



UNIVERSIDAD
DE MÁLAGA



UNIVERSIDAD DE MÁLAGA
Departamento de Matemática Aplicada

UNIVERSIDAD DE LA HABANA
Departamento de Matemática

TESIS DOCTORAL

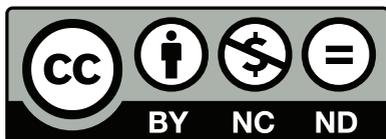
Métodos de integración geométrica para sistemas con función de Lyapunov

Yadira Hernández Solano

Málaga, 2014

AUTOR: Yadira Hernández Solano

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está sujeta a una licencia Creative Commons:

Reconocimiento - No comercial - SinObraDerivada (cc-by-nc-nd):

[Http://creativecommons.org/licenses/by-nc-nd/3.0/es](http://creativecommons.org/licenses/by-nc-nd/3.0/es)

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es

D. Miguel Alejandro Atencia Ruiz, Profesor Titular del Departamento de Matemática Aplicada de la Universidad de Málaga, y Dña. Aymée Marrero Severo, Profesora Titular del Departamento de Matemática Aplicada de la Universidad de La Habana, como directores de la Tesis Doctoral de Dña. Yadira Hernández Solano,

CERTIFICAN:

Que Dña. Yadira Hernández Solano, Licenciada en Matemáticas por la Universidad de La Habana, ha realizado dentro del Programa de Doctorado en Ingeniería Mecánica y Eficiencia Energética de la Escuela Politécnica Superior de la Universidad de Málaga en cotutela con la Universidad de La Habana, bajo su dirección, el trabajo de investigación correspondiente a su Tesis Doctoral titulado:

Métodos de integración geométrica para sistemas con función de Lyapunov

Revisado el presente trabajo, estiman que puede ser presentado al Tribunal que ha de juzgarlo. Y para que conste, autorizan la defensa de esta Tesis Doctoral en la Universidad de Málaga.

Málaga, a 21 de octubre de 2.014

Dr. Miguel Atencia Ruiz

Dra. Aymée Marrero Severo

Para mi madre

“Quien no haya experimentado la irresistible atracción de la ciencia, no podrá comprender su tiranía.”

Frankenstein
MARY SHELLEY

Agradecimientos

Cuando llega el momento de escribir los agradecimientos de una tesis, momento que indica que estás a punto de poner fin a una etapa importante de tu vida, parecería que tienes muchas cosas que decir y sin embargo te quedas en blanco. Quién me conoce sabe que quedarme sin palabras es casi imposible para mí. ¿Cómo encontrar las palabras indicadas para expresar todo mi agradecimiento a esas personas que han influido de una u otra manera a que este sueño este a punto de volverse realidad? Llegas a la conclusión de que es imposible, así que aquí les dejo el resultado de mi mayor esfuerzo.

Quisiera comenzar entonces agradeciendo a mis directores. Al Dr. Miguel Atencia Ruiz, a quién no le gusta ser protagonista pero en esta obra ha jugado un papel principal, ha sido profesor, compañero, ha sido sobre todo paciente, fijándose una y otra vez en el mínimo detalle. Gracias por aceptar el reto de trabajar conmigo, sin conocerme. Gracias por su dirección y por haber sido capaz de guiarme sin imponerme el camino. Gracias por ayudarme siempre que lo he necesitado. A la Dra. Aymée Marrero, a quién también tengo el placer de llamar amiga, por su cariño incondicional y por sus consejos.

Esta tesis es el fruto del trabajo de varios años, nacido de la colaboración entre la Universidad de Málaga y la Universidad de La Habana, es por esto que quisiera agradecer al Dr. Héctor de Arazoza y al Dr. Gonzalo Joya, artífices de esta colaboración, gracias a la que ha sido posible la realización de esta tesis en cotutela entre ambas universidades. Gracias también a ambos, por brindarme su ayuda en innumerables ocasiones.

