suficiente de convergencia (4.2.13) determina lo pequeño que pueden ser c y r . Este criterio se puede resolver para c y r . De hecho, demostraremos las ventajas en dos casos populares:

Caso 1: Método de Newton (i. e., si $c = 0, k_0 = k, k_1 = k_2$). Entonces, se puede ver fácilmente que $\{s_n\}$ (y en consecuencia $\{r_n\}$) converge siempre que (ver también [11])

$$h_2 = x^* i_2 \nu \leq 1, \quad (4.3.15)$$

donde

$$x^* i_2 = \frac{1}{4} \left(4k_0 + \sqrt{k_0 k_1} + \sqrt{k_0 k_1 + 8k_0^2} \right), \quad (4.3.16)$$

mientras que la sucesión $\{x_n\}$ converge, si

$$h_1 = k_1 \nu \leq 1 \quad (4.3.17)$$

donde

$$x^* i_1 = \frac{1}{4} \left(4k_0 + k_1 + \sqrt{k_0^2 + 8k_1 k_0} \right), \quad (4.3.18)$$

en el caso $k_0 = k_1$, obtenemos el famoso por su simplicidad y claridad criterio de suficiencia de convergencia de Kantorovich [16, 128] dado por

$$h = 2k_1 \nu \leq 1. \quad (4.3.19)$$

Observe sin embargo que

$$h \leq 1 \Rightarrow h_1 \leq 1 \Rightarrow h_2 \leq 1 \quad (4.3.20)$$

pero no necesariamente viceversa a menos que $k_0 = k_1$. Además, tenemos que

$$\frac{h_1}{h} \rightarrow \frac{1}{4}, \quad \frac{h_1}{h} \rightarrow 0, \quad \frac{h_2}{h_1} \rightarrow 0 \text{ y } \frac{k_0}{k_1} \rightarrow 0 \quad (4.3.21)$$

Caso 2: Método de la Secante Schmidt [128], Potra-Ptáček [59], Dennis [23], usaron la sucesión mayorante $\{r_n\}$ para $k_0 = k = k_1 = k_2$. Es decir, usaron la sucesión $\{t_n\}$ dada por

$$\begin{cases} t_{-1} = 0, \quad t_0 = c, \quad t_1 = c + \nu \\ t_{n+2} = t_{n+1} + \frac{k_1(t_{n+1}-t_{n-1})(t_{n+1}-t_n)}{1-k_1(t_n-t_{n+1}+c)} \end{cases} \quad (4.3.22)$$

mientras que nuestra sucesión $\{\alpha_n\}$ para $k_0 = k$ y $k_1 = k_2$ [16, 23, 21, 31, 20] se reduce a

$$\begin{cases} \alpha_{-1} = 0, \quad \alpha_0 = c, \quad \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \frac{k_1(\alpha_{n+1}-\alpha_{n-1})(\alpha_{n+1}-\alpha_n)}{1-k_0(\alpha_{n+1}-\alpha_n+c)} \end{cases} \quad (4.3.23)$$

Entonces, en el caso $k_0 < k_1$ nuestra sucesión es más precisa (ver también (4.3.12)-(4.3.14)). Observamos también que en las referencias anteriores, el criterio de convergencia suficiente asociado a $\{t_n\}$ está dado por

$$k_1 c + 2\sqrt{k_1 \nu} \leq 1 \quad (4.3.24)$$

Nuestros criterios de suficiencia de la convergencia también pueden ser más suaves en este caso (ver también los ejemplos numéricos). No vale la pena si $c = 0$ (4.3.24) ya que se reduce a (4.3.19).

Se pueden hacer observaciones similares para otras elecciones de parámetros.

4.4. Ejemplos numéricos

Para ilustrar los resultados teóricos obtenidos en este estudio y que nuestros nuevos criterios de convergencia sean realmente aplicables, vamos a aplicarlo a los siguientes ejemplos.

Ejemplo 1 Sea $\mathbb{X} = \mathbb{Y} = \mathbb{R}$ y consideremos las funciones reales

$$F(x) = x^3 - B$$

donde $B \in (-1, 1)$ y vamos a aplicar el método de la Secante para hallar la solución de $F(x) = 0$. Tomaremos el punto inicial como $x_0 = 1$ y consideramos el dominio $\Omega = B(x_0, 1)$ y dejamos x_{-1} libre a la hora de encontrar una relación entre B y x_{-1} para los cuales los viejos criterios no son satisfechos pero sí se satisfacen los nuevos criterios. En este caso, obtenemos

$$c = |1 - x_{-1}|, \quad \mathcal{A} = \frac{1}{|1 + x_{-1} + x_{-1}^2|}, \quad \nu = \frac{|(1 - B)|}{|1 + x_{-1} + x_{-1}^2|},$$

$$k_1 = \frac{8}{|1 + x_{-1} + x_{-1}^2|}, \quad k_2 = \frac{4}{|1 + x_{-1} + x_{-1}^2|},$$

$$k_0 = \frac{6}{|1 + x_{-1} + x_{-1}^2|}, \quad k = \frac{3}{|1 + x_{-1} + x_{-1}^2|},$$

Tomando todos estos datos en cuenta obtenemos los siguientes criterios:

(i) Si $0.661558\dots < B \leq 0.752883\dots$ y $\varepsilon_1 < x_{-1} \leq \varepsilon_2$

(ii) si $0.752883\dots < B \leq 0.786472\dots$ y $\varepsilon_3 < x_{-1} \leq \varepsilon_2$

(iii) si $0.786472\dots < B \leq 0.880641\dots$ y

(a) $\varepsilon_1 < x_{-1} < \varepsilon_4$ o

(b) $\varepsilon_3 < x_{-1} \leq \varepsilon_2$

(iv) si $0.880641\dots < B < 1$ y

(a) $\varepsilon_5 \leq x_{-1} < \varepsilon_4$ o

(b) $\varepsilon_3 < x_{-1} \leq \varepsilon_2$

donde ε_1 es la segunda raíz más pequeña de

$$p_1(t) = -9 + 9B + 4t + 2t^2 + 3t^3 - t^4 + t^5,$$

ε_2 es la raíz real más pequeña de

$$p_2(t) = -6 - 16B + 54B^2 + (-22 + 46B)t + (-29 + 69B)t^2 + (7 - 23B)t^3 \\ + (-14 + 23B)t^4 - 14t^5 + 11t^6 - 6t^7 + t^8$$

ε_3 es la raíz real más pequeña de

$$p_3(t) = 49 + 32B - 126t + 67t^2 - 14t^3 + t^4.$$

ε_4 es la raíz real más grande de

$$p_3(t) = 17 + 32B - 126t + 67t^2 + 18t^3 + t^4$$

y ε_5 es la cuarta raíz real más pequeña de

$$p_4(t) = -112 - 154B + 54B^2 + (-62 + 46B)t + (-89 + 69B)t^2 + (-151 + 115B)t^3 + (6 + 23B)t^4 + 46t^5 + 51t^6 + 14t^7 + t^8.$$

Por ejemplo, elegimos $B = 0.9$ y $x_{-1} = 1.6$ y obtenemos

$$c = 0.6, \quad \mathcal{A}^{-1} = 0.193798, \quad \nu = 0.0193798, \quad L_0 = 0.589041 \dots,$$

$$k_1 = 1.55039 \dots, \quad k_2 = 0.775194 \dots, \quad k = 0.581395 \dots, \quad k_0 = 1.16279 \dots,$$

$$L_0 = 0.589041 \dots$$

y

$$\alpha^* = 0.973333.$$

Darse cuenta de que la hipótesis vieja no se cumple en este caso

$$k_1 c + 2\sqrt{k_1 \nu} = 1.27691 \dots \geq 1$$

pero sí se cumplen nuestro nuevo criterio

$$\delta_1 = 0.0519103 \dots,$$

$$\delta_2 = 0.0431279 \dots$$

y

$$\delta = 0.237876 \dots$$

Ejemplo 2 Ahora, para mostrar la aplicabilidad de nuestra teoría en un problema real, vamos a considerar la siguiente ecuación cuadrática que describe la fracción de nitrógeno-hidrógeno que se convierte en amoníaco, llamada conversión fraccionaria que se muestra en [103, 202]. En la Figura 4.1 se muestra el proceso de amoníaco (Tomado de www.essentialchemicalindustry.org/chemicals/ammonia.html)

Para 250 atm y 500°C, esta ecuación toma la forma:

$$f(x) = x^4 - 7.79075x^3 + 14.7445x^2 + 2.511x - 1.674$$

En este caso, tomamos el dominio $\Omega = B(x_0, R)$, donde $x_0 = 0.4495$ y $R = 0.5595$. Eligiendo $x_{-1} = 0.446$, es fácil ver que

$$k_1 = 1.60531 \dots, \quad k_2 = 1.44842 \dots, \quad k = 1.07909 \dots, \quad k_0 = 0.873607 \dots,$$

$$\nu = 0.155169 \dots, \quad c = 0.0035, \quad L_0 = 0.996952 \dots \text{ y } \alpha^* = 0.513678 \dots$$

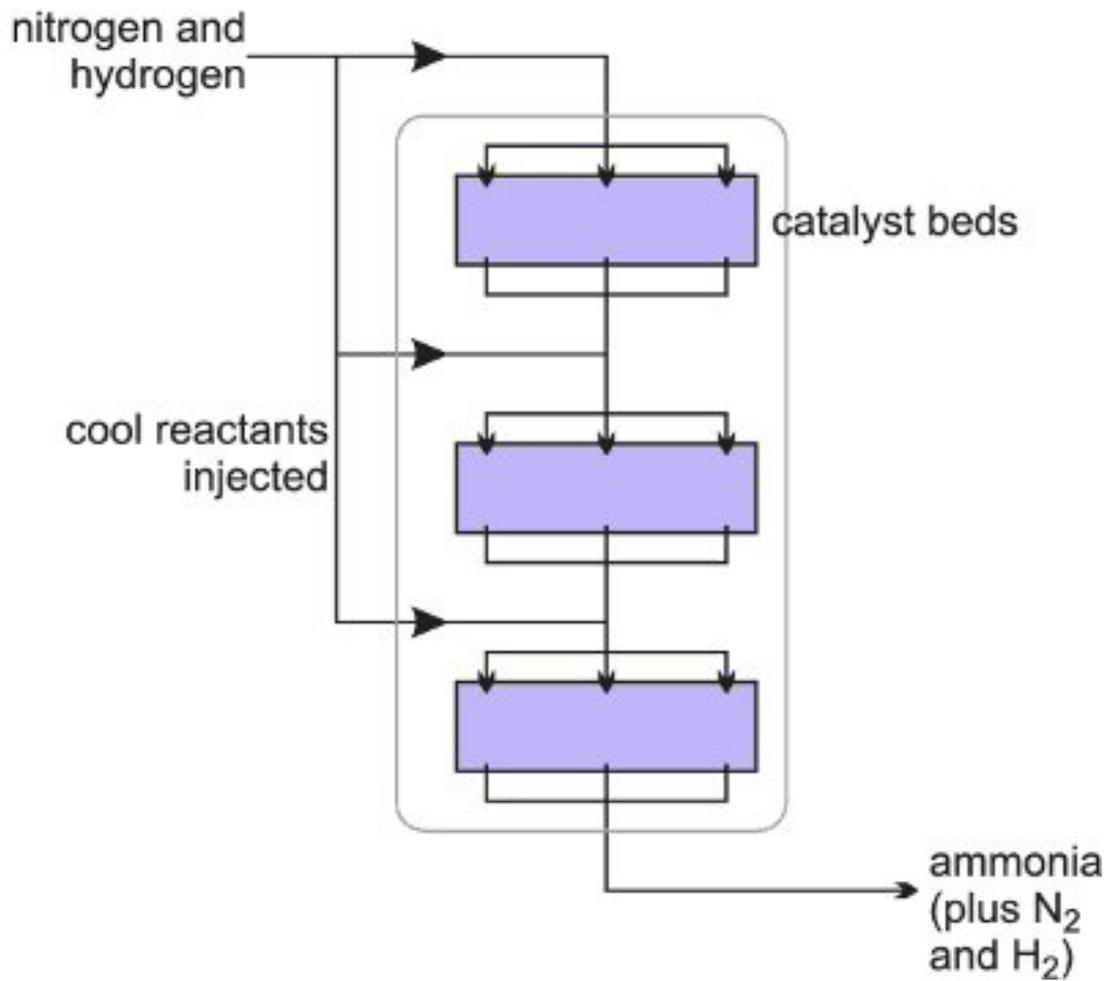


Figura 4.1: Proceso conocido como Ammonia.

Darse cuenta de que la hipótesis vieja no se cumple en este caso

$$k_1c + 2\sqrt{k_1\nu} = 1.00381\dots \geq 1$$

pero sí se cumplen nuestro nuevo criterio

$$\delta_1 = 0.311008\dots,$$

$$\delta_2 = 0.136734\dots$$

y

$$\delta = 0.311745\dots$$

Así que, podemos asegurar la convergencia del método de la Secante del Teorema 4.9.

Capítulo 5

Estudio de la convergencia y dinámica del método de la Secante II

Resumen

En este capítulo nos ocupamos del problema de aproximar una solución de ecuaciones no lineales utilizando el método de la Secante. Vamos a presentar un nuevo análisis de convergencia semilocal para el método de la Secante utilizando dominios de convergencia restringida. De acuerdo con esta idea, encontramos un dominio más preciso que los vistos en estudios anteriores donde existen las inversas de los operadores involucrados. De esta forma obtenemos constantes de Lipschitz más pequeñas que conducen a sucesiones mayorantes más precisas. Nuestros criterios de convergencia son más débiles y los límites de error son más precisos que en estudios anteriores. Aunque ya existen estudios sobre la convergencia semilocal, vamos aquí a obtener criterios de convergencia más débiles y estimaciones más precisas sobre las distancias $\|x_n - x_{n-1}\|$, $\|x_n - x^*\|$, donde x^* es solución local de la ecuación no lineal y x_n y x_{n-1} son dos aproximaciones consecutivas obtenidas mediante el método de la Secante. Bajo el mismo costo computacional en los parámetros involucrados, nuestro análisis incluye el cálculo de los límites en los puntos límite de las sucesiones mayorantes involucradas. También se presentan diferentes aplicaciones para ilustrar los resultados teóricos obtenidos en este estudio. Como puede ser el problema de la ley de radiación de Planck, que calcula la densidad de energía dentro de un cuerpo negro isotérmico.

5.1. Introducción

En este estudio nos ocupamos del problema de aproximar una solución única local x^* de la ecuación no lineal

$$F(x) = 0, \quad (5.1.1)$$

donde, F es un operador Fréchet diferenciable definido en un subconjunto no vacío \mathbb{D} de un espacio de Banach \mathbb{X} con valores en el espacio de Banach \mathbb{Y} . Varios problemas de Ciencias Aplicadas, incluyendo la Ingeniería, se pueden expresar como la ecuación (5.1.1) usando modelos matemáticos [10, 11, 13, 16, 57, 59, 74, 76, 89, 128, 151, 171, 185, 200, 215]. Pero las soluciones de estas ecuaciones solo en casos especiales pueden ser encontradas de manera exacta. Por lo que normalmente, los métodos de resolución suelen ser métodos iterativos.

En este capítulo consideramos la convergencia del método de la Secante definido como

$$x_{n+1} = x_n - \mathcal{A}_n^{-1}F(x_n), \quad \mathcal{A}_n = \delta F(x_n, x_{n-1}) \quad \text{para cada } n = 1, 2, \dots, \quad (5.1.2)$$

donde x_{-1}, x_0 son los puntos iniciales. Aquí $\mathcal{A}_n \in \mathbb{L}(\mathbb{X}, \mathbb{Y})$ es una aproximación de la derivada de Fréchet F' de F y $\mathbb{L}(\mathbb{X}, \mathbb{Y})$ representa el espacio de operadores lineales acotados de \mathbb{X} en \mathbb{Y} . Existen multitud de criterios de convergencia suficiente para el método de la Secante (5.1.2) bajo condiciones tipo Lipschitz (5.1.2) (ver [5],[200]). Es interesante darse cuenta de que aunque usamos sucesiones mayorizantes muy generales para $\{x_n\}$ nuestra técnica lleva en el caso semilocal a obtener criterios de convergencia suficientes más débiles y estimaciones más precisas sobre las distancias $\|x_n - x_{n-1}\|$, $\|x_n - x^*\|$ y al menos una información tan precisa sobre la ubicación de la solución x^* en muchos casos especiales interesantes como el método de Newton o el método de la Secante.

El resto del artículo está organizado de la siguiente manera. En la Sección 5.2 estudiamos la convergencia de las sucesiones mayorantes para $\{x_n\}$ involucradas en el método de la Secante. En la Sección 5.3 presentamos el análisis de convergencia semilocal para $\{x_n\}$. Finalmente, en la sección final 5.4 se presentan ejemplos numéricos.

5.2. Sucesiones mayorizantes para el método de la Secante

En esta sección, primero estudiaremos algunas sucesiones escalares que están relacionadas con el método de la Secante.

Tomemos los parámetros $c \geq 0$, $\nu \geq 0, k > 0$, $k_0 > 0$, $k_1 > 0$ y $k_2 \geq 0$. Definamos la sucesión escalar $\{\alpha_n\}$ por

$$\begin{cases} \alpha_{-1} = 0, \alpha_0 = c, \alpha_1 = c + \nu, \\ \alpha_{n+2} = \alpha_{n+1} + \frac{k_1(\alpha_{n+1} - \alpha_n) + k_2(\alpha_n - \alpha_{n-1})(\alpha_{n+1} - \alpha_n)}{1 - [k_0(\alpha_{n+1} - c) + k\alpha_n]}, \forall n = 0, 1, 2, \dots \end{cases} \quad (5.2.1)$$

Algunos autores han utilizado casos especiales de la sucesión $\{\alpha_n\}$ como sucesiones mayorizantes para el método de la Secante. Por ejemplo:

Caso 1 (Método de la Secante) $k_0 = k$ y $k_1 = k_2$ ha sido estudiado en [10, 16, 23, 57, 74, 79, 104, 128, 151, 171, 185, 186, 193, 215] y para $k_0 = k$, $k_1 = k_2$ y $k_0 \leq k_1$ en [27, 29, 30].

Caso 2 (Método de Newton) $k_2 = 0$, $k = 0$, $c = 0$ y $k_0 = k_1$ ha sido estudiado en [10, 16, 23, 30, 59, 76, 89, 79, 104, 128, 151, 171, 185, 186, 200, 215] y para $k_0 \leq k_1$ en [10, 11, 13].

En el presente trabajo estudiaremos la convergencia de la sucesión $\{\alpha_n\}$ simplificándola primero. De hecho, el propósito de las siguientes transformaciones es estudiar la sucesión (5.2.1) después de usar sucesiones más fáciles de estudiar definidas por (5.2.3), (5.2.6) y (5.2.8). Sean

$$\lambda = \frac{k_2}{k_1}, \quad L_0 = \frac{1}{1 + k_0c} \text{ y } L = \frac{k_1}{1 + k_0c}. \quad (5.2.2)$$

Usando (5.2.1) y (5.2.2), podemos escribir la sucesión $\{\alpha_n\}$ como

$$\begin{cases} \alpha_{-1} = 0, \alpha_0 = c, \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \frac{L(\alpha_{n+1} - \alpha_n + \lambda(\alpha_n - \alpha_{n-1}))(\alpha_{n+1} - \alpha_n)}{1 - L_0(k_0\alpha_{n+1} + k\alpha_n)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (5.2.3)$$

Debemos darnos cuenta que

$$L = k_1 L_0 \quad (5.2.4)$$

y

$$\beta_n = L_0 \alpha_n. \quad (5.2.5)$$

Entonces, podemos definir la sucesión $\{\beta_n\}$ como

$$\begin{cases} \beta_{-1} = 0, \beta_0 = L_0c, \beta_1 = L_0(c + \nu) \\ \beta_{n+2} = \beta_{n+1} + \frac{k_1(\beta_{n+1} - \beta_n + \lambda(\beta_n - \beta_{n-1}))(\beta_{n+1} - \beta_n)}{1 - (k_0\beta_{n+1} + k\beta_n)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (5.2.6)$$

Además, sea

$$\gamma_n = \frac{1}{k_0 + k} - \beta_n, \forall n = 0, 1, 2, \dots \quad (5.2.7)$$

Entonces, la sucesión $\{\gamma_n\}$ está definida como

$$\begin{cases} \gamma_{-1} = \frac{1}{k_0+k}, \gamma_0 = \frac{1}{k_0+k} - L_0c, \gamma_1 = \frac{1}{k_0+k} - L_0(c + \nu) \\ \gamma_{n+2} = \gamma_{n+1} - \frac{k_1(\gamma_{n+1} - \gamma_n + \lambda(\gamma_n - \gamma_{n-1}))(\gamma_{n+1} - \gamma_n)}{k_0\gamma_{n+1} + k\gamma_n}, \forall n = 0, 1, 2, \dots \end{cases} \quad (5.2.8)$$

Finalmente, tomemos

$$\delta_n = 1 - \frac{\gamma_n}{\gamma_{n-1}}, \forall n = 0, 1, 2, \dots \quad (5.2.9)$$

Entonces, definimos la sucesión $\{\delta_n\}$ como

$$\begin{cases} \delta_0 = 1 - \frac{\gamma_0}{\gamma_{-1}}, \delta_1 = 1 - \frac{\gamma_1}{\gamma_0} \\ \delta_{n+2} = \frac{k_1\delta_{n+1}(\lambda\delta_n + (1 - \delta_n)\delta_{n+1})}{(1 - \delta_n)(1 - \delta_{n+1})(k_0(1 - \delta_{n+1}) + k)}, \forall n = 0, 1, 2, \dots \end{cases} \quad (5.2.10)$$

Para el estudio de la convergencia de la sucesión $\{\alpha_n\}$ es conveniente definir el polinomio p como

$$p(t) = k_0t^3 - (k_1 + 3k_0 + k)t^2 + (2k + 3k_0 + k_1(\lambda + 1))t - (k_0 + k). \quad (5.2.11)$$

Tenemos que $p(0) = -(k_0 + k) < 0$ y $p(1) = k_1\lambda$ para $\lambda > 0$. Del teorema del valor medio deducimos que p tiene raíces en $(0, 1)$. Denotamos la raíz más pequeña por δ . Si $\lambda = 0$, entonces $p(t) = (t - 1)(k_0t^2 - (k_1 + k + 2k_0)t + k_0 + k)$. Por lo tanto, podemos elegir la raíz más pequeña de p dada por

$$\frac{2k_0 + k_1 + k - \sqrt{(2k_0 + k_1 + k)^2 - 4k_0(k_0 + k)}}{2k_0} \in (0, 1)$$

que es δ en este caso.

Notemos también que

$$p(t) \leq 0 \text{ para cada } t \in (-\infty, \delta]. \quad (5.2.12)$$

A continuación, estudiamos la convergencia de esta sucesión a partir de $\{\delta_n\}$.

Lema 5.1 [30] *Sea $\delta_1 > 0$, $\delta_2 > 0$ y $k_1 > 0$ parámetros dados. Supongamos que*

$$0 < \delta_2 \leq \delta_1 \leq \delta, \quad (5.2.13)$$

donde δ es la definida como la menor raíz positiva del polinomio $p(t)$ que aparece en (5.2.11). Sea $\{\delta_n\}$ la sucesión escalar definida por (5.2.10). Entonces, las siguientes afirmaciones son ciertas:

(A₁) Si

$$\delta_1 = \delta_2 \quad (5.2.14)$$

entonces,

$$\delta_n = \delta, \forall n = 1, 2, 3, \dots \quad (5.2.15)$$

(A₂) Si

$$\delta_2 < \delta_1 < \delta \quad (5.2.16)$$

entonces, la sucesión $\{\delta_n\}$ es decreciente y converge a 0.

Lema 5.2 [30] *Supongamos que se satisface la hipótesis (5.2.16). Entonces, la sucesión $\{\gamma_n\}$ es una sucesión decreciente y convergente, $\{\alpha_n\}$ y $\{\beta_n\}$ son crecientes y convergentes.*

Lema 5.3 [30] *Supongamos ciertas (5.2.13) y (5.2.14). Entonces, las siguientes afirmaciones son verdaderas para cada $n = 1, 2, \dots$*

$$\delta_n = \delta$$

$$\gamma_n = (1 - \delta)^n \gamma_0, \quad \gamma^* = \lim_{n \rightarrow \infty} \gamma_n = 0,$$

$$\beta_n = \frac{1}{k_0 + k} - (1 - \delta)^n \gamma_0, \quad \beta^* = \lim_{n \rightarrow \infty} \beta_n = \frac{1}{k_0 + k}$$

y

$$\alpha_n = \frac{1}{L_0} \left[\frac{1}{k_0 + k} - (1 - \delta)^n \gamma_0 \right], \quad \alpha^* = \lim_{n \rightarrow \infty} \alpha_n = \frac{1}{L_0(k_0 + k)}$$

Corolario 5.4 [30] *Supongamos que se cumplen las hipótesis del Lema 2.1 y del Lema 2.2. Entonces, la sucesión $\{\alpha_n\}$ definida en (8.2.1) es no decreciente y converge a*

$$\alpha^* = \beta^*(1 + k_0 c).$$

A continuación, presentamos límites inferior y superior en el punto límite α^* .

Lema 5.5 [30] *Supongamos que se satisface la condición (5.2.16). Entonces, las siguientes afirmaciones son ciertas*

$$b_1^1 \leq \alpha^* \leq b_2^1, \quad (5.2.17)$$

donde

$$b_1^1 = \frac{1 + k_0 c}{1 - k c} \left[\frac{1}{k_0 + k} - \exp \left(-2 \left(\frac{\delta_1}{2 - \delta_1} + \frac{\delta_2}{2 - \delta_2} \right) \right) \right],$$

$$b_2^1 = \frac{1 + k_0 c}{1 - k c} \left[\frac{1}{k_0 + k} - \exp(\delta^*) \right], \quad (5.2.18)$$

$$\delta^* = - \left[\frac{1}{1 - \delta_1} \left(\delta_1 + \frac{\delta_2}{1 - r} \right) + \ln \left(\frac{(k_0 + k)(1 + k_0 c)}{1 - k c} \right) \right]$$

y

$$r = k_1 \frac{\lambda \delta_1 + \delta_2 (1 - \delta_1)}{(1 - \delta_1)(1 - \delta_2)(k + k_0(1 - \delta_2))}.$$

A partir de ahora, denotaremos por (C^1) las hipótesis de los Lema 2.1 y Lema 2.2.

Nota 5.6 (a) *Vamos a introducir la siguiente notación:*

$$c^N = \alpha_{N-1} - \alpha_{N-2}, \quad \nu^N = \alpha_N - \alpha_{N-1}$$

para algún entero $N \geq 1$. Notemos que $c^1 = \alpha_0 - \alpha_{-1} = c$ y $\nu^1 = \alpha_1 - \alpha_0 = \nu$. Los resultados en los Lemas anteriores pueden hacerse más débiles de la siguiente manera. Consideramos los criterios de convergencia (C_*^N) para $N > 1$: (C^1) con c, ν y los reemplazamos por c^N, ν^N , respectivamente

$$\alpha_{-1} < \alpha_0 < \alpha_1 < \dots < \alpha_N < \alpha_{N+1},$$

$$k_0(\alpha_{N+1} - c^N) + k\alpha_N < 1.$$

Entonces, los resultados anteriores se mantienen reemplazando $c, \nu, \delta_1, \delta_2, b_1^1, b_2^1$, respectivamente por $c^N, \nu^N, \delta_N, \delta_{N+1}, b_1^N, b_2^N$.

(b) *Notemos que si*

$$k_0(\alpha_{N+1} - c^N) + k\alpha_N < 1 \text{ es cierto para cada } n = 0, 1, 2, \dots, \quad (5.2.19)$$

entonces, de (5.2.1) se sigue que la sucesión $\{\alpha_n\}$ es creciente, acotada superiormente por $\frac{1+k_0c}{k_0+k}$ y como tal, converge a su único límite superior mínimo α^* . El criterio (5.2.19) es el más débil de todos los criterios de convergencia anteriores para la sucesión $\{\alpha_n\}$. Claramente todos los criterios anteriores implican (5.2.19). Para finalizar, definimos el criterio para $N \geq 1$

$$(I^N) = \begin{cases} (C_*^N) \\ (5.2.19) \text{ si el criterio } (C_*^N) \text{ falla.} \end{cases} \quad (5.2.20)$$

Lema 5.7 [30] *Supongamos que son ciertas las condiciones (5.2.16). Entonces, la siguiente afirmación es válida*

$$b_1^1 \leq \alpha^* \leq b_2^1, \quad (5.2.21)$$

donde

$$b_1^1 = \frac{1+k_0c}{1-kc} \left[\frac{1}{k_0+k} - \exp\left(-2\left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2}\right)\right) \right],$$

$$b_2^1 = \frac{1+k_0c}{1-kc} \left[\frac{1}{k_0+k} - \exp(\delta^*) \right], \quad (5.2.22)$$

$$\delta^* = - \left[\frac{1}{1-\delta_1} \left(\delta_1 + \frac{\delta_2}{1-r} \right) + \ln \left(\frac{(k_0+k)(1+k_0c)}{1-kc} \right) \right]$$

y

$$r = k_1 \frac{\lambda\delta_1 + \delta_2(1-\delta_1)}{(1-\delta_1)(1-\delta_2)(k+k_0(1-\delta_2))}.$$

5.3. Convergencia semilocal del método de la Secante

En esta sección, primero presentamos la convergencia semilocal del método de la Secante usando $\{\alpha_n\}$ (definida en (5.2.1)) como una sucesión mayorante. Tomemos $U(x, R)$ que representa una bola abierta centrada en $x \in X$ con radio $R > 0$. Denotemos por $\bar{U}(x, R)$ su clausura. Estudiaremos el método de la Secante para las ternas $(\mathcal{F}, x_{-1}, x_0)$ pertenecientes a la clase $\mathbb{K} = \mathbb{K}(\nu, c, k, k_0, k_1, k_2)$ definidas de la siguiente manera.

Definición 5.8 Sean ν, c, k, k_0, k_1, k_2 constantes que satisfacen las hipótesis (I^N) para algún entero fijado $N \geq 1$. Una terna $(\mathcal{F}, x_{-1}, x_0)$ pertenece a la clase $\mathbb{K} = \mathbb{K}(\nu, c, k, k_0, k_1, k_2)$ si:

(\mathbb{D}_1) \mathcal{F} es un operador no lineal definido en un subconjunto convexo D de un espacio de Banach \mathbb{X} con valores en el espacio de Banach \mathbb{Y} .

(\mathbb{D}_2) x_{-1} y x_0 son dos puntos que pertenecen al interior D^0 de D y satisfacen la inecuación

$$\|x_0 - x_{-1}\| \leq c,$$

para alguna constante $c \geq 0$.

(\mathbb{D}_3) \mathcal{F} es Fréchet diferenciable en D^0 y existe un operador $\delta\mathcal{F} : D^0 \times D^0 \rightarrow \mathfrak{L}(X, Y)$ tal que $\delta\mathcal{F}(x, y)(x - y) = F(x) - F(y)$ para cada $x \neq y$, $\delta\mathcal{F}(x, x) = F'(x)$, $x \in D^0$, $F'(x_0)^{-1}$, $\mathcal{A}^{-1} = \delta\mathcal{F}(x_0, x_{-1})^{-1} \in \mathfrak{L}(Y, X)$ para todos $x, y \in D$ entonces, es cierto,

$$\|\mathcal{A}^{-1}\mathcal{F}(x_0)\| \leq \nu,$$

$$\|F'(x_0)^{-1}(\delta\mathcal{F}(x, y) - F'(x_0))\| \leq k_0\|x - x_0\| + k\|y - x_0\|$$

y para cada $x, y, z \in D_0 := U(x_0, \frac{1}{k_0+k}) \cap D$

$$\|F'(x_0)^{-1}(\delta\mathcal{F}(x, y) - F'(z))\| \leq k_1\|x - z\| + k_2\|y - z\|$$

para algunas constantes $k > 0$, $k_0 > 0$, $k_1 > 0$, $k_2 \geq 0$ y $\nu \geq 0$.

(\mathbb{D}_4)

$$\bar{U}(x_0, \alpha^* - c) \subseteq D \text{ o } U(x_0, \frac{1}{k_0+k}) \subset D.$$

donde α^* es el dado en el Lema 2.3.

A continuación, presentamos el resultado de convergencia semilocal para el método de la Secante.

Teorema 5.1 Si $(\mathcal{F}, x_{-1}, x_0) \in \mathbb{K}(\nu, c, k, k_0, k_1, k_2)$ entonces, la sucesión $\{x_n\}$ ($n \geq -1$) generada por el método de la Secante está bien definida, permanece en $\bar{U}(x_0, \alpha_0^*)$ para cada $n = 0, 1, 2, \dots$ y converge a la única solución $x^* \in \bar{U}(x_0, \alpha^* - c)$ de (5.1.1). Además, las siguientes afirmaciones son válidas para cada $n = 0, 1, 2, \dots$

$$\|x_n - x_{n-1}\| \leq \alpha_n - \alpha_{n-1} \quad (5.3.1)$$

y

$$\|x^* - x_n\| \leq \alpha^* - \alpha_n, \quad (5.3.2)$$

donde la sucesión $\{\alpha_n\}$ ($n \geq 0$) es la dada en (5.2.1). Además, si existe R tal que

$$\bar{U}(x_0, R) \subseteq D, \quad R \geq \alpha^* - c \text{ y } k_0(\alpha^* - \alpha_0) + kR < 1, \quad (5.3.3)$$

entonces, la solución x^* es única en $\bar{U}(x_0, R)$.

Demostración

Basta tener en cuenta que cada iteración se encuentra en D_0 , que es más preciso y que contiene las iteraciones $\{x_n\}$ que D usada en [30], ya que $D_0 \subseteq D$. Entonces, la prueba es exactamente la misma que la de [30]. \square

Nota 5.9 Se sigue de la demostración del Teorema 3.2 que la sucesión $\{r_n\}$ definida como

$$\begin{cases} r_{-1} = 0, \quad r_0 = c, \quad r_1 = c + \nu \\ r_2 = r_1 + \frac{(k_0(r_1 - r_0) + k(r_0 - r_{-1}))(r_1 - r_0)}{1 - (k_0(r_1 - c) + kc)} \\ r_{n+2} = r_{n+1} + \frac{(k_1(r_{n+1} - r_n) + k_2(r_n - r_{n-1}))(r_{n+1} - r_n)}{1 - (k_0(r_{n+1} - c) + kr_n)} \end{cases} \quad (5.3.4)$$

es una sucesión mayorizante más precisa para $\{x_n\}$. Claramente, la sucesión $\{r_n\}$ también converge bajo las hipótesis (I^N) .

Un argumento inductivo simple demuestra que si $k_0 < k_1$ o $k < k_2$ para cada $n = 2, 3, \dots$

$$r_n < \alpha_n \quad (5.3.5)$$

$$r_{n+1} - r_n < \alpha_{n+1} - \alpha_n \quad (5.3.6)$$

y

$$r^* = \lim_{n \rightarrow \infty} r_n \leq \alpha^* = \lim_{n \rightarrow \infty} \alpha_n. \quad (5.3.7)$$

Debemos tener en cuenta también que la sucesión $\{r_n\}$ podría converger incluso bajo hipótesis más débiles. El criterio suficiente de convergencia (5.2.20) determina la pequeñez de c y r . Este criterio se puede resolver por c y r (consulte por ejemplo (5.3.17)). De hecho, demostremos las ventajas en dos casos populares:

Caso 1: Método de Newton (i. e., si $c = 0, k_0 = k, k_1 = k_2$). Entonces, se puede ver fácilmente que $\{s_n\}$ (y en consecuencia $\{r_n\}$) converge siempre que (ver también [23])

$$h_2 = x^* i_2 \nu \leq 1, \quad (5.3.8)$$

donde

$$x^* i_2 = \frac{1}{4} \left(4k_0 + \sqrt{k_0 k_1} + \sqrt{k_0 k_1 + 8k_0^2} \right), \quad (5.3.9)$$

mientras que la sucesión $\{x_n\}$ converge, si

$$h_1 = k_1 \nu \leq 1 \quad (5.3.10)$$

donde

$$x^* i_1 = \frac{1}{4} \left(4k_0 + k_1 + \sqrt{k_0^2 + 8k_1 k_0} \right), \quad (5.3.11)$$

En el caso $k_0 = k_1$, obtenemos el famoso por su simplicidad y claridad criterio de convergencia suficiente de Kantorovich [16, 128] dado por

$$h = 2k_1 \nu \leq 1. \quad (5.3.12)$$

Observe sin embargo que

$$h \leq 1 \Rightarrow h_1 \leq 1 \Rightarrow h_2 \leq 1 \quad (5.3.13)$$

pero no necesariamente a viceversa a menos que $k_0 = k_1$. Además, tenemos que

$$\frac{h_1}{h} \rightarrow \frac{1}{4}, \quad \frac{h_2}{h} \rightarrow 0, \quad \frac{h_2}{h_1} \rightarrow 0 \text{ como } \frac{k_0}{k_1} \rightarrow 0 \quad (5.3.14)$$

Caso 2: Método de la Secante Schmidt [200], Potra-Ptáck [185], Dennis [74], Ezquerro et al. [79, 89], utilizaron la sucesión mayorante $\{r_n\}$ para $k_0 = k = k_1 = k_2$. Es decir, usaron la sucesión $\{t_n\}$ dada por

$$\begin{cases} t_{-1} = 0, \quad t_0 = c, \quad t_1 = c + \nu \\ t_{n+2} = t_{n+1} + \frac{k_1(t_{n+1} - t_{n-1})(t_{n+1} - t_n)}{1 - k_1(t_n - t_{n+1} + c)} \end{cases} \quad (5.3.15)$$

mientras que nuestra sucesión $\{\alpha_n\}$ para $k_0 = k$ y $k_1 = k_2$ [5, 16, 21, 23, 27, 29, 30, 57, 79, 89, 186, 200] se reduce a

$$\begin{cases} \alpha_{-1} = 0, \quad \alpha_0 = c, \quad \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \frac{k_1(\alpha_{n+1} - \alpha_{n-1})(\alpha_{n+1} - \alpha_n)}{1 - k_0(\alpha_{n+1} - \alpha_n + c)} \end{cases} \quad (5.3.16)$$

Entonces, en caso de que $k_0 < k_1$ nuestra sucesión es más precisa (ver también (5.3.5)-(5.3.7)). Observe también que en las referencias anteriores, el criterio de convergencia suficiente asociado a $\{t_n\}$ viene dado por

$$k_1 c + 2\sqrt{k_1} \nu \leq 1 \quad (5.3.17)$$

Nuestros criterios suficientes de convergencia son más débiles en este caso. No vale la pena si $c = 0$ ya que (5.3.17) se reduce a (5.3.12). Se pueden hacer observaciones similares para otras elecciones de parámetros.

Finalmente, observe que podemos obtener y usar constantes de Lipschitz aún más pequeñas, si simplemente trabajamos en $D_1 := U(x_1, \frac{1}{k_0+k} - \|x_1 - x_0\|)$ en vez de en D_0 , ya que todavía estamos usando los datos iniciales ($x_1 = x_0 - \mathcal{A}_0^{-1}F(x_0)$).

5.4. Aplicaciones

Aplicación 5.10 Sea $X = Y = \mathcal{C}[0, 1]$, el espacio de funciones continuas definidas en $[0, 1]$ equipado con la norma máxima. Sea $\Omega = \{x \in \mathcal{C}[0, 1]; \|x\| \leq R\}$, tal que $R > 1$ y F está definida en Ω y dada por

$$F(x)(s) = x(s) - f(s) - \Lambda \int_0^1 G(s, t)x(t)^3 dt, \quad x \in \mathcal{C}[0, 1], \quad s \in [0, 1],$$

donde $f \in \mathcal{C}[0, 1]$ es una función dada, Λ es una constante real y el núcleo G es la función de Green

$$G(s, t) = \begin{cases} (1-s)t, & t \leq s, \\ s(1-t), & s \leq t. \end{cases}$$

En este caso, para cada $x \in \Omega$, $F'(x)$ es un operador lineal definido en Ω mediante la siguiente expresión:

$$[F'(x)(v)](s) = v(s) - 3\Lambda \int_0^1 G(s, t)x(t)^2v(t) dt, \quad v \in \mathcal{C}[0, 1], \quad s \in [0, 1].$$

Si elegimos $x_0(s) = f(s) = 1$, $x_{-1} = 0.9$, se sigue que $\|I - F'(x_0)\| \leq 3|\Lambda|/8$. Así, si $|\Lambda| < 8/3$, $F'(x_0)^{-1}$ está definido y

$$\|F'(x_0)^{-1}\| \leq \frac{8}{8 - 3|\Lambda|}.$$

Además,

$$\|F(x_0)\| \leq \frac{|\Lambda|}{8}.$$

Definimos las diferencias divididas por

$$\delta F(x, y) = \int_0^1 F'(y + t(x - y))dt.$$

Tomando $\Lambda = 1$ and $R = 1.5$, tenemos

$$c = 0.1,$$

$$k_0 = 1,$$

$$k = 0.5,$$

$$\begin{aligned}
k_1 &= 1.06666\dots, \\
k_2 &= 0.533333\dots, \\
L_0 &= 0.909091\dots, \\
L &= 0.969697\dots
\end{aligned}$$

y

$$\nu = 0.2.$$

Además, como $\frac{1}{k_0+k} = 0.66666\dots$ obtenemos que

$$U(x_0, \frac{1}{k_0+k}) \subset D$$

Entonces, por la definición de δ_1, δ_2 y δ obtenemos

$$0 < \delta_2 = 0.164103\dots \leq \delta_1 = 0.315789\dots \leq \delta = 0.487078\dots$$

Entonces podemos asegurar la convergencia del método de la Secante con la solución de la ecuación. Recordar que las condiciones previas consideradas por otros autores no se cumplen ya que

$$k_1c + 2\sqrt{k_1\nu} = 1.03043\dots \geq 1$$

y la convergencia no estaba asegurada.

Aplicación 5.11 Consideramos el siguiente problema de la ley de radiación de Planck encontrado en [76]:

$$\varphi(\lambda) = \frac{8\pi cP\lambda^{-5}}{e^{\frac{cP}{\lambda BT}} - 1} \quad (5.4.1)$$

que calcula la densidad de energía dentro de un cuerpo negro isotérmico, donde

- λ es la longitud de onda de la radiación
- T es la temperatura absoluta del cuerpo negro
- B es la constante de Boltzmann
- P es la constante de Planck
- c es la velocidad de la luz.

Supongamos que nos gustaría determinar la longitud de onda λ que corresponde a la densidad de energía máxima $\varphi(\lambda)$. De (5.4.1), tenemos que

$$\varphi'(\lambda) = \left(\frac{8\pi cP\lambda^{-6}}{e^{\frac{cP}{\lambda BT}} - 1} \right) \left(\frac{(\frac{cP}{\lambda BT})e^{\frac{cP}{\lambda BT}}}{e^{\frac{cP}{\lambda BT}} - 1} - 5 \right) \quad (5.4.2)$$

El máximo de φ ocurre cuando

$$\frac{\left(\frac{cP}{\lambda BT}\right)e^{\frac{cP}{\lambda BT}-1}}{e^{\frac{cP}{\lambda BT}-1}} = 5 \quad (5.4.3)$$

Aquí, haciendo $x = \frac{cP}{\lambda BT}$, la ecuación anterior se convierte en

$$1 - \frac{x}{5} = e^{-x} \quad (5.4.4)$$

Definamos

$$f(x) = e^{-x} - 1 + \frac{x}{5}. \quad (5.4.5)$$

Como consecuencia, encontrando las raíces de (5.4.5) nos da la longitud de onda máxima de radiación (λ) por medio de la siguiente fórmula:

$$\lambda \approx \frac{cP}{x^* BT}. \quad (5.4.6)$$

Es fácil ver que la función $f(x)$ es continua y que $f(2) = -0.464665\dots$ y $f(7) = 0.400912\dots$. Entonces, del teorema del valor intermedio se deduce que $f(x)$ tiene ceros en el intervalo $(2, 7)$.

Consideramos $\mathbb{D} = [2, 7]$. Entonces, tomando $x_0 = 4$ y el método de Newton obtenemos que

$$k_0 = k = 0.161021\dots,$$

$$k_1 = k_2 = 0.372446\dots,$$

$$L_0 = 1,$$

$$L = 0.372446\dots$$

y

$$\eta = 1.$$

Además, como $\frac{1}{k_0+k} = 0.0328022\dots$ obtenemos que

$$U\left(x_0, \frac{1}{k_0+k}\right) \subset D$$

Entonces, usando las definiciones de δ_1, δ_2 y δ obtenemos

$$0 < \delta_2 = 0.210873\dots \leq \delta_1 = 0.322041\dots \leq \delta = 0.407721\dots$$

Entonces podemos asegurar la convergencia del método de Newton a la solución $x^* = 4.96511\dots$ de $f(x)$. Recordar que las condiciones previas consideradas por otros autores no se cumplen porque

$$k_1 c + 2\sqrt{k_1 \nu} = 1.22057\dots \geq 1$$

y la convergencia no estaba garantizada.

Capítulo 6

Estudio de la convergencia y dinámica de un método de orden seis libre de segundas derivadas

Resumen

En este artículo, abordamos el estudio del análisis de convergencia de la familia basada en parámetros modificados del método de continuación libre de segunda derivada para resolver ecuaciones no lineales. Obtenemos que el orden de convergencia es al menos cinco y especialmente, para el valor del parámetro $\alpha = 2$ obtenemos convergencia de sexto orden. Se discuten algunas aplicaciones como la ley de radiación de Max Planck para la cual calculamos la densidad de energía máxima dentro de un cuerpo negro isotérmico o el estudio del efecto multifactor, donde calculamos la trayectoria de un electrón en el espacio de aire entre dos placas paralelas. En éstas y otras aplicaciones vamos a demostrar la eficacia y el rendimiento de nuestro nuevo método (para $\alpha = 2$). Comparamos el valor absoluto de la función en cada iteración $|f(x_n)|$ y $|x_n - x^*|$ con nuestro método y el método de Potra y Pták [185], el método de Kou et al [135]. Observamos que nuestro método es más eficiente que los métodos existentes, en el sentido de convergencia más rápida. Además, se estudia la dinámica del método para un caso especial del parámetro en condiciones de convergencia. Mostrando las cuencas de atracción asociadas al método para diferentes valores de α . Se representan los planos dinámicos de un miembro de la familia con regiones de convergencia a un 2-ciclo atractor, el plano dinámico de un miembro de la familia con regiones de convergencia a un 3-ciclo atractor y otros casos donde no hay problemas de convergencia, ya que las únicas cuencas de atracción corresponden a las raíces de $p(z)$.

6.1. Introducción

En los últimos años, muchos académicos han intentado construir esquemas de orden superior para resolver ecuaciones no lineales de la forma

$$f(x) = 0 \tag{6.1.1}$$

donde, $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ es derivable y no lineal en un intervalo abierto D . En la literatura, conocemos métodos iterativos de orden superior para resolver la ecuación no lineal (6.1.1). Como pueden ser el método de Chebyshev, el método de Halley o el método [133, 215] de Super-Halley, etc. Además, el método de Chebyshev-Halley es ampliamente conocido como método iterativo paramétrico para resolver ecuaciones no lineales. El orden de convergencia de estos métodos es tres. En los últimos años, varios autores [18, 20, 85, 97, 98, 105, 166, 136] han estudiado variantes de los métodos de Chebyshev-Halley con métodos iterativos de orden superior. Los abordaron con el cálculo de la segunda derivada en otro punto. Estos métodos son eficientes para el caso en que la segunda derivada cueste poco calcularla. Sin embargo, el coste computacional de estos métodos es alto al depender de la segunda derivada, por lo que no son adecuados para aplicaciones prácticas. Para ello, usamos el método de Newton [173] que es cuadráticamente convergente y se usa frecuentemente para resolver ecuaciones no lineales debido a una mayor eficiencia computacional. Por esta razón, se analiza la convergencia de la familia uniparamétrica libre de segunda derivada del método iterativo por [138, 84, 86]. La convergencia de los métodos tipo Chebyshev libres de la segunda derivada es estudiada por [135]. Li et al [144] discutieron la convergencia del método de Chebyshev-Halley libre de segunda derivada. Potra y Pták (PPM) [185] definieron, una familia de métodos iterativos libres de segunda derivada de la forma:

$$x_{n+1} = x_n - \left(1 + \frac{f(y_n)}{f(x_n) - \theta f(y_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad \theta \in \mathbb{R}. \tag{6.1.2}$$

Para, $\theta = 0$ este método tiene la ventaja de que no requiere calcular la segunda derivada y puede converger de forma cúbica. Por lo tanto, utilizando métodos iterativos de tercer orden podemos obtener métodos iterativos de mayor orden.

Sin duda, con el avance de los ordenadores, la aritmética computarizada y la computación simbólica avanzan, la construcción de métodos multipunto de orden superior se vuelve vital y más popular porque el cálculo del término constante del error asintótico o el error de las ecuaciones de métodos iterativos para ceros múltiples es bastante más fácil de calcular que en tiempos anteriores. Por lo tanto, en la última década, varios académicos de todo el mundo como Li et al.[143] en (2009), Neta [165] y Li et al. [142] en (2010), y Behl et al.[42] en (2015) y Behl et al. [43] en (2016) propusieron métodos iterativos multipunto de cuarto orden.

Un método de continuación es un método iterativo basado en parámetros para resolver ecuaciones no lineales. De acuerdo con la idea básica de los métodos de

continuación, se puede definir una homotopía $\alpha h_1(x) + (1 - \alpha)h_2(x)$, donde $\alpha \in [0, 1]$, entre dos funciones h_1 y h_2 . Basándonos en esta idea, Ezquerro y Hernandez [80] diseñaron la familia uniparamétrica de métodos de continuación entre los métodos de Chebyshev y Halley en el espacio de Banach. Además, se estudia el análisis de convergencia semilocal. Sin embargo, este método depende de las segundas derivadas en el proceso de cálculo por lo que su aplicación práctica está restringida rigurosamente. Por esta razón, se desarrolla una familia de métodos iterativos libres de segunda derivada.

En este capítulo, desarrollamos la nueva variante modificada del método de continuación basado en parámetros. Discutimos el análisis de convergencia de estas variantes. El análisis muestra que el orden de convergencia de este método es cinco para el parámetro $\alpha \in \mathbb{R}$ y seis para $\alpha = 2$. Este método requiere tres evaluaciones de la función y una de sus derivadas. Algunas aplicaciones numéricas como la ley de conservación de Max Planck o el efecto multifactor se resuelven para mostrar la eficacia y la superioridad del nuevo método. El comportamiento dinámico se discute para un caso especial de su parámetro.

6.2. El método y el análisis de convergencia

El método iterativo para resolver ecuaciones discutido en [80] es

$$z_n = x_n - \left(1 + \frac{1}{2}L_f(x_n)\left(1 + \frac{\frac{\alpha}{2}L_f(x_n)}{1 - \frac{1}{2}L_f(x_n)}\right)\right) \frac{f(x_n)}{f'(x_n)}, \quad (6.2.1)$$

donde,

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}, \quad (6.2.2)$$

$$L_f(x_n) = \frac{f(x_n)f''(x_n)}{f'(x_n)^2} \quad (6.2.3)$$

Sin embargo, este método depende de las segundas derivadas en el proceso de cálculo y su aplicación práctica está rigurosamente restringida. Por esta razón, se desarrolla una familia de métodos iterativos libres de segunda derivada.

Utilizando la técnica del desarrollo en serie de Taylor, obtenemos el siguiente método iterativo uniparamétrico modificado libre de segunda derivada,

$$\left. \begin{aligned} y_n &= x_n - \frac{f(x_n)}{f'(x_n)} \\ z_n &= x_n - \left(1 + \frac{f(y_n)}{f(x_n)} \frac{(f(x_n) + (\alpha - 1)f(y_n))}{(f(x_n) - f(y_n))}\right) \frac{f(x_n)}{f'(x_n)} \end{aligned} \right\} \quad (6.2.4)$$

De manera similar con la ayuda del desarrollo en serie de Taylor,

$$f'(z_n) \approx f'(x_n) + (z_n - x_n)f''(x_n) \quad (6.2.5)$$

Consideramos, $f''(x_n) = g''(x_n)$, esto es $g''(x_n) = \frac{2f(y_n)f'(x_n)^2}{f(x_n)^2}$ entonces de (6.2.5), obtenemos una aproximación $f'(z_n)$ por

$$f'(z_n) \approx f(x_n) + (z_n - x_n)g''(x_n)$$

Por lo tanto, tenemos

$$\left. \begin{aligned} y_n &= x_n - \frac{f(x_n)}{f'(x_n)} \\ z_n &= x_n - \left(1 + \frac{f(y_n)}{f(x_n)} \frac{(f(x_n) + (\alpha - 1)f(y_n))}{(f(x_n) - f(y_n))} \right) \frac{f(x_n)}{f'(x_n)} \\ x_{n+1} &= z_n - \frac{f(z_n)}{f'(x_n) + g''(x_n)(z_n - x_n)}, \end{aligned} \right\} \quad (6.2.6)$$

Teorema 6.1 *Sea la función $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ que es diferenciable y tiene una raíz simple $x^* \in D$, donde D es un intervalo abierto. Para la aproximación inicial x_0 cercana a la solución x^* , y $\alpha \in \mathbb{R}$ el orden de convergencia del método iterativo (6.2.6) es al menos cinco y, además, si $\alpha = 2$, el orden es seis. La ecuación del error está definida como*

$$e_{n+1} = (\alpha - 2)c_2^2c_3e_n^5 + c_2 \left((\alpha - 2)^2c_2^4 + (5 - 4\alpha)c_3c_2^2 + 2(\alpha - 2)c_4c_2 + (4\alpha - 7)c_3^2 \right) e_n^6 + O(e_n)^7$$

Demostración: Sea $e_{\alpha,n} = x_{\alpha,n} - x^*$ y $d_n = z_n - x^*$, donde z_n definida en (6.2.4). Usando el desarrollo en serie de Taylor, tenemos

$$f(x_n) = f'(x^*)[e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + O(e_n^5)] \quad (6.2.7)$$

donde, $c_k = \frac{f^{(k)}(x^*)}{k!f'(x^*)}$ para $k \in \mathbb{N}$. Además, al expandir $f'(x_n)$ usando las series de Taylor alrededor de x^* obtenemos

$$f'(x_n) = f'(x^*)[1 + 2c_2e_n + 3c_3e_n^2 + 4c_4e_n^3 + O(e_n^4)]. \quad (6.2.8)$$

Además, tenemos

$$\begin{aligned} \frac{f(x_n)}{f'(x_n)} &= [e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + O(e_n^5)][1 + 2c_2e_n + 3c_3e_n^2 + 4c_4e_n^3 + O(e_n^4)]^{-1} \\ &= e_n - c_2e_n^2 + 2(c_2^2 - c_3)e_n^3 + (7c_2c_3 - 4c_2^2 - 3c_4)e_n^4 + O(e_n^5). \end{aligned} \quad (6.2.9)$$

De manera similar, tenemos

$$\frac{f'(x_n)}{f(x_n)} = \frac{1 + c_2e_n + (2c_3 - c_2^2)e_n^2}{e_n} \quad (6.2.10)$$

Usando, (6.2.9) obtenemos,

$$y_n = c_2e_n^2 + 2(-c_2^2 + c_3)e_n^3 + (-7c_2c_3 + 4c_2^3 + 3c_4)e_n^4 + O(e_n^5).$$

Al desarrollar $f(y_n)$ como serie de Taylor alrededor de x^* , obtenemos

$$f(y_n) = f'(x^*)[c_2 e_n^2 + 2(-c_2^2 + c_3)e_n^3 + (-7c_2 c_3 + 5c_2^3 + 3c_4)e_n^4 + O(e_n^5)] \quad (6.2.11)$$

De (6.2.7), (6.2.9) y (6.2.11), obtenemos

$$\begin{aligned} \left(1 + \frac{f(y_n)}{f(x_n)} \frac{(f(x_n) + (\alpha - 1)f(y_n))}{(f(x_n) - f(y_n))}\right) \frac{f(x_n)}{f'(x_n)} &= e_n + (\alpha - 2)c_2^2 e_n^3 \\ &+ c_2 \left((9 - 6\alpha)c_2^2 + (4\alpha - 7)c_3 \right) e_n^4 \\ &+ O(e_n^5) \end{aligned} \quad (6.2.12)$$

$$\frac{f(y_n)}{f(x_n)} = c_2 e_n + (2c_3 - 3c_2^2) e_n^2 + O(e_n^3) \quad (6.2.13)$$

Por lo tanto, de (6.2.6) y (6.2.12), tenemos que

$$d_n = -(\alpha - 2)c_2^2 e_n^3 - c_2 \left((9 - 6\alpha)c_2^2 + (4\alpha - 7)c_3 \right) e_n^4 \quad (6.2.14)$$

Entonces de (6.2.6), obtenemos

$$z_n - x_n = -e_n - (\alpha - 2)c_2^2 e_n^3 - c_2 \left((9 - 6\alpha)c_2^2 + (4\alpha - 7)c_3 \right) e_n^4 \quad (6.2.15)$$

De, (6.2.8), (6.2.10), (6.2.13) y (6.2.15), tenemos

$$\begin{aligned} \frac{1}{f'(x_n) + g''(x_n)(z_n - x_n)} &= \frac{1 + c_3 e_n^2 + 2 \left((\alpha - 2)c_2^3 + c_3 c_2 + c_4 \right) e_n^3}{f'(x^*)} \\ &+ \frac{O(e_n^4)}{f'(x^*)} \end{aligned} \quad (6.2.16)$$

Desarrollando en serie de Taylor $f'(z_n)$,

$$f(z_n) = f'(x^*)[d_n + O(d_n^2)] \quad (6.2.17)$$

Así, de (6.2.15), (6.2.16) y (6.2.17), obtenemos que el método definido por (6.2.6) satisface la siguiente ecuación del error

$$\begin{aligned} e_{n+1} &= (\alpha - 2)c_2^2 c_3 e_n^5 + c_2 \left((\alpha - 2)^2 c_2^4 + (5 - 4\alpha)c_3 c_2^2 + 2(\alpha - 2)c_4 c_2 + (4\alpha - 7)c_3^2 \right) e_n^6 \\ &+ O(e_n)^7 \end{aligned} \quad (6.2.18)$$

Lo que significa que el orden del método iterativo es seis para $\alpha = 2$ y es de quinto orden de convergencia para $\alpha \in \mathbb{R}$. La ecuación del error para $\alpha = 2$,

$$e_{n+1} = c_2 \left(-3c_3 c_2^2 + c_3^2 \right) e_n^6 + O(e_n)^7.$$

6.3. Aplicaciones Numéricas

Cada iteración del método (6.2.6) requiere tres evaluaciones de la función y una más de la primera derivada. Definimos índice de eficiencia como $p^{1/w}$, donde p es el orden del método y w es el número de evaluaciones de la función por iteración. El índice de eficiencia del presente método es, para $\alpha \in \mathbb{R}$ igual a $\sqrt[4]{5} \simeq 1.495$ y para $\alpha = 2$ igual a $\sqrt[4]{6} \simeq 1.565$, que son mejores que en el PPM [185] $\sqrt[3]{3} \simeq 1.442$.

Ahora, trabajaremos algunos ejemplos numéricos y los compararemos con el método de Potra y Pták (PPM) (6.3.1) y el método de Kou et al [135]. Todo el cálculo se lleva a cabo utilizando Mathematica 9. El cálculo del error residual $|f(x_n)|, |x_n - x^*|$ se expone en la Tabla 1. Están calculados usando el método de Potra y Pták (PPM), el método de Kou et al y nuestro método (6.2.6) con el mismo número total de evaluaciones de la función. De aquí, deducimos que el método iterativo (6.2.6) es mejor que el PPM y que el método de Kou et al [135].

Ejemplo 6.2 $f_1(x) = x^3 + 4x^2 - 10 = 0$, $x_0 = 1$, $x^* = 1.3652300134140969$.

Ejemplo 6.3 $f_2(x) = e^{-x} + \cos x$, $x_0 = 1$, $x^* = 1.74613953040801242$

Ejemplo 6.4 $f_3(x) = e^x \sin x + \log(x^2 + 1)$, $x_0 = 1$ $x^* = 0$

Ejemplo 6.5 $f_4(x) = x^2 - e^{-x} - 3x + 2$, $x_0 = 2.5$, $x^* = 0.25753028543986084$

Ejemplo 6.6 $f_5(x) = \sin^2 x - x^2 + 1$, $x_0 = 1$, $x^* = 1.4044916482153411$

Ejemplo 6.7 *En el estudio del efecto multifactor, la trayectoria de un electrón en el espacio de aire entre dos placas paralelas está dada por*

$$x(t) = x_0 + \left(v_0 + e \frac{E_0}{nw} \sin(\omega t_0 + \alpha) \right) (t - t_0) + e \frac{E_0}{m\omega^2} \left(\cos(\omega t + \alpha) + \sin(\omega t_0 + \alpha) \right)$$

donde e y m son la carga y la masa del electrón en reposo, x_0 y v_0 son la posición y la velocidad del electrón en el tiempo t_0 y $E_0 \sin(\omega t + \alpha)$ es el campo eléctrico RF entre las placas. Elegimos parámetros particulares en la expresión para tratar con una expresión más simple, que se define de la siguiente manera:

$$f_6(x) = x - \frac{1}{2} \cos(x) + \frac{\pi}{4}, \quad x^* = 0.3090932715417949$$

,

Ejemplo 6.8 *Consideramos el siguiente problema de ley de radiación de Planck que calcula la densidad de energía dentro de un cuerpo negro isotérmico y que está dado por:*

$$\vartheta(\lambda) = \frac{8\pi c P \lambda^{-5}}{e^{\frac{cP}{\lambda BT}} - 1}, \quad (6.3.1)$$

donde λ es la longitud de onda de la radiación, T es la temperatura absoluta del cuerpo negro, B es la constante de Boltzmann, P es la constante de Planck y c es la velocidad de la luz. Estamos interesados en determinar la longitud de onda λ que corresponde a la densidad de energía máxima $\vartheta(\lambda)$.

De (6.3.1), obtenemos

$$\vartheta'(\lambda) = \left(\frac{8\pi c P \lambda^{-6}}{e^{\frac{cP}{\lambda BT}} - 1} \right) \left(\frac{\frac{cP}{\lambda BT} e^{\frac{cP}{\lambda BT}}}{e^{\frac{cP}{\lambda BT}} - 1} - 5 \right), \quad (6.3.2)$$

para que los máximos de ϑ ocurran cuando

$$\frac{\frac{cP}{\lambda BT} e^{\frac{cP}{\lambda BT}}}{e^{\frac{cP}{\lambda BT}} - 1} = 5. \quad (6.3.3)$$

Después de esto, si $x = \frac{cP}{\lambda BT}$, entonces (6.3.3) se satisface si

$$f_7(x) = e^{-x} + \frac{x}{5} - 1 = 0. \quad (6.3.4)$$

Por lo tanto, las soluciones de $f_7(x) = 0$ dan la longitud de onda máxima de radiación λ por medio de la siguiente fórmula:

$$\lambda \approx \frac{cP}{x^* BT}, \quad (6.3.5)$$

donde x^* es una solución de (6.3.4).

Como la función (6.3.4) es siempre continua y tal que $f_7(1) = -0.432121\dots$ y $f_7(8.5) = 0.700203\dots$, es obvio que $f_7(x)$ tiene ceros en el intervalo $[1, 8.5]$. Así que, consideramos $\mathcal{D} = [1, \infty)$ y $x^* = 4.965114$.

6.4. Estudio dinámico para un caso especial del método

En primer lugar, recordemos algunos conceptos básicos relacionados con la dinámica. Dada una función racional $R: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$, donde $\hat{\mathbb{C}}$ es la esfera de Riemann, la órbita de un punto $z_0 \in \hat{\mathbb{C}}$ está definida por

$$\{z_0, R(z_0), R^2(z_0), \dots, R^n(z_0), \dots\}.$$

Un punto $z_0 \in \bar{\mathcal{C}}$, se dice que es un *punto fijo* de $R(z)$ si verifica que $R(z) = z$. Además, z_0 es un *punto periódico* de periodo $p > 1$ si es un punto tal que $R^p(z_0) = z_0$ pero $R^k(z_0) \neq z_0$, para cada $k < p$. Además, un punto z_0 se denomina *pre-periódico* si no es periódico pero existe un $k > 0$ tal que $R^k(z_0)$ es periódico.

Existen diferentes tipos de puntos fijos dependiendo de su multiplicidad asociada $|R'(z_0)|$. Teniendo en cuenta la multiplicidad asociada a un punto fijo z_0 se le denomina:

- *superatractor* si $|R'(z_0)| = 0$
- *atractor* si $|R'(z_0)| < 1$
- *repulsor* si $|R'(z_0)| > 1$
- y *parabólico* si $|R'(z_0)| = 1$.

A los puntos fijos que no corresponden con las raíces del polinomio $p(z)$ les llamamos *puntos fijos extraños*. Por otro lado, un *punto crítico* z_0 es un punto que satisface que, $R'(z_0) = 0$. Además, llamamos *punto crítico libre* a aquellos puntos críticos que no son raíces del polinomio $p(z)$.

La *cuenca de atracción* de un punto fijo atractor α se define como:

$$\mathcal{A}(\alpha) = \{z_0 \in \hat{\mathbb{C}} : R^n(z_0) \rightarrow \alpha, n \rightarrow \infty\}.$$

El *conjunto de Fatou* de la función racional R , $\mathcal{F}(R)$, es el conjunto de puntos $z \in \hat{\mathbb{C}}$ cuyas órbitas tienden a un atractor (punto fijo, órbita periódica o infinito). Su complementario en $\hat{\mathbb{C}}$ es el *conjunto de Julia*, $\mathcal{J}(R)$. Eso significa que la cuenca de atracción de cualquier punto fijo pertenece al conjunto de Fatou y los límites de estas cuencas de atracción pertenecen al conjunto de Julia.

Aplicando este operador en un polinomio cuadrático con dos raíces diferentes A y B , $p(z) = (z - A)(z - B)$ y usando la transformada de Möebius $h(z) = \frac{z-A}{z-B}$, la cual lleva la raíz A al 0, la raíz B al ∞ y el ∞ al 1, obtenemos el operador racional asociado a la familia de esquemas iterativos que es finalmente:

$$G(z, \alpha) = \frac{z^6(-\alpha + z^3 + 3z^2 + 3z + 2)^2}{(\alpha z^3 - 2z^3 - 3z^2 - 3z - 1)^2} \quad (6.4.1)$$

6.4.1. Estudio de los puntos fijos

Es claro que $z = 0$ y $z = \infty$ son puntos fijos de $G(z, \alpha)$. Además, existen algunos puntos fijos extraños que son:

- $z = 1$ relacionado con la divergencia al ∞
- Las raíces de

$$\begin{aligned} q(z) = & 2z^{10} + 7z^9 + 22z^8 + z^7(44 - 2\alpha) + z^6(65 - 8\alpha) \\ & + z^5(-\alpha^2 - 10\alpha + 73) + z^4(65 - 8\alpha) \\ & + z^3(44 - 2\alpha) + 22z^2 + 7z + 1. \end{aligned}$$

Las soluciones del polinomio dependen del valor del parámetro α y podemos verlas en la Figura 6.1.

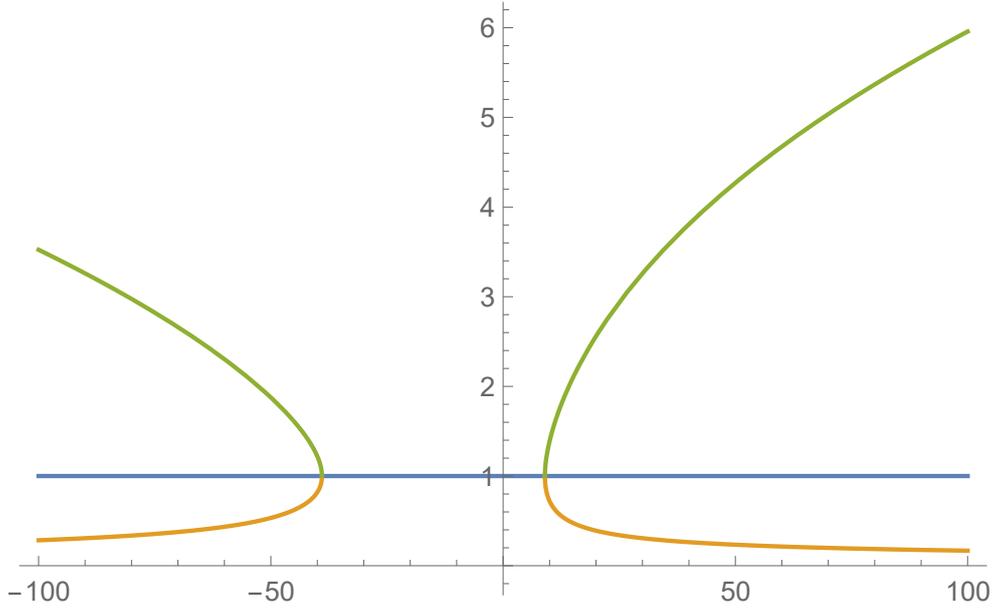


Figura 6.1: Diagrama de bifurcación del punto fijo extraño de $G(z, \alpha)$

6.4.2. Estudio de los puntos críticos y planos de parámetros

Es un hecho bien conocido que hay al menos un punto crítico asociado con cada componente invariante de Fatou [93, 139]. Los puntos críticos de la familia son las soluciones de $G'(z, \alpha) = 0$, donde

$$G'(z, \alpha) = \frac{6z^5(z+1)^2(-\alpha + z^3 + 3z^2 + 3z + 2)(\alpha z^4 + \alpha - 2z^4 - 4z^3 - 6z^2 - 4z - 2)}{(\alpha z^3 - 2z^3 - 3z^2 - 3z - 1)^3}$$

Es claro que el $z = 0$ y el $z = \infty$ son puntos críticos. Además, los puntos críticos libres son las raíces de los polinomios:

$$cp_1(z, \alpha) = 2 - \alpha + 3z + 3z^2 + z^3$$

que llamaremos $cr_i(\alpha)$ para $i = 1, 2, 3$ y las raíces del polinomio:

$$cp_2(z, \alpha) = \alpha z^4 + \alpha - 2z^4 - 4z^3 - 6z^2 - 4z - 2$$

a las que llamaremos $cr_i(\alpha)$ para $i = 4, 5, 6, 7$ y que tienen la siguiente expresión:

$$cr_4(\alpha) = \frac{1}{\alpha-2} - \frac{\sqrt{\alpha^2-\alpha}}{\sqrt{2}\sqrt{\alpha^2-4\alpha+4}}$$

$$-\frac{1}{2}\sqrt{-\frac{\left(\frac{96}{(\alpha-2)^2} + \frac{64}{(\alpha-2)^3} + \frac{32}{\alpha-2}\right)\sqrt{\alpha^2-4\alpha+4}}{4\sqrt{2}\sqrt{\alpha^2-\alpha}} - \frac{2\alpha-4}{\alpha-2} + \frac{6}{\alpha-2} + \frac{8}{(\alpha-2)^2}},$$

$$\begin{aligned}
cr_5(\alpha) &= \frac{1}{\alpha-2} - \frac{\sqrt{\alpha^2-\alpha}}{\sqrt{2}\sqrt{\alpha^2-4\alpha+4}} \\
&+ \frac{1}{2} \sqrt{-\frac{\left(\frac{96}{(\alpha-2)^2} + \frac{64}{(\alpha-2)^3} + \frac{32}{\alpha-2}\right)\sqrt{\alpha^2-4\alpha+4}}{4\sqrt{2}\sqrt{\alpha^2-\alpha}} - \frac{2\alpha-4}{\alpha-2} + \frac{6}{\alpha-2} + \frac{8}{(\alpha-2)^2}}, \\
cr_6(\alpha) &= \frac{1}{\alpha-2} + \frac{\sqrt{\alpha^2-\alpha}}{\sqrt{2}\sqrt{\alpha^2-4\alpha+4}} \\
&- \frac{1}{2} \sqrt{-\frac{\left(\frac{96}{(\alpha-2)^2} + \frac{64}{(\alpha-2)^3} + \frac{32}{\alpha-2}\right)\sqrt{\alpha^2-4\alpha+4}}{4\sqrt{2}\sqrt{\alpha^2-\alpha}} - \frac{2\alpha-4}{\alpha-2} + \frac{6}{\alpha-2} + \frac{8}{(\alpha-2)^2}}
\end{aligned}$$

y

$$\begin{aligned}
cr_7(\alpha) &= \frac{1}{\alpha-2} + \frac{\sqrt{\alpha^2-\alpha}}{\sqrt{2}\sqrt{\alpha^2-4\alpha+4}} \\
&+ \frac{1}{2} \sqrt{-\frac{\left(\frac{96}{(\alpha-2)^2} + \frac{64}{(\alpha-2)^3} + \frac{32}{\alpha-2}\right)\sqrt{\alpha^2-4\alpha+4}}{4\sqrt{2}\sqrt{\alpha^2-\alpha}} - \frac{2\alpha-4}{\alpha-2} + \frac{6}{\alpha-2} + \frac{8}{(\alpha-2)^2}}.
\end{aligned}$$

Está claro que $cr_4(\alpha) = \frac{1}{cr_5(\alpha)}$ y que $cr_6(\alpha) = \frac{1}{cr_7(\alpha)}$ así que, sin pérdida de generalidad, estudiaremos solo dos de ellos. Elegimos $cr_5(\alpha)$ y $cr_6(\alpha)$.

Siguiendo los esquemas iterativos y las técnicas usadas en [28, 150], estudiaremos la dinámica de la familia. En las figuras 6.2 - 6.3, se muestran los planos de parámetros asociados a los puntos críticos libres en los que observamos que existe una zona sin convergencia a las raíces. Un punto está pintado en cian si la iteración del método que comienza en $z_0 = cr_3(a)$ converge al punto fijo 0 (relacionado con la raíz A), o si converge a ∞ (relacionado con la raíz B) y en amarillo si la iteración converge a 1 (relacionada con el ∞). Además, aparece en rojo la convergencia, después de un máximo de 2000 iteraciones y con una tolerancia de 10^{-6} , a cualquiera de los puntos fijos extraños, en naranja la convergencia a 2-ciclos y en otro color convergencia a otros n -ciclos. Las regiones en negro corresponden a zonas de convergencia a otros ciclos, a comportamientos caóticos entre otros.. Como consecuencia, cada punto del plano que no es ni cian ni magenta no es una buena opción de α en términos de búsqueda de raíces de la familia de métodos.

Hemos detectado las anomalías, ahora vamos a mostrar algunas de ellas en los planos dinámicos. En estos planos dinámicos la convergencia a 0 aparecerá en magenta, en cian aparece la convergencia a ∞ y en negro las zonas sin convergencia, después de un máximo de 200 iteraciones y con una tolerancia de 10^{-6} a las raíces.

En las Figuras 8.9–8.8 se muestra el plano dinámico de los miembros de la familia con regiones de convergencia a algunos de los puntos fijos extraños.

En la Figura 8.6 Se muestran los planos dinámicos de un miembro de la familia con regiones de convergencia a un 2-ciclo atractor, las regiones con convergencia

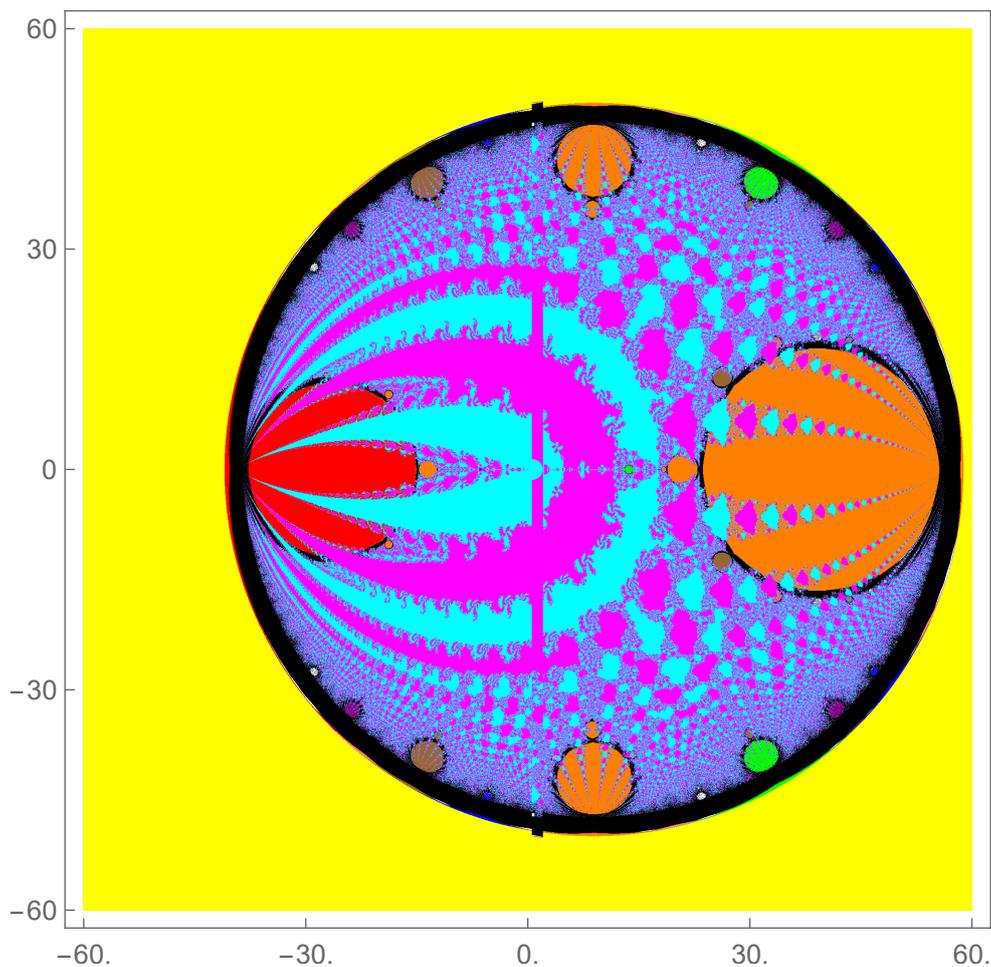


Figura 6.2: Plano de parámetros asociado al punto crítico libre $cr_5(\alpha)$.

a 2-ciclos están pintadas en negro ya que no hay convergencia con ninguna de las raíces

Además, en la Figura 6.9, se representa el plano dinámico de un miembro de la familia con regiones de convergencia a $z = 1$, relacionado con el ∞ , pintado en amarillo.

Además, en la Figura 8.12, se muestra el plano dinámico de un miembro de la familia con regiones de convergencia a un 3-ciclo atractor.

Otros casos se muestran en las figuras 8.11 y 8.7 donde no hay problemas de convergencia, ya que las únicas cuencas de atracción corresponden a las raíces de $p(z)$.

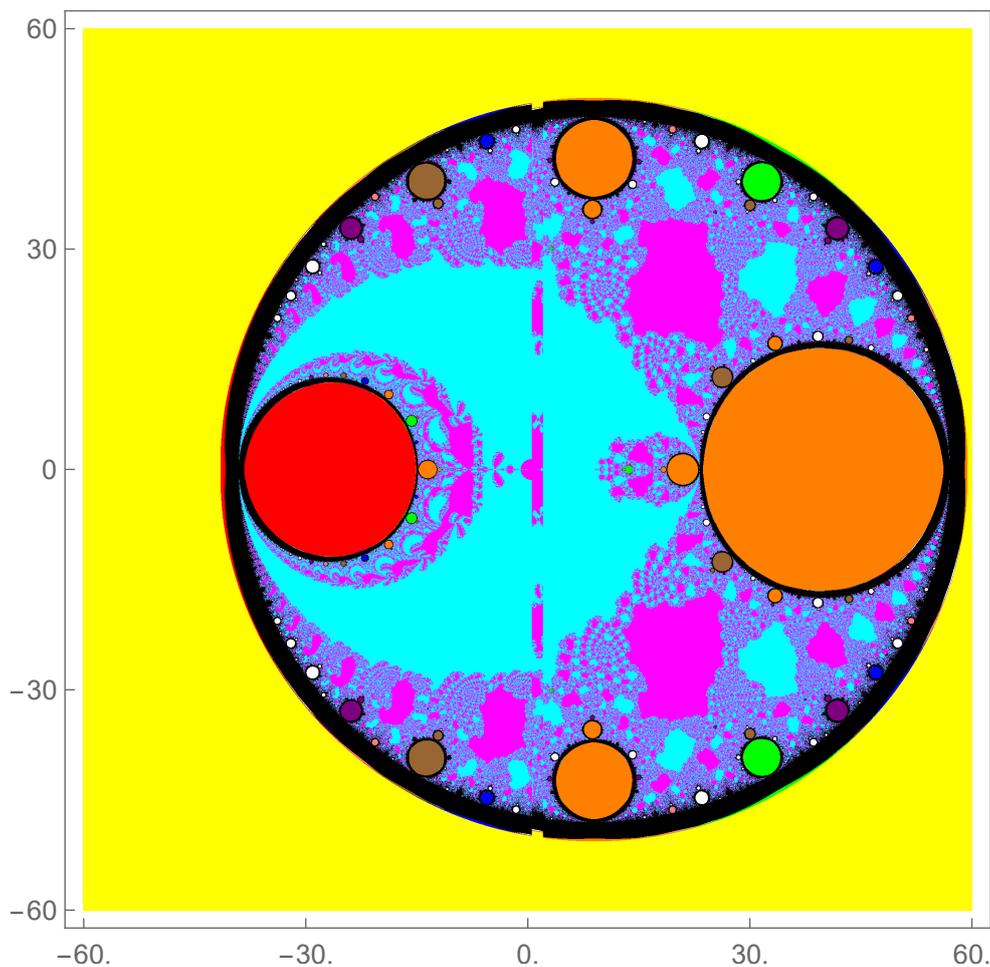


Figura 6.3: Plano de parámetros asociado al punto crítico libre $cr_6(\alpha)$.

6.5. Conclusiones

En este artículo, discutimos el análisis de convergencia del método iterativo paramétrico para resolver ecuaciones no lineales. El orden de convergencia del método iterativo es al menos cinco para el parámetro $\alpha \in \mathbb{R}$ y seis cuando $\alpha = 2$. Son resueltas algunas aplicaciones numéricas, como los modelos no lineales, y calculado el error residual $|f(x_n)|$, $|x_n - x^*|$ con nuestros métodos, el método de Potra y Pták (PPM) (6.1.2) y el método de Kou et al [135]. Finalmente, concluimos que nuestro método converge más rápidamente cuando se compara con los otros métodos. Además, estudiamos el comportamiento dinámico del método iterativo para casos especiales del parámetro.

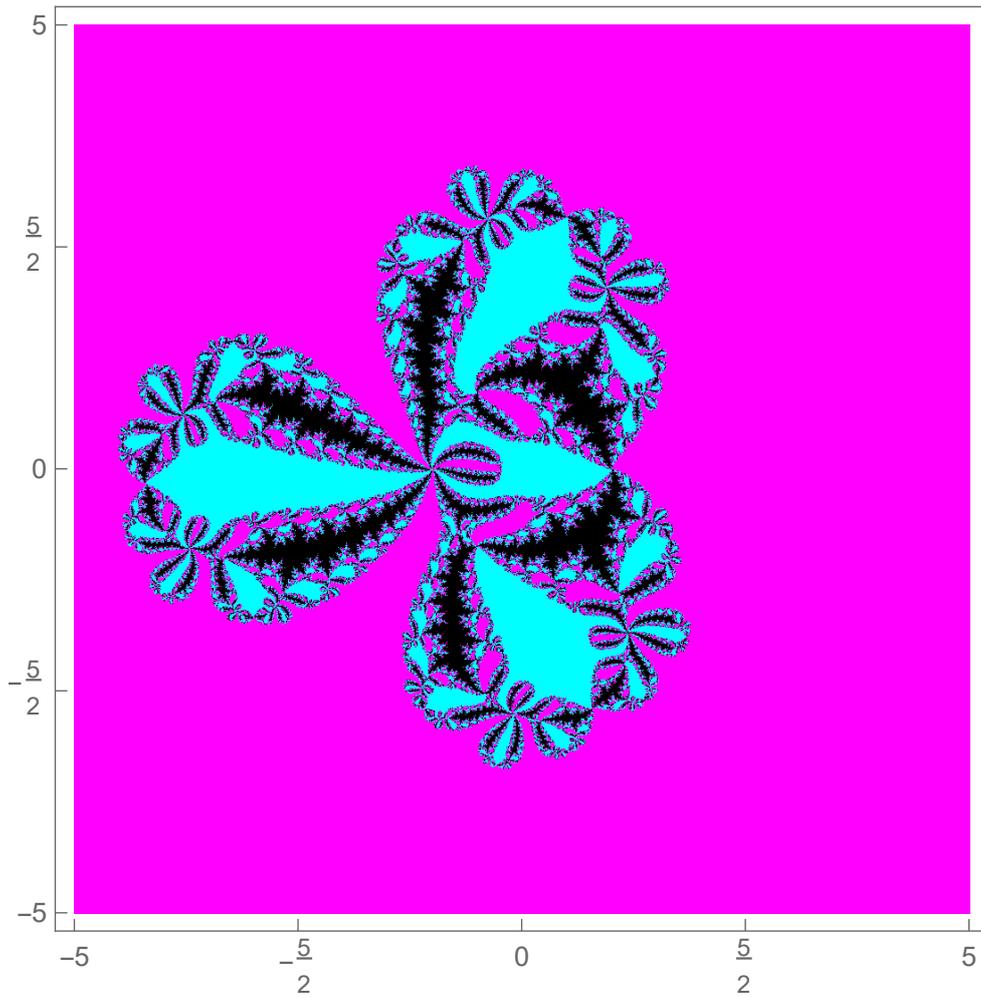


Figura 6.4: Cuencas de atracción asociadas al método con $\alpha = 13.75$.

Table 2: Comparación de métodos iterativos para funciones escalares

$f(x)$	n	PPM [185]		Kou et al [135]		Method (6.2.6)	
		$ f(x_n) $	$ x_n - x^*i $	$ f(x_n) $	$ x_n - x^*i $	$ f(x_n) $	$ x_n - x^*i $
f_1	1	8.2e-001	5.1e-002	0.21e-001	1.3e-001	9.0e-002	5.5e-003
	2	1.1e-003	6.9e-005	5.8e-002	3.5e-003	2.3e-009	1.4e-010
	3	2.6e-012	1.6e-013	6.9e-007	4.1e-008	1.0e-039	6.3e-041
	4	3.0e-038	1.8e-039	1.1e-021	6.9e-023	4.1e-161	2.5e-162
	5	4.8e-116	2.9e-117	5.1e-066	0.1e-067	1.0e-646	6.26e-6481
f_2	1	2.4e-004	2.0e-004	5.2e-004	4.5e-004	1.7e-005	1.5e-005
	2	4.5e-013	3.8e-013	9.6e-012	8.3e-012	6.8e-022	5.8e-022
	3	3.0e-039	2.6e-039	5.9e-035	5.1e-035	1.6e-087	1.4e-087
	4	8.9e-118	7.7e-118	1.4e-104	1.2e-104	5.0e-350	4.3e-350
f_3	1	3.5e-001	2.1e-001	4.8e-001	2.4e-001	1.6e-001	1.3e-001
	2	2.6e-002	2.5e-002	6.2e-002	5.4e-002	1.6e-003	1.6e-003
	3	1.0e-004	1.0e-004	1.6e-003	1.6e-003	1.3e-010	1.3e-010
	4	8.8e-12	8.8e-12	6.9e-008	6.9e-008	5.7e-039	5.7e-39
	5	5.5e-033	5.5e-033	5.3e-021	5.3e-021	1.9e-152	1.9e-152
f_4	1	2.6e-000	7.1e-001	3.7e-000	9.9e-001	6.6e-001	1.8e-001
	2	5.7e-004	1.5e-004	6.5e-002	1.7e-002	2.7e-006	7.1e-007
	3	2.3e-013	6.1e-014	6.6e-007	1.7e-007	6.3e-028	1.7e-028
	4	1.5e-041	3.9e-042	6.9e-022	1.8e-022	2.0e-114	5.2e-115
	5	3.9e-126	1.0e-126	8.2e-067	2.2e-067	1.9e-460	5.0e-461
f_5	1	7.3e-001	9.8e-002	9.9e-001	4.4e+005	2.3e-001	8.7e-002
	2	8.1e-001	1.2e-000	2.0e+011	2.5e+005	8.7e-005	3.5e-005
	3	1.0e-027	9.5e+003	3.8e+010	1.1e+005	3.1e-018	1.3e-018
	4	9.0e+007	5.9e+003	7.2e+009	4.8e+004	5.2e-072	2.1e-072
	5	1.3e+007	2.2e+003	1.4e+009	2.1e+004	3.9e-287	1.6e-287
f_6	1	7.6e-002	8.8e-002	1.2e-001	1.3e+001	4.3e-003	5.1e-003
	2	8.3e-005	9.8e-005	5.2e+004	6.2e+0045	1.2e-011	1.4e-011
	3	1.3e-013	1.5e+013	6.2e+011	7.3e+011	6.1e-046	7.1e-046
	4	4.5e+040	5.3e+040	1.1e+031	1.1e+031	4.6e-183	5.5e-183
f_7	1	4.2e-004	2.2e-003	60e+000	3.0e+002	9.6e-007	4.9e-006
	2	1.3e-012	7.0e-012	6.7e-003	3.5e-002	5.1e-028	2.6e-027
	3	4.3e-038	2.2e+037	1.0e+008	5.3e-008	4.2e-113	2.2e-112

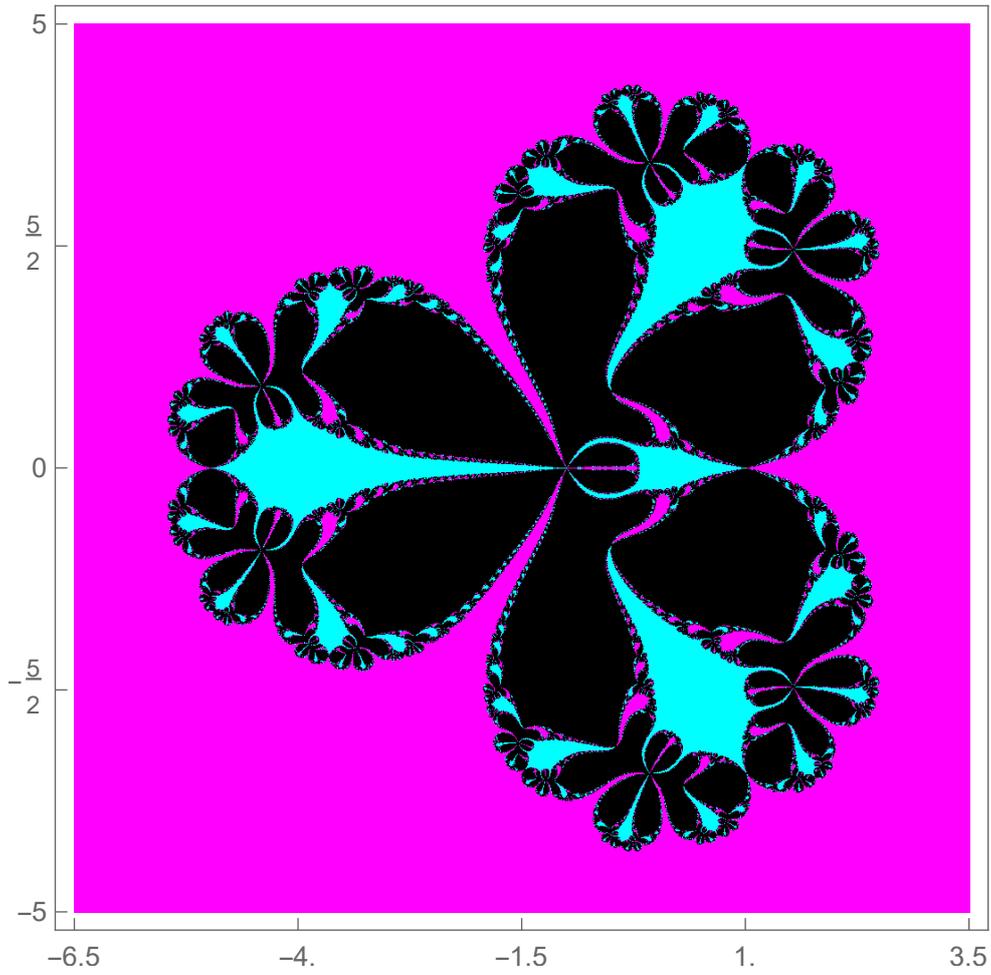


Figura 6.5: Cuencas de atracción asociadas al método con $\alpha = -30$.

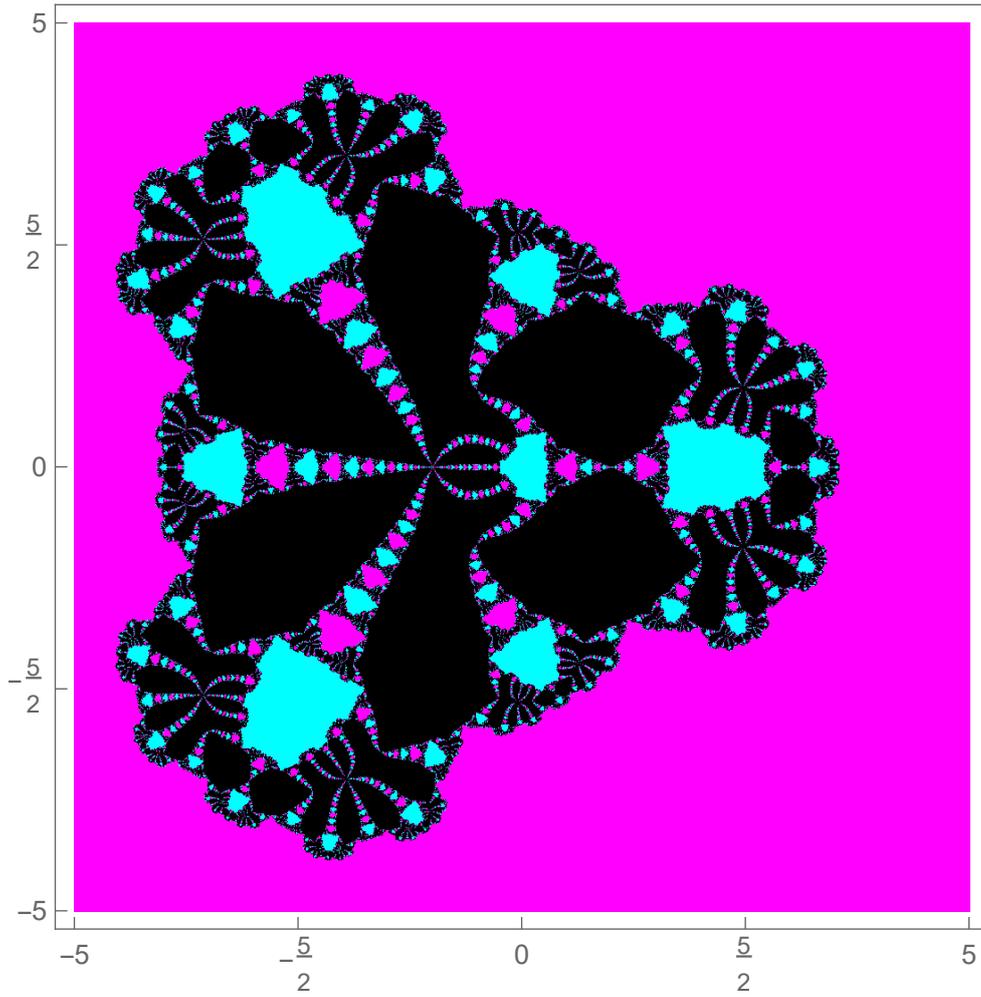


Figura 6.6: Cuencas de atracción asociadas al método con $\alpha = 35$.

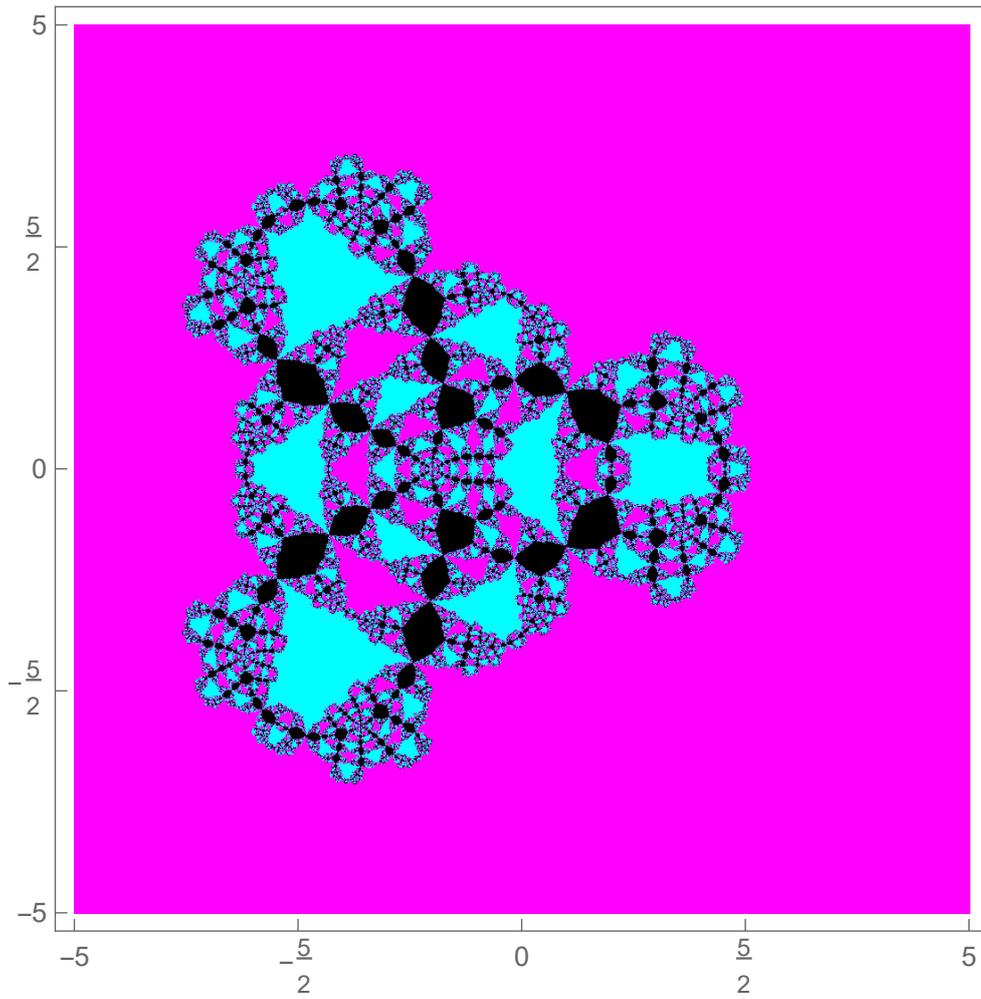


Figura 6.7: Cuencas de atracción asociadas al método con $\alpha = 20$.

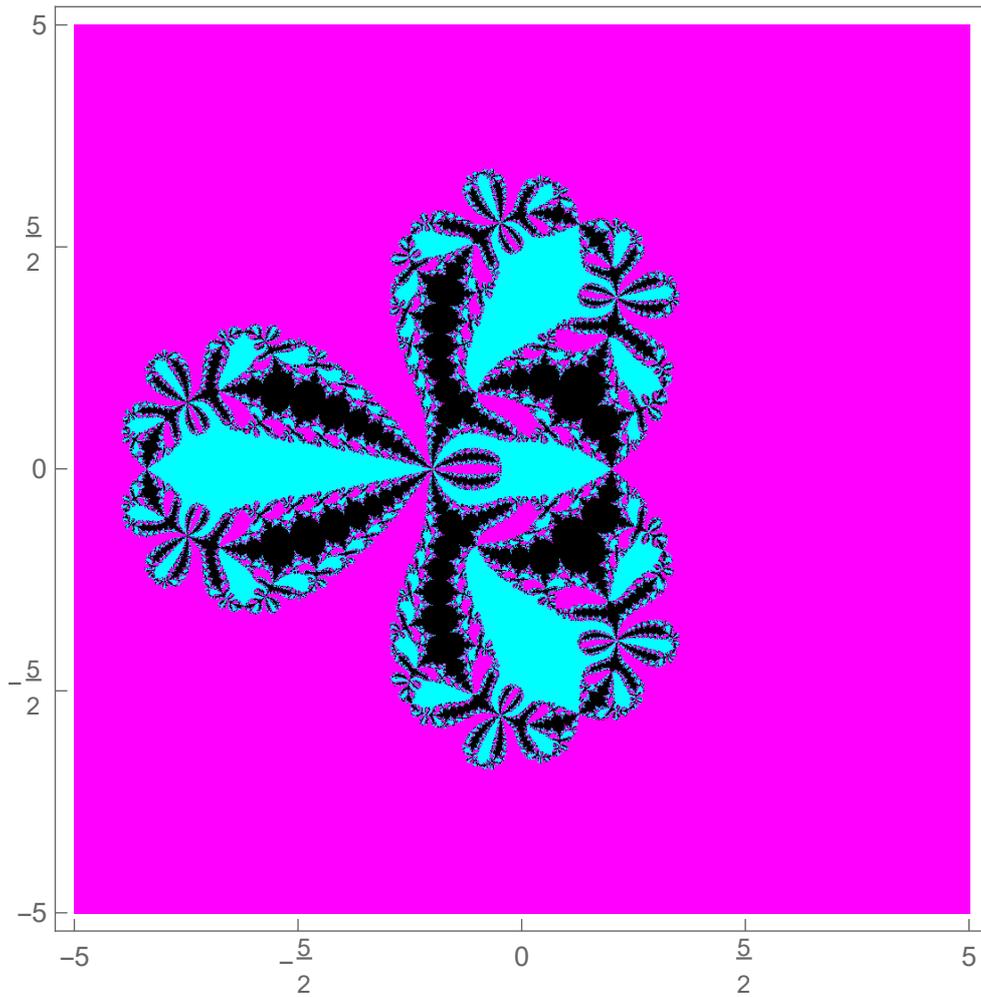


Figura 6.8: Cuencas de atracción asociadas al método con $\alpha = -15$.

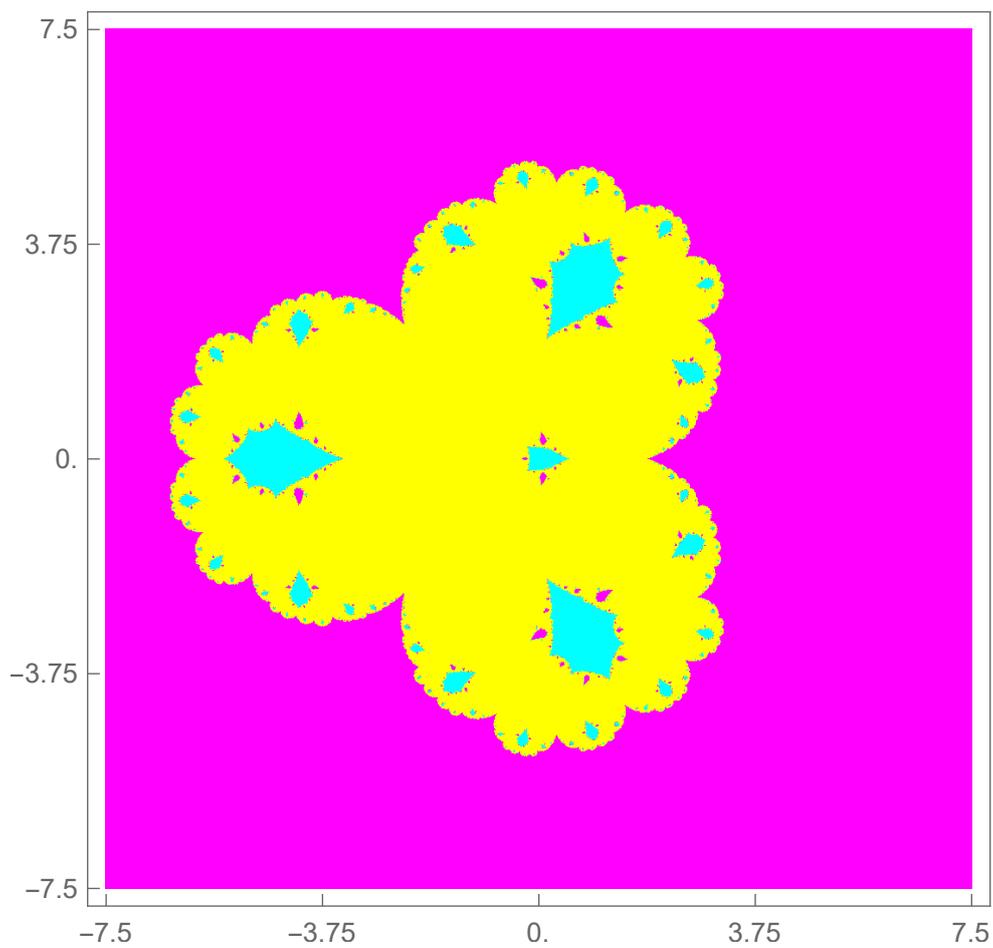


Figura 6.9: Cuencas de atracción asociadas al método con $\alpha = -50$.

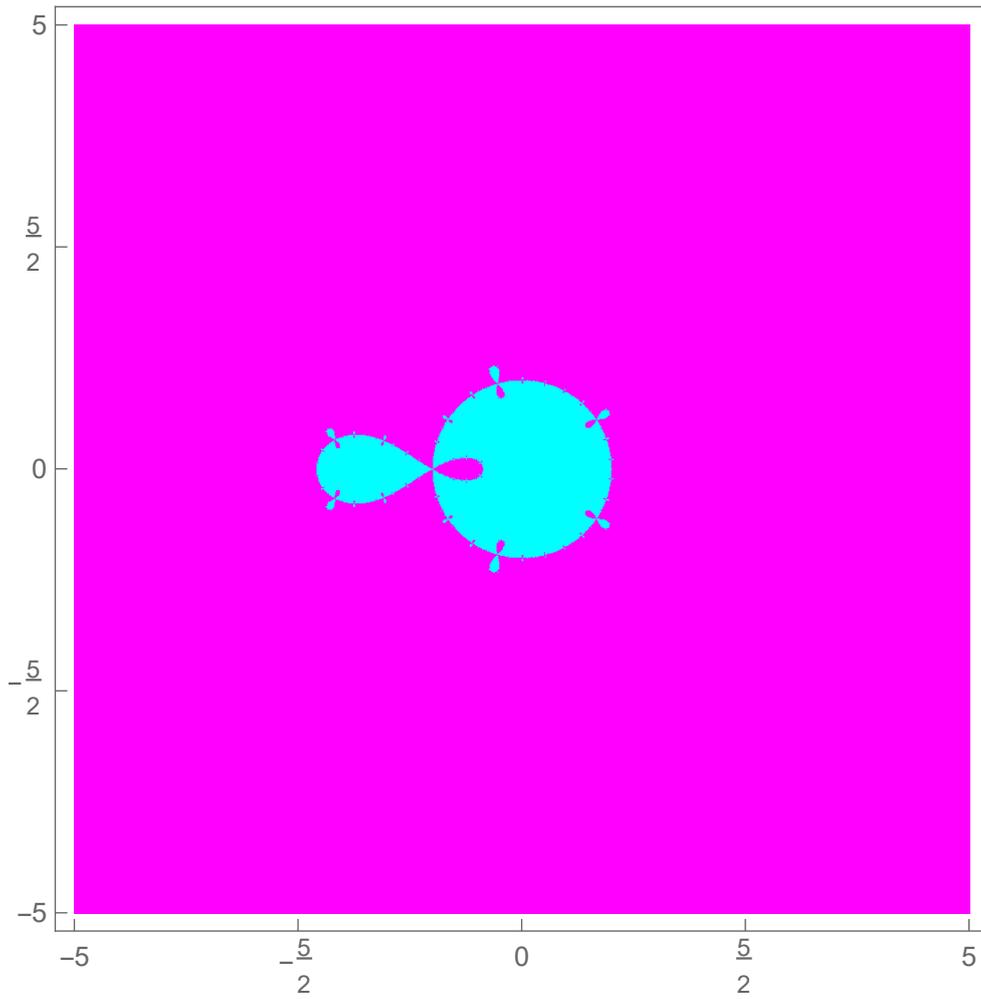


Figura 6.10: Cuencas de atracción asociadas al método con $\alpha = 0$.

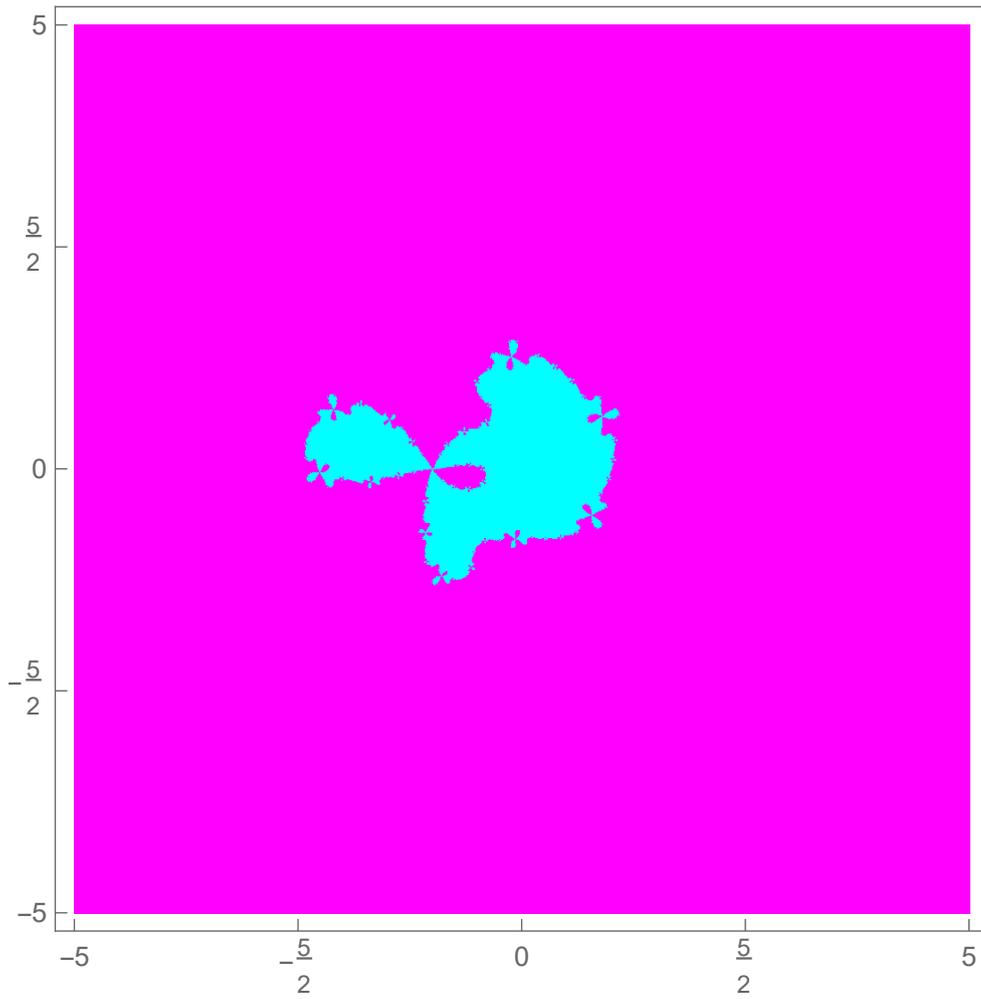


Figura 6.11: Cuencas de atracción asociadas al método con $\alpha = i$.