Quisiera agradecer también a la Agencia Española de Cooperación Internacional para el Desarrollo (AECID - proyectos D/017218/08, D/023835/09, D/030223/10), a la Asociación Universitaria Iberoamericana de Postgrado (AUIP) y al Campus de Excelencia Internacional Andalucía Tech a través de las ayudas de doctorado de la Universidad de Málaga, gracias a las que este trabajo de investigación ha sido financiado.

Muchas gracias a todos mis colegas de la Universidad de la Habana, los que una vez fueron mis profesores y contribuyeron a mi formación profesional. Un agradecimiento especial para la profesora Angela León, por brindarme todo su apoyo durante la realización de la tesis, por estar siempre dispuesta a intercambiar ideas que me fueron muy útiles, por ponerle tanto amor y dedicación a su trabajo.

A Teresita Noriega, por su cariño y su preocupación constante. A Inmaculada de las Peñas por siempre estar pendiente y seguir de cerca la evolución de esta tesis y al Dr. Francisco Villatoro por asumir su tutoría.

A mi familia, el mayor de los agradecimientos. A mi madre, por no dudar ni un minuto que este día llegaría, por ayudarme a que fuese posible y por no dejarme sola ni un momento. A mi hijo, mi mayor tesoro, por entender la distancia, por en tantas ocasiones decir: "Mamá, no te preocupes, que yo estoy bien". A mi hermana, a Eduardo y a Gaby, por ser una parte especial de mi vida. A mis abuelos, que ya no están pero recuerdo siempre, sobre todo cuando decían: "Niña, estudia que lo único que nunca te podrán quitar, es lo que sabes". A mi padre, cuyo recuerdo me inspira cada día. A Jose por hacerme ver la vida de otra manera, por compartir sus días conmigo, por regalarme eso que nunca recuperas, "Tu tiempo".

A mis amigos: Lilian, Jorge, Cueto, Celia, Adriana, Wilfredo, Salomón y Liudmila, a los nuevos, a los viejos, a los de siempre y por siempre, a los que sabes que aunque estes años sin ver, puedes contar con ellos.

Muchas Gracias!!!

Resumen

Esta tesis estudia métodos de integración numérica geométrica, específicamente, métodos numéricos diseñados para conservar la función de Lyapunov de un sistema dinámico. En otras palabras, se proponen aproximaciones numéricas cuya energía disminuye, al igual que en la ecuación diferencial de origen. Por lo tanto, se centra la atención en ecuaciones diferenciales ordinarias (EDO) que se puedan reescribir en la forma de un gradiente lineal, es decir, la parte derecha de la EDO se escribe como el producto de una matriz definida negativa y el gradiente de la función de Lyapunov del sistema. Con este objetivo, se implementan métodos de gradiente discreto para la integración numérica de un sistema de EDO. Los métodos de gradiente discreto son esquemas implícitos, que dan lugar a sistemas dinámicos discretos que poseen la misma función de Lyapunov que el sistema continuo. En este trabajo se demuestra que es posible reescribirlos de forma explícita imponiendo determinadas condiciones sobre los parámetros del método. La formulación de un método de gradiente discreto es sencilla, pues basta con establecer la relación entre la aproximación de una matriz definida negativa y un gradiente discreto, el cual tiene propiedades similares a las del gradiente de la función de Lyapunov del sistema continuo. En principio, este procedimiento produce métodos de primer orden, pero el análisis muestra el camino para el diseño de métodos de orden superior. El método propuesto se aplica a la discretización de las redes neuronales de Hopfield, con y sin autopesos. Este sistema se considera como un caso de prueba adecuado, ya que permite definir sistemas de alta dimensión fácilmente y, en la mayoría de las aplicaciones, la conservación de su función de Lyapunov es más importante que la exactitud de las trayectorias particulares. Los resultados son corroborados por medio de experimentos numéricos, en los que el método de gradiente discreto propuesto se compara con la regla de Euler, algunos métodos comerciales de Runge-Kutta y un método de proyección, diseñado específicamente para la conservación de la función de Lyapunov. Como predice la teoría, los métodos de gradiente discretos preservan la función de Lyapunov, mientras que los métodos convencionales no lo hacen, ya que al discretizar el sistema, en algunos casos aparecen soluciones periódicas y en otros la función de Lyapunov no decrece a

lo largo de las trayectorias.