Capítulo 7

Estudio de la convergencia y dinámica de una familia altamente eficiente

Resumen

En este capítulo, presentamos una nueva familia de métodos iterativos tipo Jarratt de sexto orden altamente eficiente para resolver ecuaciones no lineales junto con sus propiedades de convergencia. Además, ampliamos esta familia al caso multidimensional conservando el mismo orden de convergencia. Para ilustrar la aplicabilidad de nuestros métodos, elegimos algunos problemas del mundo real, concretamente, la síntesis cinemática descrito en forma de sistema de ecuaciones no lineales, valor límite o problema de contorno, un sistema de ecuaciones no lineales de dimensión 100×100 , la ecuación de Bratu bidimensional, la ecuación de Fisher con condiciones de frontera de Neumann homogéneas, y la ecuación integral de Hammerstein en el caso de sistemas no lineales (ver [92, 130] para más aplicaciones). Además, se realizan comparaciones numéricas, que confirman los resultados teóricos, para mostrar el rendimiento de nuestras técnicas iterativas con las ya existentes. Además, encontramos que nuestras técnicas funcionan mejor en comparación con algunos conocidos métodos ya existentes del mismo orden en términos de error residual y en términos de los errores entre dos iteraciones consecutivas. Finalmente, el análisis de estabilidad de nuestros métodos también lo respalda en gran medida.

Para finalizar el capítulo, realizamos el estudio dinámico de un caso especial del método con polinomios cuadráticos. Pintando los planos de parámetros asociados a alguno de sus puntos críticos libres, para detectar anomalías y posteriormente describir los planos dinámicos donde podemos ver regiones sin convergencia, regiones con convergencia a las raíces y regiones con convergencia a n -ciclos atractores.

7.1. Introducción

En este artículo, consideramos

$$F(x) = 0, \tag{7.1.1}$$

donde $F : \mathbb{I} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función univariante cuando $n = 1$ o función multivariante cuando $n > 1$ en un dominio abierto \mathbb{I} .

La construcción de métodos iterativos de orden superior para aproximar las soluciones de (7.1.1) es una de las más importantes y desafiantes tareas en el campo del análisis numérico. La importancia de este tema ha llevado al desarrollo de muchas técnicas numéricas. Sin embargo, la mayoría de ellas son de naturaleza iterativa porque los métodos analíticos para estos problemas son casi inexistentes. Por lo tanto, los académicos de todo el mundo están haciendo todo lo posible para utilizar métodos iterativos en las últimas décadas. Además, los métodos iterativos proporcionan una solución aproximada correcta hasta un determinado grado de precisión que depende además del método iterativo considerado y del software de programación utilizado, a saber, Maple, MATLAB, Mathematica, Fortran, etc. Además, los investigadores deben enfrentarse a varios problemas al usarlos. Algunos de ellos están relacionados con que la convergencia sea más lenta, la no convergencia, el problema de oscilación cerca de la aproximación inicial, divergencia, errores, etc. (para la explicación detallada de estos problemas, consulte [177, 215]).

Hay varios ejemplos donde podemos ver la aplicabilidad de estos métodos a problemas del mundo real. Por ejemplo, Moré [162] propuso una colección de problemas de modelos no lineales y la mayoría de ellos están redactados en términos de $F(x) = 0$. Por otro lado, Grosan y Abraham [107], también debatieron la aplicabilidad de los sistemas no lineales en la neurofisiología, la síntesis cinemática, el equilibrio químico, la combustión y los problemas de modelado económico. Además, otros problemas se resolvieron en [34, 216] al expresarlos en forma de $F(x) = 0$. Además, Lin et al. [146] también trabajaron la aplicabilidad del sistema no lineal en la teoría del transporte.

En el pasado y en los últimos años, los investigadores propusieron una gran cantidad de métodos de un punto y métodos multipunto para obtener las soluciones de (7.1.1) (podemos ver una descripción en [18, 20, 177, 215]). La mayoría de ellos son la extensión del método clásico de Newton o métodos tipo Newton a expensas de algunas evaluaciones funcionales adicionales.

Hay dos formas principales de construir métodos iterativos de orden superior para sistemas no lineales.

En la primera de ellas, los investigadores propusieron nuevos métodos iterativos para la ecuación escalar y a continuación, extendieron el mismo esquema al caso multidimensional conservando el mismo orden de convergencia.

Por ejemplo, en 2010, Cordero et al. [67], extendieron el método de Jarratt de cuarto orden [144] para ecuaciones escalares al sistema de ecuaciones no lineales.

Además, Abad et al. [1], Cordero et al. [65] y Wang y Zhang [221], presentaron una extensión de orden superior para sistemas no lineales del trabajo previamente publicado para las ecuaciones escalares. Recientemente, Narang et al. [164] también propusieron métodos de cuarto y sexto orden para sistemas no lineales que eran la extensión de los esquemas anteriores. Podemos decir que es una de las formas simples de desarrollar un nuevo esquema para el sistema de ecuaciones no lineales. Pero, no siempre es posible hacerlo de la misma forma. Una de las razones principales que subyace es el gasto computacional al calcular la evaluación de la función involucrada y su derivada, en el caso de la ecuación escalar es el mismo. Sin embargo, esto no es cierto en el caso multidimensional.

En la segunda manera, los investigadores también utilizaron otros enfoques y procedimientos para desarrollar métodos de orden superior para el sistema de ecuaciones no lineales. Recientemente, Sharma et al. [205] propusieron métodos iterativos de cuarto y sexto orden basados en la iteración ponderada de Newton. Muy recientemente, Artidiello et al. [33] propusieron métodos de cuarto orden basados en el enfoque de función de peso. Diferentes investigadores han utilizado fórmulas de cuadratura, polinomios y enfoques de diferencias divididas, etc. para construir esquemas iterativos para sistemas no lineales. Para ver los detalles de estos otros enfoques podemos referirnos a algunos libros de texto estándar y artículos de investigación [16, 26, 171, 177, 215].

En este documento, nuestro objetivo es proponer una familia de sexto orden de métodos tipo Jarratt para resolver ecuaciones escalares. Se realiza un análisis cualitativo para seleccionar algunos elementos de la familia con buenas propiedades de estabilidad. Entonces, ampliamos esta familia para el caso multidimensional preservando el mismo orden de convergencia. La eficacia de los mismos se prueba en varios ejemplos numéricos y se observa que funcionan mejor que muchos existentes en el caso escalar. Además, también hemos comprobado la aplicabilidad de ellos en el caso multidimensional en algunos ejemplos de la vida real.

El resto del capítulo está organizado de la siguiente manera: proponemos una nueva familia de sexto orden del método tipo Jarratt para ecuaciones escalares y también damos una prueba de su orden de convergencia en la Sección 2. La Sección 3 está dedicada a la generalización de la misma familia para sistemas no lineales que conservan el mismo orden de convergencia. En la Sección 4, realizamos las pruebas numéricas que confirman los resultados teóricos y nos permiten compararlos con algunos conocidos métodos existentes del mismo orden.

Ahora, vamos a recordar algunos conceptos relacionados con la convergencia de un método iterativo.

Definición 7.1 *Consideremos una sucesión $\{x_n\}$ que converge a una raíz ξ de $f(x) = 0$, donde f es una función de valores reales. Sea $e_n = x_n - \xi$ el error en la n -ésima iteración. Si existen las constantes $p \geq 1$, $c \neq 0$ tales que $e_{n+1} = ce_n^p + O(e_n^{p+1})$, conocida como la ecuación del error, entonces a p y $\eta = |c|$ les llamamos orden de convergencia y constante de error asintótico, respectivamente. A partir*

de esta definición, la constante de error asintótico es $\eta = |c| = \lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n^p|}$.

Con la ayuda de los conceptos previos, podemos definir el *orden asintótico de convergencia* como

$$p = \lim_{n \rightarrow \infty} \frac{|e_{n+1}/\eta|}{|e_n^p|}.$$

Pero, el principal inconveniente de calcular η de acuerdo a la fórmula anterior, es que involucra a la raíz exacta ξ . Existen muchas situaciones reales donde la raíz exacta no se conoce de antemano. Para superar este problema, podemos usar $(x_{n+1} - x_n)$ en vez de (e_{n+1}) en la fórmula anterior para calcular η . De manera similar, podemos considerar los mismos conceptos para el sistema de ecuaciones no lineales usando cualquier norma en \mathbb{R}^n .

7.2. Desarrollo del esquema de ecuaciones escalares

En esta sección, proponemos una nueva familia de métodos iterativos de sexto orden, que se define de la siguiente manera:

$$\begin{aligned} y_n &= x_n - \frac{2}{3} \frac{f(x_n)}{f'(x_n)}, \\ z_n &= x_n - \left[a_1 + a_2 \left(\frac{f'(x_n)}{f'(y_n)} \right)^2 \right] \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \left[\frac{f'(x_n) + b_1 f'(y_n)}{b_2 f'(x_n) + b_3 f'(y_n)} \right] \frac{f(z_n)}{f'(x_n)}, \end{aligned} \quad (7.2.1)$$

donde a_1 , a_2 , b_1 , b_2 y b_3 son parámetros de libre disposición. El siguiente teorema 7.2 demuestra que el orden de convergencia alcanza seis mediante el uso de valores particulares de los parámetros libres.

Teorema 7.2 *Sea $f : \mathbb{D} \subseteq \mathbb{R} \rightarrow \mathbb{R}$ una función suficientemente diferenciable en un intervalo \mathbb{D} conteniendo una raíz simple ξ de la ecuación $f(x) = 0$. Además, también asumimos que x_0 es una aproximación inicial suficientemente cercana a ξ garantizando la convergencia. Entonces, la familia de métodos iterativos (7.2.1) alcanza el sexto orden cuando*

$$a_1 = 1 - a_2, \quad a_2 = \frac{3}{8}, \quad b_2 = b_1 - b_3 + 1, \quad b_3 = \frac{1}{2}(5b_1 + 3), \quad (7.2.2)$$

donde b_1 es un parámetro libre.

Demostración.- : Consideremos que $e_n = x_n - \xi$ es el error en la n -ésima iteración. El desarrollo en serie de Taylor de la función $f(x_n)$ y su derivada de

primer orden $f'(x_n)$ alrededor de $x = \xi$ con la suposición de que $f'(\xi) \neq 0$ nos lleva a:

$$f(x_n) = f'(\xi) \left(e_n + c_2 e_n^2 + c_3 e_n^3 + c_4 e_n^4 + c_5 e_n^5 + c_6 e_n^6 + O(e_n^7) \right), \quad (7.2.3)$$

donde $c_l = \frac{f^{(l)}(\xi)}{l!f'(\xi)}$ para $l = 2, 3, \dots, 6$ y

$$f'(x_n) = f'(\xi) \left(1 + 2c_2 e_n + 3c_3 e_n^2 + 4c_4 e_n^3 + 5c_5 e_n^4 + 6c_6 e_n^5 + O(e_n^6) \right), \quad (7.2.4)$$

respectivamente.

Con la ayuda de las ecuaciones (7.2.3) y (7.2.4), obtenemos

$$\begin{aligned} \frac{f(x_n)}{f'(x_n)} &= e_n - c_2 e_n^2 + 2(c_2^2 - c_3) e_n^3 - (4c_2^3 - 7c_3 c_2 + 3c_4) e_n^4 \\ &+ (8c_2^4 - 20c_3 c_2^2 + 10c_4 c_2 + 6c_3^2 - 4c_5) e_n^5 + (52c_3 c_2^3 - 16c_2^5 - 28c_4 c_2^2 + \\ &(13c_5 - 33c_3^2) c_2 + 17c_3 c_4 - 5c_6) e_n^6 + O(e_n^7). \end{aligned} \quad (7.2.5)$$

Ahora, desarrollamos la función $f'(y_n) = f' \left(x_n - \frac{2}{3} \frac{f(x_n)}{f'(x_n)} \right)$ sobre ξ usando (7.2.5), y queda de la siguiente manera:

$$f'(y_n) = f'(\xi) \left[1 + \frac{2c_2 e_n}{3} + \frac{1}{3} (4c_2^2 + c_3) e_n^2 + \sum_{i=3}^6 A_i e_n^i + O(e_n^7) \right], \quad (7.2.6)$$

donde $A_i = A_i(c_2, c_3, \dots, c_6)$.

De manera similar, obtenemos la siguiente expresión haciendo el desarrollo en serie de Taylor de la función $f(z_n)$ sobre ξ ,

$$f(z_n) = f'(\xi) \left[(1 - a_1 - a_2) e_n + \frac{1}{3} (3(a_1 + a_2 - 1)^2 + 3a_1 - 5a_2) c_2 e_n^2 + \sum_{j=3}^6 B_j e_n^j + O(e_n^7) \right], \quad (7.2.7)$$

donde $B_j = B_j(a_1, a_2, c_2, c_3, \dots, c_6)$.

Insertando las expresiones (7.2.3) – (7.2.7) en el esquema (7.2.1) y después de simplificar, llegamos a

$$e_{n+1} = \frac{(a_1 + a_2 - 1)(b_1 - b_2 - b_3 + 1) e_n}{b_2 + b_3} + \sum_{k=2}^6 D_k e_n^k + O(e_n^7), \quad (7.2.8)$$

donde $D_k = D_k(a_1, a_2, b_1, b_2, b_3, c_2, c_3, \dots, c_6)$.

De la ecuación (7.2.8) es claro que obtenemos al menos convergencia cuadrática sustituyendo el siguiente valor del parámetro libre a_1

$$a_1 = 1 - a_2. \quad (7.2.9)$$

Para obtener al menos convergencia cúbica, utilizamos el valor anterior de a_1 y consideramos $D_2 = 0$ (para los detalles de D_2 por favor, revisar (7.2.8)). Entonces, obtenemos una relación independiente, que se define de la siguiente manera:

$$\frac{(8a_2 - 3)(b_1 - b_2 - b_3 + 1)c_2}{3(b_2 + b_3)} = 0, \quad (7.2.10)$$

que rinde más

$$\text{ya sea } a_2 = \frac{3}{8} \text{ o } b_2 = b_1 - b_3 + 1. \quad (7.2.11)$$

Sin pérdida de generalidad, escogemos $a_2 = \frac{3}{8}$ y utilizando este valor, D_3 se vuelve automáticamente cero (porque $(a_2 - \frac{3}{8})$ es un factor común en D_2 y D_3). Entonces, usamos a expresión (7.2.9) y $a_2 = \frac{3}{8}$ en $D_4 = 0$, obtenemos la siguiente relación independiente

$$- \frac{(b_1 - b_2 - b_3 + 1)(13c_2^3 - 9c_3c_2 + c_4)}{9(b_2 + b_3)} = 0, \quad (7.2.12)$$

de la que se desprende

$$b_2 = b_1 - b_3 + 1. \quad (7.2.13)$$

Insertando las ecuaciones (7.2.9), (7.2.13) y $a_2 = \frac{3}{8}$ en $D_5 = 0$, obtenemos

$$\frac{2(5b_1 - 2b_3 + 3)c_2(13c_2^3 - 9c_3c_2 + c_4)}{27(b_1 + 1)} = 0, \quad (7.2.14)$$

por lo que tenemos que

$$b_3 = \frac{1}{2}(5b_1 + 3). \quad (7.2.15)$$

Usando los valores de a_3 , b_3 y las expresiones (7.2.9), (7.2.13) en (7.2.8), tenemos

$$e_{n+1} = - \left(\frac{(13c_2^3 - 9c_3c_2 + c_4)(2(b_1 - 3)c_2^2 + 3(b_1 + 1)c_3)}{27(b_1 + 1)} \right) e_n^6 + O(e_n^7), \quad b_1 \neq -1. \quad (7.2.16)$$

Esta constante de error asintótico revela que el esquema propuesto (7.2.1) alcanza el sexto orden de convergencia. Esto completa la demostración. \blacksquare

7.3. Desarrollo del esquema para casos multidimensionales

El esquema previo para ecuaciones escalares (7.2.1) podemos reescribirlo para el caso multidimensional como sigue:

$$\begin{aligned} y^{(n)} &= x^{(n)} - \frac{2}{3} F'(x^{(n)})^{-1} F(x^{(n)}), \\ z^{(n)} &= x^{(n)} - \left[a_1 I + a_2 \left(F'(y^{(n)})^{-1} F'(x^{(n)}) \right)^2 \right] F'(x^{(n)})^{-1} F(x^{(n)}), \\ x^{(n+1)} &= z^{(n)} - \left(b_2 F'(x^{(n)}) + b_3 F'(y^{(n)}) \right)^{-1} \left(F'(x^{(n)} + b_1 F'(y^{(n)})) F'(x^{(n)})^{-1} F(z^{(n)}) \right), \end{aligned} \quad (7.3.1)$$

donde a_1 , a_2 , b_2 , b_3 tienen los mismos valores que los definidos en el Teorema 7.2, $b_1 \in \mathbb{R}$ es el parámetro libre e I es la matriz identidad de orden n . El siguiente Teorema 7.3 demuestra que el esquema (7.3.1) alcanza el sexto orden de convergencia. Para este propósito, utilizamos las mismas herramientas y procedimientos que se introdujeron en [67].

Teorema 7.3 *Sea $F : \mathbb{I} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función suficientemente diferenciable en un entorno abierto \mathbb{I} de su cero ξ^* . Además, suponemos que $F'(x)$ es continua y no singular para ξ^* y la aproximación inicial $x^{(0)}$ está suficientemente cerca a ξ^* para garantizar la convergencia. Entonces, el esquema (7.3.1) tiene sexto orden de convergencia cuando*

$$a_1 = 1 - a_2, \quad a_2 = \frac{3}{8}, \quad b_2 = b_1 - b_3 + 1, \quad b_3 = \frac{1}{2}(5b_1 + 3),$$

donde b_1 es un parámetro libre.

Demostración.- Asumamos que $e^{(n)} = x^{(n)} - \xi^*$ es el error en la n -ésima iteración. Además, al desarrollar $F(x^{(n)})$ en un entorno de ξ^* , tenemos

$$\begin{aligned} F(x^{(n)}) &= \\ F'(\xi^*) \left[e^{(n)} + C_2(e^{(n)})^2 + C_3(e^{(n)})^3 + C_4(e^{(n)})^4 + C_5(e^{(n)})^5 + C_6(e^{(n)})^6 \right] + O((e^{(n)})^7), \end{aligned} \quad (7.3.2)$$

donde $C_k = \frac{1}{k!} F'(\xi^*)^{-1} F^{(k)}(\xi^*)$, $k \geq 2$.

Similarmente, obtenemos

$$\begin{aligned} F'(x^{(n)}) &= \\ F'(\xi^*) \left[I + 2C_2 e^{(n)} + 3C_3 (e^{(n)})^2 + 4C_4 (e^{(n)})^3 + 5C_5 (e^{(n)})^4 + 6C_6 (e^{(n)})^5 \right] \\ + O((e^{(n)})^6). \end{aligned} \quad (7.3.3)$$

Al usar la expresión anterior (7.3.3), además tenemos

$$\begin{aligned}
F'(x^{(n)})^{-1} = \\
\left[I - 2C_2e^{(n)} + (4C_2^2 - 3C_3)(e^{(n)})^2 - (8C_2^3 - 6C_2C_3 - 6C_3C_2 + 4C_4)(e^{(n)})^3 \right] F'(\xi^*)^{-1} \\
+ O((e^{(n)})^4),
\end{aligned} \tag{7.3.4}$$

Con ayuda de la ecuación (7.3.2) y (7.3.4), tenemos

$$\begin{aligned}
F'(x^{(n)})^{-1}F(x^{(n)}) = \\
e^{(n)} - C_2(e^{(n)})^2 + 2(C_2^2 - C_3)(e^{(n)})^3 - (4C_2^3 - 4C_2C_3 - 3C_3C_2 + 3C_4)(e^{(n)})^4 \\
+ O((e^{(n)})^5),
\end{aligned} \tag{7.3.5}$$

Al usar la expresión anterior (7.3.5) en el primer paso de (7.3.1), obtenemos

$$\begin{aligned}
y^{(n)} - \xi^* = \\
\frac{1}{3}e^{(n)} + \frac{2}{3}C_2(e^{(n)})^2 - \frac{2}{3}(2C_2^2 - 2C_3)(e^{(n)})^3 + 2(4C_2^3 - 4C_2C_3 - 3C_2C_3 + 3C_4)(e^{(n)})^4 \\
+ O((e^{(n)})^5).
\end{aligned} \tag{7.3.6}$$

Con la ayuda de la expresión (7.3.6), tenemos además que

$$\begin{aligned}
F'(y^{(n)}) = \\
F'(\xi^*) \left[I + \frac{2}{3}C_2e^{(n)} + \frac{1}{3}(4C_2^2 + C_3)(e^{(n)})^2 - \frac{4}{27}(18C_2^3 - 18C_2C_3 - 9C_3C_2 - 6C_4)(e^{(n)})^3 \right] \\
+ O((e^{(n)})^4)
\end{aligned} \tag{7.3.7}$$

y

$$\begin{aligned}
F'(y^{(n)})^{-1} = \left[I - \frac{2C_2}{3}e^{(n)} - \frac{1}{9}(8C_2^2 + 3C_3)(e^{(n)})^2 + \frac{2}{27}(28C_2^3 - 33C_2C_3 - 15C_3C_2 \right. \\
\left. - 2C_4)(e^{(n)})^3 \right] F'(\xi^*)^{-1} + O((e^{(n)})^4).
\end{aligned} \tag{7.3.8}$$

Utilizando las ecuaciones (7.3.3) y (7.3.7), conseguimos

$$\begin{aligned}
F'(y^{(n)})^{-1}F'(x^{(n)}) = \\
I + \frac{4C_2}{3}e^{(n)} + \frac{4}{9}(6C_3 - 5C_2^2)(e^{(n)})^2 + \frac{8}{27}(8C_2^3 - 15C_2C_3 - 6C_3C_2 + 13C_4)(e^{(n)})^3 \\
+ O((e^{(n)})^4).
\end{aligned} \tag{7.3.9}$$

De nuevo, con (7.3.5) y (7.3.9), tenemos que

$$\begin{aligned}
& \left[a_1 I + a_2 \left(F'(y^{(n)})^{-1} F'(x^{(n)}) \right)^2 \right] F'(x^{(n)})^{-1} F(x^{(n)}) = \\
& (a_1 + a_2) e^{(n)} + \frac{1}{3} (5a_2 - 3a_1) C_2 (e^{(n)})^2 + \frac{2}{3} (3a_1 - 5a_2) (C_2^2 - C_3) (e^{(n)})^3 \\
& + \left[\left(3a_1 - \frac{7a_2}{3} \right) C_3 C_2 + \frac{4}{3} (3a_1 - 5a_2) C_2 C_3 \right. \\
& \left. + \left(\frac{76a_2}{27} - 4a_1 \right) C_2^3 + \left(\frac{127a_2}{27} - 3a_1 \right) C_4 \right] (e^{(n)})^4 + B_1 (e^{(n)})^5 + B_2 (e^{(n)})^6 \\
& + O((e^{(n)})^7),
\end{aligned} \tag{7.3.10}$$

donde B_1 y B_2 dependen de las constantes C_j .

Ahora, utilizando las expresiones (7.3.10) y los valores de los parámetros libres a_1 y a_2 (que se encuentran en el Teorema anterior 7.2), en el segundo paso del esquema (7.3.1), obtenemos

$$z^{(n)} - \xi^* = A_1 (e^{(n)})^4 + A_2 (e^{(n)})^5 + A_3 (e^{(n)})^6 + O((e^{(n)})^7), \tag{7.3.11}$$

donde $A_1 = \left(-\frac{13}{9} C_2^3 + C_3 C_2 - \frac{c_4}{9} \right)$ y A_2 y A_3 están en función de algunos C_j como A_1 .

Además, desarrollamos $F(z^{(n)})$ en un entorno de ξ^*

$$F(z^{(n)}) = F'(\xi^*) \left[A_1 (e^{(n)})^4 + A_2 (e^{(n)})^5 + A_3 (e^{(n)})^6 \right] + O((e^{(n)})^7). \tag{7.3.12}$$

Con ayuda de las expresiones (7.3.3), (7.3.5), (7.3.7), (7.3.12) y los valores de los parámetros libres b_2 y b_3 (que se encuentran en el Teorema anterior 7.2), tenemos que

$$\begin{aligned}
& \left(\frac{-(3b_1 + 1)}{2} F'(x^{(n)}) + \frac{5b_1 + 3}{2} F'(y^{(n)}) \right)^{-1} \left(F'(x^{(n)} + b_1 F'(y^{(n)})) F'(x^{(n)})^{-1} F(z^{(n)}) \right) \\
& = A_1 (e^{(n)})^4 + A_2 (e^{(n)})^5 + \frac{(A_1 (2(b_1 - 3) C_2^2 + 3(b_1 + 1) C_3) + 3A_3 (b_1 + 1))}{3(b_1 + 1)} (e^{(n)})^6 \\
& + O((e^{(n)})^7), \quad b_1 \neq 1.
\end{aligned} \tag{7.3.13}$$

Para finalizar, utilizando (7.3.11) y (7.3.13) en el último paso del esquema propuesto (7.3.1), obtenemos

$$\begin{aligned}
x^{(n+1)} - \xi^* & = z^{(n)} - \xi^* - \\
& \left[A_1 (e^{(n)})^4 + A_2 (e^{(n)})^5 \right. \\
& \left. + \frac{(A_1 (2(b_1 - 3) C_2^2 + 3(b_1 + 1) C_3) + 3A_3 (b_1 + 1))}{3(b_1 + 1)} (e^{(n)})^6 + O((e^{(n)})^7) \right] \\
& = -A_1 \frac{(2(b_1 - 3) C_2^2 + 3(b_1 + 1) C_3)}{3(b_1 + 1)} (e^{(n)})^6 + O((e^{(n)})^7).
\end{aligned} \tag{7.3.14}$$

Por lo tanto, el esquema propuesto (7.3.1) alcanza el sexto orden de convergencia. ■

7.4. Estudio dinámico de un caso especial del método (7.3.1)

En primer lugar, recordemos algunos conceptos básicos relacionados con la dinámica. Dada una función racional $R : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$, donde $\hat{\mathbb{C}}$ es la esfera de Riemann, se define la *órbita de un punto* $z_0 \in \hat{\mathbb{C}}$ como

$$\{z_0, R(z_0), R^2(z_0), \dots, R^n(z_0), \dots\}.$$

A un punto $z_0 \in \bar{\mathbb{C}}$, lo llamaremos *punto fijo* de $R(z)$ si verifica que $R(z) = z$. Además, z_0 es un *punto periódico* de periodo $p > 1$ si es un punto que cumple que $R^p(z_0) = z_0$ pero $R^k(z_0) \neq z_0$, para cada $k < p$. Además, a un punto z_0 lo llamamos *pre-periódico* si no es periódico pero existe un $k > 0$ tal que $R^k(z_0)$ es periódico.

Existen diferentes tipos de puntos fijos dependiendo de su multiplicidad asociada $|R'(z_0)|$. Tomando en cuenta la multiplicidad asociada, un punto fijo z_0 lo denominamos:

- *superatractor* si $|R'(z_0)| = 0$
- *atractor* si $|R'(z_0)| < 1$
- *repulsor* si $|R'(z_0)| > 1$
- y *parabólico* si $|R'(z_0)| = 1$.

Los puntos fijos que no corresponden a las raíces del polinomio $p(z)$ les llamaos *puntos fijos extraños*. Por otro lado, un *punto crítico* z_0 es un punto que satisface que, $R'(z_0) = 0$. Además, llamamos *punto crítico libre* a aquellos puntos críticos que no son raíces del polinomio $p(z)$.

La *cuenca de atracción* de un atractor α se define como

$$\mathcal{A}(\alpha) = \{z_0 \in \hat{\mathbb{C}} : R^n(z_0) \rightarrow \alpha, n \rightarrow \infty\}.$$

El *conjunto de Fatou* de una función racional R , $\mathcal{F}(R)$, es el conjunto de los puntos $z \in \hat{\mathbb{C}}$ cuyas órbitas tienen a un atractor (punto fijo, órbita periódica o infinito). Su complementario en $\hat{\mathbb{C}}$ es el *conjunto de Julia*, $\mathcal{J}(R)$. Eso significa que la cuenca de atracción de cualquier punto fijo pertenece al conjunto de Fatou y la frontera de estas cuencas de atracción pertenecen al conjunto de Julia.

Aplicando este operador en un polinomio cuadrático con dos raíces diferentes A y B , $p(z) = (z - A)(z - B)$ y utilizando la transformación de Möebius $h(z) = \frac{z-A}{z-B}$, la cual lleva la raíz A al 0, la raíz B al ∞ y el ∞ al 1, obtenemos que el operador racional asociado a la familia de esquemas iterativos es finalmente

$$G(z, b_1) = -\frac{z^6(9z^2+12z+13)}{(13z^2+12z+9)} \times \frac{(27b_1z^6+54b_1z^5+72b_1z^4+74b_1z^3+19b_1z^2+12b_1z-18b_1+27z^6+90z^5+192z^4+294z^3+291z^2+180z+54)}{(18b_1z^6-12b_1z^5-19b_1z^4-74b_1z^3-72b_1z^2-54b_1z-27b_1-54z^6-180z^5-291z^4-294z^3-192z^2-90z-27)}. \quad (7.4.1)$$

7.4.1. Estudio de los puntos fijos

Es obvio que $z = 0$ y $z = \infty$ son puntos fijos de $G(z, a)$. Además, existen algunos puntos fijos extraños que son:

- $z = 1$ relacionado con la divergencia al ∞
- Las raíces de

$$\begin{aligned} p(z) = & 243b_1z^{12} + 1053b_1z^{11} + 2700b_1z^{10} + 4932b_1z^9 + 6927b_1z^8 \\ & + 8459b_1z^7 + 8748b_1z^6 + 8459b_1z^5 + 6927b_1z^4 + 4932b_1z^3 \\ & + 2700b_1z^2 + 1053b_1z + 243b_1 + 243z^{12} + 1377z^{11} + 4536z^{10} \\ & + 10656z^9 + 19299z^8 + 27531z^7 + 30972z^6 + 27531z^5 + 19299z^4 \\ & + 10656z^3 + 4536z^2 + 1377z + 243 \end{aligned}$$

Las soluciones de este polinomio dependen del valor del parámetro b_1 .

7.4.2. Estudio de los puntos críticos y planos de parámetros

Es un hecho bien conocido que hay al menos un punto crítico asociado con cada componente invariante de Fatou [93, 124]. Los puntos críticos de la familia son las soluciones de $G'(z, a) = 0$, donde

$$\begin{aligned} G'(z, a) = & -\frac{12z^5(z+1)^2(3z^2+2z+3)^3}{(13z^2+12z+9)^2(18b_1z^6-12b_1z^5-19b_1z^4-74b_1z^3-72b_1z^2-54b_1z-27b_1-54z^6-180z^5-291z^4-294z^3-192z^2-90z-27)^2} \\ & \times (1053b_1^2z^8 - 972b_1^2z^7 - 1656b_1^2z^6 - 6468b_1^2z^5 - 5674b_1^2z^4 - 6468b_1^2z^3 - 1656b_1^2z^2 - 972b_1^2z \\ & + 1053b_1^2 - 2106b_1z^8 - 11628b_1z^7 - 22488b_1z^6 - 31460b_1z^5 - 32188b_1z^4 - 31460b_1z^3 \\ & - 22488b_1z^2 - 11628b_1z - 2106b_1 - 3159z^8 - 15336z^7 - 37872z^6 - 61368z^5 - 71346z^4 \\ & - 61368z^3 - 37872z^2 - 15336z - 3159) \end{aligned}$$

Es claro que $z = 0$ y $z = \infty$ son puntos críticos. Además, los puntos críticos libres son las raíces del polinomio:

$$\begin{aligned}
 q(z) = & 1053b_1^2z^8 - 972b_1^2z^7 - 1656b_1^2z^6 - 6468b_1^2z^5 - 5674b_1^2z^4 - 6468b_1^2z^3 - 1656b_1^2z^2 \\
 & - 972b_1^2z + 1053b_1^2 - 2106b_1z^8 - 11628b_1z^7 - 22488b_1z^6 - 31460b_1z^5 - 32188b_1z^4 \\
 & - 31460b_1z^3 - 22488b_1z^2 - 11628b_1z - 2106b_1 - 3159z^8 - 15336z^7 - 37872z^6 \\
 & - 61368z^5 - 71346z^4 - 61368z^3 - 37872z^2 - 15336z - 3159
 \end{aligned}$$

a los que llamamos $cr_i(a)$ para $i = 1, 2, \dots, 8$.

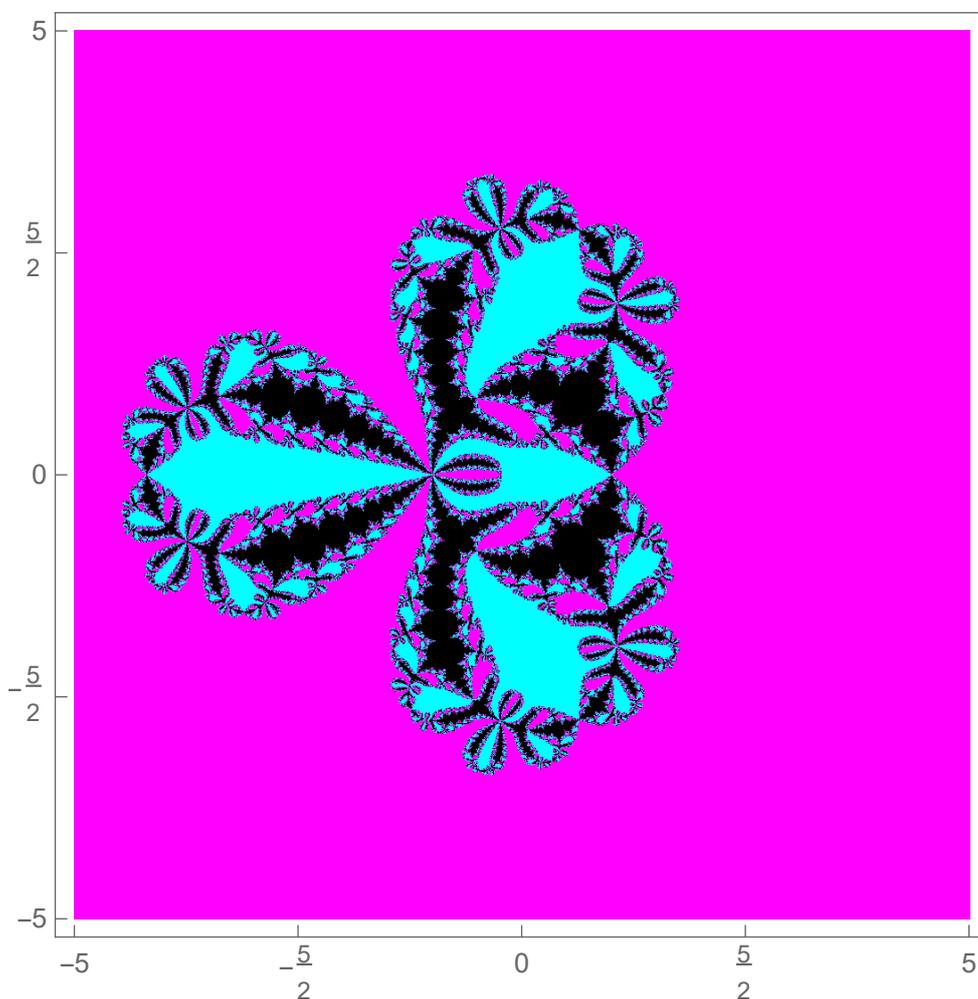


Figura 7.1: Plano de parámetros asociado al punto crítico $cr_8(b_1)$.

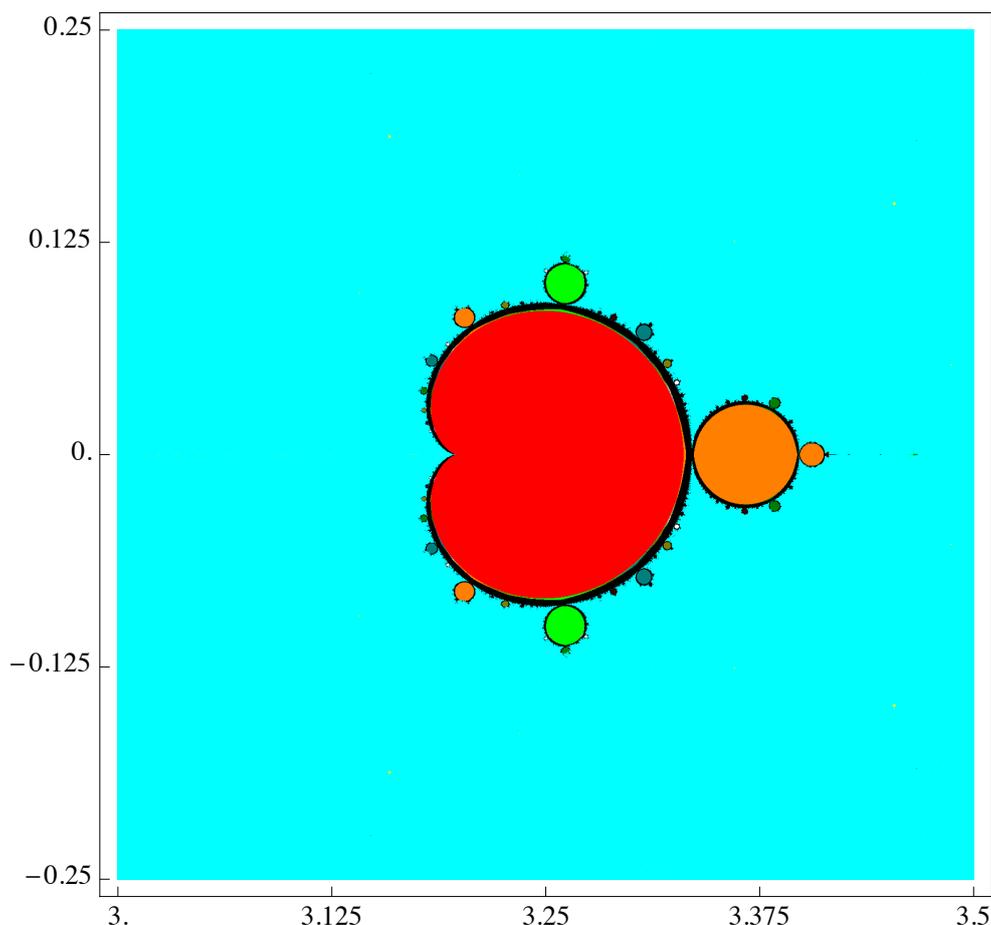


Figura 7.2: Detalle del plano de parámetros asociado al punto crítico $cr_1(b_1)$.

Como podemos ver en muchos artículos, por ejemplo [28, 150], el estudio de las órbitas de los puntos críticos da lugar al comportamiento dinámico de un método iterativo. En concreto, para determinar si existe alguna órbita periódica atractora diferente a las raíces del polinomio $p(z)$, debe responderse la siguiente pregunta: ¿Para qué valores del parámetro, las órbitas de los puntos críticos libres están atrayendo órbitas periódicas? Una forma de dar respuesta a esas preguntas es dibujar los planos de parámetros asociados a los puntos críticos libres. Existirán regiones abiertas del espacio de parámetros para las cuales la iteración de los puntos críticos libres no converge a ninguna de las raíces del polinomio $p(z)$, es decir, en estas regiones los puntos críticos convergen para atraer ciclos o a un punto fijo extraño o incluso divergen.

En las figuras 7.1–7.4, mostramos algunos de los planos de parámetros asociados a los puntos críticos libres en los que observamos que existe una zona sin convergencia a las raíces. Pintamos de cian un punto si la iteración del método eligiendo como punto inicial el punto crítico indicado en el título converge al punto fijo 0 (relacionado con la raíz A) o si converge al ∞ (relacionado con la

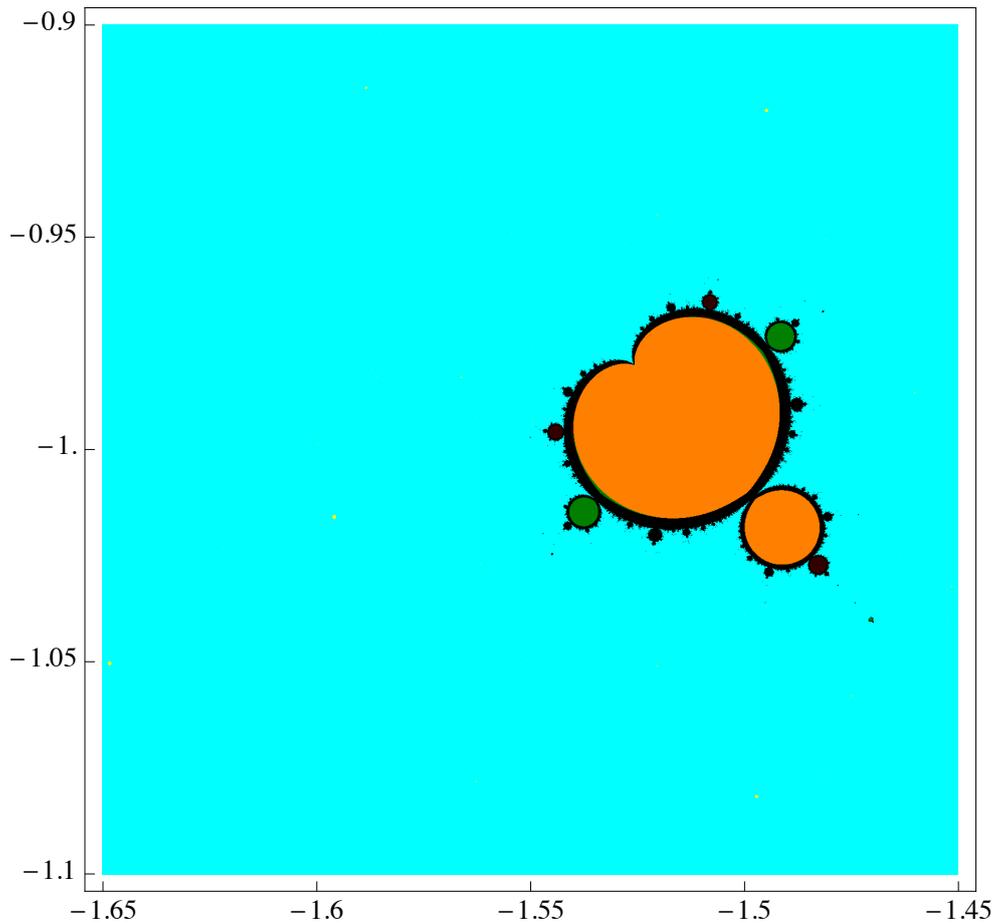


Figura 7.3: Detalle del plano de parámetros asociado al punto crítico $cr_1(b_1)$.

raíz B) y pintamos en amarillo si la iteración converge al 1 (relacionado con el ∞). Además, aparece en rojo la convergencia, después de un máximo de 2000 iteraciones y con una tolerancia de 10^{-6} , a cualquiera de los puntos fijos extraños, en naranja la convergencia a 2-ciclos, en verde claro la convergencia a 3-ciclos, en rojo oscuro a 4-ciclos, en azul oscuro a 5-ciclos, en verde oscuro a 6-ciclos, amarillo oscuro a 7-ciclos, y en blanco la convergencia a 8-ciclos. Las regiones en negro corresponden a zonas de convergencia a otros ciclos. Como consecuencia, cada punto del plano que no es cian ni magenta no es una buena opción de b_1 en términos de comportamiento numérico.

Una vez que se han detectado las anomalías, el siguiente paso consiste en describirlas en los planos dinámicos. En estos planos dinámicos la convergencia a 0 aparecerá en magenta, en cian aparece la convergencia a ∞ y en negro las zonas sin convergencia, después de un máximo de 1000 iteraciones y con una tolerancia de 10^{-6} a las raíces.

En las Figuras 7.7, 7.8, se muestran los planos dinámicos de los miembros de la familia con regiones de convergencia a algunos de los puntos fijos extraños.

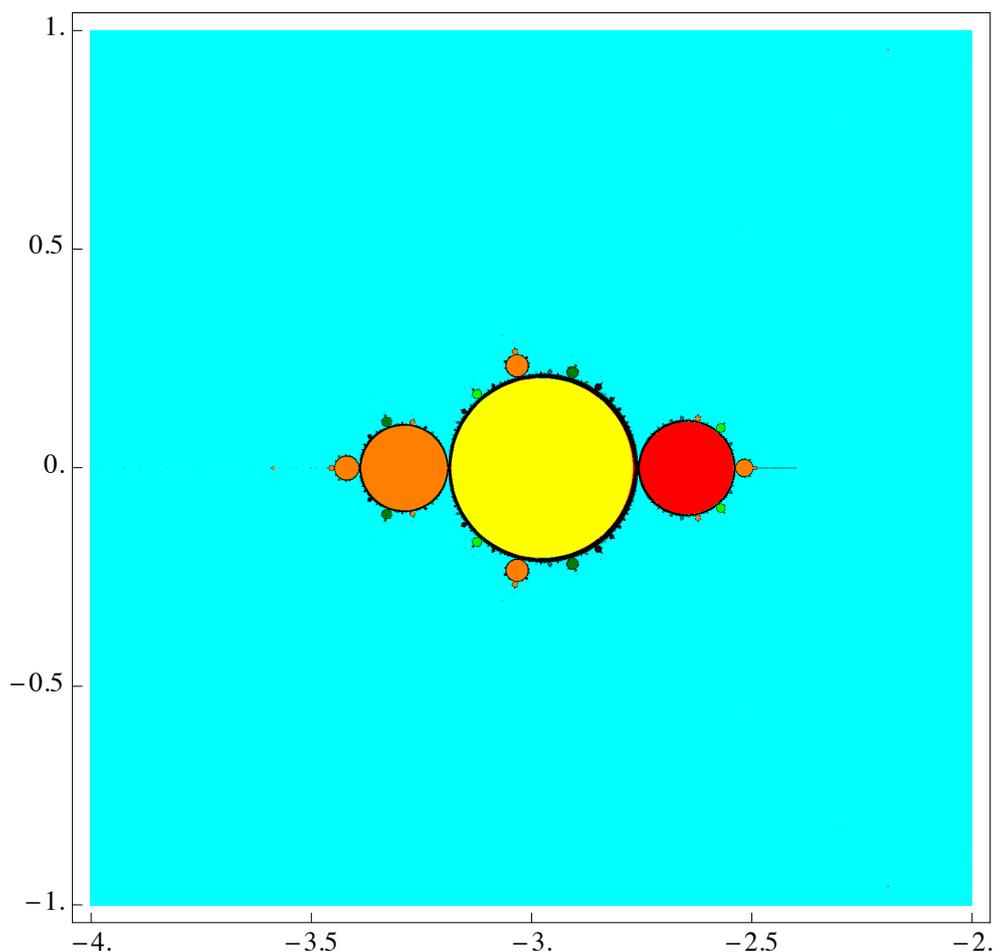


Figura 7.4: Detalle del plano de parámetros asociado al punto crítico $cr_8(b_1)$.

En las Figuras 7.11 y 7.12 mostramos otros casos donde no hay problemas de convergencia, ya que las únicas cuencas de atracción corresponden a las raíces de $p(z)$.

En consecuencia, podemos afirmar que esta familia es muy estable y que las anomalías existentes son muy raras en la práctica.

En las Figuras 8.12–7.6 se muestran los planos dinámicos de los miembros de la familia con las regiones de convergencia a 2-ciclos atractores, estas regiones con convergencia a 2-ciclos están pintadas de negro ya que no hay convergencia a ninguna de las raíces. Ambos valores del parámetro corresponden a zonas rojas en los planos de parámetros.

Por otro lado, en la Figura 7.9, presentamos el plano dinámico de un miembro de la familia con regiones de convergencia a $z = 1$, relacionado con el ∞ , pintado en amarillo. Este valor del parámetro b_1 se encuentra en una zona amarilla de los planos de los parámetros.

Por otro lado, en la Figura 7.10, mostramos el plano dinámico de un miembro

de la familia con regiones de convergencia a un 3-ciclo atractor.

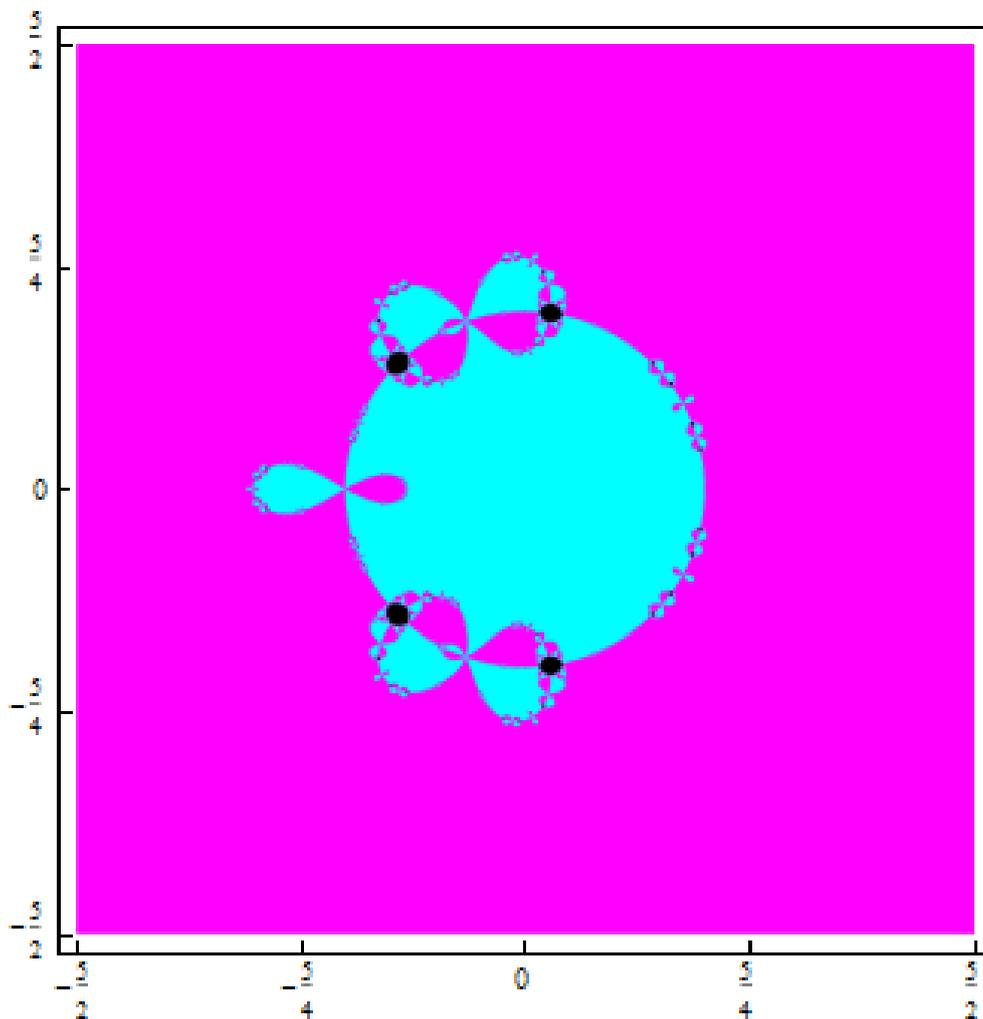


Figura 7.5: Cuenca de atracción asociada al método con $b_1 = 3.25$.

7.5. Experimentos numéricos

Esta sección está dedicada a verificar la validez y efectividad de nuestros resultados teóricos que hemos propuesto anteriormente. Además, queremos demostrar que son competitivos frente a los métodos de sexto orden existentes.

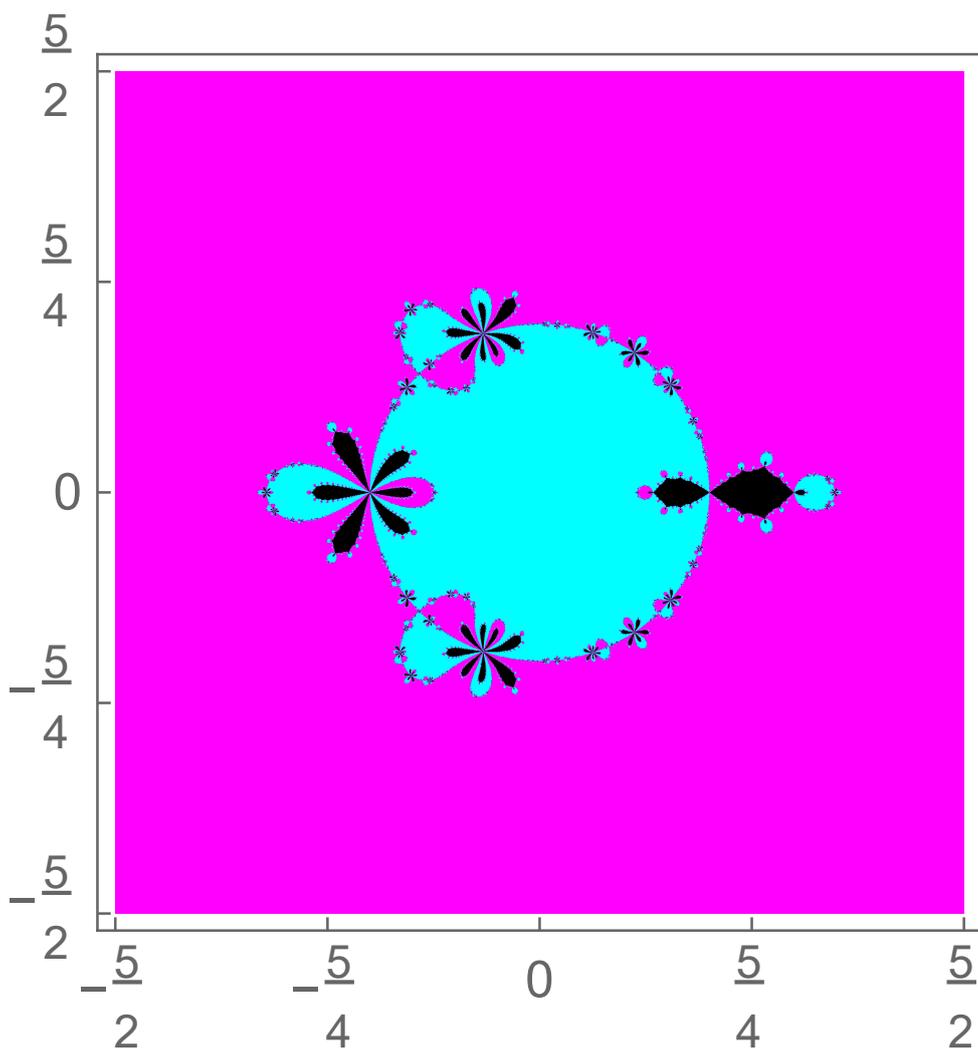


Figura 7.6: Cuenca de atracción asociada al método con $b_1 = 2.6$.

Algunos investigadores que desean afirmar que sus métodos son superiores a otros métodos existentes disponibles en la literatura, consideran únicamente algunos ejemplos ya conocidos o estándares o hechos a sí mismos y manipulan las aproximaciones iniciales para afirmar que sus métodos son superiores a otros. Para detener esta práctica, consideramos cuatro ejemplos numéricos: el primero, el segundo, el tercero y el cuarto se eligen de Kou y Li [134], Chun y Ham [63], Cordero et al. [66] y Ren et al. [191], respectivamente con las mismas aproximaciones iniciales que se mencionaron en esos artículos. Además, también queremos ver qué sucede al considerar diferentes ejemplos y con diferentes aproximaciones iniciales, que no hayan sido mencionadas en sus artículos. Por lo tanto, consideramos el quinto ejemplo de Behl et al. [41]. Los detalles de los ejemplos elegidos o las funciones de prueba están disponibles en la Tabla 7.1. Además, las funciones de prueba consideradas con sus ceros correspondientes también se dan en la

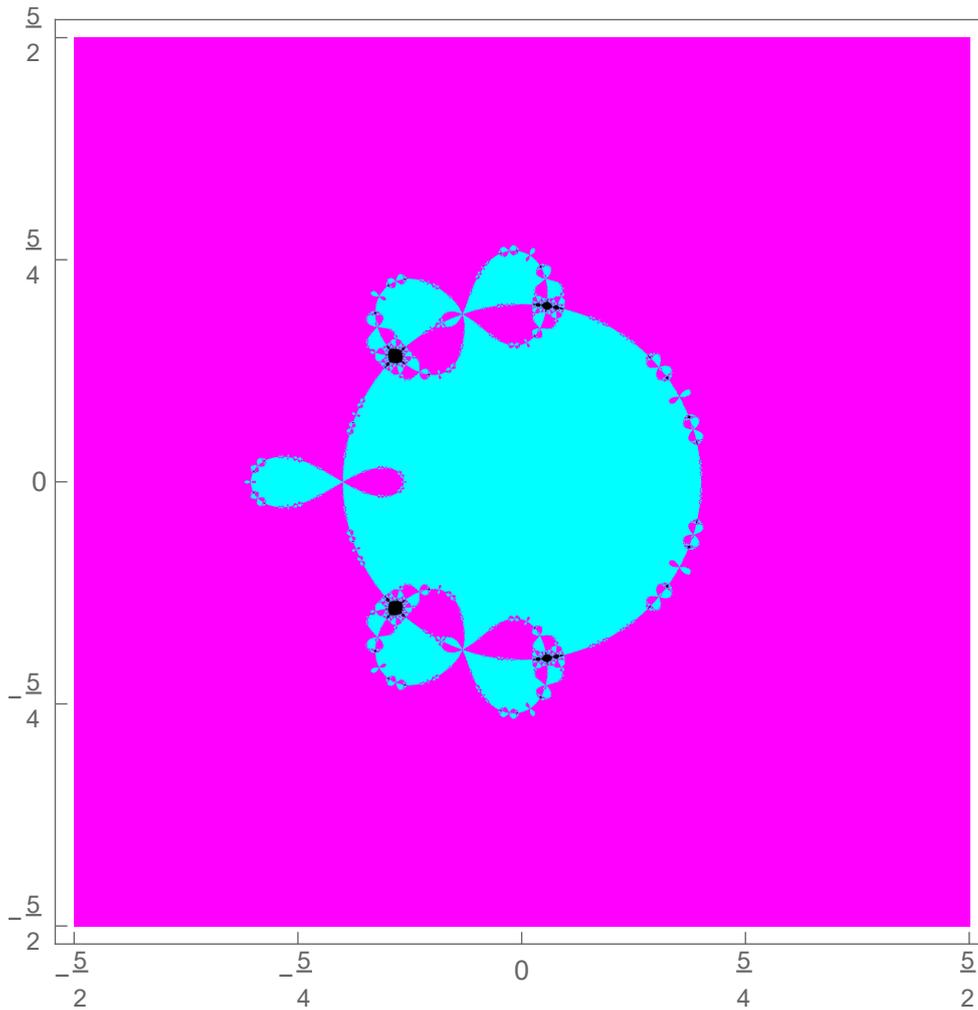


Figura 7.7: Cuenca de atracción asociada al método con $b_1 = 3.375$.

misma tabla.

Cuadro 7.1: Problemas de prueba

$f(x)$	$Raíz(\xi)$	x_0
$f_1(x) = x^3 + 4x^2 - 10$; [134]	1.36523001341409	2
$f_2(x) = x^2 - e^x - 3x + 2$; [63]	0.25753028543986	1
$f_3(x) = e^{-x} + \cos x$; [66]	1.74613953040801	2
$f_4(x) = \sin x - \frac{x}{3}$; [191]	2.27886266007582	2
$f_5(x) = \sin^{-1}(x^2 - 1) - \frac{x}{2} + 1$; [41]	0.59481096839836	0.7

Ahora, empleamos el nuevo esquema de sexto orden (7.2.1) para $b_1 = 3$,

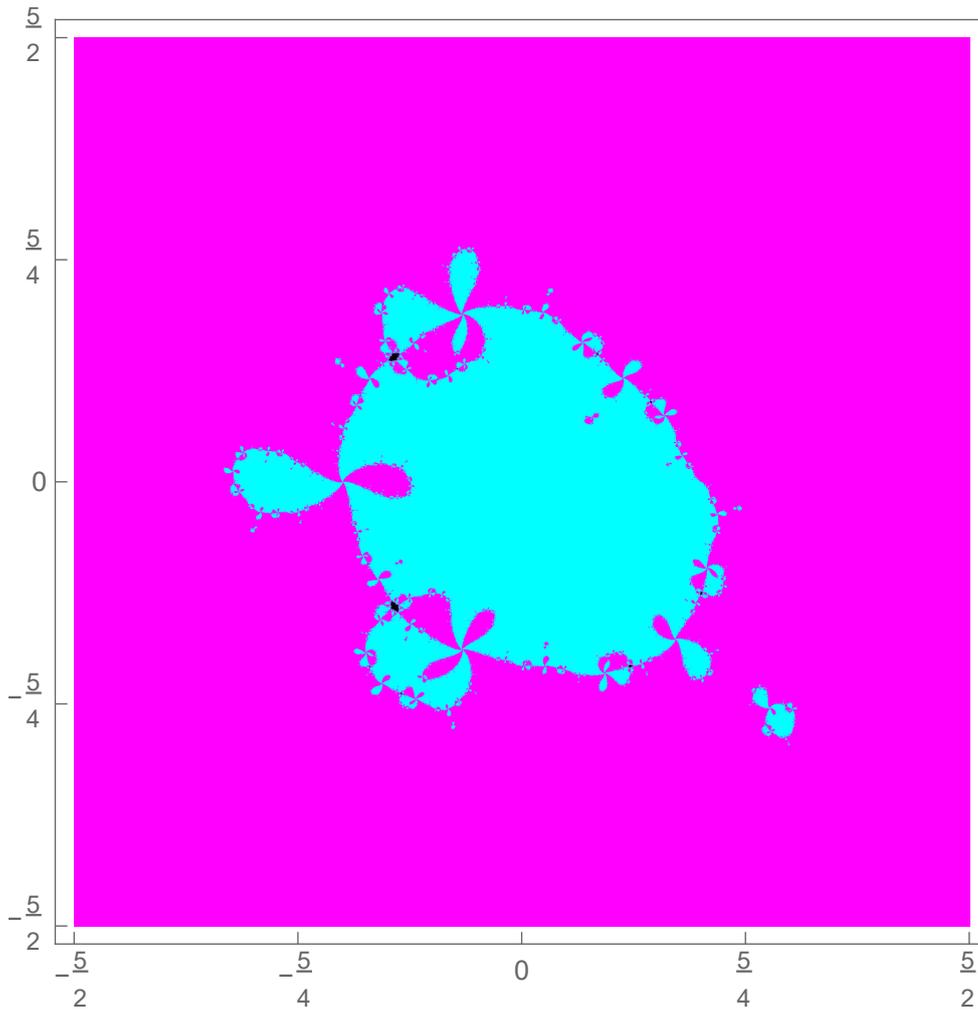


Figura 7.8: Cuenca de atracción asociada al método con $b_1 = -1.5 - i$.

$b_1 = 3.01$, $b_1 = 1$, y $b_1 = -0.6$ denotados por (OM_6^1) , (OM_6^2) , (OM_6^3) y (OM_6^4) , respectivamente para verificar la efectividad y validez de los resultados teóricos. Los comparamos con un método Jarratt de sexto orden propuesto por Kou y Li [134], de ellos consideramos el (método 8) denotado por (KLM) . Además, consideramos las variantes de Ostrowski de sexto orden propuestas por Chun y Ham en [63], y de ellas elegimos una de sus mejores expresiones (24–26) para $(\mu = 1, \gamma = 1)$, conocido como (CHM) . Además, también los comparamos con un método de sexto orden de Homeier modificado presentado por Cordero et al. [66], y de entre ellos consideramos el método (10–12), que llamaremos (CM) . Para finalizar, también los comparamos con algunas variantes nuevas del método de Jarratt teniendo una convergencia de sexto orden diseñada por Ren et al. [191], de ellos elegimos el método (54) (para $\alpha = \frac{5}{10}$, $\beta = \frac{12}{10}$, $\gamma = \frac{2}{10}$, $\delta = \frac{2}{10}$) denotado como (RM) . Para una mejor comparación de los métodos propuestos, mostramos el número de la iteración (n), los ceros aproximados (x_n), el error

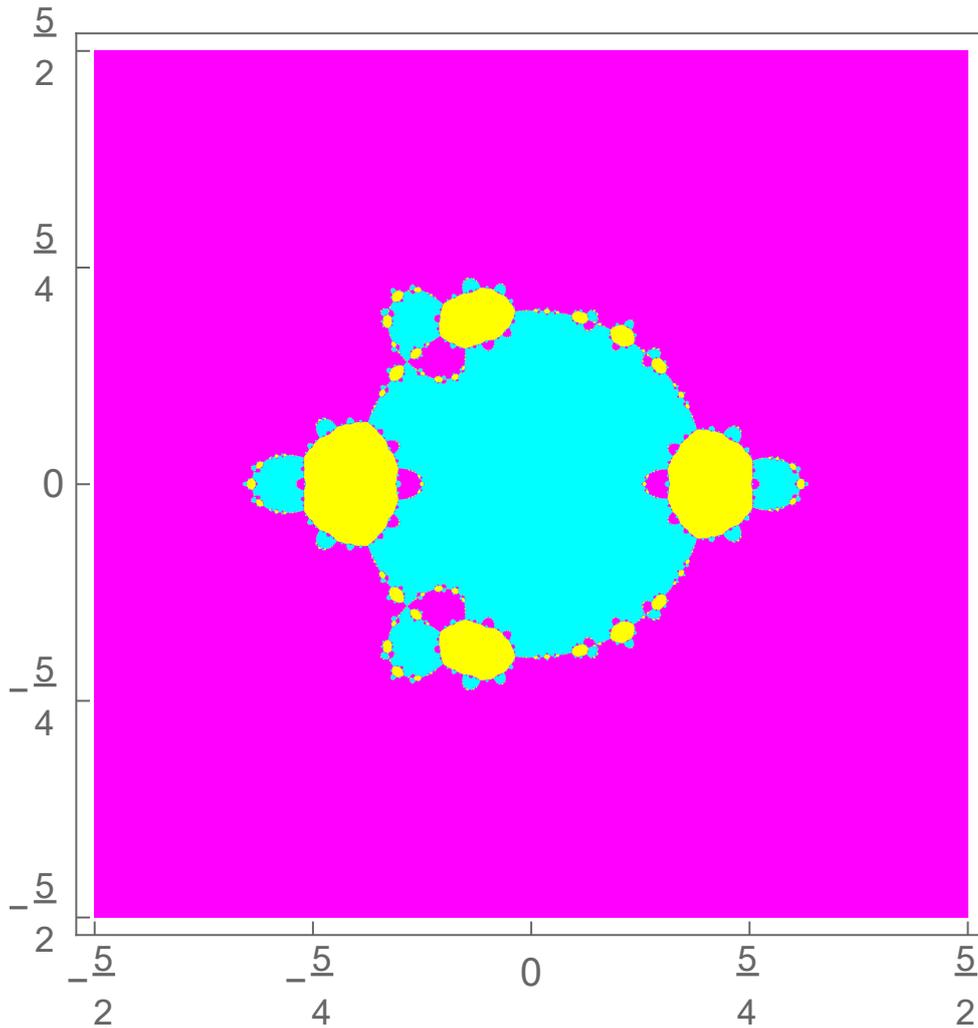


Figura 7.9: Cuenca de atracción asociada al método con $b_1 = -3$.

residual correspondiente a la función ($|f(x_n)|$), el error en cada dos iteraciones consecutivas $|x_{n+1} - x_n|$, la aproximación de la constante de error asintótico $\eta = \frac{|x_n - x_{n-1}|}{|x_{n-1} - x_{n-2}|^6}$ y la estimación del orden de convergencia $\rho = \frac{\log |(x_{n+1} - x_n)/(x_n - x_{n-1})|}{\log |(x_n - x_{n-1})/(x_{n-1} - x_{n-2})|}$ correspondiente para cada función de prueba en tablas 7.2 – 7.6. Los valores de η representados en tablas son el último valor calculado de $\frac{|x_{n+1} - x_n|}{|x_n - x_{n-1}|^6}$.

En el contexto del sistema de ecuaciones no lineales, también consideramos una variedad de ejemplos aplicados (1)–(3) para verificar adicionalmente su validez. Por lo tanto, empleamos el nuevo esquema de sexto orden (7.3.1) para $b_1 = 3, 3.01, 1, -0.6$ y $b_1 = -3$, denotado por $(\widehat{OM}_6^1), (\widehat{OM}_6^2), (\widehat{OM}_6^3), (\widehat{OM}_6^4)$, y (\widehat{OM}_6^5) , respectivamente. Los comparamos con los métodos de sexto orden (8) y (5) presentados por Abbasbandy et al. [2] y Lotfi et al. [147], conocidos como (AM) y (LM) , respectivamente. Además, los comparamos con una familia de métodos iterativos de sexto orden propuesta por Soleymani et al. [212]

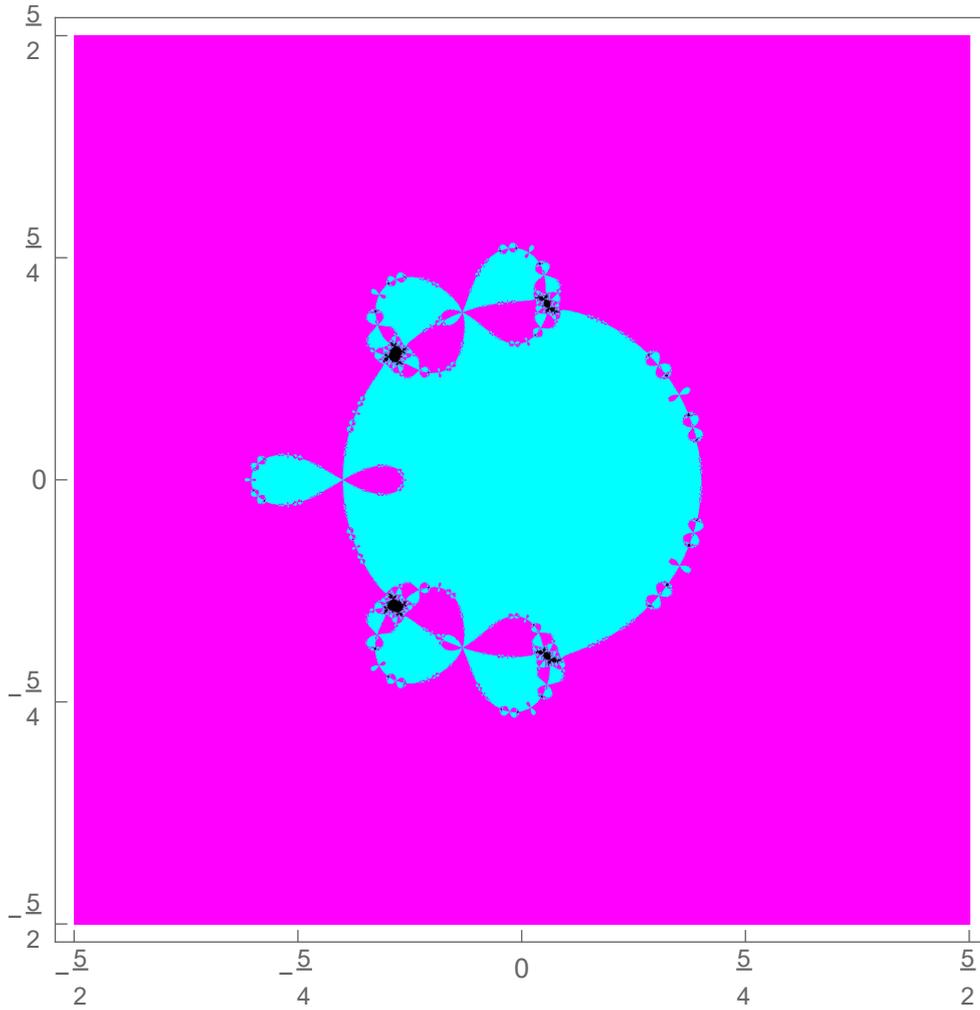


Figura 7.10: Cuenca de atracción asociada al método con $b_1 = -3.6 - \frac{i}{10}$.

y Hueso et al. [122], de ellos elegimos sus mejores expresiones (5) y (14–15) para $(t_1 = -\frac{9}{4}$ y $s_2 = \frac{9}{8})$, que llamamos (SM) y (HM) , respectivamente. Además, también los comparamos con métodos eficientes de Jarratt de sexto orden diseñados por Sharma y Arora [204], de ellos consideramos el método (18), denominado por (SAM) . Finalmente, los comparamos con los métodos de sexto orden desarrollados por Cordero et al. [67], y elegimos el método (6) denotado como (CM) . Damos el número de iteración (n), el error residual para la función correspondiente ($\|F(x^{(n)})\|$), el error en dos iteraciones consecutivas $\|x^{(n+1)} - x^{(n)}\|$, $\frac{\|x^{(n+1)} - x^{(n)}\|}{\|x^{(n)} - x^{(n-1)}\|^6}$, η^* y el orden de convergencia computacional

$$\rho^* = \frac{\log \left[\frac{\|F(x^{(n)})\|}{\|F(x^{(n-1)})\|} \right]}{\log \left[\frac{\|F(x^{(n-1)})\|}{\|F(x^{(n-2)})\|} \right]}$$

en las Tablas 7.8 – 7.12, 7.14. Los valores de η^* representados en las tablas son

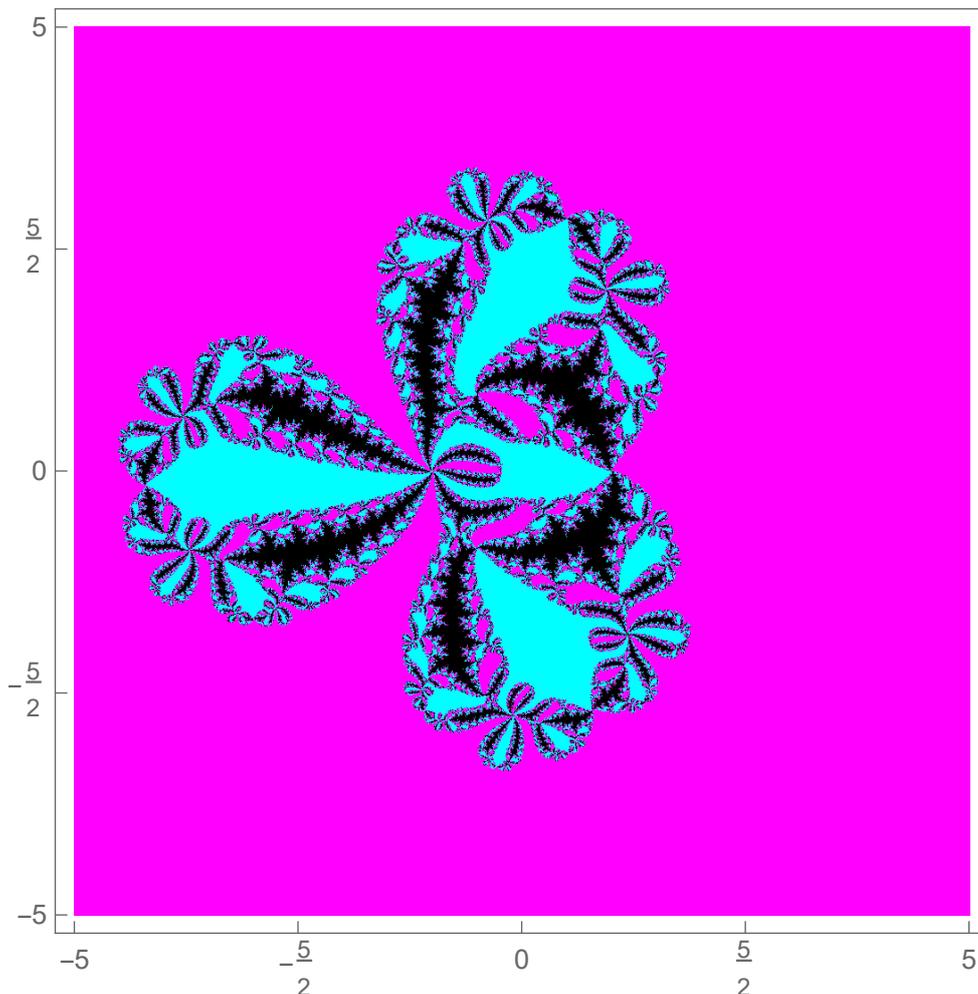


Figura 7.11: Cuenca de atracción asociada al método con $b_1 = 1$.

el último valor calculado de

$$\frac{\|x^{(n+1)} - x^{(n)}\|}{\|x^{(n)} - x^{(n-1)}\|^6}$$

Durante los experimentos numéricos actuales con el lenguaje de programación Mathematica (Versión 9), todos los cálculos se han hecho con aritmética de precisión múltiple con 1000 dígitos de mantisa, que minimizan los errores de redondeo.

El significado de $a e(\pm b)$ es $a \times 10^{(\pm b)}$ en todas las tablas. Hemos usado el comando "AbsoluteTiming[]" para calcular el tiempo de la CPU. Hemos ejecutado nuestro programa cinco veces y hemos representado el tiempo promedio de la CPU en la Tabla 7.15. A continuación damos la configuración utilizada:

Procesador: Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz
 Made: HP

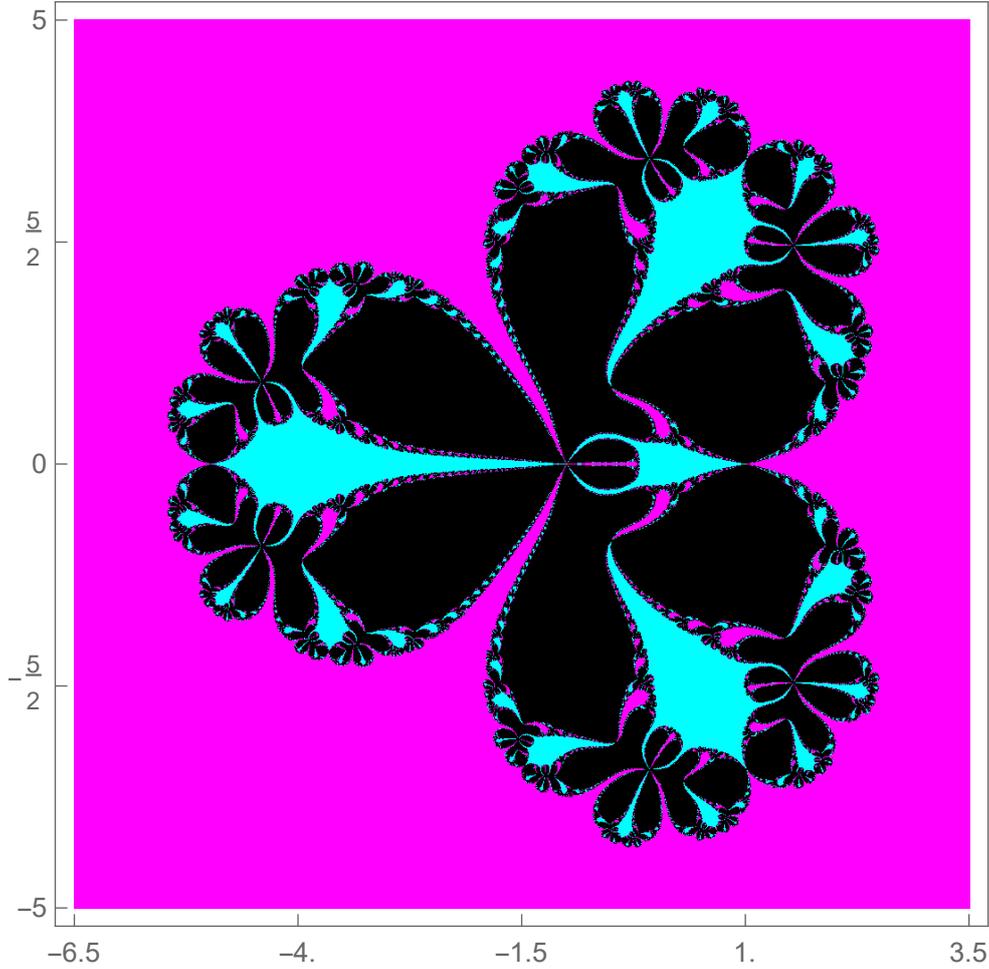


Figura 7.12: Cuenca de atracción asociada al método con $b_1 = i$.

RAM: 8:00GB

Sistema Operativo: 64-bit-Operating System, x64-based processor.

Ejemplo 7.4 Consideramos el problema de la síntesis cinemática para la dirección como se describe en [34, 216], en forma de sistema de ecuaciones no lineales, que está dado para cada $i = 1, 2, 3$ por

$$[E_i(x_2 \sin(\psi_i) - x_3) - F_i(x_2 \sin(\phi_i) - x_3)]^2 + [F_i(x_2 \cos(\phi_i) + 1) - F_i(x_2 \cos(\psi_i) - 1)]^2 - [x_1(x_2 \sin(\psi_i) - x_3)(x_2 \cos(\phi_i) + 1) - x_1(x_2 \cos(\psi_i) - x_3)(x_2 \sin(\phi_i) - x_3)]^2 = 0,$$

donde

$$E_i = -x_3 x_2 (\sin(\phi_i) - \sin(\phi_0)) - x_1 (x_2 \sin(\phi_i) - x_3) + x_2 (\cos(\phi_i) - \cos(\phi_0)),$$

y

$$F_i = -x_3 x_2 \sin(\psi_i) + (-x_2) \cos(\psi_i) + (x_3 - x_1) x_2 \sin(\psi_0) + x_2 \cos(\psi_0) + x_1 x_3.$$

Cuadro 7.2: Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_1(x)$.

Métodos	n	x_n	$ f(x_n) $	$ x_{n+1} - x_n $	$\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^6}$	η	ρ
<i>KLM</i>	1	1.3650	$3.3e(-3)$	$2.0e(-4)$			
	2	1.3652	$5.8e(-24)$	$3.5e(-25)$	$5.3379e(-3)$	$5.3375e(-3)$	
	3	1.3652	$1.6e(-148)$	$9.7e(-150)$	$5.3375e(-3)$		6.0000
<i>CHM</i>	1	1.3664	$2.0e(-2)$	$1.2e(-3)$			
	2	1.3652	$4.9e(-18)$	$3.0e(-19)$	$1.0021e(-1)$	$1.0058e(-1)$	
	3	1.3652	$1.2e(-111)$	$7.2e(-113)$	$1.0058e(-1)$		5.9999
<i>CM</i>	1	1.3651	$1.6e(-3)$	$9.6e(-4)$			
	2	1.3652	$2.8e(-26)$	$1.7e(-27)$	$2.2470e(-3)$	$2.2472e(-3)$	
	3	1.3652	$9.8e(-163)$	$5.9e(-164)$	$2.2472e(-3)$		6.0000
<i>RM</i>	1	1.3645	$1.1e(-2)$	$6.7e(-4)$			
	2	1.3652	$4.6e(-20)$	$2.8e(-21)$	$3.0801e(-2)$	$3.0758e(-)$	
	3	1.3652	$2.4e(-124)$	$e(-125)$	$3.0758e(-2)$		6.0000
<i>OM₆¹</i>	1	1.3652	$2.6e(-4)$	$1.6e(-5)$			
	2	1.3652	$2.0e(-30)$	$1.2e(-31)$	$8.5094e(-3)$	$8.5088e(-3)$	
	3	1.3652	$4.8e(-187)$	$2.9e(-188)$	$8.5088e(-3)$		6.0000
<i>OM₆²</i>	1	1.3652	$2.8e(-4)$	$1.7e(-5)$			
	2	1.3652	$3.0e(-30)$	$1.8e(-31)$	$8.5655e(-3)$	$8.5649e(-3)$	
	3	1.3652	$1.5e(-186)$	$3.4e(-187)$	$8.5649e(-3)$		6.0000
<i>OM₆³</i>	1	1.3656	$6.1e(-3)$	$3.7e(-4)$			
	2	1.3652	$6.0e(-20)$	$3.6e(-23)$	$1.4000e(-2)$	$1.4004e(-2)$	
	3	1.3652	$5.2e(-136)$	$3.1e(-137)$	$1.4004e(-2)$		6.0000
<i>OM₆⁴</i>	1	1.3671	$3.1e(-2)$	$1.9e(-3)$			
	2	1.3652	$1.4e(-16)$	$8.7e(-18)$	$1.9284e(-1)$	$1.9411e(-1)$	
	3	1.3652	$1.4e(-102)$	$8.5e(-104)$	$1.9411e(-1)$		5.9998
<i>OM₆⁵</i>	1	1.3632	$3.3e(-2)$	$2.0e(-3)$			
	2	1.3652	$7.8e(-17)$	$4.7e(-18)$	$7.6336e(-2)$	$7.6050e(-2)$	
	3	1.3652	$1.4e(-104)$	$8.4e(-106)$	$7.6050e(-2)$		6.0001

Cuadro 7.3: Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_2(x)$.

Métodos	n	x_n	$ f(x_n) $	$ x_{n+1} - x_n $	$\left \frac{x_{n+1} - x_n}{(x_n - x_{n-1})^6} \right $	η	ρ
<i>KLM</i>	1	0.2575	$3.3e(-5)$	$8.9e(-6)$			
	2	0.2575	$6.3e(-34)$	$1.7e(-34)$	$3.4817e(-4)$	$3.4817(-4)$	
	3	0.2575	$2.9e(-206)$	$7.6e(-207)$	$3.4817e(-4)$		6.0000
<i>CHM</i>	1	0.2575	$1.0e(-4)$	$2.7e(-4)$			
	2	0.2575	$8.7e(-32)$	$2.3e(-32)$	$6.0508e(-5)$	$6.0495e(-5)$	
	3	0.2575	$3.4e(-194)$	$9.1e(-195)$	$6.0495e(-5)$		6.0000
<i>CM</i>	1	0.2575	$4.3e(-4)$	$1.1e(-5)$			
	2	0.2575	$2.9e(-33)$	$7.7e(-34)$	$3.803e(-4)$	$3.803e(-4)$	
	3	0.2575	$3.1e(-202)$	$8.2e(-203)$	$3.80384e(-4)$		6.0000
<i>RM</i>	1	0.2575	$3.4e(-5)$	$9.1e(-6)$			
	2	0.2575	$8.8e(-34)$	$2.3e(-34)$	$4.1212e(-4)$	$4.1212e(-4)$	
	3	0.2575	$2.5e(-205)$	$6.5e(-206)$	$4.1212e(-4)$		6.0000
OM_6^1	1	0.2575	$3.4e(-5)$	$8.9e(-6)$			
	2	0.2575	$6.2e(-34)$	$1.6e(-34)$	$3.2747e(-4)$	$3.2747e(-4)$	
	3	0.2575	$2.3e(-206)$	$6.2e(-207)$	$3.2747e(-4)$		6.0000
OM_6^2	1	0.2575	$3.4e(-5)$	$8.9e(-6)$			
	2	0.2575	$6.2e(-34)$	$1.6e(-34)$	$3.2755e(-4)$	$3.2756e(-4)$	
	3	0.2575	$2.3e(-206)$	$6.2e(-207)$	$3.2756e(-4)$		6.0000
OM_6^3	1	0.2575	$3.0e(-5)$	$8.9e(-6)$			
	2	0.2575	$5.4e(-34)$	$1.4e(-34)$	$2.9406e(-4)$	$2.9406e(-4)$	
	3	0.2575	$9.7e(-207)$	$6.0e(-207)$	$2.9406e(-4)$		6.0000
OM_6^4	1	0.2575	$3.3e(-5)$	$8.6e(-6)$			
	2	0.2575	$4.2e(-35)$	$1.1e(-35)$	$2.6738e(-4)$	$2.6744e(-)$	
	3	0.2575	$1.9e(-214)$	$5.1e(-215)$	$2.6744e(-4)$		6.0000
OM_6^5	1	0.2575	$3.4e(-5)$	$9.0e(-6)$			
	2	0.2575	$8.6e(-34)$	$2.3e(-34)$	$4.2772e(-4)$	$4.2772e(-4)$	
	3	0.2575	$2.2e(-205)$	$5.8e(-206)$	$4.2772e(-4)$		6.0000

Cuadro 7.4: Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_3(x)$.

Métodos	n	x_n	$ f(x_n) $	$ x_{n+1} - x_n $	$\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^6}$	η	ρ
<i>KLM</i>	1	1.7461	$1.2e(-6)$	$1.0e(-6)$			
	2	1.7461	$2.8e(-39)$	$2.4e(-39)$	$2.4397e(-3)$	$2.4397e(-3)$	
	3	1.7461	$5.0e(-235)$	$5.0e(-235)$	$2.4397e(-3)$		6.0000
<i>CHM</i>	1	1.7461	$4.3e(-6)$	$3.7e(-6)$			
	2	1.7461	$1.4e(-35)$	$1.2e(-35)$	$4.8113e(-3)$	$4.8114e(-3)$	
	3	1.7461	$1.6e(-212)$	$1.3e(-212)$	$4.8114e(-3)$		6.0000
<i>CM</i>	1	1.7461	$1.4e(-6)$	$1.2e(-6)$			
	2	1.7461	$9.0e(-39)$	$7.7e(-39)$	$2.5532e(-3)$	$2.5532e(-3)$	
	3	1.7461	$6.4e(-232)$	$5.5e(-232)$	$2.5532e(-3)$		6.0000
<i>RM</i>	1	1.7461	$7.0e(-7)$	$6.1e(-7)$			
	2	1.7461	$1.1e(-40)$	$9.2e(-40)$	$1.8705e(-3)$	$1.8705e(-3)$	
	3	1.7461	$1.3e(-243)$	$1.1e(-243)$	$1.8705e(-3)$		6.0000
<i>OM₆¹</i>	1	1.7461	$1.2e(-6)$	$1.1e(-6)$			
	2	1.7461	$4.8e(-39)$	$4.1e(-39)$	$2.6162e(-3)$	$2.6162e(-3)$	
	3	1.7461	$1.5e(-233)$	$1.3e(-233)$	$2.6162e(-3)$		6.0000
<i>OM₆²</i>	1	1.7461	$1.2e(-6)$	$1.1e(-6)$			
	2	1.7461	$4.7e(-39)$	$4.1e(-39)$	$2.6154e(-3)$	$2.6154e(-3)$	
	3	1.7461	$1.4e(-233)$	$1.2e(-233)$	$2.6154e(-3)$		6.0000
<i>OM₆³</i>	1	1.7461	$1.7e(-6)$	$1.4e(-6)$			
	2	1.7461	$3.0e(-38)$	$2.6e(-38)$	$2.9553e(-3)$	$2.9553e(-3)$	
	3	1.7461	$1.0e(-228)$	$8.7e(-229)$	$2.9553e(-3)$		6.0000
<i>OM₆⁴</i>	1	1.7461	$5.6e(-6)$	$4.8e(-6)$			
	2	1.7461	$8.2e(-35)$	$7.1e(-35)$	$5.6680e(-3)$	$5.6681e(-3)$	
	3	1.7461	$8.4e(-208)$	$7.3e(-208)$	$5.6681e(-3)$		6.0000
<i>OM₆⁵</i>	1	1.7461	$9.1e(-8)$	$7.8e(-8)$			
	2	1.7461	$4.3e(-46)$	$3.7e(-46)$	$1.5989e(-3)$	$1.5989e(-3)$	
	3	1.7461	$5.0e(-276)$	$4.3e(-276)$	$1.5989e(-3)$		6.0000

Cuadro 7.5: Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_4(x)$.

Métodos	n	x_n	$ f(x_n) $	$ x_{n+1} - x_n $	$\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^6}$	η	ρ
<i>KLM</i>	1	2.2788	$6.1e(-6)$	$6.2e(-6)$			
	2	2.2788	$5.7e(-34)$	$5.8e(-34)$	$1.063e(-2)$	$1.0636e(-2)$	
	3	2.2788	$4.0e(-202)$	$4.1e(-202)$	$1.063e(-2)$		6.0000
<i>CHM</i>	1	2.2790	$1.5e(-4)$	$1.6e(-4)$			
	2	2.2788	$1.2e(-24)$	$1.3e(-24)$	$8.5586e(-2)$	$8.5646e(-2)$	
	3	2.2788	$3.4e(-145)$	$3.4e(-145)$	$8.5646e(-2)$		6.0000
<i>CM</i>	1	2.2788	$6.1e(-6)$	$6.2e(-6)$			
	2	2.2788	$3.2e(-34)$	$3.2e(-34)$	$5.8601e(-3)$	$5.8600e(-3)$	
	3	2.2788	$6.4e(-204)$	$6.5e(-204)$	$5.8600e(-3)$		6.0000
<i>RM</i>	1	2.2788	$1.8e(-5)$	$1.8e(-5)$			
	2	2.2788	$2.4e(-31)$	$2.4e(-31)$	$6.6326e(-3)$	$6.6316e(-3)$	
	3	2.2788	$1.4e(-186)$	$1.4e(-186)$	$6.6316e(-3)$		6.0000
<i>OM₆¹</i>	1	2.2788	$9.0e(-7)$	$9.1e(-7)$			
	2	2.2788	$7.6e(-39)$	$7.7e(-39)$	$1.3455e(-2)$	$1.3455e(-2)$	
	3	2.2788	$2.7e(-231)$	$2.8e(-231)$	$1.3455e(-2)$		6.0000
<i>OM₆²</i>	1	2.2788	$9.4e(-7)$	$9.6e(-7)$			
	2	2.2788	$1.0e(-38)$	$1.0e(-38)$	$1.3425e(-2)$	$1.3425e(-2)$	
	3	2.2788	$1.7e(-230)$	$1.7e(-230)$	$1.3425e(-2)$		6.0000
<i>OM₆³</i>	1	2.2788	$1.9e(-5)$	$1.9e(-5)$			
	2	2.2788	$1.2e(-30)$	$1.2e(-30)$	$2.5590e(-2)$	$2.5591e(-2)$	
	3	2.2788	$9.2e(-182)$	$9.2e(-183)$	$2.5591e(-2)$		6.0000
<i>OM₆⁴</i>	1	2.2791	$2.6e(-4)$	$2.7e(-4)$			
	2	2.2788	$4.5e(-23)$	$4.5e(-23)$	$1.2251e(-1)$	$1.2267e(-1)$	
	3	2.2788	$1.1e(-135)$	$1.1e(-135)$	$1.2267e(-1)$		6.0000
<i>OM₆⁵</i>	1	2.2788	$5.2e(-5)$	$5.2e(-5)$			
	2	2.2788	$4.7e(-28)$	$4.8e(-28)$	$2.2959e(-2)$	$2.2952e(-2)$	
	3	2.2788	$2.6e(-166)$	$2.7e(-166)$	$2.2952e(-2)$		6.0000

Cuadro 7.6: Comportamiento de la convergencia de los diferentes métodos en la función de prueba $f_5(x)$.

Métodos	n	x_n	$ f(x_n) $	$ x_{n+1} - x_n $	$\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^6}$	η	ρ
<i>KLM</i>	1	0.5948	$1.5e(-8)$	$1.4e(-8)$			
	2	0.5948	$5.6e(-50)$	$5.3e(-50)$	$6.8896e(-3)$	$6.8896e(-3)$	
	3	0.5948	$1.5e(-298)$	$1.5e(-298)$	$6.8896e(-3)$		6.0000
<i>CHM</i>	1	0.5948	$1.3e(-8)$	$1.2e(-8)$			
	2	0.5948	$1.8e(-50)$	$1.7e(-50)$	$5.0361e(-3)$	$5.0361e(-3)$	
	3	0.5948	$1.3e(-301)$	$1.2e(-301)$	$5.0361e(-3)$		6.0000
<i>CM</i>	1	0.5948	$3.8e(-8)$	$3.6e(-8)$			
	2	0.5948	$4.7e(-47)$	$4.5e(-47)$	$1.9982e(-2)$	$1.9982e(-2)$	
	3	0.5948	$1.6e(-280)$	$1.6e(-280)$	$1.9982e(-2)$		6.0000
<i>RM</i>	1	0.5948	$2.0e(-8)$	$1.9e(-8)$			
	2	0.5948	$4.4e(-49)$	$4.2e(-49)$	$9.2754e(-3)$	$9.2754e(-3)$	
	3	0.5948	$5.1e(-293)$	$4.8e(-293)$	$9.2754e(-3)$		6.0000
OM_6^1	1	0.5948	$9.1e(-9)$	$8.6e(-9)$			
	2	0.5948	$2.1e(-51)$	$2.0e(-51)$	$4.8394e(-3)$	$4.8394e(-3)$	
	3	0.5948	$3.4e(-307)$	$3.2e(-307)$	$4.8394e(-3)$		6.0000
OM_6^2	1	0.5948	$9.2e(-9)$	$8.6e(-9)$			
	2	0.5948	$2.1e(-51)$	$2.0e(-51)$	$4.8418e(-3)$	$4.8418e(-3)$	
	3	0.5948	$3.5e(-307)$	$3.3e(-307)$	$4.8418e(-3)$		6.0000
OM_6^3	1	0.5948	$7.1e(-9)$	$6.7e(-9)$			
	2	0.5948	$3.8e(-52)$	$3.6e(-52)$	$3.9084e(-3)$	$3.9084e(-3)$	
	3	0.59481	$8.7e(-312)$	$8.2e(-312)$	$3.9084e(-3)$		6.0000
OM_6^4	1	0.5948	$7.7e(-9)$	$7.3e(-9)$			
	2	0.5948	$5.4e(-52)$	$5.1e(-52)$	$3.5395e(-3)$	$3.5395e(-3)$	
	3	0.5948	$6.9e(-311)$	$6.6e(-311)$	$3.5395e(-3)$		6.0000
OM_6^5	1	0.5948	$1.6e(-8)$	$1.5e(-8)$			
	2	0.5948	$8.1e(-50)$	$7.7e(-50)$	$7.6324e(-3)$	$7.6324e(-3)$	
	3	0.5948	$1.6e(-297)$	$1.5e(-297)$	$7.6324e(-3)$		6.0000

Cuadro 7.7: Los valores de ψ_i y ϕ_i (en radianes) para el ejemplo 7.4.

i	ψ_i	ϕ_i
0	1.395417004174709	1.746175649415084
1	1.744482854573574	2.036469112791960
2	2.065623436940531	2.239097786826597
3	2.460067847891250	2.460067840980934

Mostramos los valores de ψ_i y ϕ_i (en radianes) en la Tabla 7.7. Consideramos la aproximación inicial $x^{(0)} = (0.7, 0.7, 0.7)$. Obtenemos la siguiente solución aproximada

$$\xi = (0.9051567\dots, 0.6977417\dots, 0.6508335\dots)^T.$$

Ejemplo 7.5 Consideremos el problema de contorno (para más detalles ver Ortega y Rheinboldt [171]), de se define de la siguiente manera:

$$y'' = \frac{1}{2}y^3 + 3y' - \frac{3}{2-x} + \frac{1}{2}, \quad y(0) = 0, \quad y(1) = 1. \quad (7.5.1)$$

Además, consideramos la siguiente partición del intervalo $[0, 1]$, dada por

$$x_0 = 0 < x_1 < x_2 < x_3 < \dots < x_n, \quad \text{where } x_{i+1} = x_i + h, \quad h = \frac{1}{n}.$$

Además, también consideramos que $y_0 = y(x_0) = 0$, $y_1 = y(x_1)$, \dots , $y_{n-1} = y(x_{n-1})$, $y_n = y(x_n) = 1$. Ahora, discretizamos el problema anterior (7.5.1) con la ayuda de la siguiente fórmula numérica para la primera y segunda derivada

$$y'_j = \frac{y_{j+1} - y_{j-1}}{2h}, \quad y''_j = \frac{y_{j-1} - 2y_j + y_{j+1}}{h^2}, \quad j = 1, 2, \dots, n-1.$$

De esta manera, obtenemos $(n-1) \times (n-1)$ un sistema de ecuaciones no lineales, definido como:

$$y_{j+1} - 2y_j + y_{j-1} - \frac{h^2}{2}y_j^3 - \frac{3}{2-x_j}h^2 - \frac{1}{h^2} = 0, \quad j = 1, 2, \dots, n-1.$$

Consideremos la aproximación inicial $y_k^{(0)} = \left(\frac{3}{2}, \frac{3}{2}, \frac{3}{2}, \frac{3}{2}, \frac{3}{2}, \frac{3}{2}\right)^T$. En particular, resolvemos este problema para $n = 7$ para que podamos obtener un sistema 6×6 de ecuaciones no lineales. La solución de este problema es

$$\xi = \left(0.076543\dots, 0.16587\dots, 0.27152\dots, 0.39845\dots, 0.55388\dots, 0.74868\dots\right)^T$$

y los resultados obtenidos en las estimaciones numéricas aparecen en la Tabla 7.9.

Cuadro 7.8: Comportamiento de la convergencia de los diferentes métodos en el problema de la síntesis cinemática 7.4.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$1.1e(-3)$	$5.3e(-3)$			
	2	$9.1e(-9)$	$1.3e(-6)$	$5.6e(+7)$	$2.3e(+6)$	
	3	$2.0e(-31)$	$1.0e(-29)$	$2.3e(+6)$		4.5
LM	1	$1.2e(-4)$	$2.7e(-3)$			
	2	$4.8e(-13)$	$3.0e(-11)$	$7.9e(+4)$	$2.2e(+11)$	
	3	$6.9e(-53)$	$1.6e(-52)$	$2.2e(+11)$		4.7
SM	1	$4.3e(-5)$	$5.1e(-3)$			
	2	$3.8e(-11)$	$2.1e(-9)$	$1.1e(+5)$	$1.0e(+4)$	
	3	$2.0e(-50)$	$1.0e(-48)$	$1.0e(+4)$		6.4
HM	1	$3.0e(-4)$	$7.0e(-3)$			
	2	$8.7e(-9)$	$9.3e(-7)$	$7.8e(+6)$	$8.1e(+9)$	
	3	$1.4e(-28)$	$5.2e(-27)$	$8.1e(+9)$		4.4
SAM	1	$7.5e(-4)$	$4.9e(-3)$			
	2	$3.3e(-9)$	$6.5e(-7)$	$4.8e(+7)$	$1.0e(+6)$	
	3	$1.7e(-33)$	$7.7e(-32)$	$1.0e(+6)$		4.5
CM	1	$9.9e(-5)$	$4.9e(-3)$			
	2	$8.2e(-11)$	$5.2e(-9)$	$3.7e(+5)$	$2.4e(+4)$	
	3	$1.2e(-47)$	$4.9e(-46)$	$2.4e(+4)$		6.1
\widehat{OM}_6^1	1	$7.0e(-5)$	$5.7e(-3)$			
	2	$6.1e(-12)$	$6.8e(-10)$	$2.0e(+4)$	$5.3e(+3)$	
	3	$4.2e(-54)$	$5.1e(-52)$	$5.3(+3)$		5.9
\widehat{OM}_6^2	1	$7.0e(-5)$	$5.7e(-3)$			
	2	$6.0e(-12)$	$6.7e(-12)$	$2.0e(+4)$	$5.3e(+3)$	
	3	$4.1e(-54)$	$5.0e(-52)$	$5.3e(+3)$		5.9
\widehat{OM}_6^3	1	$4.1e(-5)$	$5.6(-3)$			
	2	$5.1e(-11)$	$2.9e(-9)$	$9.5e(+4)$	$8.9e(+3)$	
	3	$9.5e(-50)$	$5.1e(-48)$	$8.9e(+3)$		6.6
\widehat{OM}_6^4	1	$3.8e(-4)$	$4.6e(-3)$			
	2	$5.3e(-10)$	$7.3e(-8)$	$7.5e(+6)$	$2.1e(+5)$	
	3	$7.9e(-40)$	$3.3e(-38)$	$2.1e(+5)$		5.1

Cuadro 7.9: Comportamiento de la convergencia de los diferentes métodos en el problema de valores de contorno 7.5.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$1.9e(-4)$	$6.1e(-4)$			
	2	$8.7e(-27)$	$2.8e(-26)$	$5.1e(-7)$	$5.9e(-7)$	
	3	$8.1e(-161)$	$2.6e(-160)$	$5.9e(-7)$		6.0
LM	1	$2.9e(-4)$	$9.1e(-4)$			
	2	$1.7e(-23)$	$8.6e(-24)$	$1.5e(-5)$	$1.6e(+17)$	
	3	$2.3e(-122)$	$6.6e(-122)$	$1.6e(+17)$		5.1
SM	1	$1.5e(-4)$	$6.3e(-4)$			
	2	$2.3e(-26)$	$9.1e(-26)$	$1.4e(-6)$	$1.7e(-6)$	
	3	$2.7e(-157)$	$9.5e(-157)$	$1.7e(-6)$		6.0
HM	1	$1.3e(-4)$	$5.7e(-4)$			
	2	$8.1e(-23)$	$2.8e(-22)$	$8.2e(-3)$	$2.0e(+16)$	
	3	$3.5e(-114)$	$1.0e(-113)$	$2.0e(+16)$		5.0
SAM	1	$7.3e(-5)$	$2.7e(-4)$			
	2	$8.2e(-29)$	$2.7e(-28)$	$7.8e(-7)$	$9.0e(-6)$	
	3	$1.1e(-172)$	$3.6e(-172)$	$9.0e(-6)$		6.0
CM	1	$1.0e(-4)$	$3.8e(-4)$			
	2	$9.3e(-28)$	$3.8e(-27)$	$1.3e(-6)$	$1.6e(-6)$	
	3	$1.3e(-165)$	$4.7e(-165)$	$1.6e(-6)$		6.0
\widehat{OM}_6^1	1	$1.8e(-4)$	$7.6e(-4)$			
	2	$7.7e(-26)$	$3.0e(-25)$	$1.5e(-6)$	$1.7e(-6)$	
	3	$3.5e(-154)$	$1.2e(-153)$	$1.7e(-6)$		6.0
\widehat{OM}_6^2	1	$1.8e(-4)$	$7.6e(-4)$			
	2	$7.7e(-26)$	$3.0e(-25)$	$1.5e(-6)$	$1.7e(-6)$	
	3	$3.6e(-154)$	$1.3e(-153)$	$1.7e(-6)$		6.0
\widehat{OM}_6^3	1	$1.4e(-4)$	$5.9e(-4)$			
	2	$1.5e(-26)$	$6.1e(-26)$	$1.4e(-6)$	$1.7e(-6)$	
	3	$2.5e(-158)$	$8.8e(-158)$	$1.7e(-6)$		6.0
\widehat{OM}_6^4	1	$9.7e(-5)$	$3.3e(-4)$			
	2	$4.4e(-28)$	$1.4e(-27)$	$1.1e(-6)$	$1.1e(-6)$	
	3	$3.0e(-168)$	$9.6e(-168)$	$1.1e(-6)$		6.0

Ejemplo 7.6 Consideremos el siguiente sistema no lineal de ecuaciones no lineales [203]

$$\begin{cases} x_j^2 x_{j+1} - 1 = 0, & 1 \leq j \leq n, \\ x_n^2 x_1 - 1 = 0. \end{cases} \quad (7.5.2)$$

Para obtener un gran sistema de 100×100 , tomamos $n = 100$ y la aproximación inicial $x^{(0)} = (1.25, 1.25, 1.25, \dots, (100\text{veces}))^T$ para este problema. La solución requerida de este problema es $\xi = (1, 1, 1, \dots, (100\text{veces}))^T$. Los resultados obtenidos se pueden observar en la Tabla 7.10.

Ejemplo 7.7 El problema bidimensional de Bratu [129] se define como:

$$\begin{aligned} u_{xx} + u_{tt} + Ce^u &= 0, \text{ en} \\ \Omega : (x, t) \in 0 \leq x \leq 1, 0 \leq t \leq 1, & \\ \text{con condiciones frontera } u &= 0 \text{ en } \Omega. \end{aligned} \quad (7.5.3)$$

La solución de una ecuación diferencial parcial no lineal se puede encontrar usando discretización de diferencias finita que se reduce a resolver un sistema de ecuaciones no lineales. Sea $w_{i,j} = u(x_i, t_j)$ su solución aproximada en los puntos de la cuadrícula de la malla. Sean M y N el número de pasos en las direcciones x y t , h y k son respectivamente el tamaño del paso. Para resolver el PDE dado, aplicamos diferencias centrales a u_{xx} y u_{tt} es decir, $u_{xx}(x_i, t_j) = (w_{i+1,j} - 2w_{i,j} + w_{i-1,j})/h^2$, $C = 0.1, t \in [0, 1]$. Hemos considerado la solución del sistema para $M = 11$ y $N = 11$ de tamaño 100, con el vector inicial $x_0 = 0.1(\sin(\pi h)\sin(\pi k), \sin(2\pi h)\sin(2\pi k), \dots, \sin(10\pi h)\sin(10\pi k))^T$ que ha sido evaluado según los diferentes métodos y presentados en la Tabla 7.11.

Ejemplo 7.8 Consideramos otro problema no lineal típico que es la ecuación de Fisher [198] con condiciones de frontera de Neumann homogéneas y coeficiente de difusión D :

$$\begin{aligned} u_t &= Du_{xx} + u(1 - u) = 0, \\ u(x, 0) &= 1.5 + 0.5\cos(\pi x), 0 \leq x \leq 1, \\ u_x(0, t) &= 0, \forall t \geq 0, \\ u_x(1, t) &= 0, \forall t \geq 0. \end{aligned} \quad (7.5.4)$$

Usando de nuevo la discretización de diferencias finitas, la ecuación (7.5.4) se reduce a un sistema de ecuaciones no lineales. Consideramos $w_{i,j} = u(x_i, t_j)$ su solución aproximada en los puntos de la cuadrícula de la malla. Sean M y N el número de pasos en las direcciones x y t , y sean h y k los respectivos tamaños de paso. Aplicando diferencias centrales a $u_{xx}(x_i, t_j) = (w_{i+1,j} - 2w_{i,j} + w_{i-1,j})/h^2$, diferencias regresivas a $u_t(x_i, t_j) = (w_{i,j} - w_{i,j-1})/k$ y diferencias progresivas a $u_x(x_i, t_j) = (w_{i+1,j} - w_{i,j})/h, t \in [0, 1]$. Para la solución del sistema, hemos considerado $M = 9$ y $N = 9$, lo cual se reduce a un sistema no lineal de tamaño 64, donde el vector inicial $x_0 = (i/(M-1)^2)^T, i = 1, 2, \dots, M-1$ ha sido evaluado usando los diferentes métodos y los datos expuestos en la Tabla 7.12.