También como parte de la tesis se propone la extensión de alto orden del método de gradiente discreto anteriormente mencionado. En esencia, se considera la composición del método de gradiente discreto de primer orden con su adjunto, dando como resultado un método de gradiente discreto de segundo orden. El método adjunto es la aplicación inversa del método original con el tamaño de paso negativo. Por último, se diseñan experimentos numéricos con el objetivo de validar el funcionamiento del método propuesto, en los que se compara con el método de gradiente discreto de primer orden y la regla trapezoidal. En estos experimentos se muestra la capacidad de este método tanto para conservar la función de Lyapunov del sistema continuo como para aproximarla con precisión. En este caso, también se utilizan las redes de Hopfield como caso de estudio, considerando la ausencia de autopesos.

Teniendo en cuenta los resultados de las simulaciones en cada caso, se concluye que tanto el método de gradiente discreto como su extensión son una opción válida y prometedora para la discretización de sistemas de gradiente con función de Lyapunov.

Palabras clave

Integración Numérica Geométrica; Sistemas Dinámicos; Función de Lyapunov; Estabilidad; Métodos Numéricos para EDOs; Gradientes Discretos.

Abstract

This thesis studies numerical methods that preserve a Lyapunov function of a dynamical system, i.e numerical approximations whose energy decreases, just like in the original differential equation. Therefore we will focus on Ordinary Differential Equations (ODEs) that can be rewritten in the form of a linear gradient, i.e, the right-hand side of the ODE is written as the product of a negative-definite matrix and the gradient of the Lyapunov function of the system. With this aim, a discrete gradient methods is implemented for numerical integration of a system of ordinary differential equations. Discrete gradient methods are implicit methods, which result in discrete dynamical systems that possess the same Lyapunov function as the continuous system. In this work it is shown that they can be computed explicitly for particular families of systems that fulfil some mild assumptions. The formulation of discrete gradient methods is straightforward, since they can be written by means of the approximation of a negative-definite matrix and a discrete gradient, which has properties that are similar to those of the gradient of the Lyapunov function. In principle, this procedure yields first order methods, but the analysis paves the way to designing higher-order methods. The proposed methods are applied to continuous Hopfield neural networks, both with and without self-weights, which are regarded as a suitable test case since high-dimensional systems can easily be defined and preservation of the Lyapunov function is more important than accuracy of particular trajectories in most applications. The results are illustrated by means of numerical experiments, where the discrete gradient methods is compared to the Euler rule, standard Runge-Kutta methods and a projection method specifically designed for the conservation of the Lyapunov function. As predicted by the theory, discrete gradient methods preserve the Lyapunov function, whereas conventional methods fail to do so, since either periodic solutions appear or the energy does not decrease.

Also as part of this thesis, the extension of discrete gradient methods to higher order is proposed. In essence we are considering a discrete first order gradient together with its adjoint, giving as a result a discrete gradient method of second order. The adjoint method is the inverse map of the original method with negative step si-

ze. Finally, the proposed method is implemented in numerical experiments, where it is compared to the first order discrete gradient methods and to the trapezoidal rule. These experiments show the method characteristic feature, namely preservation of the Lyapunov function, as well as its ability to approximate the continuous system precisely. In this case, the proposed methods are also applied to the discretization of the Hopfield neural network, for which only the case without self-weights has been considered.

Given the simulation results in each case, it is concluded that both the discrete gradient method and its extension are a valid and promising choice for the discretization of gradient systems with Lyapunov function.

Keywords

Geometric Numerical Integration; Dynamical Systems; Lyapunov Function; Stability; Numerical Methods for ODEs; Discrete Gradients.

Índice general

Resumen	IX
Abstract	XI
Índice general	XIII
1. Preliminares	1
1.1. Introducción	1
1.2. Sistemas dinámicos	8
1.2.1. Introducción a las ecuaciones diferenciales	8
1.2.2. Estabilidad de sistemas de ecuaciones diferenciales lineales	11
1.2.3. Retrato de fases de sistemas de ecuaciones diferenciales li-	
neales con matriz de coeficientes constantes	12
1.2.4. Análisis de la estabilidad por linealización	16
1.2.5. Formalización de los sistemas dinámicos generales	18
1.2.6. Función de Lyapunov	20
1.2.7. Formulación de gradiente para sistemas con función de Lya-	
punov	21
1.3. Métodos numéricos para ecuaciones diferenciales ordinarias	23
1.3.1. Aproximación numérica y sus propiedades	23
1.3.2. Método de Euler	28
1.3.3. Formulación implícita y explícita de métodos numéricos	29
1.3.4. Métodos de paso simple o multipaso	30

1.3.5. Métodos de Runge-Kutta	31
1.4. Conclusiones	33
2. Integración geométrica	35
2.1. Integración numérica geométrica	36
2.2. Adjunto de un método numérico	38
2.3. Métodos de composición	39
2.4. Método de proyección que conserva la función de Lyapunov	41
2.4.1. Definición del método de proyección	41
2.4.2. Ilustración del funcionamiento del método	46
2.5. Conclusiones	51
3. Gradientes discretos y discretización de sistemas de gradiente lineal	53
3.1. Formulaciones de gradiente para sistemas dinámicos y gradientes discretos.	54
3.2. Algunos gradientes discretos.	57
3.3. Método de integración	58
3.4. Formulación de un método de gradiente discreto	60
3.4.1. Métodos de gradiente discreto para sistemas escalares simples	60
3.4.2. Formulación de un método de gradiente discreto como un método numérico explícito	66
3.5. Análisis del orden del método	68
3.5.1. Análisis del orden del método de gradiente discreto	69
3.5.2. Análisis del orden del método, para sistemas escalares simples	73
3.6. Experimentos numéricos	75
3.7. Conclusiones	84
4. Discretización de las redes de Hopfield con gradiente discreto	87
4.1. Las redes de Hopfield	88
4.2. Aplicación del gradiente discreto a las redes de Hopfield	93
4.2.1. Formulación del sistema en forma de gradiente lineal	94
4.2.2. Definición de los parámetros del método	96

4.2.3. Discretización mediante gradiente discreto	97
4.3. Generalización del gradiente discreto al sistema de Hopfield de dimensión n	99
4.4. Análisis del orden de la aplicación al sistema de Hopfield	101
4.5. Método de gradiente discreto para el sistema de Hopfield con autopesos	104
4.6. Método de gradiente discreto adjunto	110
4.7. Conclusiones	111
5. Experimentos numéricos	113
5.1. Método de gradiente discreto	114
5.1.1. Enfoque cualitativo. Comparación con el método de Euler .	116
5.1.2. Enfoque cuantitativo: análisis del error	122
5.1.3. Comparación con métodos de Runge-Kutta	125
5.1.4. Comparación con el método de proyección	128
5.2. Método de composición	131
5.3. Conclusiones	137
6. Conclusiones	141
Bibliografía	145

Capítulo 1

Preliminares

Resumen del capítulo

Este capítulo está compuesto de tres secciones. En la primera se hace una breve introducción al tema en el que se enmarca la tesis, la integración geométrica, haciendo un breve recorrido por los antecedentes y los principales resultados que aparecen en la literatura hasta la fecha. Luego, se introduce el sistema que se utilizará como caso de estudio, las redes de Hopfield. Posteriormente y con el objetivo de que la tesis sea autocontenida, se describen también en este capítulo los fundamentos de las técnicas matemáticas usadas en diferentes momentos del trabajo, estos resultados se recogen en las otras dos secciones. Una de estas secciones estará dedicada al estudio de las ecuaciones diferenciales ordinarias vistas como un caso particular de los sistemas dinámicos; mientras que la otra, se dedicará al estudio de métodos numéricos diseñados para la resolución de ecuaciones diferenciales ordinarias.