Cuadro 7.10: Comportamiento de la convergencia de los diferentes métodos en el ejemplo 7.6.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$3.9e(-2)$	$1.3e(-2)$			
	2	$5.8e(-15)$	$1.9e(-15)$	$4.2e(-4)$	$4.3e(-4)$	
	3	$6.7e(-92)$	$2.2e(-92)$	$4.3e(-4)$		6.0
LM	1	$2.4e(-2)$	$8.0e(-3)$			
	2	$1.6e(-16)$	$5.4e(-17)$	$2.1e(-4)$	$2.1e(-4)$	
	3	$1.6e(-101)$	$5.4e(-102)$	$2.1e(-4)$		6.0
SM	1	$1.9e(-3)$	$6.3e(-4)$			
	2	$8.3e(-25)$	$2.8e(-25)$	$4.4e(-6)$	$4.4e(-6)$	
	3	$6.2e(-153)$	$2.1e(-153)$	$4.4e(-6)$		6.0
HM	1	$2.1e(-2)$	$7.1e(-3)$			
	2	$1.4e(-14)$	$4.7e(-15)$	$3.7e(+2)$	$5.6(+10)$	
	3	$1.9e(-75)$	$6.3e(-76)$	$5.6(+10)$		5.0
SAM	1	$2.7e(-1)$	$9.1e(-2)$			
	2	$4.5e(-16)$	$1.5e(-16)$	$2.6e(-4)$	$2.6e(-4)$	
	3	$9.1e(-99)$	$3.0e(-99)$	$2.6e(-4)$		6.0
CM	1	$3.3e(-3)$	$1.1e(-3)$			
	2	$5.9e(-23)$	$2.0e(-23)$	$1.1e(-5)$	$1.1e(-5)$	
	3	$1.9e(-141)$	$6.2e(-142)$	$1.1e(-5)$		6.0
\widehat{OM}_6^1	1	$5.3e(-4)$	$1.8e(-4)$			
	2	$3.5e(-28)$	$1.2e(-28)$	$3.7e(-6)$	$3.7e(-6)$	
	3	$2.9e(-173)$	$9.7e(-174)$	$3.7e(-6)$		6.0
\widehat{OM}_6^2	1	$5.4e(-4)$	$1.8e(-4)$			
	2	$3.8e(-28)$	$1.3e(-28)$	$3.7e(-6)$	$3.7e(-6)$	
	3	$4.4e(-173)$	$1.5e(-173)$	$3.7e(-6)$		6.0
\widehat{OM}_6^3	1	$1.6e(-3)$	$5.3e(-4)$			
	2	$2.6e(-25)$	$8.6e(-26)$	$3.7e(-6)$	$3.7e(-6)$	
	3	$4.5e(-156)$	$1.5e(-156)$	$3.7e(-6)$		6.0
\widehat{OM}_6^4	1	$1.1e(-2)$	$3.6e(-3)$			
	2	$4.1e(-19)$	$1.4e(-19)$	$6.2e(-5)$	$6.2e(-5)$	
	3	$1.2e(-117)$	$4.1e(-118)$	$6.2e(-5)$		6.0

Cuadro 7.11: Comportamiento de la convergencia de los diferentes métodos en el problema bidimensional de Bratu 7.7.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$4.4e(-15)$	$2.4e(-14)$			
	2	$6.9e(-95)$	$3.5e(-94)$	$1.4e(-12)$	$1.9e(-12)$	
	3	$7.9e(-574)$	$3.9e(-573)$	$1.9e(-12)$		6.0
LM	1	$6.4e(-16)$	$2.4e(-15)$			
	2	$2.9e(-87)$	$5.3e(-87)$	$1.4e(-13)$	$2.8e(+1)$	
	3	$9.1e(-445)$	$3.3e(-444)$	$2.8e(+1)$		5.0
SM	1	$4.2e(-15)$	$2.3e(-14)$			
	2	$5.5e(-95)$	$2.7e(-94)$	$1.3e(-12)$	$1.9e(-12)$	
	3	$1.9e(-574)$	$9.0e(-574)$	$1.9e(-12)$		6.0
HM	1	$2.1e(-13)$	$1.2e(-12)$			
	2	$2.1e(-71)$	$1.2e(-70)$	$7.3e(-11)$	$3.4e(+1)$	
	3	$1.7e(-361)$	$9.3e(-361)$	$3.4e(+1)$		5.0
SAM	1	$4.4e(-15)$	$2.4e(-14)$			
	2	$7.1e(-95)$	$3.6e(-94)$	$1.4(-12)$	$1.9e(-12)$	
	3	$9.2e(-574)$	$4.5e(-573)$	$1.9e(-12)$		6.0
CM	1	$4.3e(-15)$	$2.3e(-14)$			
	2	$5.7e(-95)$	$2.9e(-94)$	$1.3e(-12)$	$1.9e(-12)$	
	3	$2.5e(-574)$	$1.2e(-573)$	$1.9e(-12)$		6.0
\widehat{OM}_6^1	1	$4.2e(-15)$	$2.3e(-14)$			
	2	$5.2e(-95)$	$2.6e(-94)$	$1.3e(-12)$	$1.9e(-12)$	
	3	$1.4e(-574)$	$6.7e(-574)$	$1.9e(-12)$		6.0
\widehat{OM}_6^2	1	$4.2e(-15)$	$2.3e(-14)$			
	2	$5.2e(-95)$	$2.6e(-94)$	$1.3e(-12)$	$1.9e(-12)$	
	3	$1.4e(-574)$	$6.7e(-574)$	$1.9e(-12)$		6.0
\widehat{OM}_6^3	1	$4.2e(-15)$	$2.3e(-14)$			
	2	$5.4e(-95)$	$2.7e(-94)$	$1.3e(-12)$	$1.9e(-12)$	
	3	$1.7e(-574)$	$8.2e(-574)$	$1.9e(-12)$		6.0
\widehat{OM}_6^4	1	$4.4e(-15)$	$2.4e(-14)$			
	2	$7.0e(-95)$	$3.5e(-94)$	$1.4e(-12)$	$1.9e(-12)$	
	3	$8.4e(-574)$	$4.1e(-573)$	$1.9e(-12)$		6.000

Cuadro 7.12: Comportamiento de la convergencia de los diferentes métodos en el problema de Fisher 7.8.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$1.4e(+1)$	2.9			
	2	$2.4e(-3)$	$3.1e(-4)$	$4.2e(-6)$	$5.3e(-7)$	
	3	$1.5e(-26)$	$1.4e(-27)$	$5.3e(-7)$		5.9
LM	1	6.8	1.3			
	2	$2.6e(-5)$	$2.7e(-6)$	$4.7e(-6)$	$6.2e(-7)$	
	3	$2.7e(-35)$	$2.6e(-36)$	$6.2e(-7)$		5.3
SM	1	$2.7e(-1)$	$6.8e(-2)$			
	2	$4.5e(-14)$	$4.8e(-15)$	$4.4e(-7)$	$4.9e(-8)$	
	3	$4.8e(-93)$	$4.7e(-94)$	$4.9e(-8)$		6.0
HM	1	4.0	$8.9e(-1)$			
	2	$1.7e(-5)$	$1.8e(-6)$	$8.9e(-6)$	$3.7e(-6)$	
	3	$8.0e(-34)$	$7.6e(-35)$	$3.7e(-6)$		5.0
SAM	1	$2.0e(+1)$	3.5			
	2	$4.7e(-3)$	$4.9e(-4)$	$4.4e(-6)$	$2.6e(-7)$	
	3	$1.5e(-25)$	$1.5e(-26)$	$2.6e(-7)$		5.8
CM	1	$4.8e(-1)$	$1.1e(-1)$			
	2	$1.5e(-12)$	$1.6e(-13)$	$7.3e(-7)$	$7.2e(-8)$	
	3	$1.5e(-83)$	$1.4e(-84)$	$7.2e(-8)$		6.0
\widehat{OM}_6^1	1	$2.0e(-1)$	$4.8e(-2)$			
	2	$4.4e(-15)$	$4.8e(-16)$	$3.1e(-7)$	$3.8e(-8)$	
	3	$2.0e(-100)$	$2.0e(-101)$	$3.8e(-8)$		6.1
\widehat{OM}_6^2	1	$2.0e(-1)$	$4.8e(-2)$			
	2	$4.4e(-15)$	$4.8e(-16)$	$3.1e(-7)$	$3.8e(-8)$	
	3	$1.8e(-100)$	$1.8e(-101)$	$3.8e(-8)$		6.1
\widehat{OM}_6^3	1	$3.2e(-1)$	$7.9e(-2)$			
	2	$1.3e(-13)$	$1.4e(-14)$	$5.1e(-7)$	$5.7e(-8)$	
	3	$3.2e(-90)$	$3.1e(-91)$	$5.7e(-8)$		6.0
\widehat{OM}_6^4	1	3.0	$6.2e(-1)$			
	2	$1.1e(-7)$	$1.2e(-8)$	$3.1e(-6)$	$2.0e(-7)$	
	3	$8.2e(-54)$	$7.8e(-55)$	$2.0e(-7)$		6.0

Ejemplo 7.9 En este ejemplo, consideramos uno de los más famosos problemas de la ciencia aplicada que se conoce como ecuación integral de Hammerstein (ver [171, pp. 19-20]) para verificar la efectividad y la aplicabilidad de nuestros métodos propuestos en comparación con otros métodos existentes. Podemos escribirlo de la siguiente manera:

$$x(s) = 1 + \frac{1}{5} \int_0^1 F(s, t)x(t)^3 dt$$

donde $x \in C[0, 1]$, $s, t \in [0, 1]$ y el núcleo F es

$$F(s, t) = \begin{cases} (1-s)t, & t \leq s, \\ s(1-t), & s \leq t. \end{cases}$$

Transformaremos la ecuación anterior en un problema de dimensión finita mediante el uso de la fórmula de cuadratura de Gauss Legendre dada como $\int_0^1 f(t)dt \simeq \sum_{j=1}^8 w_j f(t_j)$, donde las abscisas t_j y los pesos w_j vienen determinados por $t = 8$ por la fórmula de cuadratura de Gauss Legendre. Denotando las aproximaciones de $x(t_i)$ por x_i ($i = 1, 2, \dots, 8$), obtenemos el sistema de ecuaciones no lineales $5x_i - 5 - \sum_{j=1}^8 a_{ij}x_j^3 = 0$, donde $i = 1, 2, \dots, 8$

$$a_{ij} = \begin{cases} w_j t_j (1 - t_i), & j \leq i, \\ w_j t_i (1 - t_j), & i < j. \end{cases}$$

Donde las abscisas t_j y los pesos w_j son conocidos y representados en la Tabla 7.13 cuando $t = 8$.

La convergencia de los métodos hacia la raíz

$$\xi^* = (1.002 \dots, 1.009 \dots, 1.019 \dots, 1.026 \dots, 1.026 \dots, 1.019 \dots, 1.009 \dots, 1.002 \dots)^T,$$

se prueba en la Tabla 7.14 sobre la base de la aproximación inicial

$$x^{(0)} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right)^T.$$

7.5.1. Resultados y discusiones

Sobre la base de los resultados representados en las Tablas 7.2–7.6 es sencillo decir que nuestros métodos funcionan mucho mejor que los métodos existentes del mismo orden, aun cuando los comparamos con los mismos problemas de prueba y las mismas conjeturas iniciales en el caso de la ecuación escalar. En lo que respecta al sistema de ecuaciones no lineales, también encontramos que nuestros métodos tienen errores residuales más pequeños y menor error entre dos

Cuadro 7.13: (Abscisas y pesos de la fórmula de cuadratura de Gauss Legendre para $t = 8$)

j	t_j	w_j
1	0.01985507175123188...	0.050614268145188129...
2	0.10166676129318663...	0.111190517226687235...
3	0.23723379504183550...	0.156853322938943643...
4	0.40828267875217509...	0.181341891689180991...
5	0.59171732124782490...	0.181341891689180991...
6	0.76276620495816449...	0.156853322938943643...
7	0.89833323870681336...	0.111190517226687235...
8	0.98014492824876811...	0.050614268145188129...

iteraciones consecutivas que los errores de los métodos existentes como se puede ver en las Tablas 7.8–7.14. Sin lugar a dudas, la Tabla 7.15 demuestra que los métodos LM y CM consumen una pequeña cantidad de tiempo en comparación con los otros métodos mencionados.

En realidad, sabemos que existen dos formas de verificar la eficacia de los métodos iterativos: la primera es el índice de eficiencia y la segunda se basa en dos restricciones que son, la precisión de la aproximación a la raíz y el tiempo. Aquí, vamos a considerar que la segunda forma es la mejor a la hora de verificar la eficiencia de los métodos.

Pero estas dos restricciones no suelen darse juntas. Por ejemplo, el método A consume menos cantidad de tiempo que el método B pero este método B puede tener mayor precisión a la hora de dar aproximación a la raíz requerida. También encontramos veces que los métodos tienen comportamientos similares como LM y CM . Ellos están consumiendo una cantidad mínima de tiempo en comparación con otros, pero no tienen mayor precisión a la raíz requerida que nuestros métodos. Además, si ejecutamos nuestro programa en una gran estación de trabajo, el tiempo puede ser casi igual para todos los métodos iterativos mencionados. Finalmente, llegamos a la conclusión de que nuestros métodos pueden haber tardado unos segundos más, pero tienen errores residuales más pequeños, errores más pequeños entre dos iteraciones consecutivas y un orden computacional de convergencia estable en comparación con el existente tanto en el caso de ecuaciones escalares como en el caso del sistema de ecuaciones no lineales.

Cuadro 7.14: Comportamiento de la convergencia de los diferentes métodos en el problema integral de Hammerstein 7.9.

Métodos	n	$\ F(x^{(n)})\ $	$\ x^{(n+1)} - x^{(n)}\ $	$\frac{\ x^{(n+1)} - x^{(n)}\ }{\ x^{(n)} - x^{(n-1)}\ ^6}$	η	ρ^*
AM	1	$1.1e(-5)$	$2.4e(-6)$			
	2	$5.4e(-39)$	$1.2e(-39)$	$6.5e(-6)$	$7.0e(-6)$	
	3	$8.0e(-239)$	$1.7e(-239)$	$7.0e(-6)$		6.0
LM	1	$8.6e(-6)$	$1.8e(-6)$			
	2	$8.4e(-37)$	$1.7e(-37)$	$4.4e(-3)$	$1.2e(+31)$	
	3	$1.4e(-189)$	$3.0e(-190)$	$1.2e(+31)$		4.9
SM	1	$5.4e(-6)$	$1.2e(-6)$			
	2	$1.9e(-41)$	$4.1e(-42)$	$1.6e(-6)$	$1.7e(-6)$	
	3	$3.7e(-254)$	$8.0e(-255)$	$1.7e(-6)$		6.0
HM	1	$3.0e(-5)$	$6.5e(-6)$			
	2	$6.9e(-31)$	$1.5e(-31)$	1.9	$9.2e(+25)$	
	3	$4.4e(-159)$	$9.4e(-160)$	$9.2e(+25)$		5.0
SAM	1	$9.4e(-6)$	$2.0e(-6)$			
	2	$1.5e(-39)$	$3.2e(-40)$	$4.9e(-6)$	$5.3e(-6)$	
	3	$2.7e(-242)$	$5.7e(-243)$	$5.3e(-6)$		6.0
CM	1	$5.8e(-6)$	$1.2e(-6)$			
	2	$3.3e(-41)$	$7.1e(-42)$	$1.8e(-6)$	$2.0e(-6)$	
	3	$1.3e(-252)$	$2.7e(-253)$	$2.0e(-6)$		6.0
\widehat{OM}_6^1	1	$5.2e(-6)$	$1.1e(-6)$			
	2	$1.3e(-41)$	$2.8e(-42)$	$1.4e(-6)$	$1.5e(-6)$	
	3	$3.6e(-255)$	$7.7e(-256)$	$1.5e(-6)$		6.0
\widehat{OM}_6^2	1	$5.2e(-6)$	$1.1e(-6)$			
	2	$1.3e(-41)$	$2.8e(-42)$	$1.4e(-6)$	$1.5e(-6)$	
	3	$3.6e(-255)$	$7.7e(-256)$	$1.5e(-6)$		6.0
\widehat{OM}_6^3	1	$5.5e(-6)$	$1.2e(-6)$			
	2	$2.0e(-41)$	$4.3e(-42)$	$1.6e(-6)$	$1.8e(-6)$	
	3	$5.4e(-254)$	$1.2e(-254)$	$1.8e(-6)$		6.0
\widehat{OM}_6^4	1	$7.5e(-6)$	$1.6e(-6)$			
	2	$2.6e(-40)$	$5.6e(-41)$	$3.1e(-6)$	$3.4e(-6)$	
	3	$5.1e(-247)$	$1.1e(-247)$	$3.4e(-6)$		6.0

Cuadro 7.15: Tiempo de CPU tomado por diferentes métodos para el ejemplo correspondiente.

<i>Método</i>	<i>Ej1</i>	<i>Ej2</i>	<i>Ej3</i>	<i>Ej4</i>	<i>Ej5</i>	<i>Ej6</i>	<i>Media Time</i>
<i>AM</i>	0.87812	0.270195	37.94686	139.30154	14.15401	28.34705	36.8162975
<i>LM</i>	0.52488	0.151101	19.23762	72.57634	8.17179	14.07195	19.1222840
<i>SM</i>	0.91315	0.26168	24.64143	84.84902	10.26124	27.74261	24.7781923
<i>HM</i>	1.19985	0.328744	38.71340	140.02104	17.19316	28.34703	37.6338749
<i>SAM</i>	2.39269	0.675990	41.03603	86.42112	9.52973	84.93808	37.4989419
<i>CM</i>	0.935165	0.52478	15.63204	48.09502	7.69843	28.84539	16.9551404
\widehat{OM}_6^1	1.006211	0.274	27.75462	95.77575	11.87340	28.10860	27.4654650
\widehat{OM}_6^2	1.00871	0.275193	27.64554	96.13399	11.99347	28.77534	27.6387122
\widehat{OM}_6^3	1.01421	0.27420	27.84768	95.94844	12.01449	29.46183	27.7601466
\widehat{OM}_6^4	1.003771	0.276201	28.17594	96.43620	12.21464	29.00251	27.85154

Capítulo 8

Estudio de la convergencia y dinámica de una familia de métodos de orden alto

Resumen

Presentamos el análisis de convergencia local y el estudio de la dinámica de un método iterativo óptimo de decimosexto orden para aproximar una solución localmente única de multiplicidad mayor que uno de una ecuación no lineal. La convergencia se obtiene mediante el uso de una condición centrada de Hölder, en la que la bola de convergencia es mayor que en estudios previos. En el análisis de convergencia local del método, utilizaremos hipótesis que involucran sólo a la primera derivada de la función F , frente a otros análisis de convergencia anteriores que utilizaban al menos la tercera derivada. Con ello damos un radio computable de convergencia y unas estimaciones del error basadas en las constantes de Lipschitz. Además, también se presenta la dinámica del método. Para ello se estudian los puntos críticos libres e independientes del método dibujando los planos de parámetros asociados, donde encontramos miembros de la familia con un comportamiento dinámico complicado. Continuamos mostrando los planos dinámicos de la familia con regiones de convergencia a cualquiera de los puntos fijos extraños, a regiones de convergencia a un 2-ciclo atractor o a regiones de convergencia a un 3-ciclo atractor. Para finalizar, se proporciona una aplicación del mundo real relacionada con la química, más concretamente con la fracción de nitrógeno-hidrógeno que se convierte en amoníaco, que valida los resultados teóricos de convergencia obtenidos en este capítulo.

8.1. Introducción

En este capítulo nos ocupamos del problema de aproximar una solución única local x^* de la ecuación

$$F(x) = 0, \tag{8.1.1}$$

donde F es una función diferenciable definida sobre un subconjunto convexo no vacío D de S con valores en S , donde S es \mathbb{R} o \mathbb{C} .

Muchos problemas de las Ciencias Aplicadas, incluyendo la ingeniería, pueden resolverse mediante la búsqueda de soluciones de ecuaciones en una forma similar a (8.1.1) utilizando Modelado Matemático [16, 26, 92, 130, 193]. Excepto en casos especiales, las soluciones de estas ecuaciones no se pueden encontrar de forma exacta. Esta es la razón principal por la cual los métodos de resolución más comúnmente utilizados suelen ser iterativos. El análisis de la convergencia de métodos iterativos se divide generalmente en dos categorías: análisis de convergencia semilocal y análisis de convergencia local. El análisis de convergencia semilocal se basa en la información alrededor de un punto inicial para proporcionar criterios que garanticen la convergencia de los procedimientos de iteración, mientras que el análisis de convergencia local se basa en la información alrededor de una solución para encontrar estimaciones de los radios calculados de las bolas de convergencia. Los resultados locales son importantes ya que nos proporcionan el grado de dificultad para elegir los puntos iniciales. Un problema muy importante en el estudio de los procedimientos iterativos es el dominio de convergencia. En general, el dominio de convergencia es pequeño. Por lo tanto, es importante ampliar el dominio de convergencia sin tener que añadir hipótesis adicionales. Otro problema importante es encontrar estimaciones de error más precisas en las distancias $\|x_{n+1} - x_n\|$, $\|x_n - x^*\|$. Estos son, junto con el estudio del comportamiento dinámico, nuestros objetivos en este capítulo.

Las propiedades dinámicas relacionadas con un método iterativo aplicado a polinomios brindan información importante sobre su estabilidad y exactitud. En estudios recientes, autores como Amat et al [8, 9, 12], Chun et al. [60], Gutiérrez et al. [112], Magreñán [134, 144, 150], y muchos otros [50, 52, 53, 54, 87, 88, 94, 112, 118, 120, 123, 137, 191, 193, 215, 220] han encontrado planos dinámicos interesantes, que incluyen comportamientos periódicos y otras anomalías. Uno de nuestros principales intereses en este trabajo es el estudio de los espacios de parámetros asociados a una familia de métodos iterativos, que nos permitan distinguir entre los métodos buenos y malos en términos de sus propiedades numéricas.

Presentamos la convergencia local y la dinámica del método de tres pasos

definido para cada $n = 0, 1, 2, \dots$ por

$$\begin{aligned}
y_n &= x_n - \alpha F'(x_n)^{-1} F(x_n) \\
z_n &= y_n - A_1(x_n) F'(x_n)^{-1} F(y_n) \\
v_n &= z_n - A_2(x_n) F'(x_n)^{-1} F(z_n) \\
x_{n+1} &= z_n - A_3(x_n) F'(x_n)^{-1} F(v_n),
\end{aligned} \tag{8.1.2}$$

donde x_0 es un punto inicial, $\alpha \in \mathbb{R}$ es un parámetro y $A_i : \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, 2, 3$ son funciones continuas dadas. Muchos métodos multipaso son casos especiales de método (8.1.2).

▪ Método de Artidiello et al [32] ($\alpha = 1$)

$$\begin{aligned}
y_n &= x_n - \alpha F'(x_n)^{-1} F(x_n) \\
z_n &= y_n - \bar{A}_1 \left(\frac{F(y_n)}{F(x_n)} \right) F'(x_n)^{-1} F(y_n) \\
v_n &= z_n - \bar{A}_2 \left(\frac{F(y_n)}{F(x_n)}, \frac{F(z_n)}{F(y_n)} \right) F'(x_n)^{-1} F(z_n) \\
x_{n+1} &= z_n - \bar{A}_3 \left(\frac{F(y_n)}{F(x_n)}, \frac{F(z_n)}{F(y_n)}, \frac{F(v_n)}{F(z_n)} \right) F'(x_n)^{-1} F(v_n).
\end{aligned} \tag{8.1.3}$$

El método (8.1.3) puede reducirse aún más a

▪ Método de Dzunic et al [190]

$$\begin{aligned}
y_n &= x_n - \alpha F'(x_n)^{-1} F(x_n) \\
x_{n+1} &= z_n - \bar{A}_1 \left(\frac{F(y_n)}{F(x_n)} \right) F'(x_n)^{-1} F(y_n).
\end{aligned}$$

Para $A_1(u) = \frac{1}{(1-u)^2}$, obtenemos el método de Kung-Traub [220]; para $A_1(u) = \frac{1+\alpha u}{2+(\beta-2)u}$, obtenemos la familia de cuarto orden de King y para $A_1(u) = \frac{1+2u+\alpha u^2}{2+(\beta-5)u}$, obtenemos el método de cuarto orden dado por Zhao et al en [234], donde $u = \frac{F(y_n)}{F(x_n)}$. Finalmente, si ponemos $\alpha = 1$ en el método (8.1.3), saltamos el último subpaso y hacemos $x_{n+1} = v_n$, entonces obtenemos el método de octavo orden estudiado por Dzunic et al [77].

Sin embargo, la convergencia en todos los métodos después del método (8.1.2) se ha demostrado utilizando el desarrollo de Taylor y bajo hipótesis en al menos la tercera derivada, aunque solo la primera derivada aparece en estos métodos. En este trabajo presentamos el análisis de convergencia local del método (8.1.2)

usando hipótesis solo en la primera derivada de la función F . También proporcionamos un radio computable de convergencia y estimaciones de error basadas en las constantes de Lipschitz. Observe también que, en particular, hay una gran cantidad de métodos iterativos para aproximar soluciones de ecuaciones no lineales definidas en \mathbb{R} o \mathbb{C} [62, 68, 81, 151, 177]. Estos resultados muestran que si el punto inicial x_0 está lo suficientemente cerca de la solución x^* , entonces la sucesión $\{x_n\}$ converge a x^* . Pero, ¿cómo de cerca de la solución x^* debe de estar la aproximación inicial? Estos resultados locales no proporcionan información sobre el radio de la bola de convergencia para el método correspondiente. Abordamos esta cuestión para el método (8.1.2) en la Sección 2. La misma técnica se puede usar para otros métodos.

8.2. Convergencia local

Sean $U(v, \rho)$, $\bar{U}(v, \rho)$ bolas abiertas y cerradas en S , respectivamente con centro $v \in S$ y de radio $\rho > 0$. Presentamos el análisis de convergencia local del método (8.1.2) en esta sección bajo las condiciones (C):

(C₁) $F : D \subset S \rightarrow S$ es una función diferenciable.

Existe $x^* \in D$ constante, $L_0 > 0$, tal que para cada $x \in D$

(C₂) $F(x^*) = 0$, $F'(x^*) \neq 0$.

(C₃) $\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0\|x - x^*\|$

Sea $D_0 := D \cap U(x^*, \frac{1}{L_0})$. Existen constante $L > 0$, $M \geq 1$ tales que para cada $x, y \in D_0$

(C₄) $\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq L\|x - y\|$

(C₅) $\|F'(x^*)^{-1}F'(x)\| \leq M$.

Existen parámetros γ_i y funciones continuas no decrecientes $\psi_i : [0, \gamma_i) \rightarrow \mathbb{R}$ tales que $i = 0, 1, 2, 3$:

(C₆) $\gamma_{i+1} \leq \gamma_i \leq \frac{1}{L_0}$

y

(C₇) $\psi_i(t) \rightarrow a$ número positivo o $+\infty$ como $t \rightarrow \gamma_i^{-1}$. Para $\alpha \in \mathbb{R}$, definimos las funciones

$$q_j : [0, \gamma_j) \rightarrow \mathbb{R} \quad j = 0, 1, 2, 3 \text{ por}$$

$$q_j(t) = \begin{cases} M|1 - \alpha|, & j = 0 \\ M^{i+j}|1 - \alpha| \prod_{i=0}^j \psi_i(t) \cdots \psi_j(t), & j = 1, 2, 3 \end{cases}$$

(C₈) $p_j := q_j(0) < 1$, $j = 0, 1, 2, 3$,

(C₉) $A_i : S \rightarrow S$ son funciones continuas tales que para cada $x \in D_0$, $\|A_i(x)\| \leq \psi_i(\|x - x^*\|)$ y

(C₁₀) $\bar{U}(x^*, r) \subset D$ para algún $r > 0$ que se determinará más tarde.

También es conveniente para el análisis de la convergencia local del método (8.1.2) introducir algunas funciones y parámetros. Definimos la función g_0 en el intervalo $[0, \frac{1}{L_0})$ por

$$g_0(t) = \frac{1}{2(1 - L_0 t)}(Lt + 2M|1 - \alpha|)$$

y los parámetros r_0, r_A por

$$r_0 = \frac{2(1 - M|1 - \alpha|)}{2L_0 + L}, \quad r_A = \frac{2}{2L_0 + L}.$$

Entonces, como $p_0 = M|1 - \alpha| < 1$ por (C₈), tenemos que $0 < r_0 < r_A$, $g_0(r_1) = 1$ y para cada $t \in [0, r_1)$ $0 \leq g_0(t) < 1$. Las funciones g_i, h_i definidas en el intervalo $[0, \gamma_i)$ por

$$g_i(t) = \left(1 + \frac{M\psi_i(t)}{1 - L_0 t}\right)g_{i-1}(t)$$

y

$$h_i(t) = g_i(t) - 1$$

para $i = 1, 2, 3$. Por (C₈) tenemos que $h_i(0) = p_j - 1 < 0$ y por (C₆) y (C₇) $h_i(t) \rightarrow a$ número positivo o $+\infty$. Por el teorema del valor intermedio deducimos que la función h_i tiene ceros en el intervalo $[0, \gamma_i)$. Denotamos por r_i el más pequeño de esos ceros. Sea

$$r = \min\{r_j\}, \quad j = 0, 1, 2, 3. \quad (8.2.1)$$

Entonces, tenemos que

$$0 \leq r < r_A \quad (8.2.2)$$

y para cada $j = 0, 1, 2, 3$, $t \in [0, r)$

$$0 \leq g_j(t) < 1. \quad (8.2.3)$$

A continuación, usando la notación anterior y las condiciones (C) presentamos el resultado de convergencia local para el método (8.1.2).

Teorema 8.1 *Supongamos que se dan las condiciones (C) con el radio r en (C₁₀) definida para (8.2.1). Entonces, la sucesión $\{x_n\}$ generada por $x_0 \in U(x^*, r) \setminus$*

$\{x^*\}$ por el método (8.1.2) está bien definida, permanece en $U(x^*, r)$ para cada $n = 0, 1, 2, \dots$ y converge a x^* . Además, las siguientes estimaciones se mantienen:

$$\|y_n - x^*\| \leq g_0(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\| < r, \quad (8.2.4)$$

$$\|z_n - x^*\| \leq g_1(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|, \quad (8.2.5)$$

$$\|v_n - x^*\| \leq g_2(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\| \quad (8.2.6)$$

y

$$\|x_{n+1} - x^*\| \leq g_3(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|, \quad (8.2.7)$$

donde las funciones “ g ” son las definidas antes del Teorema 8.1. Además, para

$$T \in [r, \frac{2}{L_0}) \quad (8.2.8)$$

el punto límite x^* es la única solución a la ecuación $F(x) = 0$ en $\bar{U}(x^*, T) \cap D$.

Demostración.- Vamos a demostrar las estimaciones (8.2.4) y (8.2.8) usando inducción matemática. De la hipótesis $x_0 \in U(x, r) \setminus \{x^*\}$, las condiciones (C_1) , (C_3) y (8.2.1), obtenemos que

$$\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x_0 - x^*\| < L_0r < 1. \quad (8.2.9)$$

Se sigue del lema de Banach sobre funciones invertibles [16, 26, 190] que $F'(x_0)^{-1} \in L(S, S)$ y

$$\|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|}. \quad (8.2.10)$$

Por lo tanto, y_0 está bien definido para el primer paso del método (8.1.2) para $n = 0$. De las condiciones (C_1) y (C_2) podemos escribir que

$$F(x_0) = F(x_0) - F(x^*) = \int_0^1 F'(x^* + \theta(x_0 - x^*))(x_0 - x^*)d\theta. \quad (8.2.11)$$

Notemos que $\|x^* + \theta(x_0 - x^*) - x^*\| = \theta\|x_0 - x^*\| < r$, así que $x^* + \theta(x_0 - x^*) \in U(x^*, r)$. Entonces, usando (8.2.11) y la condición (C_5) , tenemos que

$$\|F'(x^*)^{-1}F(x_0)\| \leq \left\| \int_0^1 F'(x^*)^{-1}F'(x^* + \theta(x_0 - x^*))(x_0 - x^*)d\theta \right\| \leq M\|x_0 - x^*\|. \quad (8.2.12)$$

En vista de las condiciones (C_2) , (C_4) , (8.2.1), (8.2.3) (para $j = 0$), (8.2.10) y (8.2.12), obtenemos que

$$\begin{aligned}
\|y_0 - x^*\| &= \|x_0 - x^* - F'(x_0)^{-1}F(x_0) + (1 - \alpha)F'(x_0)^{-1}F(x_0)\| \\
&\leq \|x_0 - x^* - F'(x_0)^{-1}F(x_0)\| + |1 - \alpha| \|F'(x_0)^{-1}F'(x^*)\| \|F'(x^*)^{-1}F(x_0)\| \\
&\leq \|F'(x_0)^{-1}F'(x^*)\| \|F'(x^*)^{-1}F(x_0)\| \\
&\quad \times \left\| \int_0^1 F'(x^*)^{-1}(F'(x^* + \theta(x_0 - x^*)) - F'(x_0))(x_0 - x^*)d\theta \right. \\
&\quad \left. + \frac{|1 - \alpha|M\|x_0 - x^*\|}{1 - L_0\|x_0 - x^*\|} \right\| \\
&\leq \frac{L\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)} + \frac{|1 - \alpha|M\|x_0 - x^*\|}{1 - L_0\|x_0 - x^*\|} \\
&= g_0(\|x_0 - x^*\|)\|x_0 - x^*\| < \|x_0 - x^*\| < r,
\end{aligned} \tag{8.2.13}$$

lo cual demuestra (8.2.4) para $n = 0$ y $y_0 \in U(x^*, r)$. Entonces, por la condición (C_9) , (8.2.1), (8.2.3) (para $j = 1$), (8.2.10), (8.2.12) (para $y_0 = x_0$) y (8.2.13), tenemos que

$$(8.2.14)$$

lo cual demuestra (8.2.5) para $n = 0$ y $z_0 \in U(x^*, r)$. Similarmente, se demuestran las estimaciones (8.2.6)-(8.2.7) para $n = 0$ y $v_0, x_1 \in U(x^*, r)$. Simplemente reemplazando x_0, y_0, z_0, v_0, x_1 por $x_k, y_k, z_k, v_k, x_{k+1}$ en las estimaciones anteriores llegamos a (8.2.4)-(8.2.7). Usando las estimaciones $\|x_{k+1} - x^*\| \leq c\|x_k - x^*\| < r$, $c = g_3(\|x_0 - x^*\|) \in [0, 1)$ deducimos que $\lim_{k \rightarrow \infty} x_k = x^*$ y $x_{k+1} \in U(x^*, r)$. Finalmente, para demostrar la unicidad, sea $y^* \in \bar{U}(x^*, T)$ tal que $F(y^*) = 0$. Definimos $B = \int_0^1 F'(y^* + \theta(x^* - y^*))d\theta$. Usando la condición (C_2) , obtenemos que

$$\|F'(x^*)^{-1}(B - F'(x^*))\| \leq \frac{L_0}{2}\|y^* - x^*\| \leq \frac{L_0}{2}T < 1. \tag{8.2.15}$$

Por lo tanto, $B \neq 0$. Usando la identidad $0 = F(y^*) - F(x^*) = B(y^* - x^*)$, concluimos que $x^* = y^*$. ■

Nota 8.2 1. A la vista de (8.2.8) y la estimación

$$\begin{aligned}
\|F'(x^*)^{-1}F'(x)\| &= \|F'(x^*)^{-1}(F'(x) - F'(x^*)) + I\| \\
&\leq 1 + \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \\
&\leq 1 + L_0\|x_0 - x^*\|
\end{aligned}$$

la condición (8.2.8) puede ser desechada y M puede ser reemplazada por

$$M(t) = 1 + L_0t \text{ o simplemente por } M(t) = M = 2, \text{ ya que } t \in [0, \frac{1}{L_0}).$$

2. Los resultados obtenidos aquí se pueden usar para operadores F que satisfacen la ecuación diferencial autónoma [16, 26] de la forma

$$F'(x) = P(F(x)),$$

donde P es un operador continuo conocido. Como $F'(x^*) = P(F(x^*)) = P(0)$, podemos aplicar los resultados sin saber realmente la solución x^* . Sea por ejemplo $F(x) = e^x - 1$. Entonces, podemos elegir $P(x) = x + 1$.

3. Se ha demostrado en [16], [26] que el radio r_A es el radio de convergencia del método de Newton (8.1.3) bajo las condiciones (8.2.8) y (8.2.9). De (8.2.6) y de la definición de r_1 se sigue que el radio de convergencia r del método (8.1.2) no puede ser más grande que el radio de convergencia r_A del método de segundo orden de Newton (8.1.3). Como ya se señaló el r_A es como mínimo la bola de convergencia dada por Rheinboldt [193]

$$r_R = \frac{2}{3L}. \quad (8.2.16)$$

En particular, para $L_0 < L$ tenemos que

$$r_R < r_A$$

y

$$\frac{r_R}{r_A} \rightarrow \frac{1}{3} \quad \text{como} \quad \frac{L_0}{L} \rightarrow 0.$$

Nuestra bola de convergencia r_1 es como mucho tres veces más grande que la de Rheinboldt. El mismo valor para r_R dado por Traub [215].

4. Vale la pena notar que el método (8.2.1) no varía si usamos las condiciones del Teorema 2.1 en lugar de las condiciones más fuertes dadas en [68, 144]. Además, en la práctica, para los límites de error, podemos usar el orden computacional de convergencia (COC) [16, 26]

$$\xi = \frac{\ln \frac{\|x_{n+2} - x_{n+1}\|}{\|x_{n+1} - x_n\|}}{\ln \frac{\|x_{n+1} - x_n\|}{\|x_n - x_{n-1}\|}}, \quad \text{para cada } n = 1, 2, \dots$$

o el orden computacional aproximado de convergencia (ACOC) [68]

$$\xi^* = \frac{\ln \frac{\|x_{n+2} - x^*\|}{\|x_{n+1} - x^*\|}}{\ln \frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|}}, \quad \text{para cada } n = 0, 1, 2, \dots$$

De esta manera obtenemos en la práctica el orden de convergencia de una manera que evita los límites que implican estimaciones más altas que la primera derivada de Fréchet.

Nota 8.3 Vamos a demostrar cómo elegir las funciones en el caso del método (8.1.3). Tenemos que

$$\begin{aligned}\overline{A_1}\left(\frac{F(y_n)}{F(x_n)}\right) &= A_1(x_n), & \overline{A_2}\left(\frac{F(y_n)}{F(x_n)}\right), \\ & & \frac{F(z_n)}{F(y_n)} &= A_2(x_n), \\ \overline{A_3}\left(\frac{F(y_n)}{F(x_n)}, \frac{F(z_n)}{F(y_n)}, \frac{F(v_n)}{F(z_n)}\right) &= A_3(x_n)\end{aligned}$$

En primer lugar vemos que la condición (C_3) puede ser obviada, ya que $\alpha = 1$. Entonces, para $x_n \neq x^*$, tenemos

$$\begin{aligned}& \| (F'(x^*)(x_n - x^*))^{-1} [F(x_n) - F(x^*) - F'(x^*)(x_n - x^*)] \| \\ & \leq \|x_n - x^*\|^{-1} \frac{L_0}{2} \|x_n - x^*\| = \frac{L_0}{2} \|x_n - x^*\| < \frac{L_0}{2} r < 1.\end{aligned}$$

Por lo tanto, tenemos que

$$\|F'(x_n)^{-1}F(x^*)\| \leq \frac{1}{\|x_n - x^*\|(1 - \frac{L_0}{2}\|x_n - x^*\|)}.$$

En consecuencia, obtenemos que

$$\begin{aligned}\left\|\frac{F(y_n)}{F(x_n)}\right\| &= \|F'(x_n)^{-1}F'(x^*)\| \|F'(x^*)^{-1}F(y_n)\| \\ &\leq \frac{M\|y_n - x^*\|}{\|x_n - x^*\|(1 - \frac{L_0}{2}\|x_n - x^*\|)} \\ &\leq \frac{Mg_0(\|x_n - x^*\|)}{1 - L_0\|x_n - x^*\|}.\end{aligned}\tag{8.2.17}$$

Similarmente, obtenemos

$$\begin{aligned}\|F(y_n)^{-1}F'(x^*)\| &\leq \frac{1}{\|y_n - x^*\|(1 - \frac{L_0}{2}\|y_n - x^*\|)}, \\ \left\|\frac{F(z_n)}{F(y_n)}\right\| &\leq \frac{M(1 + \frac{M\psi_1(\|x_n - x^*\|)}{1 - L_0\|x_n - x^*\|})}{1 - \frac{L_0}{2}g_0(\|x_n - x^*\|)\|x_n - x^*\|}, \\ \|F(z_n)^{-1}F'(x^*)\| &\leq \frac{1}{\|z_n - x^*\|(1 - \frac{L_0}{2}\|y_n - x^*\|)},\end{aligned}\tag{8.2.18}$$

y

$$\left\| \frac{F(z_n)}{F(y_n)} \right\| \leq \frac{M(1 + \frac{M\psi_2(\|x_n - x^*\|)}{1 - L_0\|x_n - x^*\|})}{1 - \frac{L_0}{2}g_0(\|x_n - x^*\|)\|x_n - x^*\|}, \quad (8.2.19)$$

Elijamos A_i , $i = 1, 2, 3, 4$ como en

$$A_1(a) = 1 + 2a + 4a^3 - 3a^4 \quad (8.2.20)$$

$$A_2(a, b) = 1 + 2a + b + a^2 + 4ab + 3a^2b + 4ab^2 + 4a^3b - 4a^2b^2 \quad (8.2.21)$$

y

$$A_3(a, b, c) = 1 + 2a + b + c + a^2 + 4ab + 2ac + 4a^2b + a^2c + 6ab^2 + 8abc - b^3 + 2bc. \quad (8.2.22)$$

Estas funciones satisfacen las condiciones impuestas en el Teorema 8.1, así que el orden de convergencia del método (8.1.3) será al menos 16.

Sea

$$a = a(t) = \frac{Mg_0(t)}{1 - L_0t}, \quad (8.2.23)$$

$$b = b(t) = \frac{M(1 + \frac{M\psi_1(t)}{1 - L_0t})}{1 - \frac{L_0}{2}t}, \quad (8.2.24)$$

$$c = c(t) = \frac{M(1 + \frac{M\psi_2(t)}{1 - L_0t})}{1 - \frac{L_0}{2}t}, \quad (8.2.25)$$

y

$$\gamma_i = \frac{1}{L_0}, \quad i = 0, 1, 2, 3.$$

Entonces se sigue de (8.2.17)-(8.2.22) que las funciones ψ_i pueden ser definidas por

$$\psi_1(t) = 1 + 2a + 4a^3 + 3a^4 \quad (8.2.26)$$

$$\psi_2(t) = 1 + 2a + b + a^2 + 4ab + 3a^2b + 4ab^2 + 4a^3b + 4a^2b^2 \quad (8.2.27)$$

y

$$\psi_3(t) = 1 + 2a + b + c + a^2 + 4ab + 2ac + 4a^2b + a^2c + 6ab^2 + 8abc + b^3 + 2bc. \quad (8.2.28)$$

8.3. Estudio dinámico de un caso especial del método (8.1.2)

En esta sección vamos a estudiar la dinámica compleja de un caso especial del método (8.1.2), que consiste en elegir:

$$A_1(x_n) = F'(y_n)^{-1}F'(x_n),$$

$$A_2(x_n) = F'(z_n)^{-1}F'(x_n)$$

y

$$A_3(x_n) = F'(y_n)^{-1}F'(x_n).$$

Aplicando este operador en un polinomio cuadrático con dos raíces diferentes A y B , $p(z) = (z-A)(z-B)$. Usando la transformación de Möebius $h(z) = \frac{z-A}{z-B}$, que lleva la raíz A al 0, la raíz B al ∞ y el ∞ al 1, obtenemos que el operador racional asociado a la familia de esquemas iterativos es finalmente

$$G(z, \alpha) = \frac{z^8(1 - \alpha + z)^8}{(-1 - z + \alpha z)^8}. \quad (8.3.1)$$

8.3.1. Estudio de los puntos fijos y su estabilidad

Es claro que el $z = 0$ y el $z = \infty$ son puntos fijos de $G(z, \alpha)$. Además, existen algunos puntos fijos extraños que son:

- $z = 1$ relativo a la divergencia al ∞

- Las raíces del polinomio de grado 14:

$$\begin{aligned} & z^{14} - 8\alpha z^{13} + 9z^{13} + 28\alpha^2 z^{12} - 64\alpha z^{12} + 37z^{12} - 56\alpha^3 z^{11} + 196\alpha^2 z^{11} \\ & - 232\alpha z^{11} + 93z^{11} + 70\alpha^4 z^{10} - 336\alpha^3 z^{10} + 616\alpha^2 z^{10} - 512\alpha z^{10} + 163z^{10} \\ & - 56\alpha^5 z^9 + 350\alpha^4 z^9 - 896\alpha^3 z^9 + 1176\alpha^2 z^9 - 792\alpha z^9 + 219z^9 + 28\alpha^6 z^8 \\ & - 224\alpha^5 z^8 + 770\alpha^4 z^8 - 1456\alpha^3 z^8 + 1596\alpha^2 z^8 - 960\alpha z^8 + 247z^8 - \alpha^8 z^7 \\ & + 56\alpha^6 z^7 - 336\alpha^5 z^7 + 980\alpha^4 z^7 - 1680\alpha^3 z^7 + 1736\alpha^2 z^7 - 1008\alpha z^7 + 254z^7 \\ & + 28\alpha^6 z^6 - 224\alpha^5 z^6 + 770\alpha^4 z^6 - 1456\alpha^3 z^6 + 1596\alpha^2 z^6 - 960\alpha z^6 + 247z^6 \\ & - 56\alpha^5 z^5 + 350\alpha^4 z^5 - 896\alpha^3 z^5 + 1176\alpha^2 z^5 - 792\alpha z^5 + 219z^5 + 70\alpha^4 z^4 \\ & - 336\alpha^3 z^4 + 616\alpha^2 z^4 - 512\alpha z^4 + 163z^4 - 56\alpha^3 z^3 + 196\alpha^2 z^3 - 232\alpha z^3 \\ & + 93z^3 + 28\alpha^2 z^2 - 64\alpha z^2 + 37z^2 - 8\alpha z + 9z + 1 \end{aligned}$$

Las soluciones de este polinomio dependen del valor del parámetro α .

En la Figura 8.1 se muestra el diagrama de bifurcación de los puntos fijos

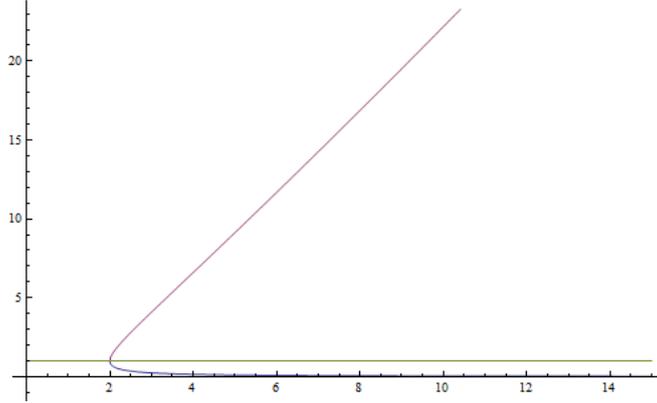


Figura 8.1: Diagrama de bifurcación de los puntos fijos.

8.3.2. Estudio de los puntos críticos y espacios de parámetros

Es un hecho bien conocido que hay al menos un punto crítico asociado con cada componente invariante de Fatou. Los puntos críticos de la familia son las soluciones de $G'(z, \alpha) = 0$, donde

$$G'(z, \alpha) = -\frac{8(-1 + \alpha - x)^7 x^7 (-1 + \alpha - 2x - x^2 + \alpha x^2)}{(-1 - x + \alpha x)^9}$$

Es claro que el $z = 0$ y el $z = \infty$ son puntos críticos. Además, los puntos críticos libres son:

$$cr_1(\alpha) = 2(-1 + \alpha)$$

$$cr_2(\alpha) = \frac{1 - \sqrt{-(-2 + \alpha)\alpha}}{-1 + \alpha}$$

$$cr_3(\alpha) = \frac{1 + \sqrt{-(-2 + \alpha)\alpha}}{-1 + \alpha}$$

Establecemos las relaciones entre los puntos críticos libres en el siguiente lema.

Lema 8.4 a) Si $\alpha = 0$

$$(i) \quad cr_1(\alpha) = cr_2(\alpha) = cr_3(\alpha) = -1.$$

b) Si $\alpha = 2$

$$(i) \quad cr_1(\alpha) = cr_2(\alpha) = cr_3(\alpha) = 1.$$

También está claro que por cada valor de α ,

$$cr_2(\alpha) = \frac{1}{cr_3(\alpha)}.$$

Entonces, solo hay un punto crítico libre e independiente, sin pérdida de generalidad, consideramos en este documento el punto crítico libre $cr_2(\alpha)$. Ahora, vamos a buscar a los mejores miembros de la familia mediante el uso del espacio de parámetros asociado a los puntos críticos libres.

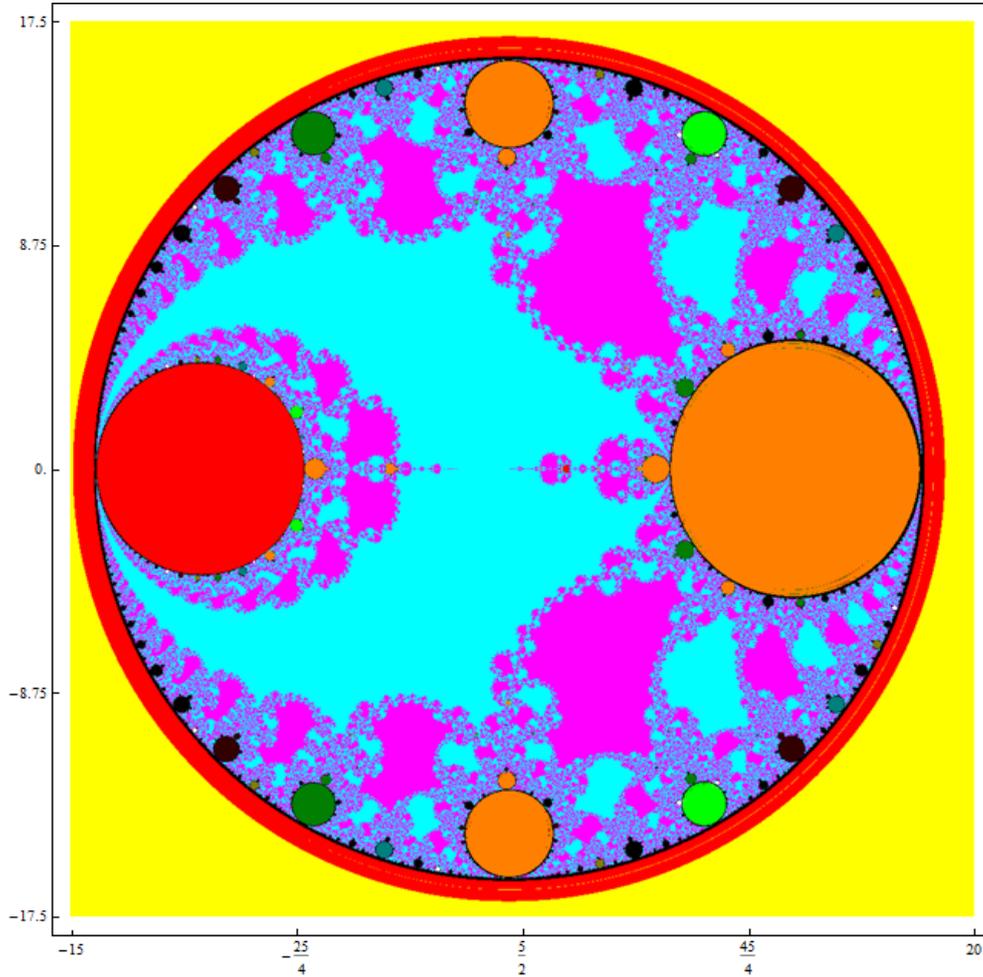


Figura 8.2: Espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$.

Mostramos en las Figuras 8.2–8.4 los planos de parámetros asociados a $cr_2(\alpha)$. Un punto está pintado en cian si la iteración del método que comienza en $z_0 = cr_1(\alpha)$ converge al punto fijo 0 (relativo a la raíz A), o si converge al ∞ (relativo a la raíz B) y en amarillo si la iteración converge a 1 (relativo al ∞). Además, aparece en rojo la convergencia, después de un máximo de 2000 iteraciones y con una tolerancia de 10^{-6} , a cualquiera de los puntos fijos extraños, en naranja la convergencia a 2-ciclos, en verde claro la convergencia a 3-ciclos, en rojo oscuro la convergencia a 4-ciclos, en azul oscuro la convergencia a 5-ciclos, en verde oscuro

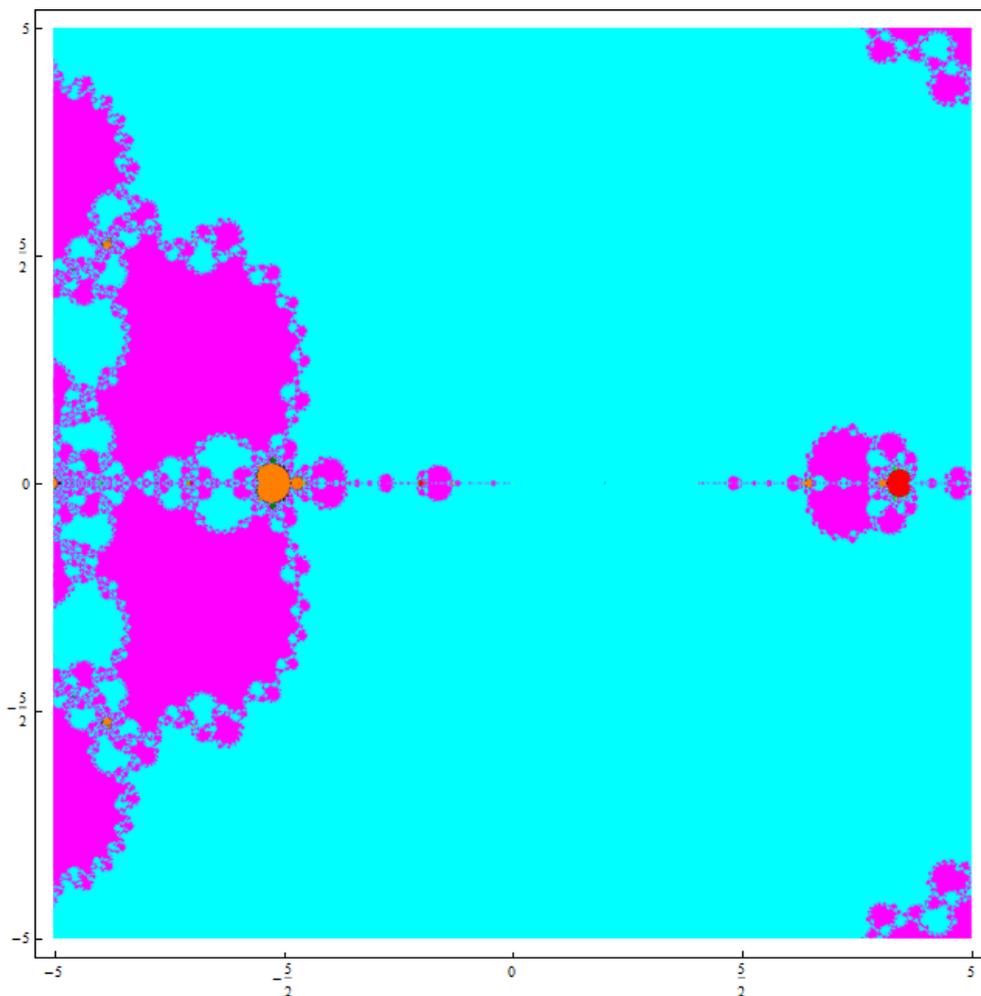


Figura 8.3: Detalle del espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$.

a 6-ciclos, en amarillo oscuro a 7-ciclos, y en blanco la convergencia a 8-ciclos. Las regiones en negro corresponden a zonas de convergencia a otros ciclos. Como consecuencia, cada punto del plano que no es ni cian ni magenta no es una buena opción de α en términos de comportamiento numérico.

En estos planos dinámicos hemos pintado en magenta la convergencia a 0, en cian la convergencia a ∞ y en negro las zonas sin convergencia a las raíces.

Luego, enfocando la atención en la región que se muestra en la Figura 8.3 es evidente que existen miembros de la familia con un comportamiento complicado. En las Figuras 8.5–8.6, se muestran los planos dinámicos de un miembro de la familia con regiones de convergencia a cualquiera de los puntos fijos extraños.

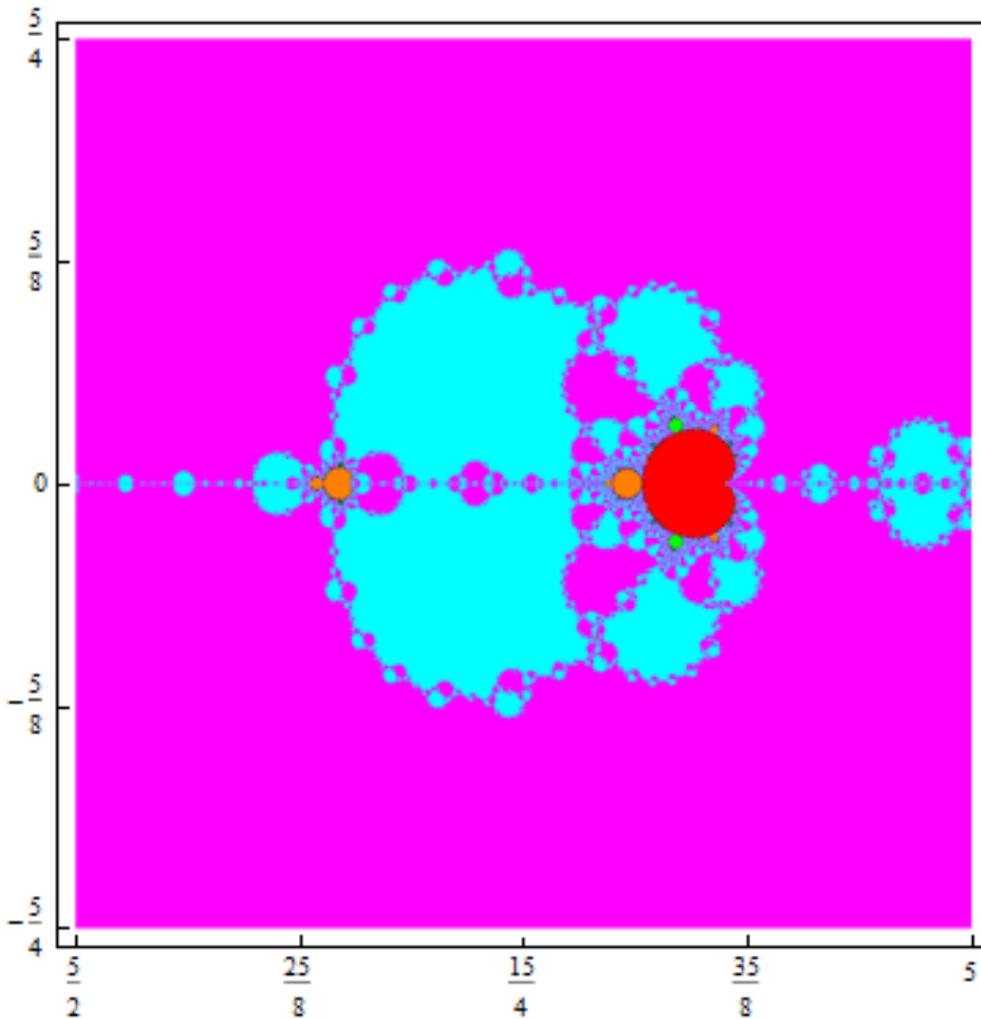


Figura 8.4: Detalle del espacio de parámetros asociado al punto crítico libre $cr_2(\alpha)$.

En las Figuras 8.7–8.8 se muestran los planos dinámicos de los miembros de la familia con las regiones de convergencia a un 2-ciclo atractor.

Por otro lado, en la Figura 8.9, presentamos el plano dinámico de un miembro de la familia con regiones de convergencia a un 3-ciclo atractor.

Mostramos otros caso especiales en las Figuras 8.10 y 8.11.

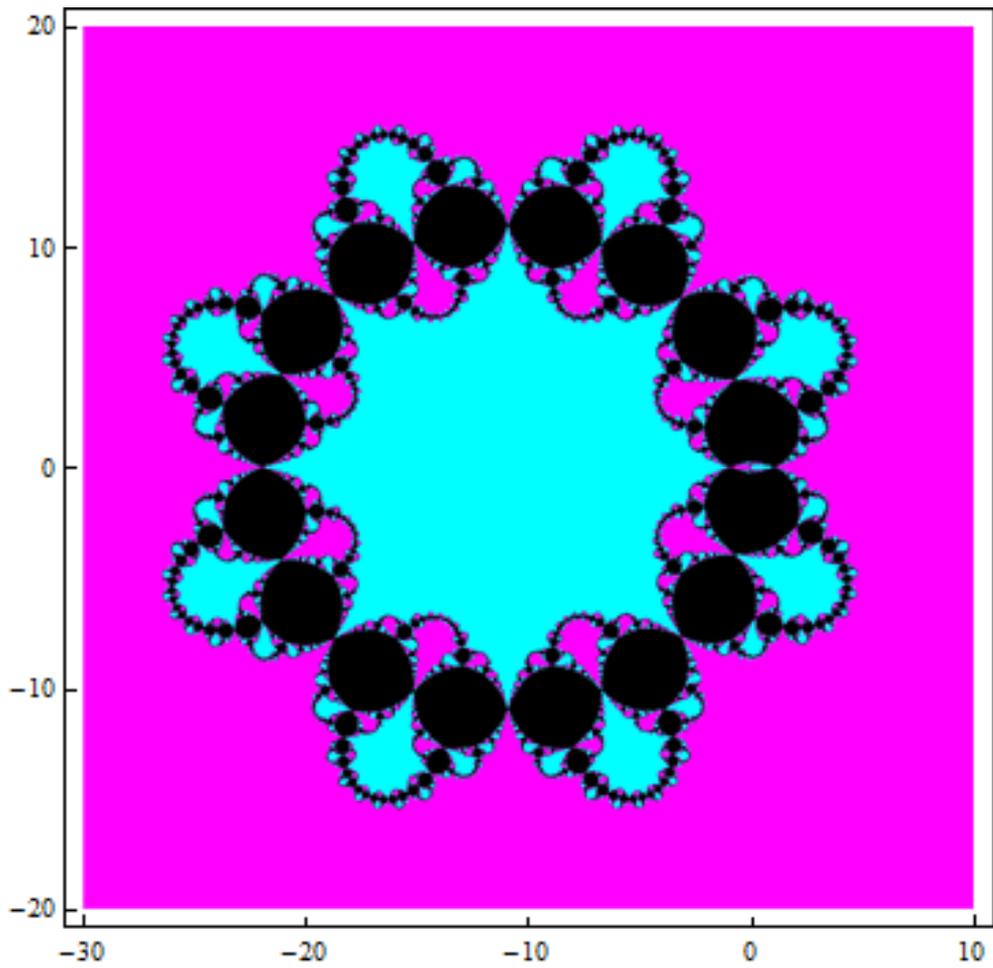


Figura 8.5: Cuencas de atracción asociadas al método con $\alpha = -10$.

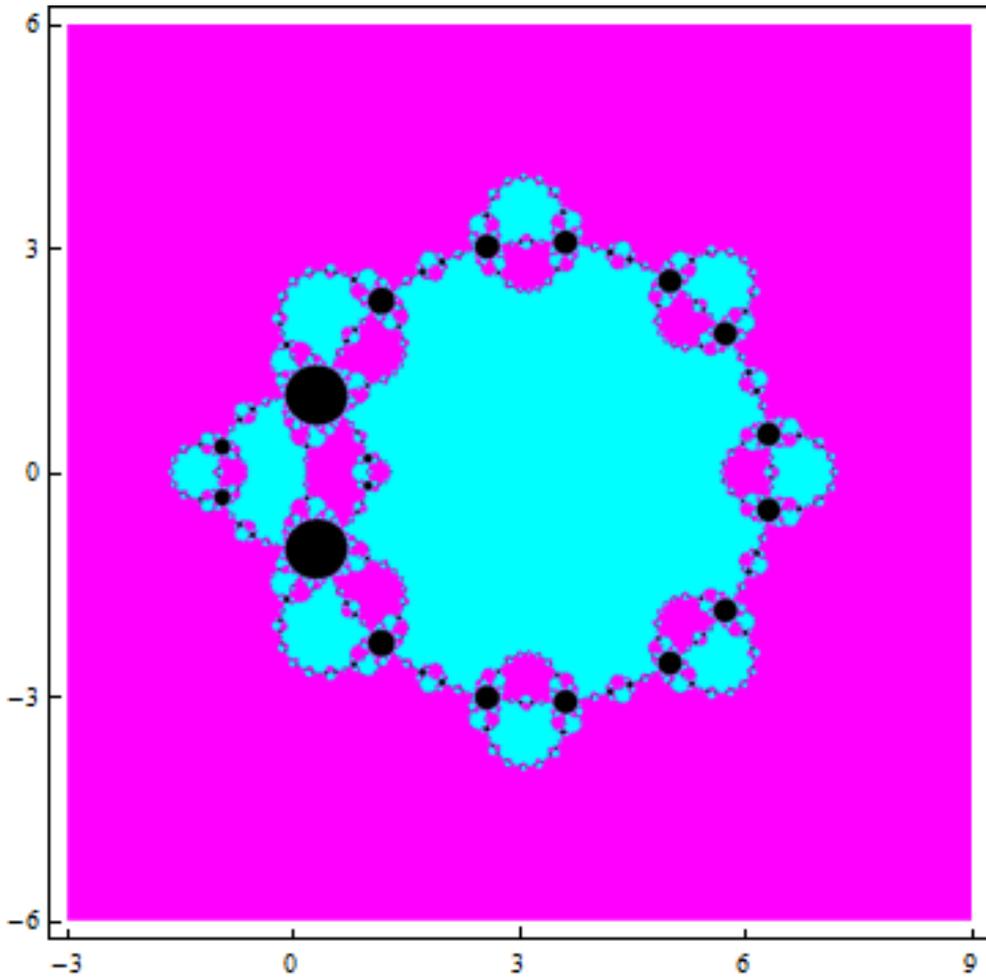


Figura 8.6: Cuencas de atracción asociadas al método con $\alpha = 4.25$.

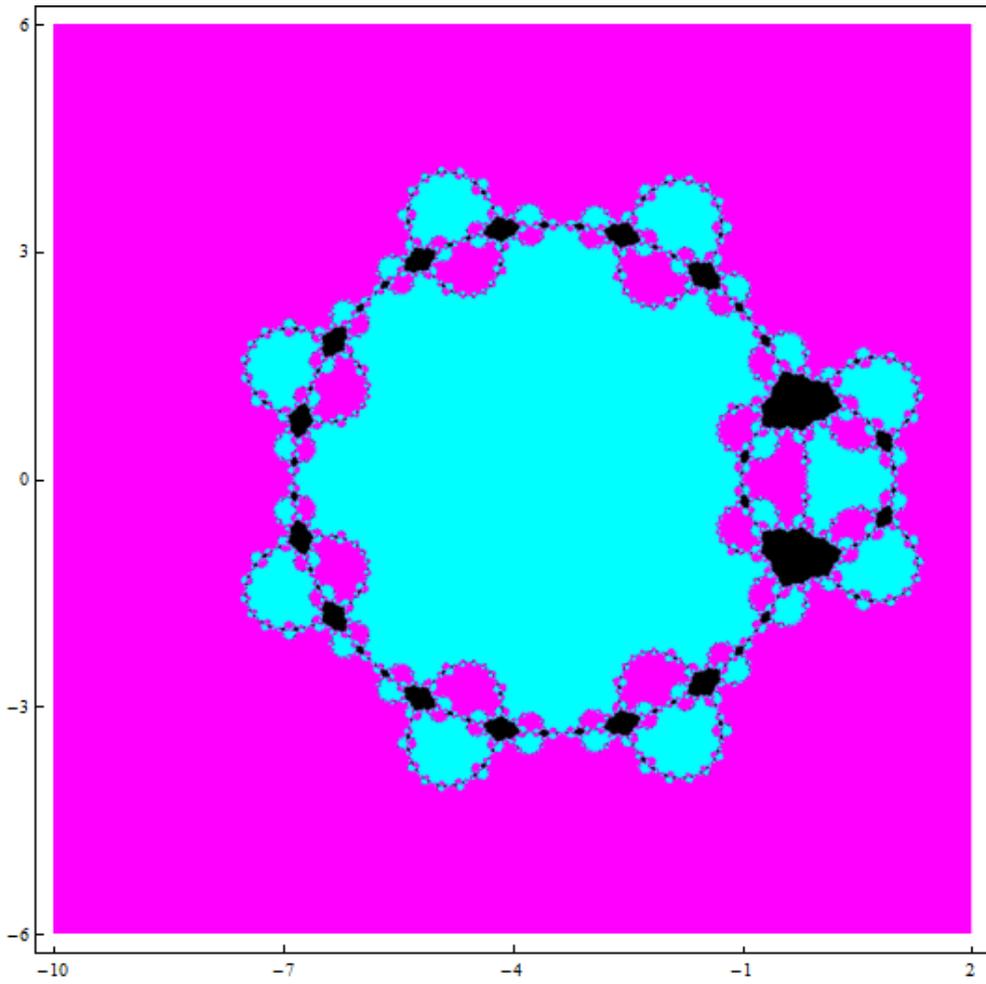


Figura 8.7: Cuencas de atracción asociadas al método con $\alpha = -2.5$.

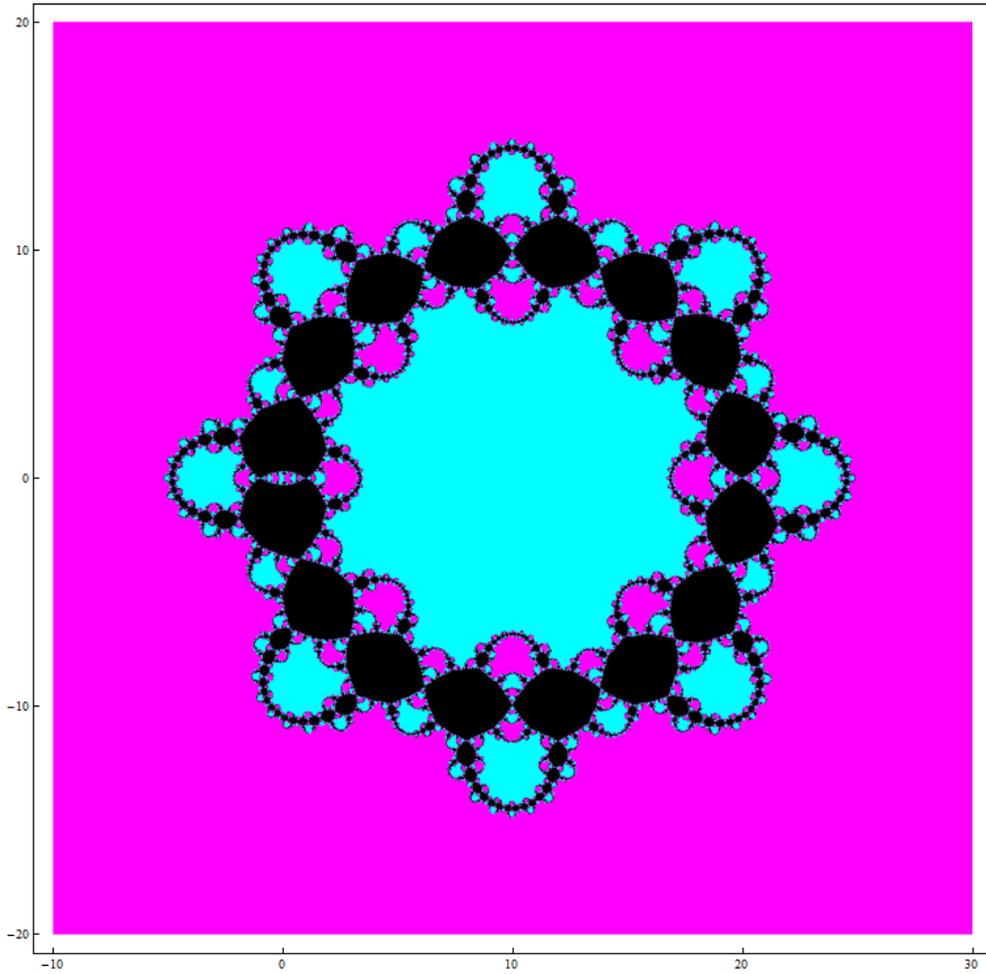


Figura 8.8: Cuencas de atracción asociadas al método con $\alpha = 11$.

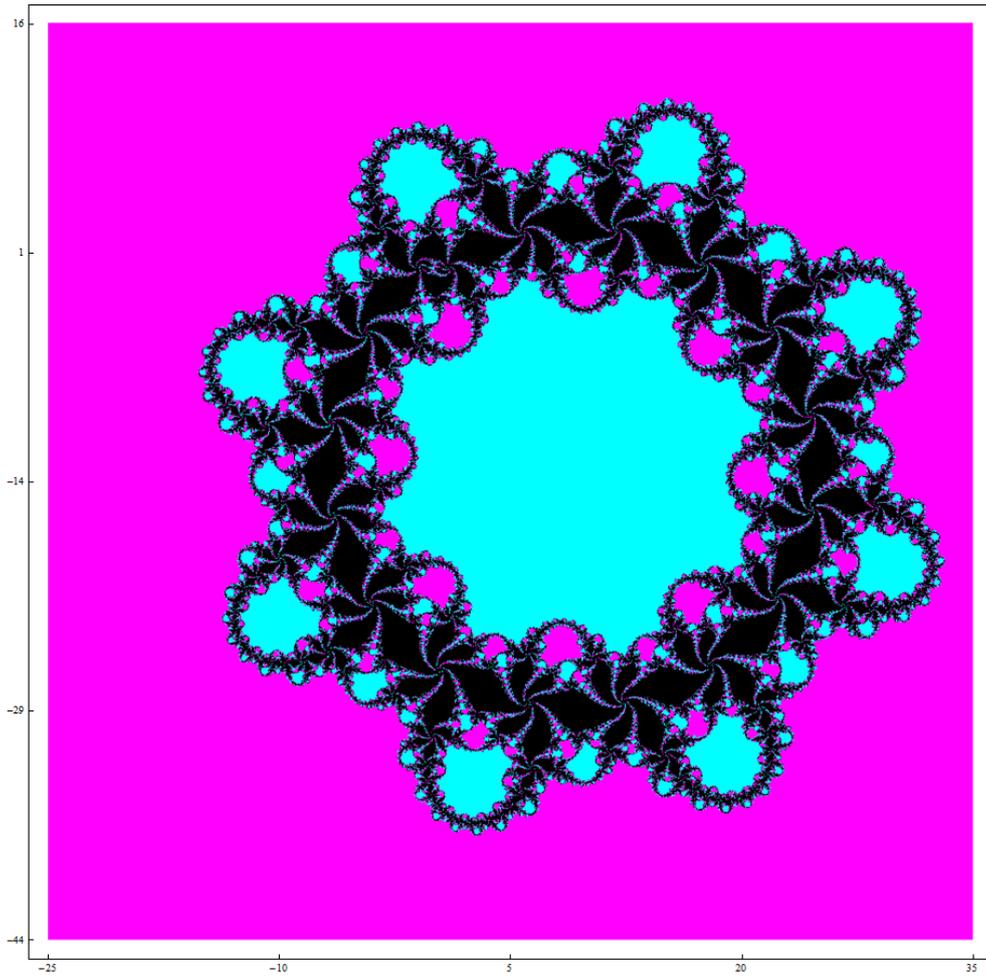


Figura 8.9: Cuencas de atracción asociadas al método con $\alpha = 10 - 13i$.

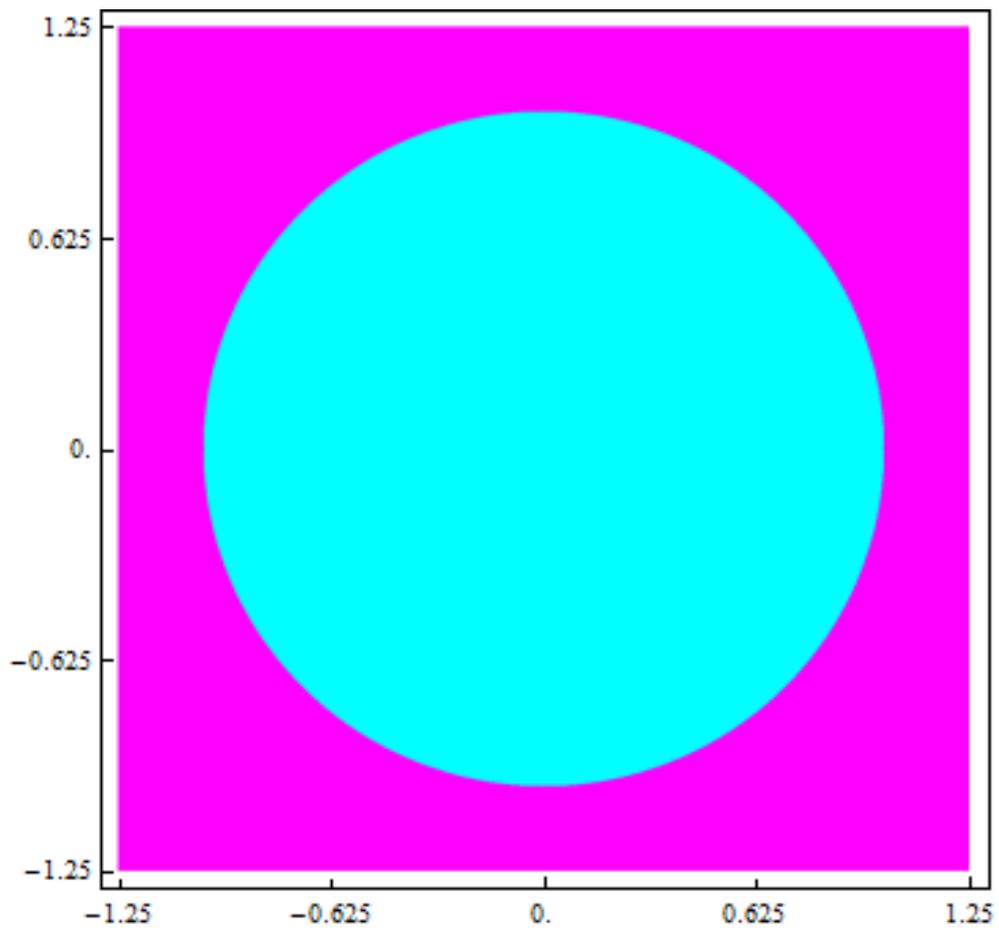


Figura 8.10: Cuencas de atracción asociadas al método con $\alpha = 0.5$.

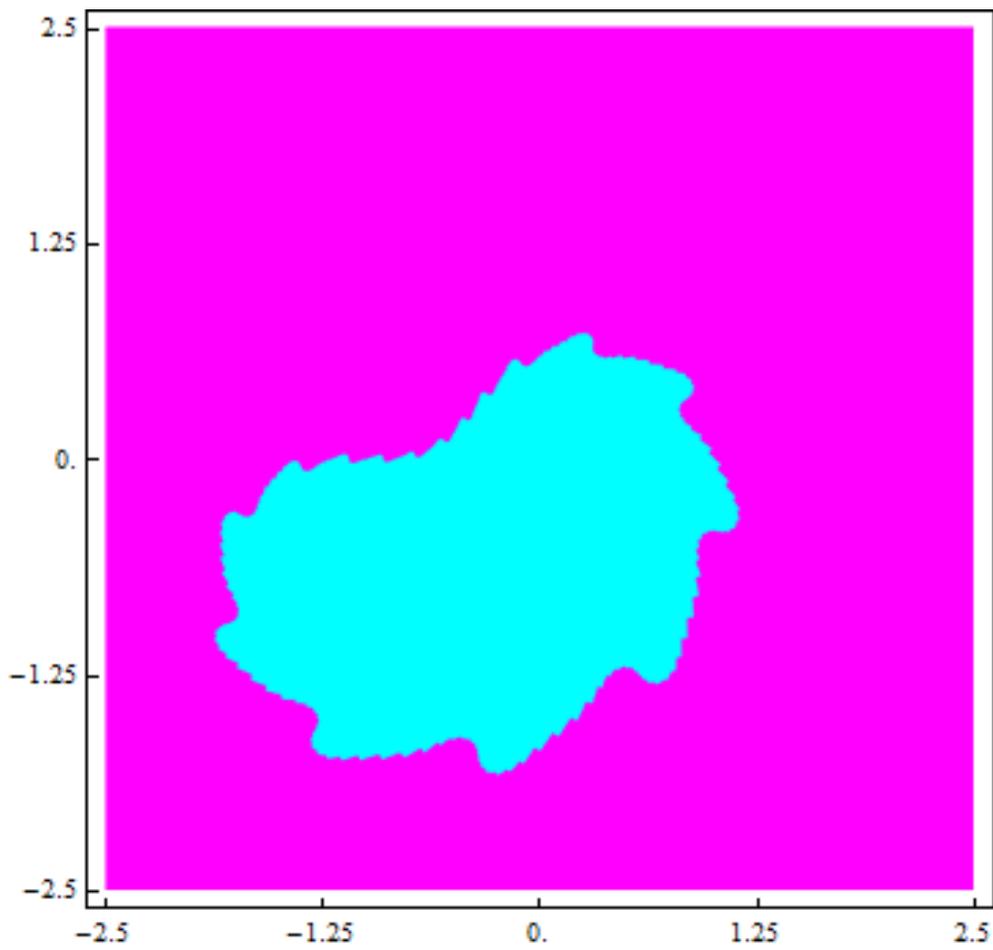


Figura 8.11: Cuencas de atracción asociadas al método con $\alpha = -0.5i$.

8.4. Aplicación

Ahora, para mostrar la aplicabilidad de nuestra teoría en un problema real, vamos a considerar la siguiente ecuación cuadrática que describe la fracción de la alimentación de nitrógeno-hidrógeno que se convierte en amoníaco, llamada conversión fraccional que se muestra en [103, 202]. Mostramos el proceso del amonio en la Figura 8.12.

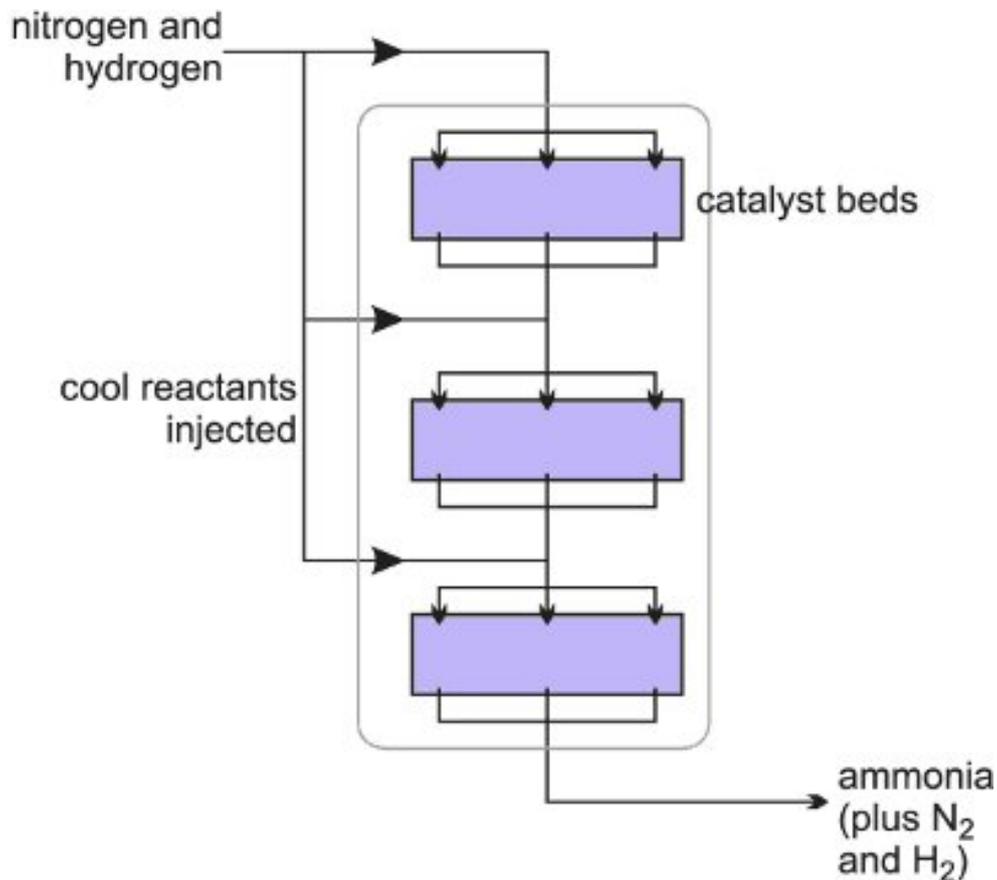


Figura 8.12: Obtenido de www.essentialchemicalindustry.org/chemicals/ammonia.html. Proceso del Amonio

Para 250 atm y $500^{\circ}C$, esta ecuación toma la forma:

$$f(x) = x^4 - 7.79075x^3 + 14.7445x^2 + 2.511x - 1.674$$

Consideremos $S = \mathbb{R}$, $D = [0, 1]$, $\xi = 0$. Definimos la función F en D . Entonces, escogiendo las funciones $\psi_i(t)$ para $i = 1, 2, 3$ y M como en el primer apartado de la Nota 8.2 y $\alpha = 1.025$ obtenemos que

$$L_0 = 2.59403$$

$$L = 3.28225$$

y

$$M = 1.44197.$$

Entonces, por la definición de las funciones “g” obtenemos

$$r_0 = 0.227607 \dots$$

$$R_A = 0.236119 \dots$$

$$r_1 = 0.0820633 \dots,$$

$$r_2 = 0.155458 \dots,$$

$$r_3 = 0.24533 \dots$$

y en consecuencia

$$r = r_1 = 0.0820633 \dots$$

Entonces podemos asegurar la convergencia del método 8.1.2 con $\alpha = 1.025$ debido al Teorema 8.1.

Capítulo 9

Herramienta Pedagógica basada en el desarrollo de una aplicación informática para la mejora del aprendizaje en matemática avanzada

Resumen

El estudio dinámico de los métodos iterativos ha aumentado en las últimas décadas debido al desarrollo de los computadores, aspecto por el cual se ha visto la necesidad de incluir la enseñanza de estos métodos en los planes de estudio. En la actualidad hay varios tipos de software cuya aplicación didáctica en las aulas es de gran utilidad, pero no se han diseñado atendiendo a las dificultades que los alumnos presentan en relación al aprendizaje de la dinámica de los métodos iterativos. Cabe, asimismo, destacar que no existe un software diseñado exclusivamente para la enseñanza de métodos iterativos y este hecho, junto con las dificultades encontradas en los alumnos en esta temática, han llevado a que muchos alumnos no entiendan los conceptos fundamentales ya que se trata de una materia que tiene un alto componente visual. Teniendo en cuenta todos los anteriores factores, se ha diseñado un software que sirve para ayudar a los alumnos en la comprensión de esta materia permitiendo a los profesores realizar simulaciones en el aula y a la vez evitar que los alumnos puedan utilizar los planos de parámetros o los planos dinámicos, que no correspondan en cada situación a la que se enfrenten. El presente artículo aborda el desarrollo de una propuesta metodológica en la que se emplea el software diseñado en una muestra de 30 alumnos de la asignatura Sistemas Dinámicos Discretos y Continuos del Máster en Ingeniería Matemática y Computación de La Universidad Internacional de Rioja.

9.1. Introducción

A mediados de los años ochenta surge en España la necesidad de conocer los motivos de la problemática en el aprendizaje de las Matemáticas en términos de adquisición de competencias y habilidades, intentado dejar atrás el hecho de valorar únicamente los aspectos meramente cognitivos (Azcárate y Camacho, 2003). Centrándonos principalmente en el pensamiento matemático avanzado, relacionado con contenidos matemáticos universitarios, a fin de poder desarrollar metodologías de enseñanza y aprendizaje que fomenten el aprendizaje significativo en los alumnos, las investigaciones se centran en la inclusión de procesos de definición, prueba y demostración de forma que se puedan establecer modelos para comprender los procesos cognitivos de los estudiantes.

Algunos de los problemas que se observaron en relación al pensamiento matemático avanzado tienen que ver con el papel que pueden jugar los conceptos en los propios procesos u objetos matemáticos. Por tal motivo se distinguen dos tipos de concepciones para un mismo concepto matemático, las operacionales y las estructurales (Sfard, 1991). Mientras las primeras tratan los conceptos desde una perspectiva dinámica, como algoritmos, las segundas los tratan desde un punto de vista estático, como objetos abstractos.

Vinner (1991) comenta que otro de los aspectos sobre los que se observan problemas en los alumnos para adquirir un pensamiento matemático avanzado tiene que ver con las definiciones ya que generan un conflicto entre las concepciones estructurales concebidas por los matemáticos y los procesos cognitivos que usan los alumnos a la hora de adquirir los conceptos al tenderse a pensar que son tales definiciones las que permiten que los alumnos sean capaces de resolver los problemas. Pero tal y como comentan Azcárate y Camacho (2003) las definiciones juegan un papel muy importante cuando los alumnos tienen que realizar sus tareas cognitivas para desarrollar los esquemas conceptuales. Por tales motivos, es necesario educar a los alumnos en el uso de estas definiciones para la elaboración de los esquemas conceptuales, empleando para ello situaciones didácticas adecuadas.

Tal y como comenta Rodríguez-Vásquez (2010), en relación al pensamiento matemático avanzado, las investigaciones se centran principalmente en el estudio de la relevancia que los procesos cognitivos implicados en el aprendizaje de las matemáticas van tomando a medida que avanzamos a cursos posteriores, en el estudio epistemológico sobre los conceptos fundamentales de análisis y en el papel de las herramientas tecnológicas en el aprendizaje de estos conceptos.

En relación a la construcción de conceptos matemáticos, desde el punto de vista del constructivismo, resulta necesaria la interacción del sujeto con todas las representaciones posibles del objeto. En este sentido, desde el punto de vista epistemológico, la visualización es fundamental para el aprendizaje de las matemáticas ya que se basa en una técnica que permite a los alumnos generar ideas para llegar a un fin que es el aprendizaje de los conceptos, por lo que no ha de

ser solo centrada desde se aceptación de ver u observar. Rodríguez-Vásquez (2003) propone una clasificación de la visualización en base a lo que supone en relación a la adquisición del pensamiento matemático avanzado. En ella habla de:

- La visualización como medio que sirve de enlace entre la intuición y el razonamiento.
- La visualización como la capacidad de articular las distintas representaciones de un objeto para obtener imágenes mentales de él y por lo tanto darle significación.
- La visualización como la acción que lleva a cabo el ser humano para conectar las diferentes representaciones del objeto.
- La visualización como proceso mental que nos ayuda a representar, transformar, etc. la información visual.

En este contexto, las aplicaciones tecnológicas suponen una gran ventaja a fin de que los alumnos sean capaces de aprender los conceptos a partir de la visualización, siempre y cuando sean usadas de forma adecuada, siempre como un medio de enseñanza que favorezca el aprendizaje significativo a partir de metodologías novedosas y no se vea como un fin en sí mismo, meros dispositivos técnicos empleados en el aula.

Tal y como comentan Martins, Fracchia, Allan y Parra (2010), el uso de las nuevas tecnologías en las aulas necesita que se aborde el diseño de nuevas actividades de aprendizaje a fin de que los alumnos desarrollen capacidades que les permitan afrontar las situaciones que se presentan en la sociedad. Por tal motivo, resulta crucial el desarrollo de escenarios que favorezcan el procesamiento y la modelización de la información.

Estudios como el realizado por De Faria (2001) concluyen que los dispositivos tecnológicos aplicados a la enseñanza de las Matemáticas son un factor que determina el alcance y la limitación que puede haber a la hora de tratar un concepto matemático. Los estudios de Castro (2001) sobre el uso de calculadoras gráficas en las clases de matemáticas concluyen que fomenta en los alumnos la capacidad de análisis, de abstracción y de desarrollo del pensamiento lógico.

En la actualidad es muy común el uso de softwares, en particular los CAS (Computer Algebra Systems) los cuáles son de gran utilidad en las aulas ya que permiten al alumno explorar los conceptos matemáticos fomentando la toma de decisiones en el control de las estrategias a desarrollar para resolver los problemas relacionados con esos conceptos.

Se puede decir, por lo tanto, que las herramientas tecnológicas ayudan a la tarea de visualización y por lo tanto a conectar contenidos y significados. Este aspecto es crucial, tal y como comenta Santos (2003), ya que permite que los alumnos establezcan conexiones de los conceptos matemáticos no solo entre ellos,

si no en relación con otras áreas. Tal y como comenta Santos (2003), el uso de softwares dinámicos en las aulas establece grandes oportunidades didácticas para los alumnos, permitiéndoles explorar concepciones matemáticas a través de la interacción con las construcciones.

El modelado y la simulación de sistemas resultan muy adecuados ya que permiten la creación de ambientes virtuales que imitan el comportamiento de cualquier tipo de sistema. En la actualidad hay varios tipos de softwares cuya aplicación didáctica en las aulas es de gran utilidad. El uso de estos softwares es de gran utilidad para los alumnos ya que les permiten interpretar métodos numéricos y adaptarlos a otras situaciones que podrían resultarles de utilidad en su futuro profesional, ya que les confieren la capacidad de afianzar sus conocimientos en programación (Ascheri y Pizarro, 2006). En este sentido, el rol de docente juega un papel muy importante ya que hay que fomentar que los alumnos vean la tecnología como una herramienta a partir de la cual puedan ampliar sus capacidades cognitivas, lo que implica, hacer un uso reflexivo de la misma (Hitt, 2003). Algunos de estos softwares son: SAGE (empleado en álgebra, cálculo, teoría de grupos, criptografía, etc.), Genius (permite trabajar muchos conceptos y su lenguaje está diseñado para parecerse a la sintaxis matemática normal) o Scilab (para simulaciones matemáticas, visualizaciones 2D y 3D, optimización, diseño de sistemas control, procesamiento de señales, etc.).

Son muchas las situaciones prácticas en las que se ven involucradas soluciones a través de métodos numéricos de ecuaciones no lineales (Amat, Busquier y Legaz, 2015; Arís y Orcos, 2015; Jiménez, Mediavilla, Portús, López y San Vicente, 2015). Por tal motivo, el trabajo de esta investigación está relacionado con el concepto de ecuaciones no lineales y el empleo de métodos iterativos para la resolución de las mismas.

Los métodos iterativos fueron desarrollados por los griegos, babilonios y árabes, pero no fue hasta el Siglo XVII cuando se desarrollaron en Europa. Importantes investigadores de esa época como Leonardo da Vinci, Galileo, Descartes, Newton y Leibniz mostraron interés por el estudio de las ecuaciones diferenciales, término acuñado por Leibniz, para lo cual consideraron necesario el desarrollo de nuevos métodos.

El motivo de este estudio es debido a las dificultades cognitivas que se observan en los alumnos universitarios en relación a este concepto y que es tan necesario dominar en el ámbito tanto de la matemática aplicada como de la ingeniería, ya que la inmensa mayoría de problemas a los que van a tener que enfrentarse serán resueltos de forma numérica mediante procesos iterativos.

9.1.1. Conceptos básicos de métodos iterativos

Dentro de la Matemática Aplicada existen diferentes ramas, pero una de las más estudiadas es la de encontrar las soluciones de una ecuación o sistema de ecuaciones no lineales. En este contexto, aparecen los métodos iterativos, ya

que son unos métodos que, partiendo de un punto inicial y bajo unas ciertas condiciones, van a generar una sucesión que va a converger a la solución buscada. A partir del uso de métodos iterativos para resolver ecuaciones no lineales se puede aplicar el concepto de deflación para la obtención de todas las raíces de un polinomio aplicando la división sintética del método de Horner. Concretamente, cuando se trata de ecuaciones no lineales, los métodos de Newton, de la Secante, de Raphson o de Müller solo permiten obtener una raíz a la vez.

Queda de manifiesto que es interesante observar que la técnica de la función reducida puede utilizarse para cualquier tipo de función, sin importar si es una función polinómica o no, permitiendo hallar todas y cada una de las raíces de las ecuaciones no lineales. Uno de los problemas que aparecen en el proceso es la determinación de raíces múltiples, así como las indeterminaciones o cuando las funciones presentan discontinuidades. A pesar de que este razonamiento resulta trivial, no se ha encontrado en la revisión bibliográfica mencionado este caso. En este sentido, los autores Curtis Gerald y Patrick Wheatley ya escribieron sobre el potencial de aplicar este concepto, pero, sin embargo, no profundizaron en la puesta en práctica de la misma. En el caso que aquí nos atañe, sí que se ha hecho hincapié en ello y se ha enseñado a los alumnos, cómo funciona el algoritmo, así como todas las hipótesis que se deben cumplir para llegar a ponerlo en práctica. De hecho, una de las tareas propias de la materia es la elaboración de algoritmos eficientes que sirvan en la búsqueda de soluciones de ecuaciones no lineales. Existen muchos métodos iterativos famosos, como, por ejemplo:

- Método de Newton

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

- Método de Halley

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

- Método de Chebyshev

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

Como vemos estos métodos iterativos llevan el nombre de sus descubridores. Durante las últimas décadas los esfuerzos de los investigadores en el área se han centrado en el diseño de nuevos métodos con diferentes características (órdenes de convergencia altos, libres de derivadas, etc.) y en suavizar las condiciones que garantizan la convergencia de la sucesión generada a la solución buscada.

Por otro lado, en los últimos años y debido a la irrupción de los ordenadores, se ha comenzado a estudiar el comportamiento dinámico de estos métodos iterativos, ya que, si estudiamos dicho comportamiento, podemos conocer de primera

mano qué va a suceder cuando lo apliquemos a diferentes funciones de las cuales queremos calcular sus raíces.

Si nos centramos en el método de Newton, debido a su relevancia y sus buenas características, podemos decir que el estudio de la dinámica de este método en el plano complejo tiene una importancia histórica desde que, primero E. Schöder en 1870 ([206]) y, 9 años más tarde A. Cayley propusieran utilizar dicho método para resolver ecuaciones definidas en el plano complejo.

En este sentido, uno de los problemas más famosos es el conocido como problema de Cayley que consiste en estudiar las cuencas de atracción (es decir, aquellos puntos que al iterar el método convergen a la solución) de cada una de las raíces del polinomio complejo al cual aplicamos el método de Newton, ver [149] para entender más el problema. En un principio para polinomios de la forma

$$p(z) = z^2 - a,$$

siendo a un número complejo cualquiera, el problema resultó ser sencillo, ya que las cuencas de atracción se corresponden con cada uno de los semiplanos separados por la mediatriz del segmento que une ambas raíces. Sin embargo, cuando dio el paso a polinomios de la forma

$$p(z) = z^3 - a,$$

se encontró con que no era capaz de resolverlo. Hoy en día y con la ayuda de los potentes ordenadores podemos entender la razón por la que no fue capaz. En las Figuras 8.12–8.9, vemos la diferencia existente entre la caracterización de cada zona para polinomios de grado 2 y 3.

Es fácil ver cómo en la parte derecha, sin la presencia de ordenadores, sería imposible caracterizar cada zona, más aún sabiendo que su comportamiento es fractal. En 1977, el famoso matemático, Benoit Mandelbrot en su libro *The fractal nature of geometry* ([152]) presentó numerosas aplicaciones de este tipo de estructuras en la naturaleza. Además, en 1986, ([153]) dio una definición intuitiva de lo que significaba para él el término conjunto fractal: un conjunto en que las partes son similares al total, en algún sentido. De forma coloquial, entendemos que un objeto es fractal si cumple, alguna de las siguientes condiciones (hemos omitido las más técnicas):

- Tiene detalles a todas las escalas.
- Es autosemejante.

Si hiciéramos zoom en las las intersecciones que se producen en la Figura 2, veríamos cómo efectivamente vuelven a aparecer las mismas estructuras y, es por esta razón por la que sin ordenador es imposible el caracterizar las cuencas de atracción de un polinomio cúbico. Por otro lado, los matemáticos están centrandó esfuerzos en el diseño de familias de métodos iterativos, que dependan de un

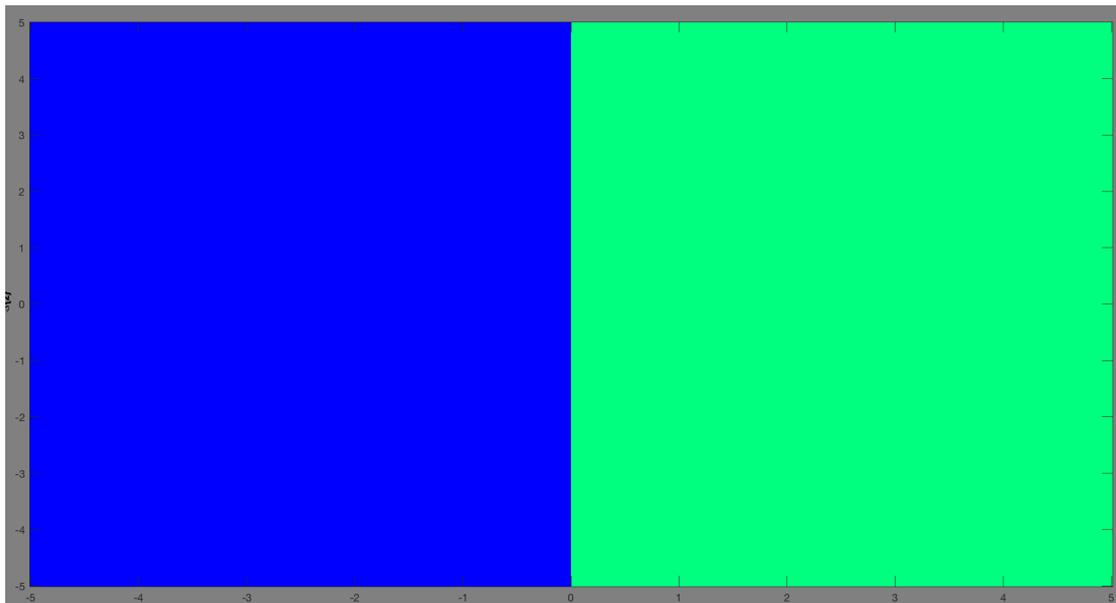


Figura 9.1: Caracterización de las cuencas de atracción asociadas al método de Newton aplicado al polinomio $p(z) = z^2 - 1$.

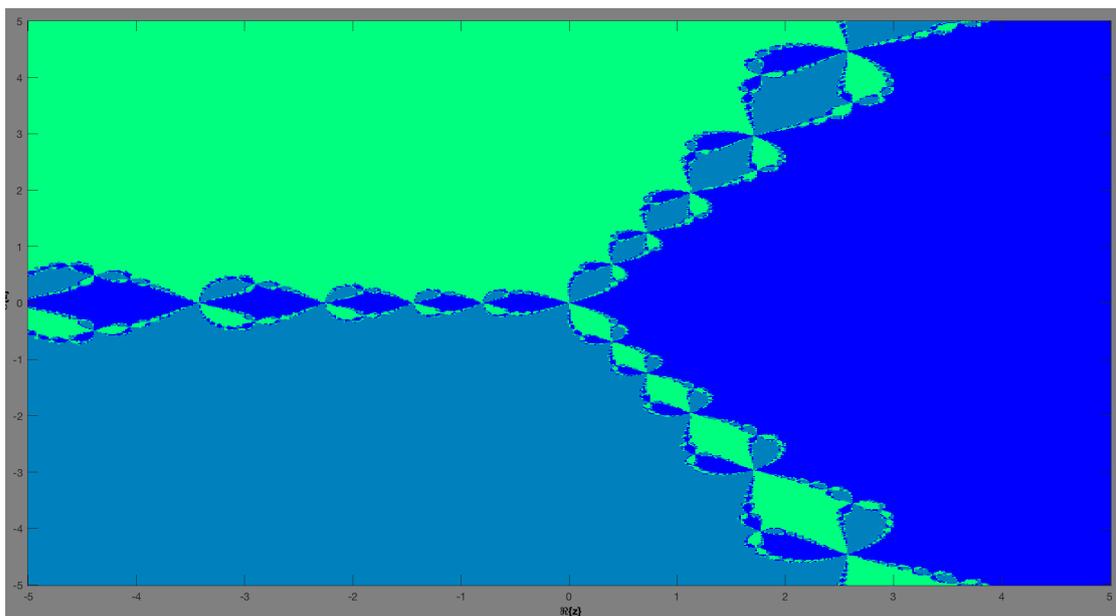


Figura 9.2: Caracterización de las cuencas de atracción asociadas al método de Newton aplicado al polinomio $p(z) = z^3 - 1$.

(o más) parámetro, ya que estas familias van a permitir tener una infinidad de nuevos métodos iterativos o incluso observar el comportamiento de varios métodos a la vez. Algunas de los métodos o familias uniparamétricos más conocidos y que dependen de un parámetro son :

- Método de Newton amortiguado

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

- Familia de Chebyshev-Halley

$$x(n+1) = x_n - \frac{(L_f(x_n))}{(2 * (1 - \alpha L_f(x_n)))} \frac{f(x_n)}{f'(x_n)}$$

Es evidente que el método de Newton queda englobado dentro del método de Newton amortiguado si $\lambda = 1$ y la familia de Chebyshev-Halley engloba a Halley ($\alpha = 1/2$), Chebyshev ($\alpha = 0$) y Newton Halley ($\alpha \rightarrow \pm\infty$).

Para proceder al estudio dinámico de este tipo de familias, lo primero que se hace es el estudio de los planos de parámetros (ver Figura 3) para determinar qué valores del parámetro, van a tener un comportamiento dinámico que nos interese y cuales no, para obtener una información más detallada de cada uno de los conceptos involucrados ver [149]. Una vez detectados los valores del parámetro para los que existen anomalías, dibujaremos las cuencas de atracción (también conocido su gráfico como plano dinámico) para corroborar dichas anomalías y caracterizarlas. Por supuesto, la situación ideal sería que nos encontráramos con un plano de parámetros sin anomalías, ya que en tal caso tendríamos una familia de métodos iterativos que se comporta, en términos de convergencia, de forma perfecta para cada miembro.

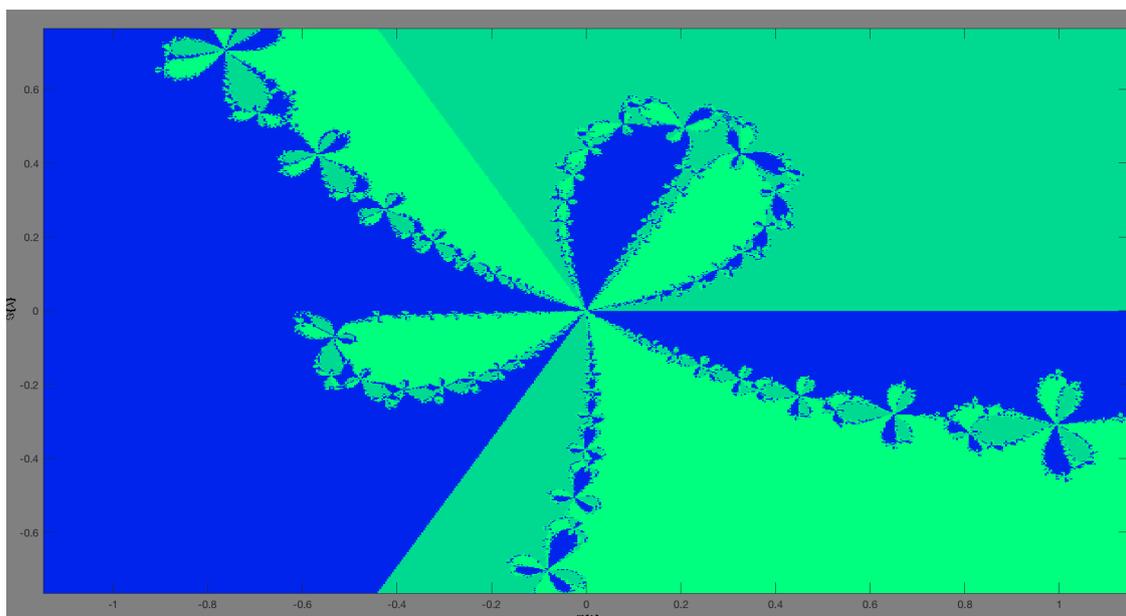


Figura 9.3: Plano de parámetros en el que existen diferentes anomalías.

9.2. Dificultades encontradas

Para este estudio hemos seleccionado un total de 34 alumnos de la asignatura Sistemas Dinámicos Discretos y Continuos del Máster Universitario en Ingeniería Matemática y Computación impartido en la Universidad Internacional de La Rioja. Estos alumnos se corresponden con las dos primeras promociones que han accedido a dicha asignatura y es la primera vez que se enfrentan al estudio dinámico de métodos y familias de métodos iterativos. En este sentido se ha llevado a cabo una investigación tanto cuantitativa como cualitativa, en la que se han llevado tanto entrevistas personales con los alumnos, como teniendo en cuenta sus resultados en los trabajos entregados. En la Tabla 9.1, se puede observar el porcentaje de alumnos que ha presentado o no los trabajos vinculados con el estudio dinámico de métodos iterativos.

A nivel general se puede decir que los alumnos presentan dificultades a la hora de distinguir qué tipo de gráfico, de plano dinámico o paramétrico, tienen que usar. Los planos dinámicos se obtienen a partir de la aplicación de métodos simples mientras que los métodos paramétricos requieren que se fije el valor del parámetro mediante el uso del plano paramétrico para poder, una vez fijado el valor, dibujar el plano dinámico. De esta manera, cuando hay que trabajar con familias de métodos paramétricas, primero hay que ir al plano de parámetros para ver qué valores del parámetro son buenos, en términos de convergencia y una vez que este esté seleccionado, se dibuja el plano dinámico para comprobar que el método funciona correctamente. La confusión entre los planos de parámetros con los planos dinámicos conlleva a que los alumnos no puedan continuar con el estudio de la familia de métodos, por tal motivo, se hace necesaria la aplicación de vías alternativas que permitan que los alumnos palien esta confusión y logren un aprendizaje significativo de estos conceptos matemáticos. Este estudio se basa en la aplicación de un software que permite obtener el plano que corresponda en base a la función que se esté estudiando, ya sea paramétrico o dinámica. De esta manera no se podrá obtener nunca un plano paramétrico a partir de la aplicación de un método dinámico y viceversa. Esto conlleva a que los alumnos comprendan la fundamentación de los métodos iterativos y sepan aplicarlos ya que, a partir del trabajo autónomo con el software, serán ellos los protagonistas activos de su proceso de aprendizaje, aprenderán a su ritmo y de forma autónoma comprendiendo en todo momento la aplicabilidad de los mismos.

Tal y como se puede observar en la Tabla 9.2, el número de alumnos que no han entregado los trabajos (NP, no presentado) es muy elevado, lo que lleva a pensar que no entienden la fundamentación de los métodos iterativos ni saben aplicarlos. De hecho, podemos observar cómo un tercio de los alumnos no ha llegado a entregar ninguno de los dos trabajos.

Como resumen de los datos, obtenemos que el porcentaje total de trabajos entregados es inferior al 50 % en todos los casos, y salvo, el segundo trabajo que lo entrega un 50 %, el resto no llega al 61.76 %, por lo tanto, es evidente que

Cuadro 9.1: Resultados del primer grupo

Alumno	Trabajo 1	Trabajo 2	Trabajo 3
1	NP	9.5	5
2	9	NP	NP
3	0	NP	NP
4	0	8	7
5	0	NP	NP
6	NP	5	NP
7	0	8.5	3
8	NP	NP	NP
9	6	10	NP
10	NP	NP	NP
11	NP	NP	NP
12	9	9	NP
13	2	8	10
14	NP	NP	NP
15	10	NP	NP
16	5	0	0
17	8	0	0
18	NP	0	0
19	NP	NP	NP
20	5	9.5	7
21	NP	8	8
22	NP	5	NP
23	0	NP	NP
24	4	6	NP
25	0	NP	NP
26	10	0	0
27	NP	NP	NP
28	NP	0	0
29	NP	NP	5
30	NP	NP	NP
31	NP	NP	NP
32	0	0	0
33	NP	NP	NP
34	NP	NP	NP

existen problemas vinculados con la materia.

Cuadro 9.2: Resultados del primer grupo

Columna1	Trabajo 1	Trabajo 2	Trabajo 3
NP	17	17	21
Suspensos	9	6	7
Aprobados	8	11	6
Total	34	34	34
Total sin NP	17	17	13
% presentados con NP	50 %	50 %	61.76 %
% aprobados con NP	23.53 %	32.35 %	17.65 %
% Aprobados sin NP	47.06 %	64.71 %	46.15 %
Nota Media con NP	2.00	2.54	1.32
Nota Media sin NP	4.00	5.09	3.46

9.3. Trabajos

Tal y como se puede observar en la Tabla 9.2, el porcentaje de aprobados es mayor en el segundo trabajo, mientras que el porcentaje de presentados es superior en el tercer trabajo. Por otro lado, vemos cómo la nota media del primer y tercer trabajo es menor de 5 puntos, lo cual deja patente que las dificultades que presentan los alumnos son importantes y deben ser tenidas en cuenta, mientras que en el segundo trabajo la nota media apenas llega al 5, por lo que, también son videntes los problemas.