1.1. Introducción

Los conceptos de derivada e integral nacieron conjuntamente con las ecuaciones diferenciales y el cálculo infinitesimal [20] en el siglo XVIII y están estrechamente relacionados con la solución de problemas mecánicos y geométricos, entre otros. En

muchas ocasiones, resolver estos problemas no se reduce a encontrar una primitiva o una integral sino a resolver una ecuación diferencial. Resulta de vital importancia señalar que prácticamente cada ley física que rige nuestro universo puede ser modelada matemáticamente mediante una ecuación diferencial, reflejando cada una de las posibles evoluciones del sistema en cuestión, mediante las infinitas soluciones que presente dicha ecuación diferencial. Si se trabaja con un problema de valores iniciales, también llamado problema de Cauchy, se estaría identificando realmente la evolución del sistema físico que se está estudiando. Si se tiene en cuenta el papel fundamental que juegan las ecuaciones diferenciales en la aplicación de la matemática a otras ciencias, resulta lógico el gran esfuerzo que han dedicado grandes matemáticos de la talla de Newton, Leibniz entre otros [18] a la tarea de resolverlas o integrarlas. La incapacidad de los métodos elementales de integración para resolver algunos de los problemas más importantes motivó la introducción de técnicas de mayor potencia y generalidad, entre las que se puede citar el trabajo con series y los métodos numéricos [25] de los que se hablará posteriormente en este capítulo. A pesar de ser incuestionable la importancia y utilidad del análisis numérico para la resolución aproximada de sistemas de ecuaciones diferenciales, se ha demostrado que incluso los métodos numéricos más sofisticados son incapaces de respetar la dinámica del sistema continuo, a menos que se trabaje con un tamaño de paso suficientemente pequeño, siendo este un inconveniente a tener en cuenta cuando de la resolución de problemas reales se trata. Esta incapacidad de los métodos numéricos tradicionales, por llamarlos de alguna manera, ha dado lugar al surgimiento de una nueva línea de investigación, conocida como *Integración Geométrica*.

La integración geométrica [17], es una línea activa de investigación desde diversos puntos de vista, la cual, enlaza la metodología del análisis de sistemas dinámicos con el diseño de métodos numéricos [51] que preserven las propiedades cualitativas del sistema continuo. La motivación de la creación y desarrollo de algoritmos que preserven estas propiedades, en diferentes tipos de problemas, viene dada independientemente en cada una de las áreas de investigación, como por ejemplo: la astronomía, la dinámica molecular, la mecánica y la física teórica. En otras palabras, se puede decir que se diseña el método numérico teniendo en cuenta cual o cuales

propiedades del sistema continuo es interesante conservar, teniendo en cuenta el área de investigación y las principales aplicaciones del sistema en cuestión. Asimismo, los resultados obtenidos en las distintas áreas aplicadas contribuyen a su vez al desarrollo de técnicas más precisas en el campo del análisis numérico.

En las últimas décadas, ha surgido un creciente interés por el estudio de métodos numéricos que preserven las propiedades geométricas del flujo de una ecuación diferencial. En este sentido, el objetivo fundamental es considerar las características cualitativas de las trayectorias del sistema dinámico, por ejemplo: conservación de la energía, estabilidad, conservación del Hamiltoniano, entre otras. El problema sería entonces el diseño de métodos numéricos, de manera tal que la trayectoria discreta del método, tenga también las mismas propiedades.

Dentro del campo de la integración geométrica, se pueden encontrar un gran número de resultados relacionados con el estudio de sistemas Hamiltonianos [45]. Estos sistemas pueden considerarse desde dos puntos de vista. Por una parte, son sistemas que preservan una determinada magnitud, específicamente, la energía y por otra, pueden verse como sistemas cuyas soluciones evolucionan o están definidas en una cierta variedad. Además de los resultados relacionados con sistemas Hamiltonianos, métodos simplécticos y de proyección, existen una gran variedad de resultados relacionados con la conservación de invariantes cuadráticos [47]. Sin embargo, cuando se refiere a la conservación de la función de Lyapunov de un sistema de gradiente los resultados son escasos, centrándose principalmente en dos direcciones: métodos de gradiente discreto [39] y métodos de proyección [12]. Resulta sorprendente la escases de resultados referentes a la conservación de la estabilidad de los sistemas dinámicos, si se tiene en cuenta que el análisis dinámico de ecuaciones diferenciales ordinarias dista de ser nueva en el campo del análisis numérico. De hecho, el concepto de A-estabilidad [29] equivale a la preservación de la estabilidad de la solución de sistemas escalares simples como casos de prueba. Desde este punto de vista, se puede ver la conservación de la función de Lyapunov de un sistema de ecuaciones diferenciales como una generalización del concepto de A-estabilidad para sistemas no lineales. El objeto de esta investigación es diseñar métodos numéricos para la conservación de la función de Lyapunov de un sistema de gradiente.

Los métodos de gradiente discreto fueron propuestos inicialmente con el objetivo de conservar primeras integrales [41]. Posteriormente, se ampliaron para incluir sistemas de gradiente [37], teniendo en cuenta que la ecuación que define un sistema de gradiente puede reescribirse en la forma de gradiente lineal. Para reescribir un sistema en la forma de gradiente lineal, basta con expresarlo como el producto de una matriz definida negativa y el gradiente de la función de Lyapunov del sistema. Basándose en esta idea se puede afirmar que el diseño de los métodos de gradiente discreto es relativamente sencillo: definir una aproximación de la matriz definida negativa y del gradiente de la función de Lyapunov. A la aproximación del gradiente de la función de Lyapunov se le define como gradiente discreto.

En el próximo capítulo donde se realiza un estudio detallado de estos métodos se verá que al definir el gradiente discreto a partir del gradiente de la función de Lyapunov del sistema, ambos tendrán propiedades similares. Por otra parte, teniendo en cuenta su metodología de construcción, los métodos de gradiente discreto son métodos numéricos implícitos que considerados como sistemas dinámicos discretos, conservan la función de Lyapunov del sistema continuo. Se pueden encontrar otros resultados en [38] pero la experiencia práctica sobre estos métodos es limitada y la aplicación de estos a la discretización de sistemas reales aparece escasamente referenciada en la literatura.

Por su parte, los métodos de proyección están diseñados de manera tal que al proyectar la solución del sistema sobre una determinada variedad, la función de Lyapunov evaluada en este valor proyectado, disminuye su valor. De esta manera se garantiza que la trayectoria discreta de dicha función sea decreciente a lo largo de las trayectorias. Se sabe que los métodos de proyección se derivan de esquemas de Runge-Kutta explícitos siempre que los coeficientes que definen el método sean positivos [11], aunque esta restricción se elimina más adelante [12]. Es válido señalar que en principio la formulación de los métodos de proyección responde a un esquema explícito, que son menos complejos computacionalmente que los implícitos. Sin embargo, su formulación requiere de la resolución de una ecuación escalar no lineal en cada paso del método por lo que se complejiza grandemente su implementación práctica. Por otra parte, hay trabajos que han estudiado la aplicación de métodos de

Runge-Kutta a la integración de sistemas de gradiente [16]. En este trabajo se demuestra que algunos esquemas implícitos de Radau, originalmente propuestos para la discretización de sistemas *stiff* y sistemas hamiltonianos, también conservan la función de Lyapunov de sistema continuo imponiendo condiciones y restricciones al tamaño de paso. Tanto los métodos de proyección como los esquemas implícitos de Runge-Kutta anteriormente mencionados están sustentados por teoremas no constructivos, por lo que la conservación de la función de Lyapunov solo se garantiza para determinados tamaños de paso. En otras palabras, con el fin de conservar la función de Lyapunov del sistema continuo es necesario realizar un ajuste *ad-hoc* del tamaño de paso. En resumen, aunque hay que reconocer que estos métodos son prometedores, su implementación es complicada y puede elevar considerablemente el coste computacional, por lo que no son adecuados para todos los casos. Es por esto que se propone el uso de métodos de gradiente discreto en lo que a la conservación de la función de Lyapunov se refiere. Aunque en el desarrollo de la tesis se realizan estudios comparativos de estos con algunos de los métodos anteriormente mencionados.