A continuación, vamos a pasar a describir los trabajos en los que se detectaron los problemas y donde los alumnos tienen las dificultades que han sustentado este estudio y la puesta en marcha del software.

Primer trabajo

En el primer trabajo, el objetivo fundamental es que comparen la dinámica compleja de un método iterativo comparado con el método de Newton y se les facilitan las pautas de dicha comparativa, al aplicarlo a un polinomio genérico de grado dos con un parámetro. Las principales dificultades existentes para los alumnos han sido problemas a la hora de calcular analíticamente los puntos fijos y su dinámica y los puntos críticos libres y sobretodo problemas con los gráficos. Por otro lado, ha habido incluso alumnos que no han sabido calcular el índice de eficiencia ni comprobar si el método es óptimo respecto a la conjetura de Kung-Traub.

Como se puede observar en la Tabla 9.2, sólo se han presentado 17 trabajos de 34 alumnos, y de ellos sólo 9 han aprobado, es decir, poco más del 25 % de los alumnos, por lo que es evidente que existen grandes dificultades en la resolución de los mismos ejercicios. De hecho, en un primer momento, se les

Trabajo: 1

El método X es un método basado en Newton, en el que la derivada se sustituye por las diferencias finitas centrales, de convergencia cuadrática. El operador de punto fijo del método es:

$$X(z) = z - \frac{2p^2(z)}{p(z + p(z)) - p(z - p(z))}.$$

El objetivo de este ejercicio es la descripción de las características de este método iterativo, al aplicarle polinomios de grado dos del tipo $p(z) = z^2 + \lambda$. Para ello, es necesario que en la descripción se incluyan los siguientes ítems:

- » Características cuantificadas del método:
 - Índice de eficiencia.
 - Método óptimo según conjetura de Kung-Traub.

- » Características analíticas:
 - Puntos fijos.
 - Dinámica de los puntos fijos.
 - Puntos críticos libres.

- » Características gráficas del método:
 - Cuencas de atracción inmersas e implicaciones.
 - Simetrías.
 - Estabilidad (región de convergencia).

- » Comparativa con el método de Newton.

Figura 9.4: Primer Trabajo.

debió ampliar el plazo y permitir el reenvío de trabajos, ya que presentaban dificultades muy grandes en la elaboración del mismo. A continuación, pasamos a enumerar algunas de dichas dificultades:

- ¿Qué valor le doy al parámetro?
- ¿Cómo calculo los puntos fijos se hay un parámetro?
- ¿Por qué no pide que calculemos los planos de parámetros?
- ¿Qué significa cuencas de atracción?

Como puede observarse, todas estas dificultades están vinculadas con el hecho de tener un polinomio con un parámetro y de no saber cómo tratarlo.

Segundo trabajo

En el segundo trabajo, el objetivo fundamental es que comparen la dinámica compleja de un método iterativo que depende de un parámetro, por lo que, como se ha visto con anterioridad, primero deben estudiar los planos de parámetros para determinar los valores de α para los que el método presenta o no problemas de convergencia, para pasar después a estudiar los planos dinámicos y detectar, en su caso, dichos problemas.

Como se puede observar en la Tabla 9.2, sólo se han presentado 17 trabajos de 34 alumnos, y de ellos sólo 11 han aprobado, es decir, menos del 33 % de los alumnos, por lo que es evidente que existen grandes dificultades en la resolución de los mismos ejercicios. A continuación, pasamos a enumerar algunas de dichas dificultades:

- No entiendo cuando debo usar el plano dinámico y cuando el plano de parámetros
- ¿No es lo mismo el plano de parámetros que el dinámico?
- ¿Qué valores del parámetro debo utilizar en los planos dinámicos?
- ¿Qué dibujo primero, el plano de parámetros o el dinámico?
- ¿Por qué no me grafica bien los planos dinámicos el software y me sale el dibujo negro?
- ¿Con qué valor del parámetro debo dibujar el plano de parámetros?
- ¿Cómo puedo saber si algo es periódico?
- ¿La tolerancia influye en algo?

Trabajo 2

Ejercicio 1

La función `orbita.sci`, cuyo flujoograma ha sido presentado en la figura 1 y su código queda reflejado en la figura 3, sale del programa cuando la diferencia entre iterados está por debajo de un umbral determinado por la variable `t` o cuando se ha alcanzado el número máximo de iteraciones determinado por `maxIter`.

En este segundo caso es posible que la órbita haya divergido o que haya convergido y se haya quedado en una órbita periódica. En este ejercicio se generará un programa `orbitaWtu.sci`, basado en `orbita.sci`, que tenga las mismas variables de entrada y de salida que `orbita.sci`, pero que cuando acabe lance un mensaje en la consola indicando el motivo por el que ha terminado. La estructura del programa será la que muestra la figura E1.

```
01 function [iter, d, x]=orbitaWtu(f, x0, t, maxIter)
02 iter=1; d=1; x=x0;
03 while iter<maxIter&&d($)>t
04     xk=fava(x($),f);
05     d=|d-abs(xk-x($))|;
06     x=Lx xk];
07     iter=iter+1;
08 end
09 if // converge
10     disp('Converge');
11 else if //diverge
12     disp('Órbita periódica');
13 else
14     disp('Diverge');
15 end
16 end
17 endfunction;
```

Figura E1. Esquema del programa `orbitaWtu.sci`

Hay que modificar los comentarios de las líneas 9 y 11 para adaptarlos al propósito del ejercicio.

Ejercicio 2

Realiza el estudio dinámico del sistema $f(x) = -\frac{3}{2}x^2 + \frac{5}{2}x + 1$. Para ello, sigue los siguientes pasos:

- Calcula de forma analítica los puntos fijos.
- Obtén el período que tiene el sistema. Para visualizarlo, obtén una tabla que represente 7 iteraciones partiendo de la semilla $x_0 = 0$, de forma que concluyas determinando el período.
- Representa con la función `verbuja.sci` el diagrama, con `casoX=tipos(1,3)` y el resto de parámetros igual que en los ejemplos del tema.

Ejercicio 3

Realiza el estudio dinámico de la familia cuadrática $f_\lambda(x) = x^2 + \lambda$. Para ello, sigue los siguientes pasos:

- Determina los puntos candidatos a ser puntos de bifurcación.
- Dibuja el diagrama de bifurcación tomando como valores del parámetro $\lambda \in [-1,4,0]$ y de la semilla $x_0 \in [-1,6,0,8]$.
- Dibuja tres diagramas de `verbuja` en los que el comportamiento sea diferente, en función del resultado obtenido en el diagrama de bifurcación comentando el resultado.

Como puede observarse, todas estas dificultades están vinculadas con el hecho de no diferenciar entre el plano de parámetros y el plano dinámico y no entender que primero deben dibujar el plano de parámetros y después, una vez dibujado, seleccionar los valores del parámetro que deben fijar para dibujar los planos dinámicos y encontrar las anomalías existentes para ese valor. Otra de las dificultades que han presentado los alumnos es el no entender qué deben exigir para obtener un comportamiento periódico y poder garantizar que algo es convergente o divergente.

Por lo tanto, estamos detectando problemas en todos y cada uno de los ejercicios de este trabajo en concreto, a pesar, de ser en el que mejor nota media se ha obtenido.

Tercer trabajo

En el tercer trabajo, el objetivo fundamental es primero que calculen los planos dinámicos de unos polinomios concretos, es decir, ellos ven un polinomio que depende de un parámetro, pero se les facilita valores del parámetro para los que deben realizar los estudios. Haciendo este ejercicio, se pone de manifiesta si entienden la diferencia entre el plano de parámetros y el plano dinámico y cuándo deben utilizar cada uno de ellos, ya que en este caso sólo deben usar los planos dinámicos. En el segundo apartado, es un estudio complejo completo, es decir, primero usando el plano de parámetros y posteriormente los planos dinámicos.

Como se puede observar en la Tabla 9.2, sólo se han presentado 13 trabajos de 34 alumnos y, de ellos, sólo han aprobado 6, por lo tanto, no llegan al 18% de los alumnos. Se observa que existen grandes dificultades en la resolución del primer apartado, ya que es en el que se demuestra realmente dominio del tema. A continuación, pasamos a enumerar algunas de dichas dificultades:

- No entiendo por qué debo usar el plano dinámico y no el plano de parámetros en el primer ejercicio ya que existe un parámetro
- Si pinto el plano dinámico, ¿qué pasa con el parámetro? ¿se queda fijo?
- No entiendo la diferencia entre ambos ejercicios. ¿no son iguales? En las dos hay un parámetro.
- ¿Qué dibujo primero, el plano de parámetros o el dinámico?
- ¿Por qué me da error al dibujar los planos de parámetros en el primer apartado?
- ¿Con qué valor del parámetro debo dibujar el plano de parámetros?
- ¿Debo aplicar el método de Newton también en el segundo apartado u otro método?

Trabajo 3

Ejercicio 1

Obtén el plano dinámico del método de Newton cuando se aplica sobre polinomios cúbicos del tipo $f(z) = z^3 + \lambda$ particularizándolos para $\lambda = \{1, i, -1, -i\}$, con $z_0 \in [-3,3] + i[-3,3]$. En la entrega deberás incluir:

- » Expresión del operador de punto fijo del método de Newton cuando se aplica sobre polinomios cúbicos de la forma $f(z) = z^3 + \lambda$. Obtención de puntos fijos.
- » El código en SciLab del programa. Solo se permitirá un programa para la ejecución de los cuatro valores de λ .
- » Los cuatro planos dinámicos generados.

Ejercicio 2

Obtén el estudio de la dinámica compleja asociada a la función logística $f(z) = \lambda z(1 - z)$ cuando se aplica sobre variable compleja. Para ello, sigue los siguientes pasos:

- » Obtención de los puntos fijos y críticos del sistema.
- » Modificación del código del programa de la figura 8 para generar el plano de parámetros. Incluir código en la memoria.
- » Representación del plano de parámetros, con $\lambda \in [-2,4] + i[-1,1]$ que tenga 201x201 puntos.
- » Representación de tantos planos dinámicos como regiones haya en el plano de parámetros. Incluir código de los planos dinámicos en la memoria.
- » Relación entre los valores de λ , la cuenca de atracción en la que está el punto crítico y los puntos fijos.

Figura 9.6: Tercer Trabajo.

Como puede observarse, de nuevo todas las dificultades están vinculadas con el hecho de no diferenciar entre el plano de parámetros y el plano dinámico y no entender que primero deben dibujar el plano de parámetros y después una vez dibujado, seleccionan los valores del parámetro que deben fijar para dibujar los planos dinámicos y encontrar las anomalías existentes para ese valor.

Como ya hemos visto, las dificultades que tienen están vinculadas con diferenciar qué tipo de herramienta gráfica deben utilizar y también con diferencias el estudio dinámico de un método, de una familia y de una función en concreto. Como el problema existente en cada trabajo es no saber bien qué herramienta utilizar en cada caso y, por ello, el objetivo del software que se va a diseñar en este estudio es ayudarles en este sentido.

9.4. Descripción herramienta

El paquete computacional fue construido en MATLAB (Matrix Laboratory), versión R2017b y consta de varios archivos entre los cuales, 22 fueron desarrollados en lenguaje M, el lenguaje de programación de MATLAB. Los 4 restantes, son los contenedores de las características básicas de las interfaces gráficas de usuario construidas (GUI por sus siglas en inglés). Estos últimos tienen información como: número de botones y/o número de ejes de gráficos, así como sus posiciones, tamaño, color, tipo de letra, etc. Dichos archivos se diferencian de los demás, por la extensión “.fig” al final del nombre del archivo.

Lo anterior implica que, dentro de los 22 archivos, 4 de ellos están relacionados con las GUI, ya que, al diseñar la apariencia de una GUI, MATLAB crea dos archivos, uno “.fig”, descrito arriba, y el otro “.m”, para definir las tareas a realizar, cada vez que el usuario oprime un botón, activa un radiobutton, introduce texto o pasa el ratón sobre una gráfica.

Los archivos relacionados con las GUI son:

- myfirst.m
- myfirst.fig
- DynamicTool.m
- DynamicTool.fig
- FPandCP.m
- FPandCP.fig
- SelectFCP.m
- SelectFCP.fig

Los 26 archivos creados deben estar contenidos en una misma carpeta, en el disco duro del computador. Una vez iniciado MATLAB R2017b, es necesario navegar, por medio de sus botones, a la carpeta contenedora de los archivos. El software arranca con la GUI myfirst, por lo tanto, el primer archivo a correr, desde la consola de MATLAB, es myfirst.m.

Una vez la GUI myfirst ha arrancado, es necesario introducir la función polinómica $P(z)$. Todos los pormenores para introducir $P(z)$ se encuentran en el ítem 1 y 2, Botón entrar Función Polinómica y Ventana de texto para introducir la Función Polinómica $P(z)$, respectivamente. En la Figura 9.4, puede

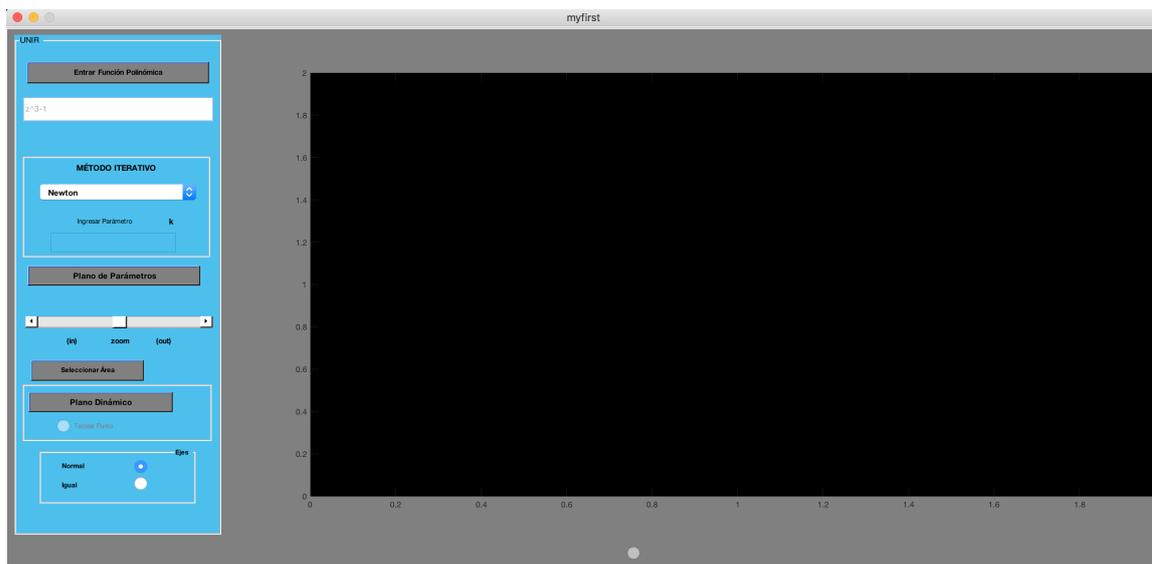


Figura 9.7: Primera pantalla de la aplicación.

Ahora, pasaremos a mostrar lo que hace cada una de las GUIs.

9.4.1. GUI myfirst.m

La GUI, myfirst.m, permite:

- Introducir una función polinómica,
- Escoger un método iterativo de entre una amplia variedad, donde están las más clásicos.
- Dibujar el plano de parámetros: esta opción llama la GUI “SelectFCP”, con la cual se escoge un punto crítico libre (PCL), necesario para dibujar el PP. Una vez dibujado, es posible:
 - Seleccionar una porción de área del gráfico y representarla en toda la pantalla,

- Realizar Zoom In o Zoom Out, en el Plano de Parámetros.
- Llamar la GUI “DynamicTool”, que dibuja el plano dinámico. Como para pintar un plano dinámico, necesitamos fijar el valor del parámetro, la aplicación permite la introducción del valor pinchando con el ratón o introducirlo mediante el teclado.
- Personalizar los valores de los ejes, para obtener un gráfico a medida.

Además, la aplicación proporciona los valores tanto de los puntos fijos como de los puntos críticos. Estos datos se consiguen llamando a la GUI “FPandCP”.

9.4.2. GUI DynamicTool

La GUI, DynamicTool, proporciona el gráfico del plano dinámico y dentro de él permite

- Hacer Zoom In y Zoom Out.
- Personalizar los valores de los ejes.
- Dibujar la órbita de un punto seleccionado sobre el plano. Dicha órbita se dibuja sobre el mismo PD.

Además, proporciona información sobre el valor del parámetro así como los valores tanto de los puntos fijos como de los puntos críticos. En la Figura 9.4.2 puede ver la GUI “DynamicTool”.

9.4.3. GUI SelectFCP

La GUI SelectFCP, proporciona los puntos críticos libres y permite elegir uno para proceder a dibujar el plano de parámetros asociado a dicho punto crítico libre. En la Figura 9.4.3 puede ver la GUI “SelectFCP”.

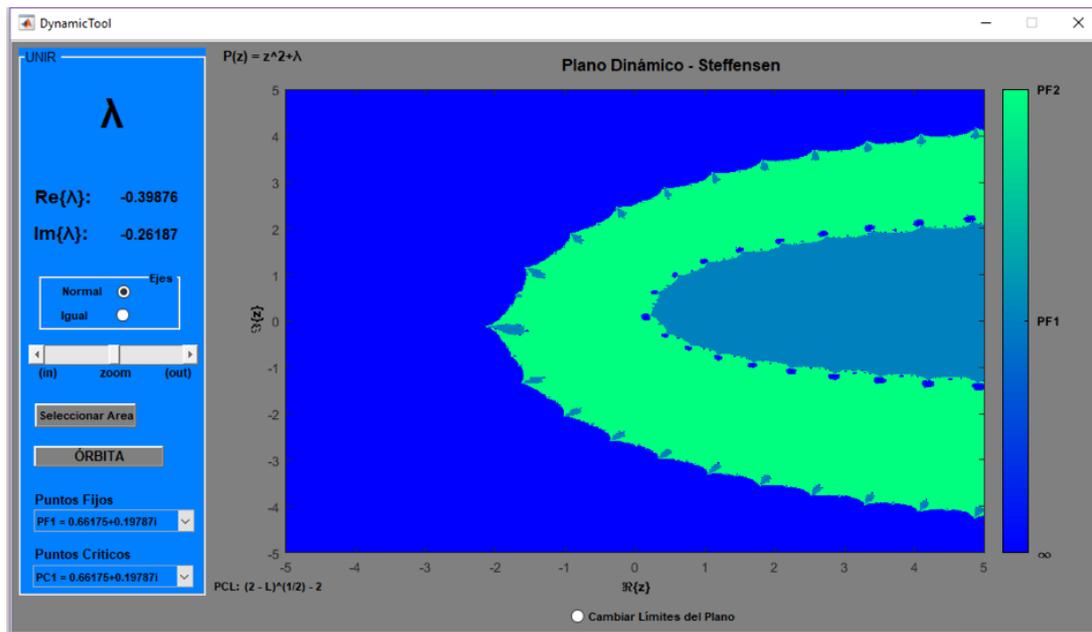


Figura 9.8: GUI DynamicTool

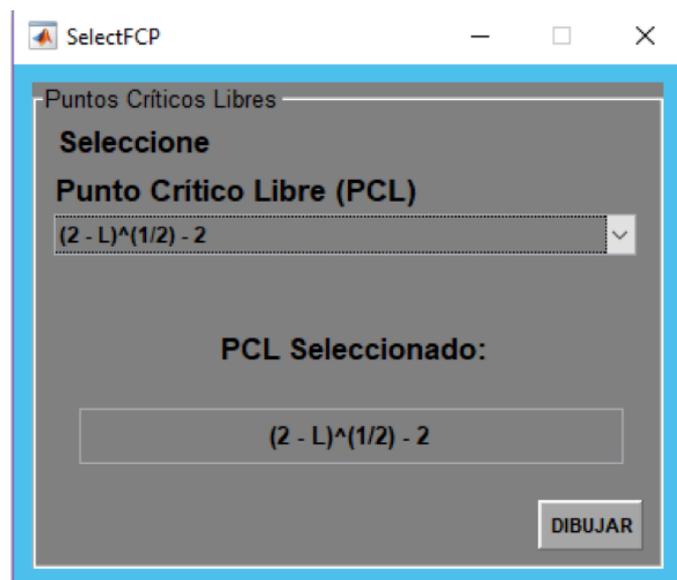


Figura 9.9: GUI SelectFCP.

9.4.4. GUI FPandCP

Por último, La GUI, FPandCP, presenta los Puntos fijos y los puntos críticos asociados al método iterativo aplicado a la función que hemos escrito.

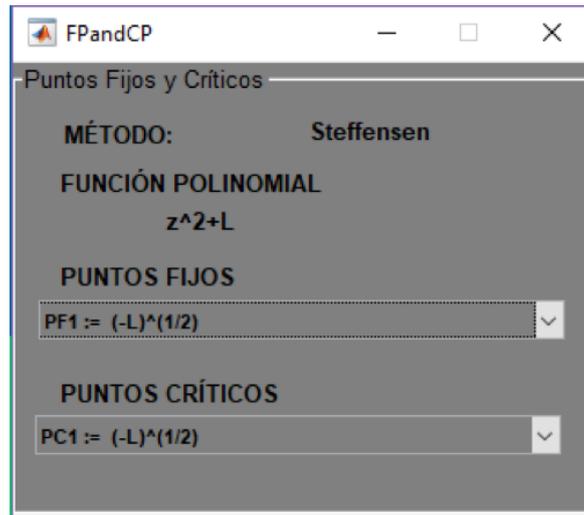


Figura 9.10: GUI FPandCP

9.4.5. Ejemplo de uso

En esta sección vamos a ver paso a paso en las Figuras 9.4.5–9.4.5, algunas de las opciones que nos da la aplicación.

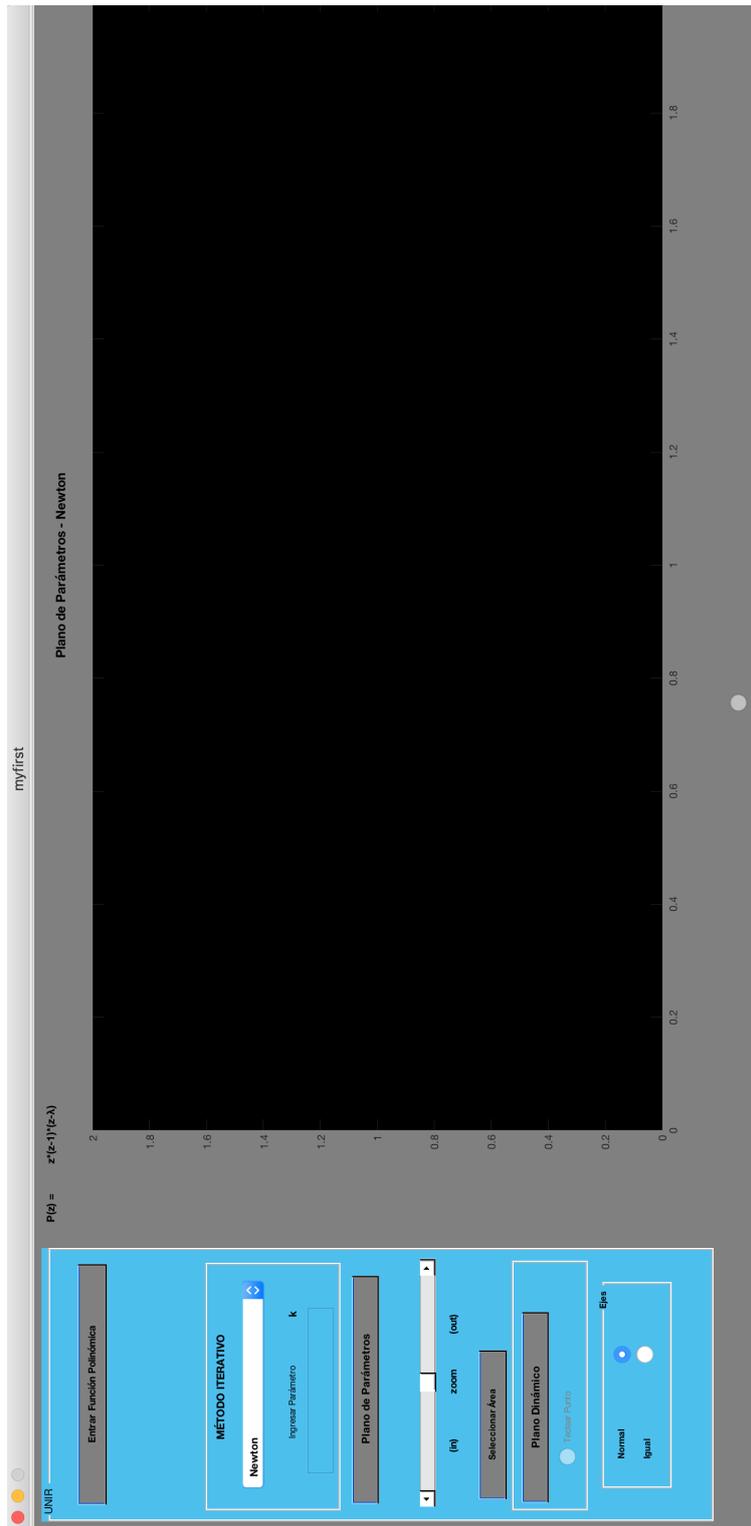


Figura 9.11: Pantalla inicial. Introducimos el polinomio y seleccionamos el método de Newton.

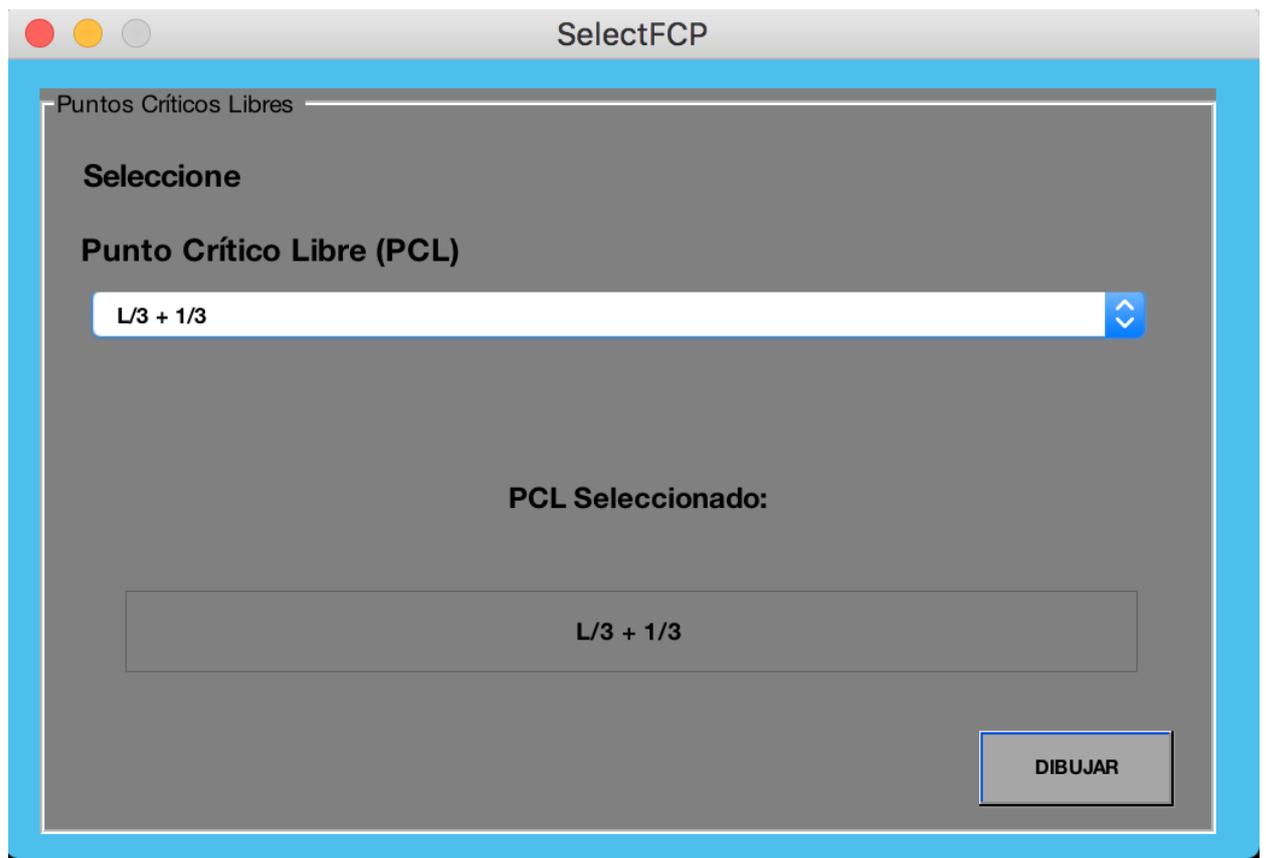


Figura 9.12: Seleccionar el punto crítico libre para pintar el plano de parámetros.

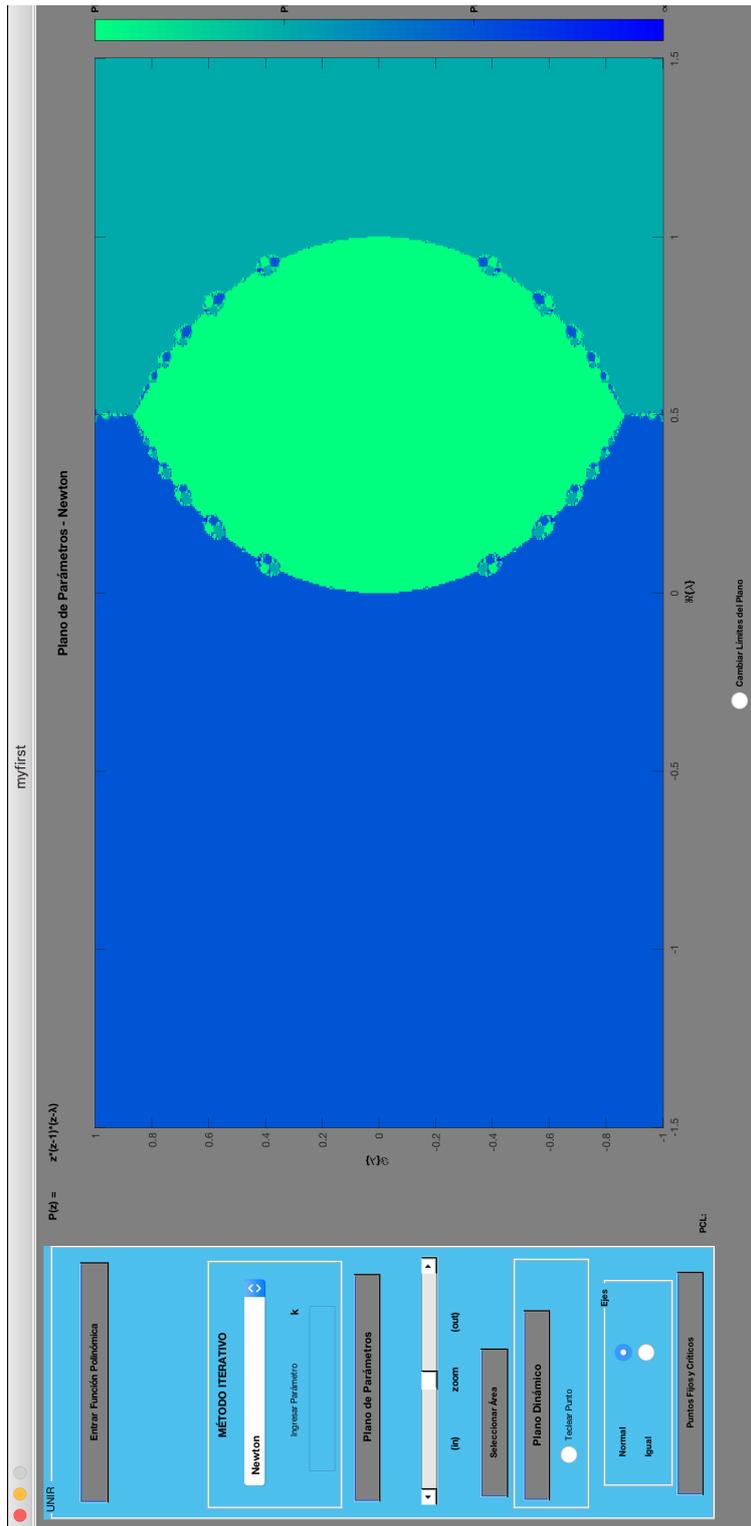


Figura 9.13: Plano de parámetros.

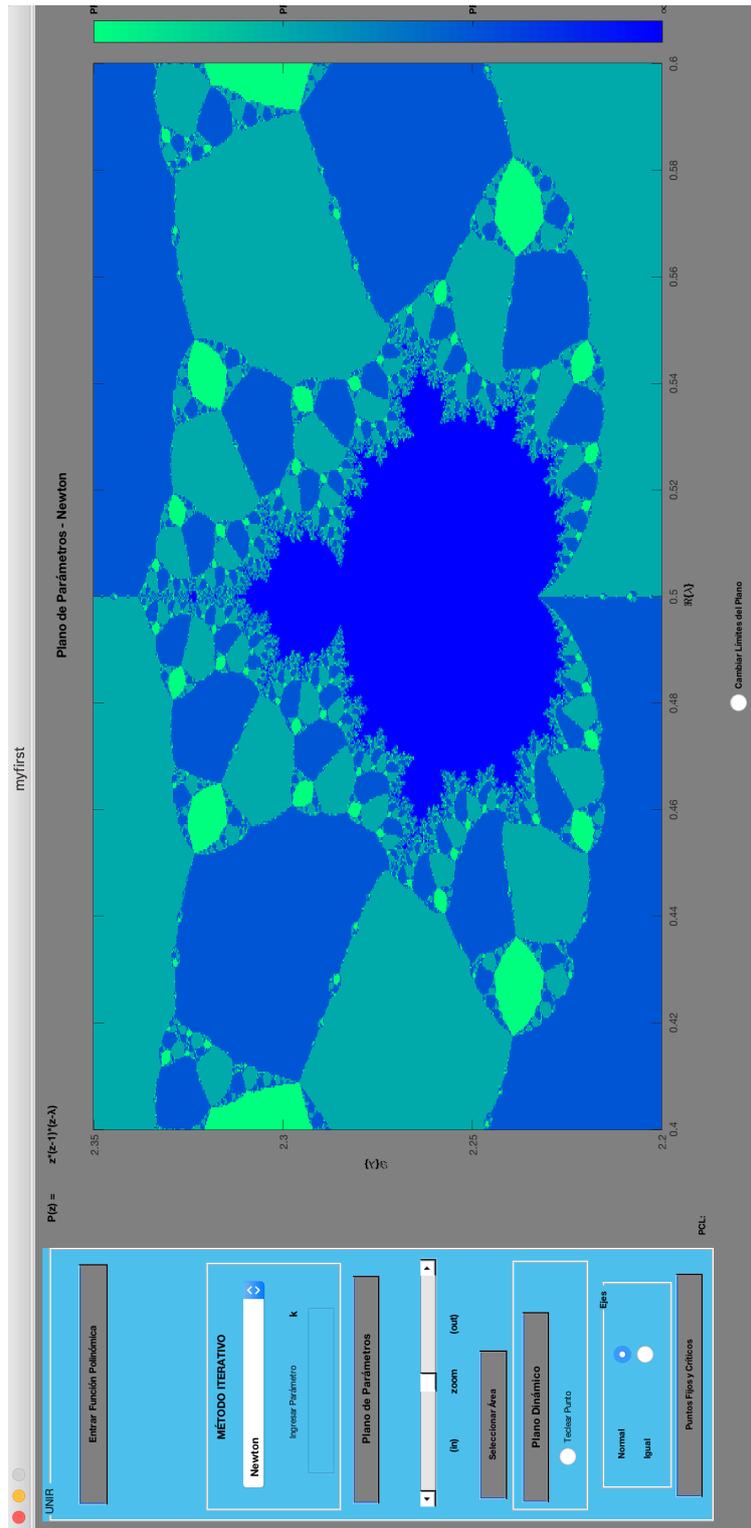


Figura 9.14: Zoom del plano de parámetros.

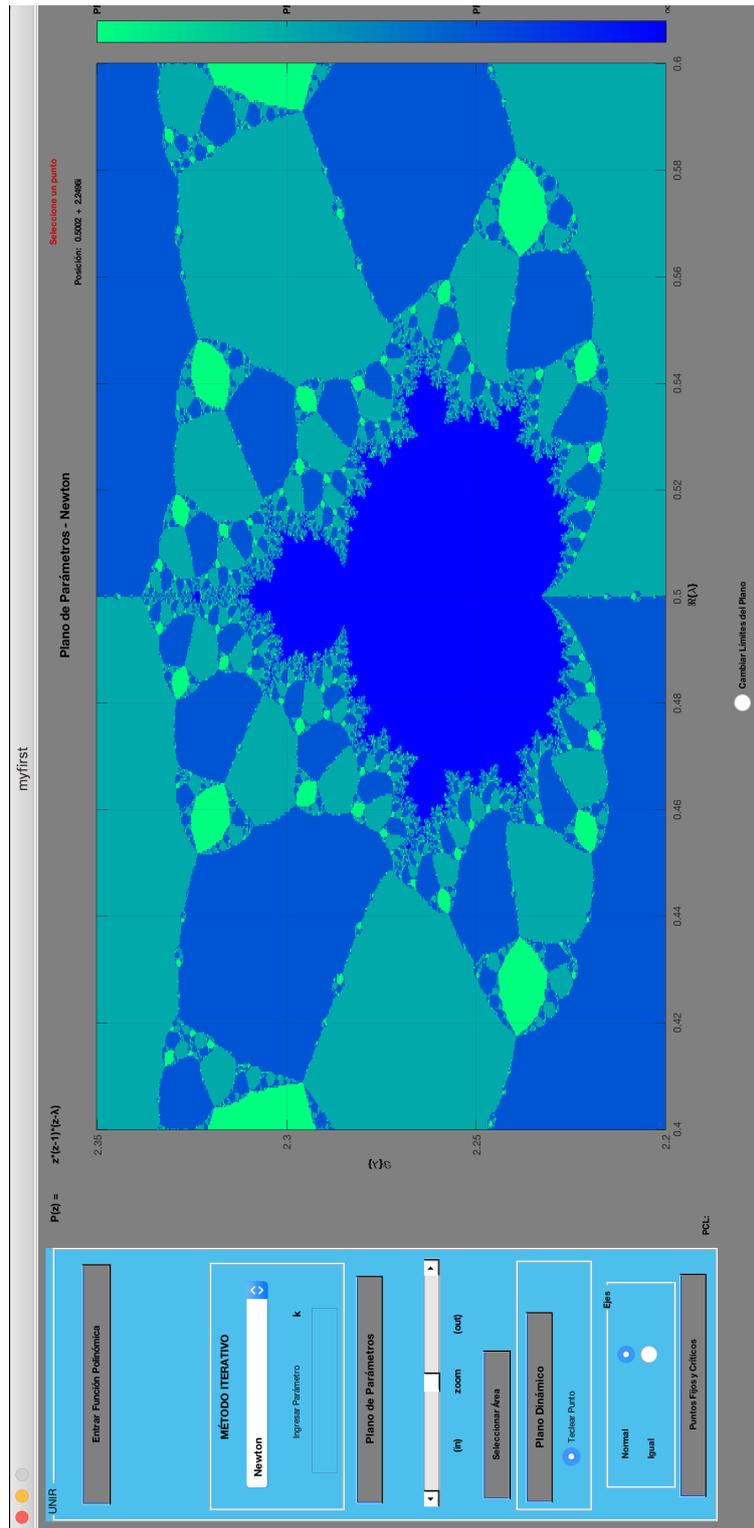


Figura 9.15: Fijamos el valor del parámetro para pintar los planos dinámicos.

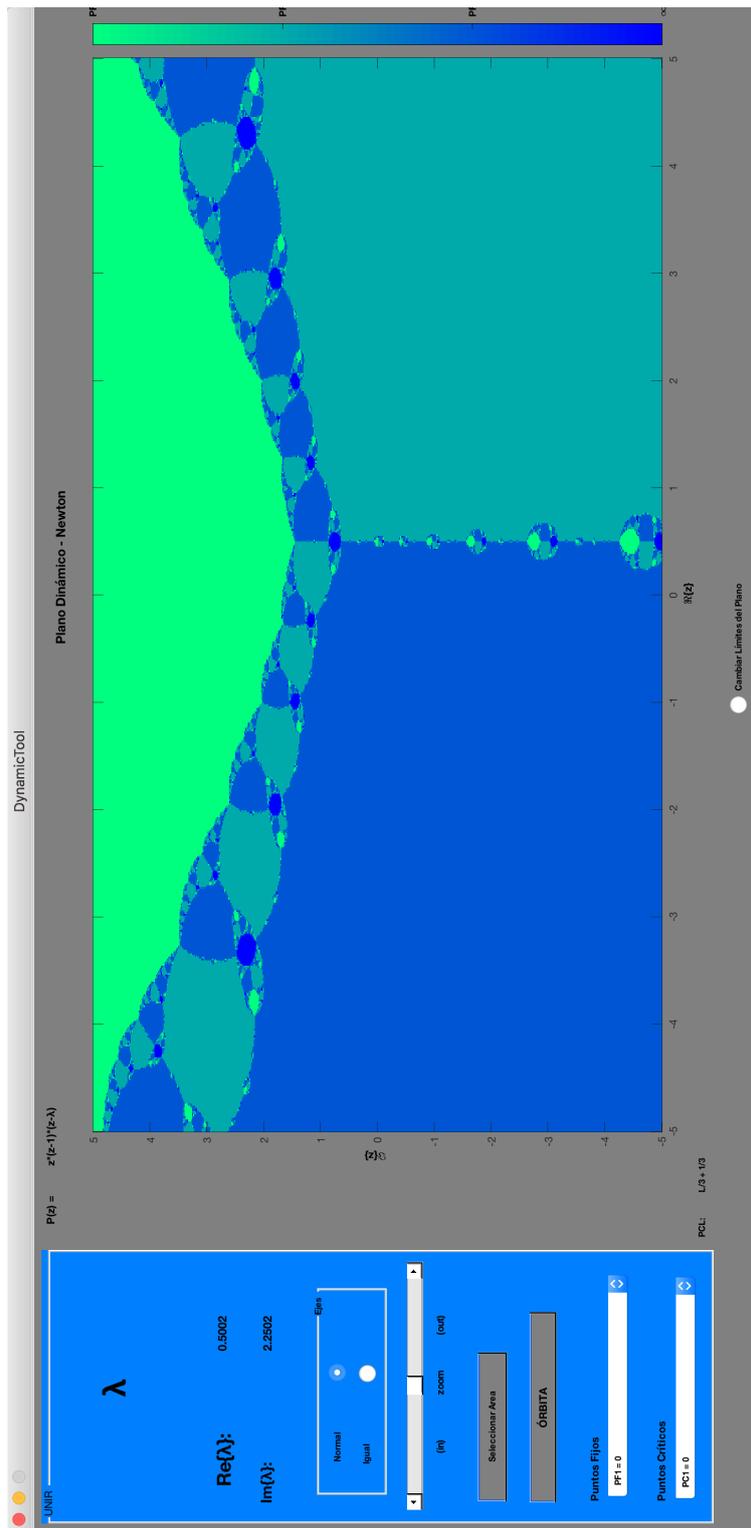


Figura 9.16: Plano dinámico.

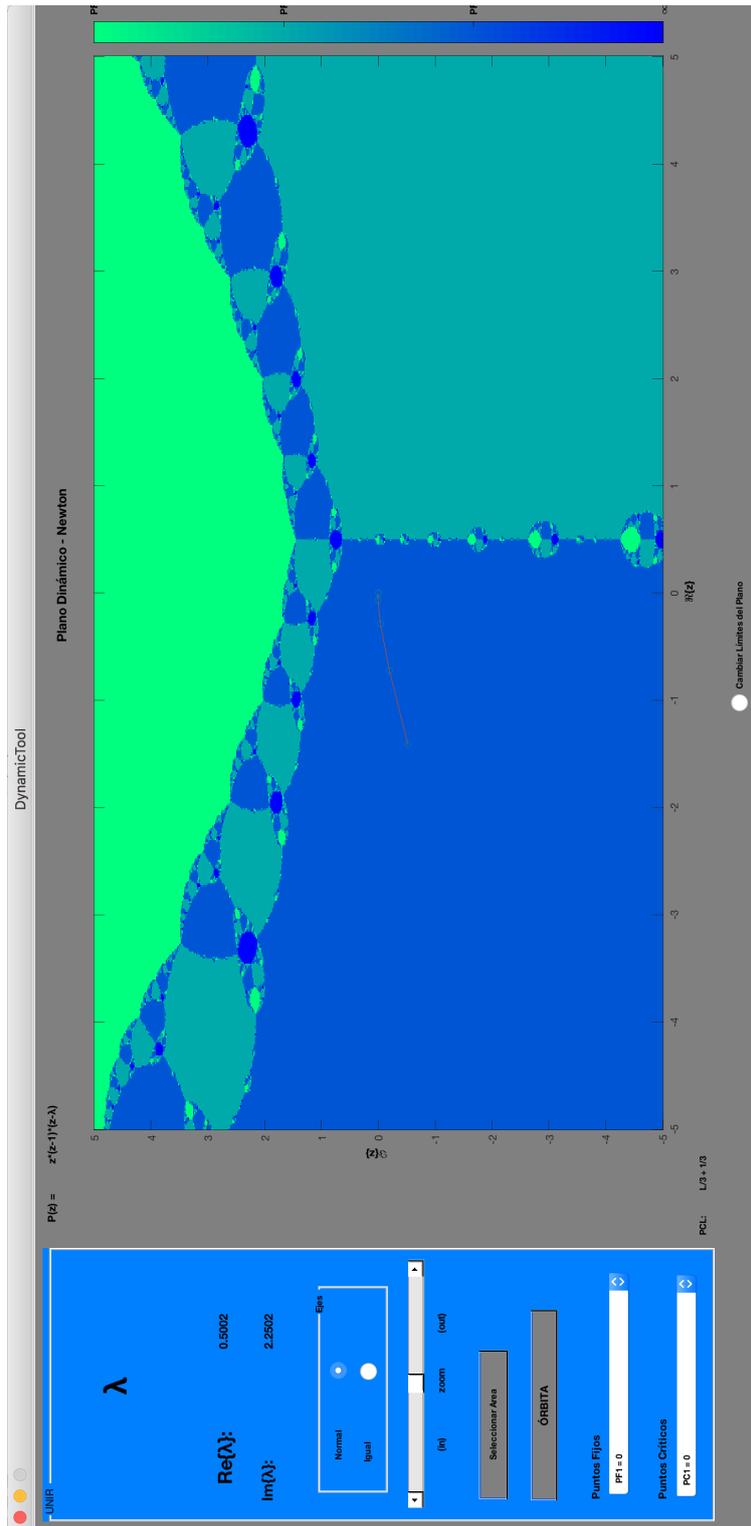


Figura 9.17: Plano dinámico con una órbita dibujada.

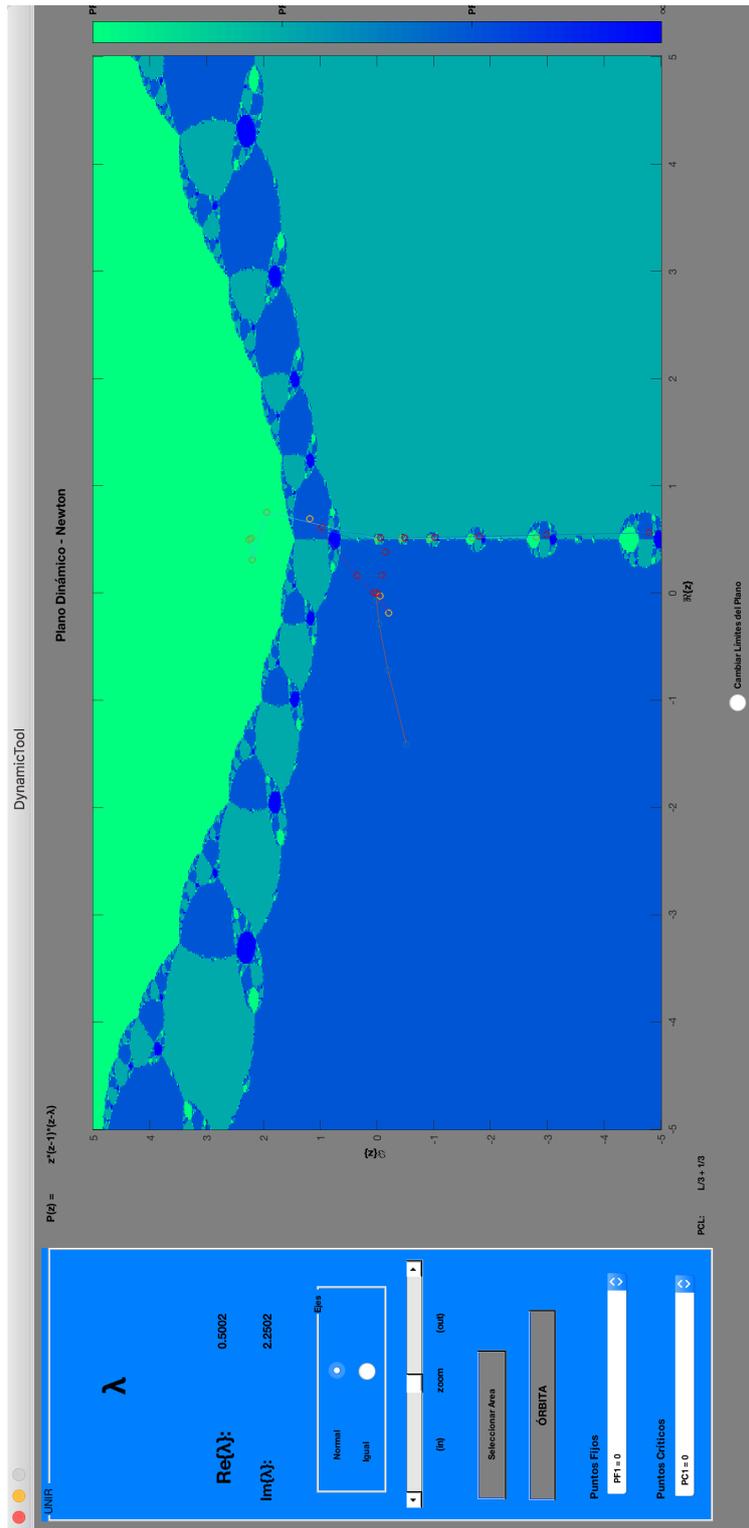


Figura 9.18: Plano dinámico con varias órbitas dibujadas.

9.5. Puesta en marcha de la herramienta

Se ha diseñado la asignatura de tal forma que se comienza las primeras semanas con la explicación teórica de los contenidos fundamentales, sin las cuales sería inviable asimilar los contenidos y adquirir las competencias atribuidas a ella, y posteriormente están programadas una serie de actividades que exigen cierto grado de interactividad por parte de los estudiantes en su proceso de formación. Conceptualmente, la asignatura comienza con contenidos teóricos, se continúa con una breve introducción de los comandos elementales del software diseñado, debidamente ilustrados y con numerosos ejemplos para facilitar el aprendizaje de la herramienta. Continuamos con ejercicios sencillos para que los alumnos puedan asimilar rápidamente la utilidad de la herramienta y que no suponga una carga de trabajo extraordinaria de dedicación por su parte. En una segunda etapa, se abordan contenidos teóricos avanzados, y se plantean problemas, en orden creciente de dificultad, que el estudiante debe resolver con la ayuda de la herramienta creada para tal fin.

Logramos afianzar así los conocimientos genéricos ya adquiridos del mismo modo que conseguimos que los estudiantes descubran nuevos comandos o instrucciones de la herramienta, diseñadas para tareas más específicas. La interrelación entre teoría y práctica permite al alumno consolidar las nociones más abstractas y desarrollar su capacidad de resolver problemas de forma autónoma. Para finalizar, serán formulados problemas similares a los que los estudiantes deben enfrentarse en el examen final, con el mismo índice de dificultad que les será exigido. A medida que los estudiantes hagan uso de la herramienta, la potencia gráfica y el gran componente visual de este software, estimulará su curiosidad e interés por el programa y por tanto su utilización y facilitará que los usuarios puedan adquirir los conocimientos suficientes como para realizar las entregas de las actividades planificadas.

9.6. Resultados

En las Tablas 9.3-9.4 pueden verse los resultados obtenidos por los alumnos, después de aplicar los cambios que se han comentado en la sección anterior.

Por otro lado, en la Tabla 9.5, podemos ver las estadísticas relacionadas con los resultados obtenidos por el segundo grupo.

Cuadro 9.3: Resultados del segundo grupo

Alumno	Trabajo 1	Trabajo 2	Trabajo 3
1	9	1.5	5
2	NP	NP	NP
3	NP	NP	NP
4	NP	NP	NP
5	NP	NP	NP
6	NP	NP	NP
7	7	2	4
8	3	5	NP
9	10	8	10
10	NP	NP	NP
11	10	8	8.5
12	7.5	NP	NP
13	NP	NP	NP
14	3	9.5	2.5
15	7	7.5	5.5
16	NP	NP	NP
17	NP	4	8
18	NP	NP	NP
19	8.5	8	5
20	NP	NP	NP
21	9	8	10
22	9	8	5.5
23	9	6	9
24	2	3	0
25	NP	NP	NP
26	10	8	7
27	NP	NP	NP
28	5	5.5	2
29	9.5	10	10
30	9.5	7	10
31	NP	NP	NP
32	10	10	9
33	9	NP	8
34	8	5	4

Para este estudio hemos seleccionado un total de 50 alumnos de la asignatura Sistemas Dinámicos Discretos y Continuos del Máster Universitario en Ingeniería Matemática y Computación impartido en la Universidad Internacional de La

Cuadro 9.4: Resultados del segundo grupo

Alumno	Trabajo 1	Trabajo 2	Trabajo 3
35	0	0	0
36	9	9	8
37	10	8	9
38	9	7	10
39	9	8	8
40	NP	NP	NP
41	8	8	9
42	NP	NP	NP
43	8	8	5.5
44	7	6.5	3
45	8	9	4
46	8	8	8
47	6	10	2
48	7	3	7.5
49	9	9	10
50	3	5.5	2

Cuadro 9.5: Estadísticas del segundo grupo

	Trabajo 1	Trabajo 2	Trabajo 3
NP	17	18	18
Suspensos	4	5	9
Aprobados	29	27	23
Total	50	50	50
Total sin NP	33	32	32
% presentados	66.00 %	64.00 %	64.00 %
% aprobados con NP	58.00 %	54.00 %	46.0 %
% Aprobados sin NP	87.88 %	84.38 %	71.88 %
Nota Media con NP	5.12	4.46	4.18
Nota Media sin NP	7.76	6.99	6.53

Rioja. Estos alumnos se corresponden con la tercera promoción que ha accedido a dicha asignatura y es la primera vez que se enfrentan al estudio dinámico de métodos y familias de métodos iterativos. En este sentido nos hemos centrado en una medición cuantitativa teniendo en cuenta sus resultados en los trabajos entregados.

A nivel general, podemos ver que tanto los porcentajes de presentados, como

de aprobados así como las notas medias aumentan sustancialmente, pero vamos a pasar a estudiar cada uno de los trabajos por separado y tomando en consideración dos opciones, la primera descartando como datos los no presentados y la segunda, considerando que los no presentados supone un 0 en la nota.

9.6.1. Resultados Trabajo 1 sin NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el primer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.1-9.6.1.

Grupo = PER1-2

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 1	17	10,0	,0	10,0	68,0
N válido (por lista)	17				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 1	4,000	,9777	4,0311	16,250	,324
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 1	,550	-1,614	1,063
N válido (por lista)			

a. Grupo = PER1-2

Figura 9.19: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 1

N	Válido	17
	Perdidos	17

a. Grupo = PER1-2

Trabajo 1^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	7	20,6	41,2	41,2
	2,0	1	2,9	5,9	47,1
	4,0	1	2,9	5,9	52,9
	5,0	2	5,9	11,8	64,7
	6,0	1	2,9	5,9	70,6
	8,0	1	2,9	5,9	76,5
	9,0	2	5,9	11,8	88,2
	10,0	2	5,9	11,8	100,0
	Total	17	50,0	100,0	
Perdidos	Sistema	17	50,0		
Total		34	100,0		

a. Grupo = PER1-2

Figura 9.20: Descriptivos asociados al primer grupo.

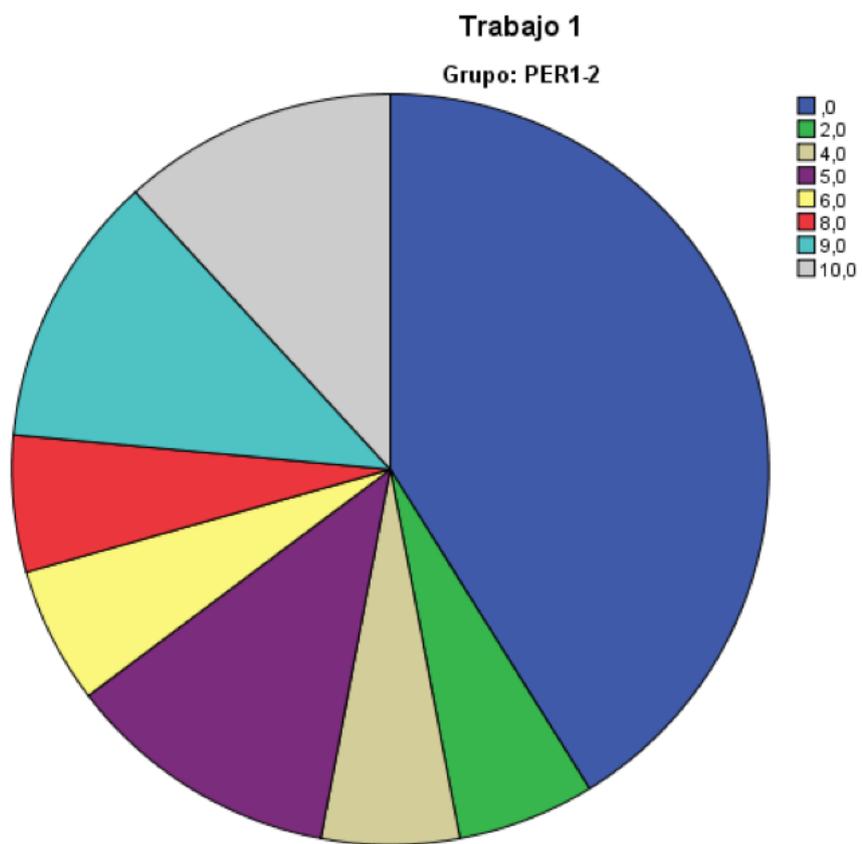


Figura 9.21: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.1-9.6.1.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 1	33	8,0	2,0	10,0	256,0
N válido (por lista)	33				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 1	7,758	,3893	2,2365	5,002	-1,360
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 1	,409	1,023	,798
N válido (por lista)			

a. Grupo = PER3

Figura 9.22: Descriptivos asociados al segundo grupo.

Frecuencias

Grupo = PER3

Estadísticos^a

Trabajo 1

N	Válido	33
	Perdidos	17

a. Grupo = PER3

Trabajo 1^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	2,0	1	2,0	3,0	3,0
	3,0	3	6,0	9,1	12,1
	5,0	1	2,0	3,0	15,2
	6,0	1	2,0	3,0	18,2
	7,0	4	8,0	12,1	30,3
	7,5	1	2,0	3,0	33,3
	8,0	5	10,0	15,2	48,5
	8,5	1	2,0	3,0	51,5
	9,0	9	18,0	27,3	78,8
	9,5	2	4,0	6,1	84,8
	10,0	5	10,0	15,2	100,0
	Total	33	66,0	100,0	
Perdidos	Sistema	17	34,0		
	Total	50	100,0		

a. Grupo = PER3

Figura 9.23: Descriptivos asociados al segundo grupo.

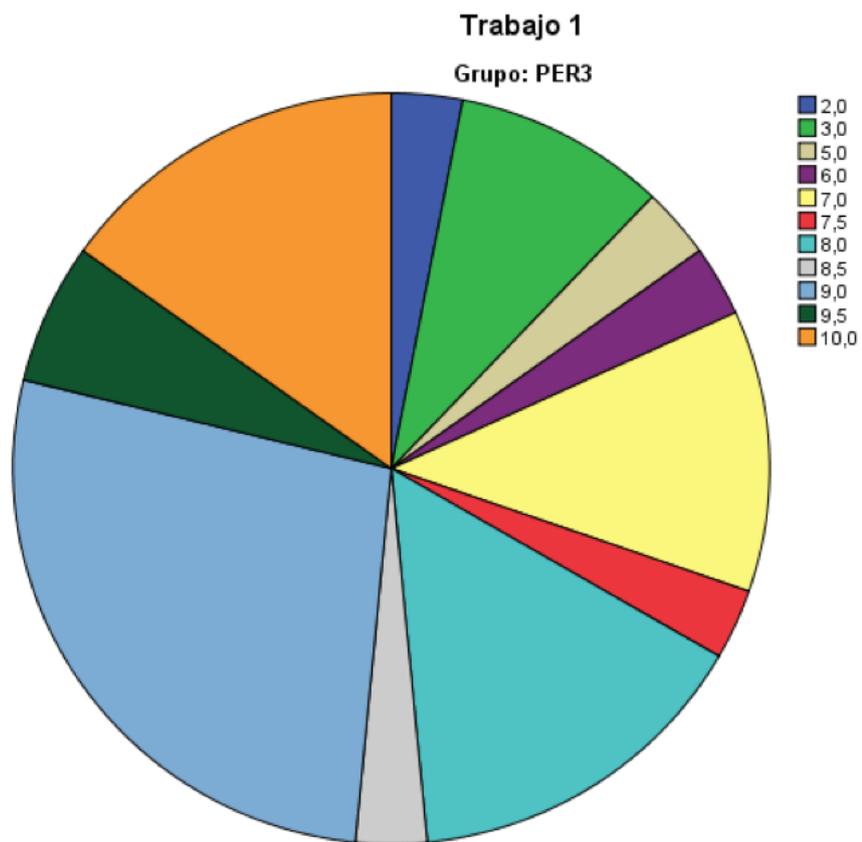


Figura 9.24: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que el primer grupo, sólo dispone de 17 casos válidos utilizaremos la prueba U de Mann-Withney que nos da SPSS.

Prueba de Mann-Whitney

Rangos				
	Grupo	N	Rango promedio	Suma de rangos
Trabajo 1	PER3	33	30,24	998,00
	PER1-2	17	16,29	277,00
	Total	50		

Estadísticos de prueba ^a	
	Trabajo 1
U de Mann-Whitney	124,000
W de Wilcoxon	277,000
Z	-3,223
Sig. asintótica (bilateral)	,001

a. Variable de agrupación: Grupo

Figura 9.25: Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el primer trabajo.

El resultado puede verse en la Figura 9.6.1. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 3.223 superior al grupo impartido con la primera metodología y además este resultado es significativo.

9.6.2. Resultados Trabajo 1 con NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el primer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.2-9.6.2.

Grupo = PER1-2

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 1	34	10,0	,0	10,0	68,0
N válido (por lista)	34				

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
	Trabajo 1	2,000	,5941	3,4641	12,000
N válido (por lista)					

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 1	,403	,521	,788
N válido (por lista)			

a. Grupo = PER1-2

Figura 9.26: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 1

N	Válido	34
	Perdidos	0

a. Grupo = PER1-2

Trabajo 1^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	24	70,6	70,6	70,6
	2,0	1	2,9	2,9	73,5
	4,0	1	2,9	2,9	76,5
	5,0	2	5,9	5,9	82,4
	6,0	1	2,9	2,9	85,3
	8,0	1	2,9	2,9	88,2
	9,0	2	5,9	5,9	94,1
	10,0	2	5,9	5,9	100,0
	Total	34	100,0	100,0	

a. Grupo = PER1-2

Figura 9.27: Descriptivos asociados al primer grupo.

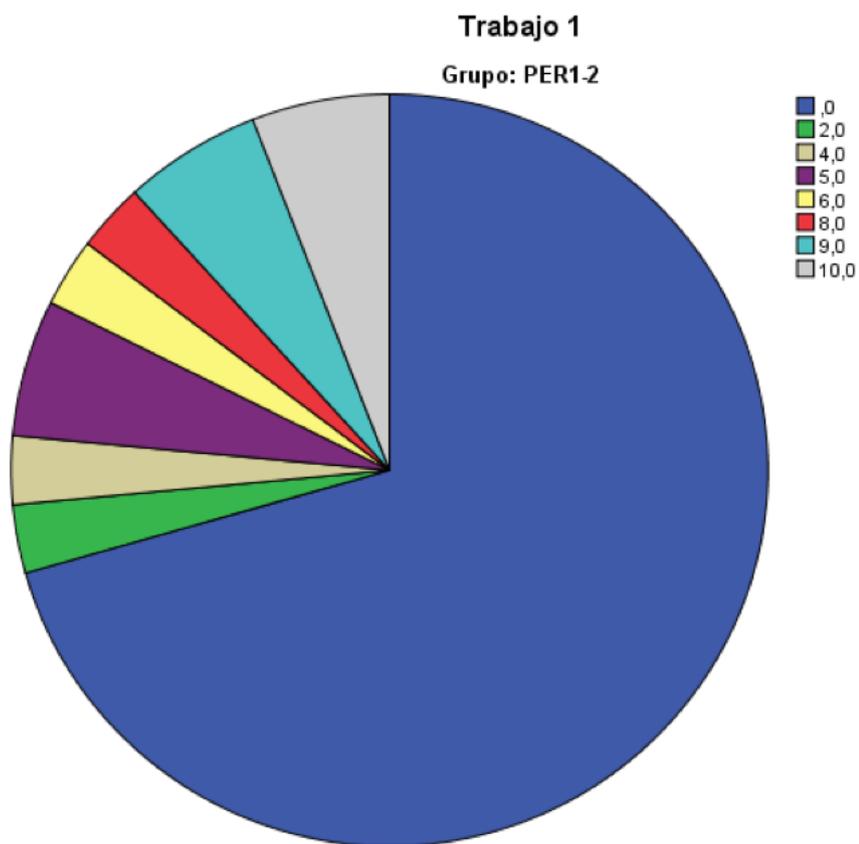


Figura 9.28: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.2-9.6.4.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 1	50	10,0	,0	10,0	256,0
N válido (por lista)	50				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 1	5,120	,5839	4,1287	17,047	-,262
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 1	,337	-1,763	,662
N válido (por lista)			

a. Grupo = PER3

Figura 9.29: Descriptivos asociados al segundo grupo.

Grupo = PER3

Estadísticos^a

Trabajo 1

N	Válido	50
	Perdidos	0

a. Grupo = PER3

Trabajo 1^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	17	34,0	34,0	34,0
	2,0	1	2,0	2,0	36,0
	3,0	3	6,0	6,0	42,0
	5,0	1	2,0	2,0	44,0
	6,0	1	2,0	2,0	46,0
	7,0	4	8,0	8,0	54,0
	7,5	1	2,0	2,0	56,0
	8,0	5	10,0	10,0	66,0
	8,5	1	2,0	2,0	68,0
	9,0	9	18,0	18,0	86,0
	9,5	2	4,0	4,0	90,0
	10,0	5	10,0	10,0	100,0
	Total	50	100,0	100,0	

a. Grupo = PER3

Figura 9.30: Descriptivos asociados al segundo grupo.

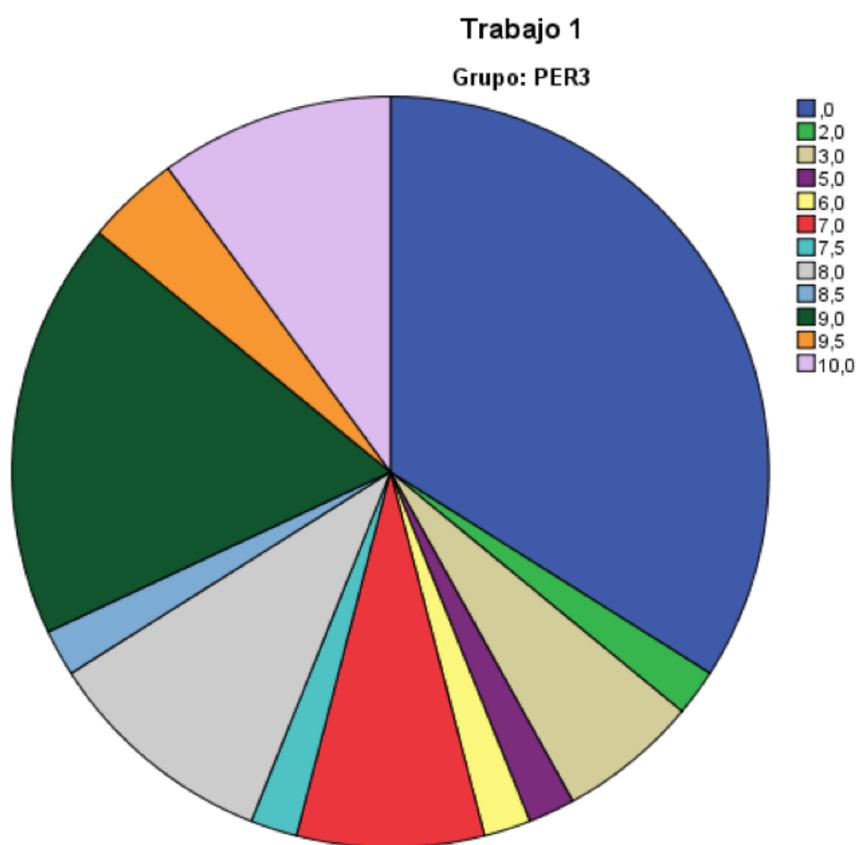


Figura 9.31: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que ahora ambos grupos tienen más de 30 observaciones, podemos usar la prueba T de Student.

Prueba T

Estadísticas de grupo				
Grupo	N	Media	Desviación estándar	Media de error estándar
Trabajo 1 PER3	50	5,120	4,1287	,5839
PER1-2	34	2,000	3,4641	,5941

Prueba de muestras independientes					
		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias	
		F	Sig.	t	gl.
Trabajo 1	Se asumen varianzas iguales	8,457	,005	3,622	82
	No se asumen varianzas iguales			3,746	78,325

Prueba de muestras independientes				
		prueba t para la igualdad de medias		
		Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar
Trabajo 1	Se asumen varianzas iguales	,001	3,1200	,8614
	No se asumen varianzas iguales	,000	3,1200	,8330

Prueba de muestras independientes				
		prueba t para la igualdad de medias		
		95% de intervalo de confianza de la diferencia		
		Inferior	Superior	
Trabajo 1	Se asumen varianzas iguales	1,4065	4,8335	
	No se asumen varianzas iguales	1,4618	4,7782	

Figura 9.32: Prueba T de Student para la comparación de medias de notas obtenidas en el primer trabajo.

El resultado puede verse en la Figura 9.6.2. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 3.120 superior al grupo impartido con la primera metodología y además este resultado es significativo.

9.6.3. Resultados Trabajo 2 sin NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el primer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.3-9.6.3.

Grupo = PER1-2

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 2	17	10,0	,0	10,0	86,5
N válido (por lista)	17				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
	Trabajo 2	5,088	,9993	4,1202	16,976
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
	Trabajo 2	,550	-1,778
N válido (por lista)			

a. Grupo = PER1-2

Figura 9.33: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 2

N	Válido	17
	Perdidos	17

a. Grupo = PER1-2

Trabajo 2^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	6	17,6	35,3	35,3
	5,0	2	5,9	11,8	47,1
	6,0	1	2,9	5,9	52,9
	8,0	3	8,8	17,6	70,6
	8,5	1	2,9	5,9	76,5
	9,0	1	2,9	5,9	82,4
	9,5	2	5,9	11,8	94,1
	10,0	1	2,9	5,9	100,0
	Total	17	50,0	100,0	
Perdidos	Sistema	17	50,0		
Total		34	100,0		

a. Grupo = PER1-2

Figura 9.34: Descriptivos asociados al primer grupo.

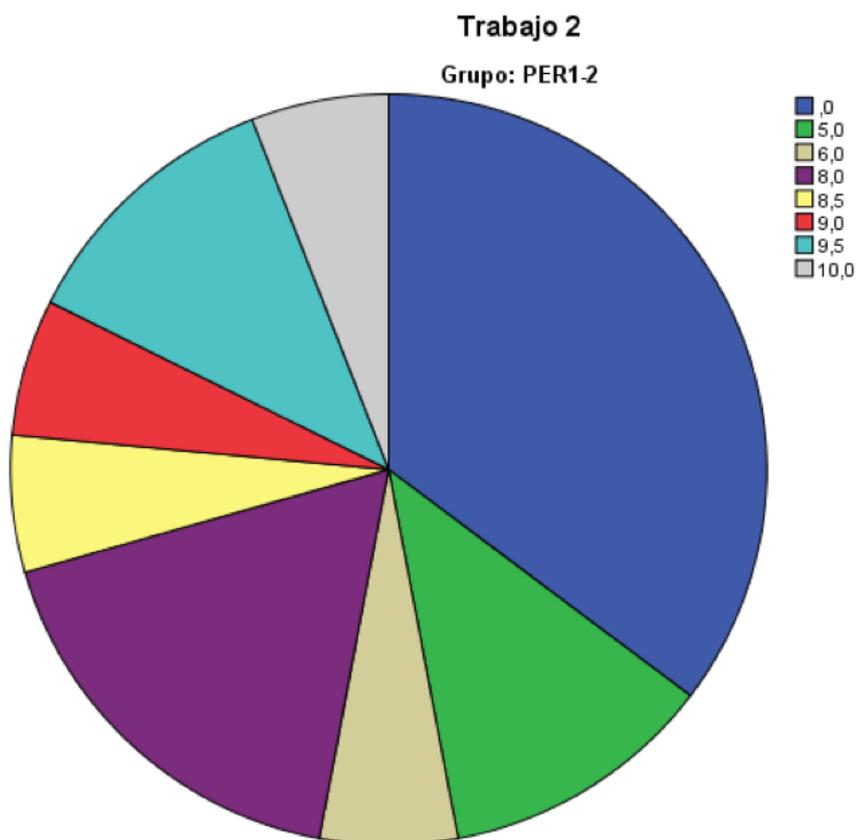


Figura 9.35: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.3-9.6.3.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 2	32	8,0	2,0	10,0	224,0
N válido (por lista)	32				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 2	7,000	,3972	2,2469	5,048	-,799
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 2	,414	-,222	,809
N válido (por lista)			

a. Grupo = PER3

Figura 9.36: Descriptivos asociados al segundo grupo.

Grupo = PER3

Estadísticos^a

Trabajo 2

N	Válido	32
	Perdidos	18

a. Grupo = PER3

Trabajo 2^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	2,0	1	2,0	3,1	3,1
	2,5	1	2,0	3,1	6,3
	3,0	2	4,0	6,3	12,5
	4,0	1	2,0	3,1	15,6
	5,0	2	4,0	6,3	21,9
	5,5	2	4,0	6,3	28,1
	6,0	1	2,0	3,1	31,3
	6,5	1	2,0	3,1	34,4
	7,0	2	4,0	6,3	40,6
	7,5	1	2,0	3,1	43,8
	8,0	11	22,0	34,4	78,1
	9,0	3	6,0	9,4	87,5
	9,5	1	2,0	3,1	90,6
	10,0	3	6,0	9,4	100,0
	Total	32	64,0	100,0	
Perdidos	Sistema	18	36,0		
Total		50	100,0		

a. Grupo = PER3

Figura 9.37: Descriptivos asociados al segundo grupo.

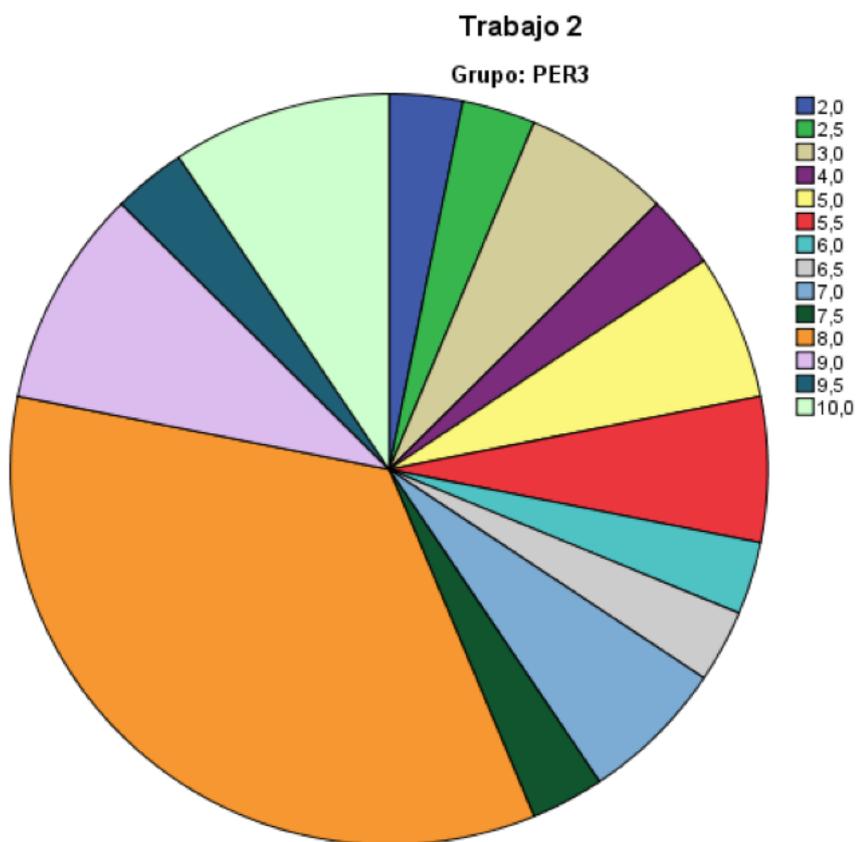


Figura 9.38: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que el primer grupo, sólo dispone de 17 casos válidos utilizaremos la prueba U de Mann-Withney que nos da SPSS.

Prueba de Mann-Whitney

Rangos				
	Grupo	N	Rango promedio	Suma de rangos
Trabajo 2	PER3	32	26,66	853,00
	PER1-2	17	21,88	372,00
	Total	49		

Estadísticos de prueba ^a	
	Trabajo 2
U de Mann-Whitney	219,000
W de Wilcoxon	372,000
Z	-1,129
Sig. asintótica (bilateral)	,259

a. Variable de agrupación: Grupo

Figura 9.39: Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el segundo trabajo.

El resultado puede verse en la Figura 9.6.3. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 3.223 superior al grupo impartido con la primera metodología, sin embargo, este resultado no es significativo.

9.6.4. Resultados Trabajo 2 con NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el primer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.4-9.6.4.

Grupo = PER1-2

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 2	34	10,0	,0	10,0	86,5
N válido (por lista)	34				

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 2	2,544	,6620	3,8600	14,900	,978
N válido (por lista)					

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 2	,403	-,902	,788
N válido (por lista)			

a. Grupo = PER1-2

Figura 9.40: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 2

N	Válido	34
	Perdidos	0

a. Grupo = PER1-2

Trabajo 2^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	23	67,6	67,6	67,6
	5,0	2	5,9	5,9	73,5
	6,0	1	2,9	2,9	76,5
	8,0	3	8,8	8,8	85,3
	8,5	1	2,9	2,9	88,2
	9,0	1	2,9	2,9	91,2
	9,5	2	5,9	5,9	97,1
	10,0	1	2,9	2,9	100,0
	Total	34	100,0	100,0	

a. Grupo = PER1-2

Figura 9.41: Descriptivos asociados al primer grupo.

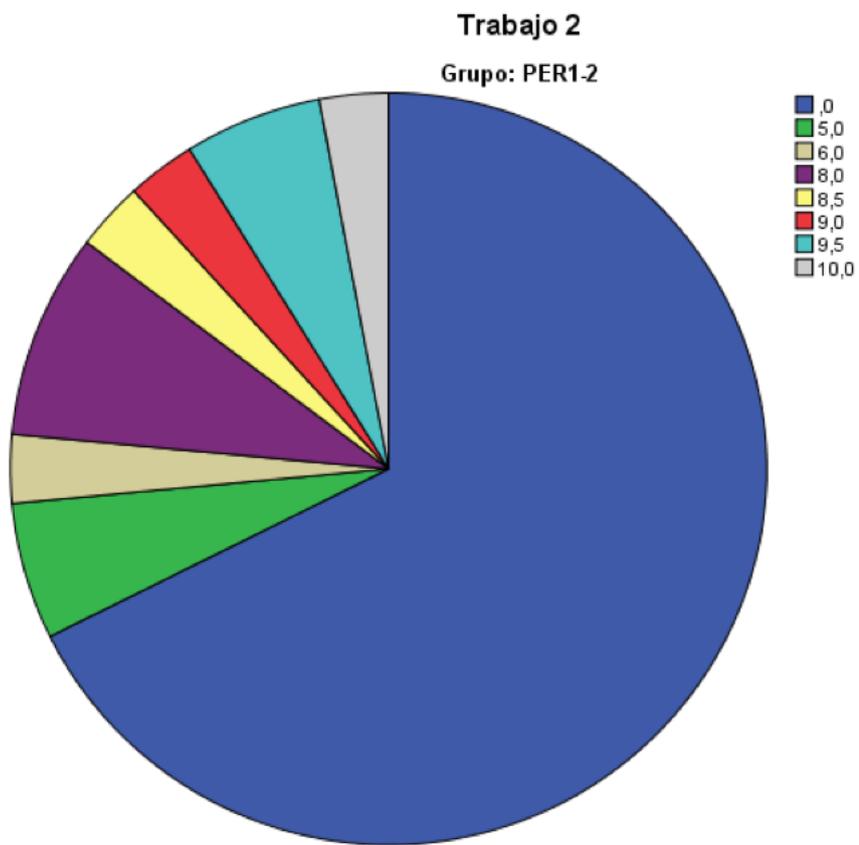


Figura 9.42: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.4-??.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 2	50	10,0	,0	10,0	224,0
N válido (por lista)	50				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 2	4,480	,5425	3,8359	14,714	-,070
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 2	,337	-1,716	,662
N válido (por lista)			

a. Grupo = PER3

Figura 9.43: Descriptivos asociados al segundo grupo.

Grupo = PER3

Estadísticos^a

Trabajo 2

N	Válido	50
	Perdidos	0

a. Grupo = PER3

Trabajo 2^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	18	36,0	36,0	36,0
	2,0	1	2,0	2,0	38,0
	2,5	1	2,0	2,0	40,0
	3,0	2	4,0	4,0	44,0
	4,0	1	2,0	2,0	46,0
	5,0	2	4,0	4,0	50,0
	5,5	2	4,0	4,0	54,0
	6,0	1	2,0	2,0	56,0
	6,5	1	2,0	2,0	58,0
	7,0	2	4,0	4,0	62,0
	7,5	1	2,0	2,0	64,0
	8,0	11	22,0	22,0	86,0
	9,0	3	6,0	6,0	92,0
	9,5	1	2,0	2,0	94,0
	10,0	3	6,0	6,0	100,0
	Total	50	100,0	100,0	

a. Grupo = PER3

Figura 9.44: Descriptivos asociados al segundo grupo.

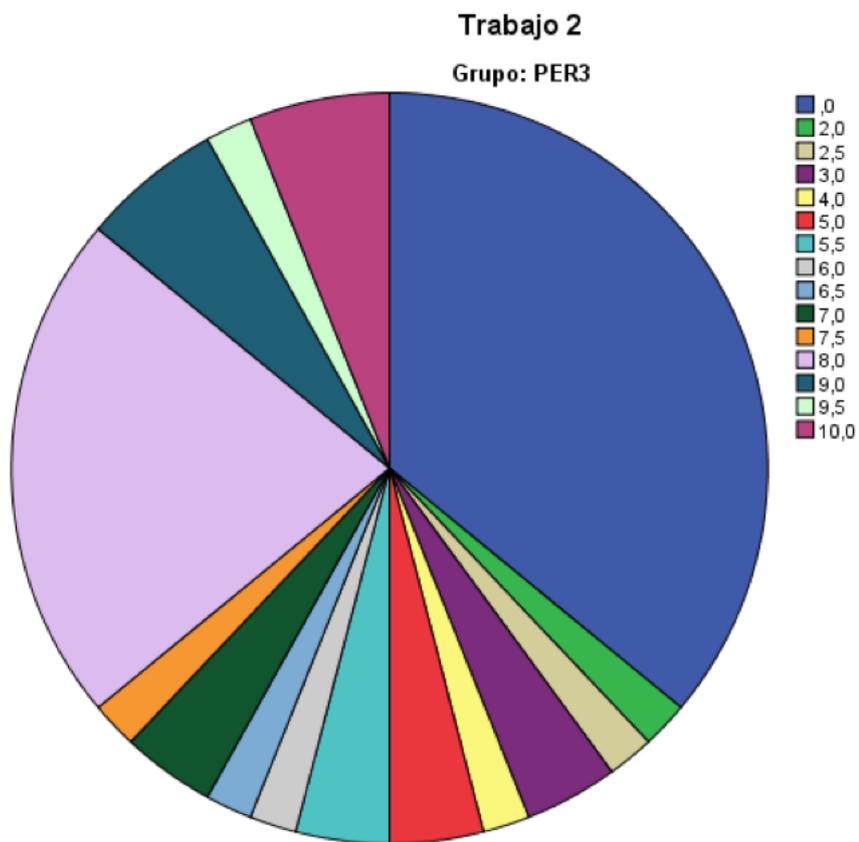


Figura 9.45: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que ahora ambos grupos tienen más de 30 observaciones, podemos usar la prueba T de Student.

Prueba T

Estadísticas de grupo					
Grupo		N	Media	Desviación estándar	Media de error estándar
Trabajo 2	PER3	50	4,480	3,8359	,5425
	PER1-2	34	2,544	3,8600	,6620

Prueba de muestras independientes					
		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias	
		F	Sig.	t	gl.
Trabajo 2	Se asumen varianzas iguales	,090	,765	2,265	82
	No se asumen varianzas iguales			2,262	70,723

Prueba de muestras independientes					
		prueba t para la igualdad de medias			
		Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	
Trabajo 2	Se asumen varianzas iguales	,026	1,9359	,8548	
	No se asumen varianzas iguales	,027	1,9359	,8559	

Prueba de muestras independientes					
		prueba t para la igualdad de medias			
		95% de intervalo de confianza de la diferencia			
		Inferior	Superior		
Trabajo 2	Se asumen varianzas iguales	,2354	3,6364		
	No se asumen varianzas iguales	,2292	3,6425		

Figura 9.46: Prueba T de Student para la comparación de medias de notas obtenidas en el segundo trabajo.

El resultado puede verse en la Figura 9.6.4. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 1,9359 superior al grupo impartido con la primera metodología y además este resultado es significativo.

9.6.5. Resultados Trabajo 3 sin NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el tercer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.5-9.6.5.

Grupo = PER1-2

Estadísticos descriptivos ^a					
	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 3	13	10,0	,0	10,0	45,0
N válido (por lista)	13				

Estadísticos descriptivos ^a					
	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
	Trabajo 3	3,462	1,0292	3,7107	13,769
N válido (por lista)					

Estadísticos descriptivos ^a				
	Asimetría		Curtosis	
	Error estándar		Estadístico	Error estándar
	Trabajo 3			-1,427
N válido (por lista)	,616			

a. Grupo = PER1-2

Figura 9.47: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 3

N	Válido	13
	Perdidos	21

a. Grupo = PER1-2

Trabajo 3^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	6	17,6	46,2	46,2
	3,0	1	2,9	7,7	53,8
	5,0	2	5,9	15,4	69,2
	7,0	2	5,9	15,4	84,6
	8,0	1	2,9	7,7	92,3
	10,0	1	2,9	7,7	100,0
	Total	13	38,2	100,0	
Perdidos	Sistema	21	61,8		
Total		34	100,0		

a. Grupo = PER1-2

Figura 9.48: Descriptivos asociados al primer grupo.

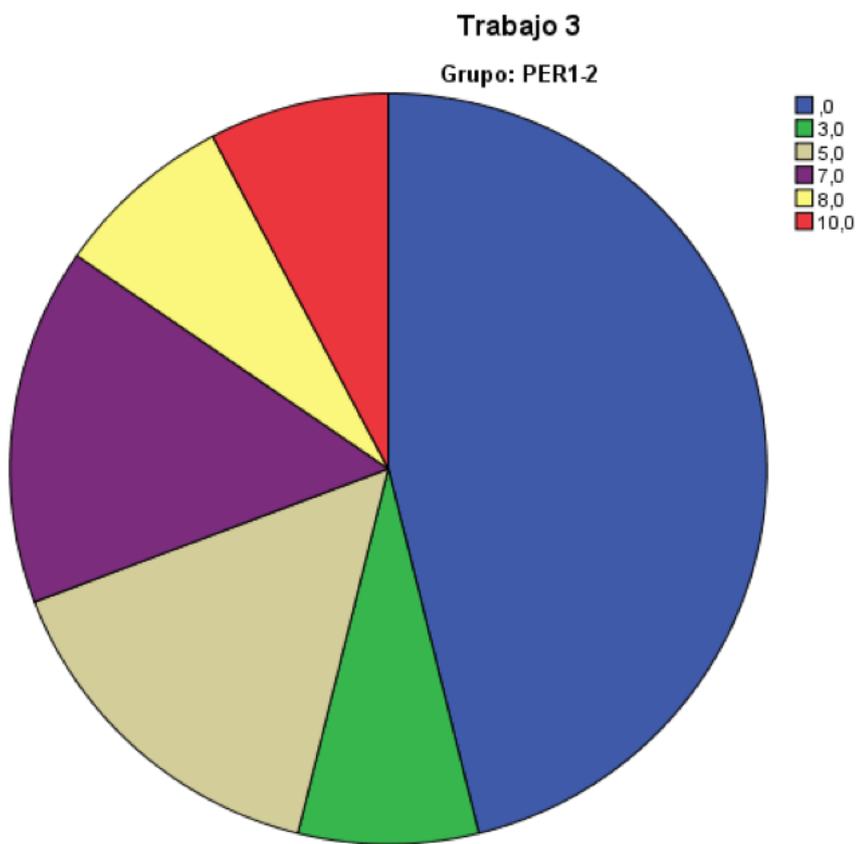


Figura 9.49: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.5-9.6.5.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 3	32	10,0	,0	10,0	209,0
N válido (por lista)	32				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 3	6,531	,5241	2,9647	8,789	-,517
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 3	,414	-,933	,809
N válido (por lista)			

a. Grupo = PER3

Figura 9.50: Descriptivos asociados al segundo grupo.

Grupo = PER3

Estadísticos^a

Trabajo 3

N	Válido	32
	Perdidos	18

a. Grupo = PER3

Trabajo 3^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	1	2,0	3,1	3,1
	2,0	3	6,0	9,4	12,5
	2,5	1	2,0	3,1	15,6
	3,0	1	2,0	3,1	18,8
	4,0	3	6,0	9,4	28,1
	5,0	2	4,0	6,3	34,4
	5,5	3	6,0	9,4	43,8
	7,0	1	2,0	3,1	46,9
	7,5	1	2,0	3,1	50,0
	8,0	5	10,0	15,6	65,6
	8,5	1	2,0	3,1	68,8
	9,0	4	8,0	12,5	81,3
	10,0	6	12,0	18,8	100,0
	Total	32	64,0	100,0	
Perdidos	Sistema	18	36,0		
Total		50	100,0		

a. Grupo = PER3

Figura 9.51: Descriptivos asociados al segundo grupo.

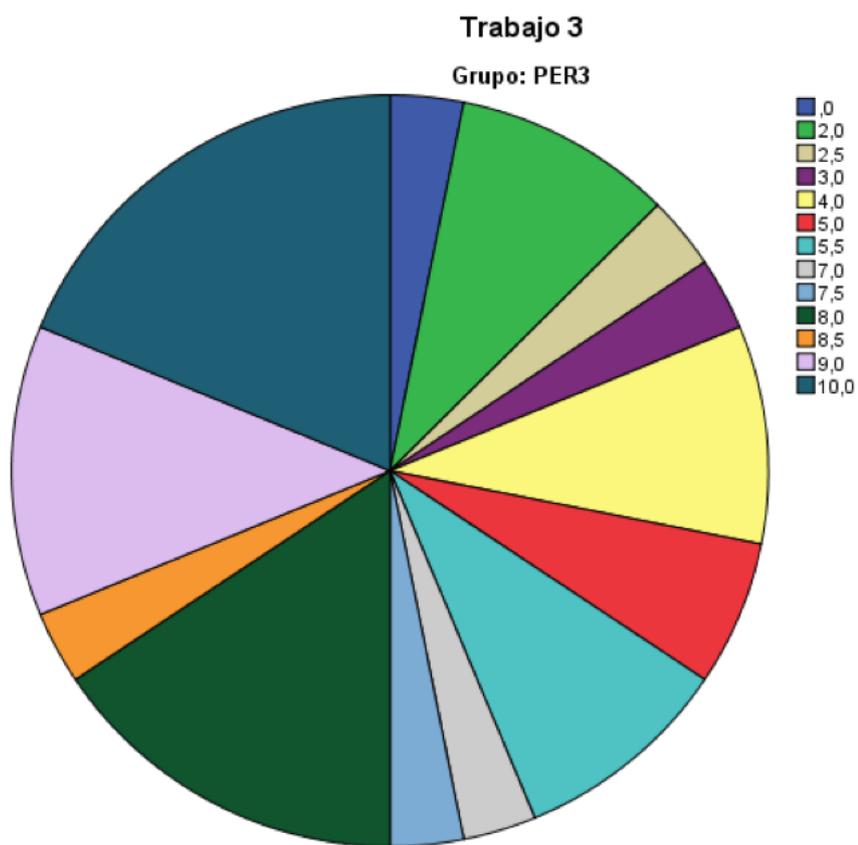


Figura 9.52: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que el primer grupo, sólo dispone de 17 casos válidos utilizaremos la prueba U de Mann-Withney que nos da SPSS.

Prueba de Mann-Whitney

Rangos				
	Grupo	N	Rango promedio	Suma de rangos
Trabajo 3	PER3	32	26,28	841,00
	PER1-2	13	14,92	194,00
	Total	45		

Estadísticos de prueba ^a	
	Trabajo 3
U de Mann-Whitney	103,000
W de Wilcoxon	194,000
Z	-2,641
Sig. asintótica (bilateral)	,008

a. Variable de agrupación: Grupo

Figura 9.53: Prueba U de Mann-Withney para la comparación de medias de notas obtenidas en el tercer trabajo.

El resultado puede verse en la Figura 9.6.5. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 3.223 superior al grupo impartido con la primera metodología y además este resultado es significativo.

9.6.6. Resultados Trabajo 3 con NP

En esta sección, comenzamos viendo los estadísticos descriptivos que arroja el programa SPSS relacionados con el tercer grupo, denotado por “PER1-2”, que aparecen en las Figuras 9.6.6-9.6.6.

Grupo = PER1-2

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 3	34	10,0	,0	10,0	45,0
N válido (por lista)	34				

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 3	1,324	,4827	2,8147	7,922	1,961
N válido (por lista)					

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 3	,403	2,607	,788
N válido (por lista)			

a. Grupo = PER1-2

Figura 9.54: Descriptivos asociados al primer grupo.

Grupo = PER1-2

Estadísticos^a

Trabajo 3

N	Válido	34
	Perdidos	0

a. Grupo = PER1-2

Trabajo 3^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	27	79,4	79,4	79,4
	3,0	1	2,9	2,9	82,4
	5,0	2	5,9	5,9	88,2
	7,0	2	5,9	5,9	94,1
	8,0	1	2,9	2,9	97,1
	10,0	1	2,9	2,9	100,0
	Total	34	100,0	100,0	

a. Grupo = PER1-2

Figura 9.55: Descriptivos asociados al primer grupo.

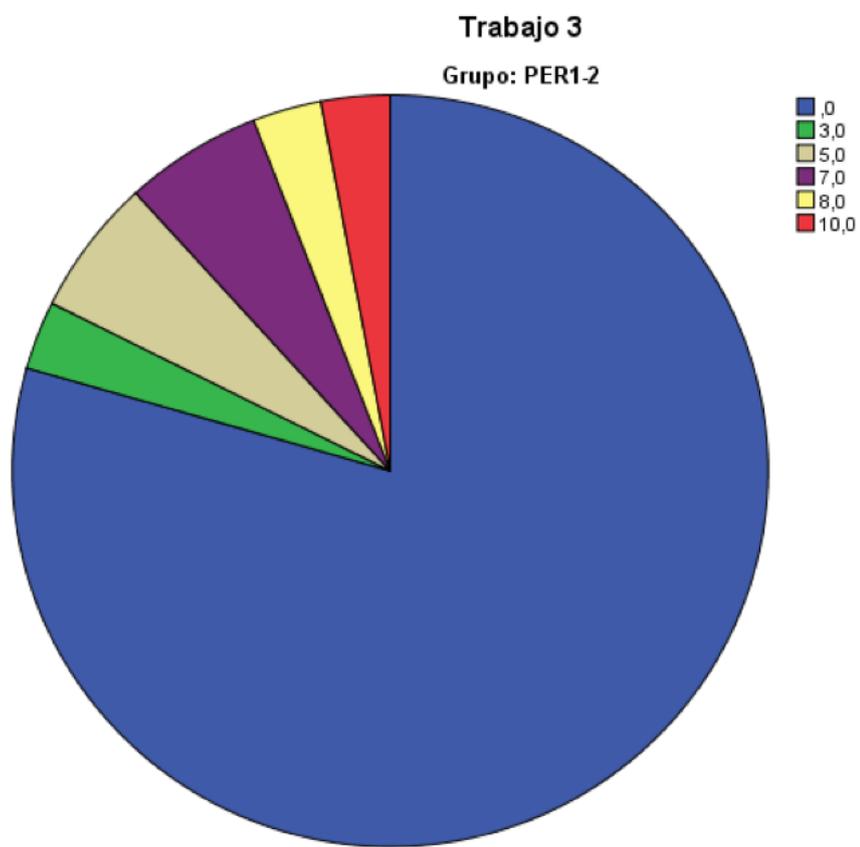


Figura 9.56: Descriptivos asociados al primer grupo.

Por otro lado, los estadísticos descriptivos que arroja el programa SPSS relacionados con el segundo grupo, denotado por “PER3”, que aparecen en las Figuras 9.6.6-9.6.6.

Grupo = PER3

Estadísticos descriptivos^a

	N	Rango	Mínimo	Máximo	Suma
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Trabajo 3	50	10,0	,0	10,0	209,0
N válido (por lista)	50				

Estadísticos descriptivos^a

	Media		Desviación estándar	Varianza	Asimetría
	Estadístico	Error estándar	Estadístico	Estadístico	Estadístico
Trabajo 3	4,180	,5584	3,9483	15,589	,218
N válido (por lista)					

Estadísticos descriptivos^a

	Asimetría	Curtosis	
	Error estándar	Estadístico	Error estándar
Trabajo 3	,337	-1,624	,662
N válido (por lista)			

a. Grupo = PER3

Figura 9.57: Descriptivos asociados al segundo grupo.

Grupo = PER3

Estadísticos^a

Trabajo 3

N	Válido	50
	Perdidos	0

a. Grupo = PER3

Trabajo 3^a

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	,0	19	38,0	38,0	38,0
	2,0	3	6,0	6,0	44,0
	2,5	1	2,0	2,0	46,0
	3,0	1	2,0	2,0	48,0
	4,0	3	6,0	6,0	54,0
	5,0	2	4,0	4,0	58,0
	5,5	3	6,0	6,0	64,0
	7,0	1	2,0	2,0	66,0
	7,5	1	2,0	2,0	68,0
	8,0	5	10,0	10,0	78,0
	8,5	1	2,0	2,0	80,0
	9,0	4	8,0	8,0	88,0
	10,0	6	12,0	12,0	100,0
	Total	50	100,0	100,0	

a. Grupo = PER3

Figura 9.58: Descriptivos asociados al segundo grupo.

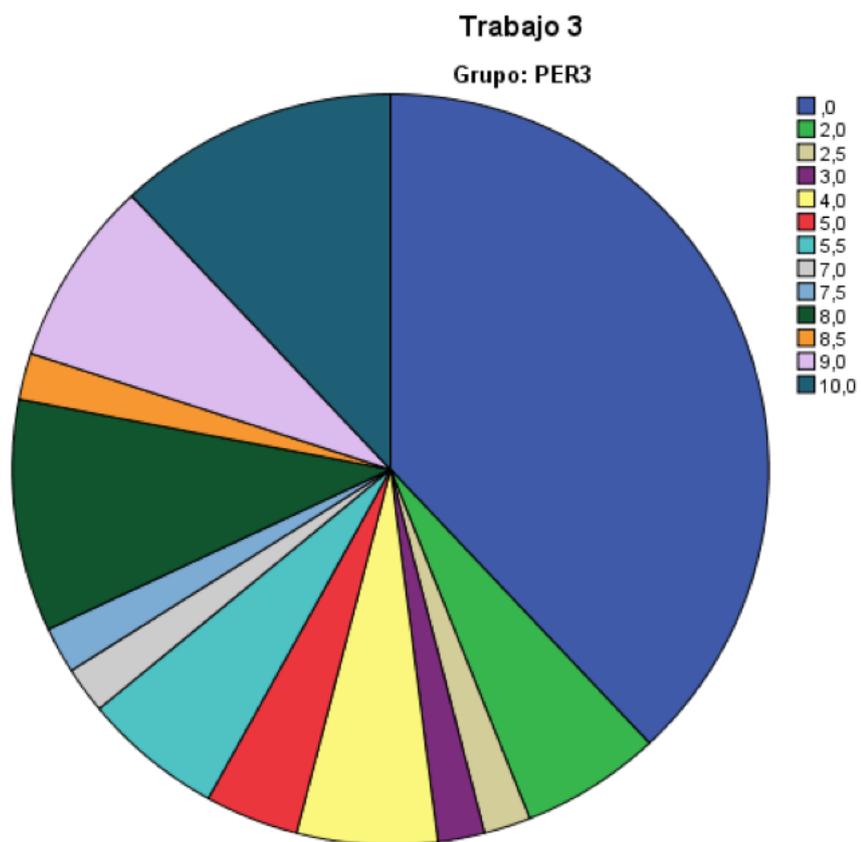


Figura 9.59: Descriptivos asociados al segundo grupo.

Por último, queremos comprobar si las diferencias que vemos con las medias son significativas y, para ello, y teniendo en cuenta que ahora ambos grupos tienen más de 30 observaciones, podemos usar la prueba T de Student.

Prueba T

Estadísticas de grupo					
Grupo		N	Media	Desviación estándar	Media de error estándar
Trabajo 3	PER3	50	4,180	3,9483	,5584
	PER1-2	34	1,324	2,8147	,4827

Prueba de muestras independientes					
		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias	
		F	Sig.	t	gl.
Trabajo 3	Se asumen varianzas iguales	15,226	,000	3,634	82
	No se asumen varianzas iguales			3,870	81,783

Prueba de muestras independientes				
		prueba t para la igualdad de medias		
		Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar
Trabajo 3	Se asumen varianzas iguales	,000	2,8565	,7860
	No se asumen varianzas iguales	,000	2,8565	,7381

Prueba de muestras independientes				
		prueba t para la igualdad de medias		
		95% de intervalo de confianza de la diferencia		
		Inferior	Superior	
Trabajo 3	Se asumen varianzas iguales	1,2928	4,4201	
	No se asumen varianzas iguales	1,3881	4,3249	

Figura 9.60: Prueba T de Student para la comparación de medias de notas obtenidas en el primer trabajo.

El resultado puede verse en la Figura 9.6.6. El resultado que se desprende es, que en el segundo grupo, es decir, el que ha seguido el nuevo método obtiene una nota en media de 2.8565 superior al grupo impartido con la primera metodología y además este resultado es significativo.

Parte IV
Conclusiones

Capítulo 10

Conclusiones y trabajo futuro

Resumen

En este capítulo se presentan las conclusiones finales de la tesis. Para ello se evalúan los distintos objetivos obtenidos de la hipótesis de partida de la tesis, para comprobar si se han cumplido las expectativas iniciales. A continuación, se exponen las líneas de investigación futuras abiertas por esta tesis. Por último, se presentan los trabajos derivados de las investigaciones realizadas a lo largo de la tesis, para dar a conocer a la comunidad científica los avances llevados a cabo.

Verificación, contraste y evaluación de los objetivos

En el primer capítulo de este trabajo se planteó la hipótesis de partida de esta tesis:

La adquisición, de una manera eficaz y eficiente, de competencias complejas, dentro del área de los sistemas dinámicos continuos y discretos es factible. El desarrollo de una nueva metodología y el soporte otorgado por el uso de las tecnologías de la información, gracias al desarrollo de aplicativos orientados a diferentes métodos iterativos, permitirán que el objetivo de aprendizaje de los estudiantes sea alcanzado.

Y un objetivo principal derivado de esta hipótesis de partida:

Diseñar una metodología basada en la aplicación de una herramienta para tratar de paliar los problemas detectados y aplicarla a diferentes métodos iterativos

Este objetivo se ha cumplido, tal y como se ve en el apartado 9.6 “Resultados”.

Para lograr este objetivo se procedió a subdividirlo en objetivos parciales sencillos:

- Objetivo 1.- Diseño y puesta en marcha de una modificación en temporalización como en impartición de las clases en la asignatura “Sistemas dinámicos continuos y discretos”.
- Objetivo 2.- Diseño de una herramienta, que será desarrollada en lenguaje MATLAB, y su aplicación en la asignatura “Sistemas dinámicos continuos y discretos”.
- Objetivo 3.- Mejorar los resultados tanto a nivel cualitativo como cuantitativo en la asignatura “Sistemas dinámicos continuos y discretos” utilizando la herramienta y la nueva forma de impartir la asignatura.
- Objetivo 4.- Encontrar condiciones de convergencia para el método de la secante más suaves que la ya existentes.
- Objetivo 5.- Estudiar el comportamiento dinámico de la secante para diferentes polinomios.
- Objetivo 6.- Encontrar condiciones de convergencia para diferentes métodos y familias de métodos de órdenes altos.
- Objetivo 7.- Estudiar el comportamiento dinámico de diferentes métodos y familias de métodos de órdenes altos.

Otro objetivo, en última instancia, fue por lo tanto responder a la pregunta:

¿Se ha resuelto el problema del alto nivel de abstención en las entregas y se han mejorado los resultados obtenidos por los estudiantes?

Una vez finalizada la investigación, el cumplimiento de los objetivos ha sido correcto, tal y como se describe a continuación:

- Objetivo 1.- Cumplido. La información aportada en el capítulo nueve del documento, confirma la modificación de la temporalización de las clases en la asignatura “Sistemas dinámicos continuos y discretos” para la tercera edición del Máster en Ingeniería Matemática y Computación.
- Objetivo 2.- Cumplido. En el capítulo noveno se puede ver una descripción de la herramienta junto con las características básicas de las interfaces gráficas de usuario construidas y un manual para su utilización. A partir del punto 9.5. se especifica la puesta en marcha de la herramienta en la asignatura.
- Objetivo 3.- Cumplido. En el capítulo nueve se realiza un análisis estadístico exhaustivo de los resultados obtenidos por los estudiantes tanto con anterioridad como posteriormente a la utilización de la herramienta, donde queda patente la mejoría de los resultados después de la puesta en marcha de la herramienta con la nueva forma de impartir la asignatura.
- Objetivo 4.- Cumplido. Durante los capítulos cuarto y quinto de esta tesis se ha investigado para descubrir nuevos criterios de convergencia menos exigentes que los conocidos y estimaciones más precisas sobre las distancias entre dos aproximaciones consecutivas a la solución local de una ecuación no lineal.
- Objetivo 5.- Cumplido. Durante los capítulos cuarto y quinto de esta tesis se ha procedido a un estudio del método de la Secante para polinomios de grados dos y tres respectivamente.
- Objetivo 6.- Cumplido. Especialmente en el capítulo octavo encontramos condiciones de convergencia favorables para un método de tres pasos de orden dieciséis.
- Objetivo 7.- Cumplido. En los capítulos sexto, séptimo y octavo se estudia el comportamiento dinámico de diferentes métodos y familias de métodos de órdenes altos.

Una vez verificada la hipótesis de partida, cumplidos los objetivos parciales y cumplido el objetivo principal, se puede responder a la pregunta ¿Se ha resuelto el problema del alto nivel de abstención en las entregas y se han mejorado los resultados obtenidos por los estudiantes?, de forma AFIRMATIVA. Efectivamente se ha logrado reducir la abstención de las entregas de los tres trabajos planteados en la asignatura y aumentado las notas obtenidas por los estudiantes significativamente. Tanto a nivel de trabajo a calificar como a nivel de nota total obtenida.

Conclusiones obtenidas en el ámbito educativo

Tal y como comentan Ascheri y Pizarro (2006), el uso de software en cálculo numérico es de gran utilidad para los alumnos ya que les permiten interpretar métodos numéricos y adaptarlos a otras situaciones que podrían resultarles de utilidad en su futuro profesional, ya que les confieren la capacidad de afianzar sus conocimientos en programación. En este contexto, consideramos que el aprendizaje se vuelve mucho más dinámico y activo por parte del alumno ya que puede interactuar con el software a su propio ritmo de aprendizaje a la par que obtiene una respuesta más inmediata.

Resulta evidente destacar que consideramos necesario que, para el correcto uso de este tipo de software, el alumno tiene que tener una formación previa de programación y del mismo modo, ha de tener unos conocimientos previos de métodos numéricos. Del mismo modo el rol del docente cambia al incluir este tipo de herramientas en el aula ya que pasa de ser el trasmisor del conocimiento al mediador del proceso de aprendizaje de los alumnos, es decir, el guía que indica a los alumnos cómo han de seguir en el caso de no saber hacerlo. Resulta clave, por lo tanto, la formación del profesorado en el uso de este tipo de software ya que su papel en el proceso de aprendizaje de los alumnos se verá afectado por el dominio que tenga en el empleo de los mismos.

Del mismo modo, consideramos que la importancia del software reside no solo en que permita a los alumnos resolver actividades relacionadas con el cálculo numérico si no en que le sirva de utilidad para ser capaz de afrontar y comprender otras situaciones y conocer los errores que se van cometiendo y dónde se cometen.

Así mismo, se considera que es importante hacer especial hincapié en que los alumnos sean conscientes de que lo importante del uso de software para el cálculo numérico es que tengan siempre muy en cuenta cómo funciona éste a medida que se acerca a la solución que se desea. Es decir, es importante que valoren la importancia de la gráfica que se obtiene de la función que se está analizando, así como el intervalo de análisis o las condiciones de convergencia y no se centren únicamente en el resultado final del programa.

No obstante, tanto la metodología empleada como la herramienta utilizada, han permitido validar la hipótesis de esta tesis.

Trabajo futuro

Esta tesis sienta las bases para la creación de una herramienta que mejore la comprensión y los resultados obtenidos por los estudiantes del Máster en Ingeniería Matemática y Computación de la Universidad Internacional de La Rioja. A continuación, se exponen algunas de las líneas de investigación abiertas:

- Objetivo 1.- Aumentar el número de métodos iterativos estudiados con la herramienta software.
- Objetivo 2.- Realizar un empaquetado de la herramienta software para poder garantizar su libre difusión a la comunidad científica. Para ello sería conveniente exportar el código MATLAB a otro lenguaje de programación y alojar la aplicación en un servidor.
- Objetivo 3.- Continuar con el desarrollo de la herramienta para continuar y asentar la nueva metodología desarrollada durante esta tesis doctoral.
- Objetivo 4.- Encontrar condiciones de convergencia para diferentes métodos y familias de métodos.
- Objetivo 5.- Estudiar la dinámica de funciones. No quedarse sólo con la dinámica de los polinomios.