Como campo de aplicación y caso de estudio de las técnicas propuestas, se adopta la familia de sistemas denominados redes de Hopfield [26]. La redes de Hopfield son sistemas dinámicos, inspirados en las redes neuronales biológicas. Estos sistemas han sido propuestos como algoritmos computacionales para resolver problemas de memoria asociativa [24], optimización combinatoria [5] y estimación de parámetros [6], entre otras numerosas aplicaciones. Durante el desarrollo del trabajo se diseña un método de gradiente que por construcción, preserva la función de Lyapunov que presenta el sistema de Hopfield, motivo por el cual desde el punto de vista dinámico se considera un sistema estable.

Existen algunos trabajos que han abordado ya el problema desde este punto de vista. A partir del trabajo previo realizado sobre el sistema continuo, se formula la red discreta como un método numérico no estándar que resuelve la ecuación continua, dependiendo del tamaño de paso de discretización. Desde el punto de vista numérico, la red discreta es una aproximación de primer orden de la continua. El análisis dinámico muestra que se preservan los puntos fijos del sistema, así como su estabilidad o inestabilidad. Sin embargo, en general, no puede garantizarse la conser-

vación de la función de Lyapunov, lo que se demuestra mediante la construcción de soluciones periódicas. En resumen, la destrucción de la estructura de gradiente, dificulta la capacidad de estos sistemas como algoritmos de optimización. Ya en estos trabajos previos, mediante experimentos numéricos, se exploró la posibilidad de utilizar métodos numéricos alternativos que preserven la función de Lyapunov así como las demás propiedades favorables de la red continua.

El trabajo propone la búsqueda de métodos numéricos para sistemas de ecuaciones diferenciales que preserven la función de Lyapunov. El hallazgo de métodos numéricos que preserven la función de Lyapunov es un tema muy interesante, complicado y prácticamente inexplorado, existiendo muy pocos estudios profundos y experimentos computacionales realizados en este sentido, por lo que resulta interesante profundizar el estudio en esta línea.

En este trabajo se abordará el problema de la conservación de la función de Lyapunov mediante el estudio de los métodos de gradiente. Se centra la atención en los métodos de gradiente discreto, trabajando principalmente con el gradiente discreto del incremento de las coordenadas, introducido en [38].

Algunos resultados preliminares de este trabajo se muestran en [23] aunque luego se extienden los resultados obtenidos a sistemas dinámicos más generales. Se sigue el siguiente esquema, inicialmente se muestra la metodología del diseño de métodos de gradiente discreto y su funcionamiento al integrar sistemas escalares simples. Luego, se introducen las redes de Hopfield y se centra la atención en la discretización de éstas. Para esto se diseña un método de gradiente discreto, escogiéndolo los parámetros de manera tal que el método obtenido se puede reescribir explícitamente. Posteriormente se generaliza este resultado, demostrando que bajo ciertas condiciones de los parámetros, siempre es posible reescribir el método de gradiente discreto, que por definición es implícito, en un método explícito. Por otra parte, se realiza un estudio detallado del método de proyección propuesto en [12], con vistas a realizar una comparación justa entre este y el método de gradiente discreto propuesto. En el próximo capítulo, se integra un sistema escalar simple mediante el método de proyección anteriormente mencionado, de esta manera se evidencian algunas deficiencias de este método, que se interpretan como desventajas respecto al método de gradiente dis-

creto. Se realiza un análisis teórico y computacional de los resultados. Se valida el funcionamiento del método de gradiente discreto obtenido comparándolo con la regla de Euler explícita, algunos esquemas comerciales de Runge-Kutta y el método de proyección antes mencionado, los resultados se muestran en el capítulo 5. Posteriormente, se propone la extensión a alto orden del método de gradiente discreto. Se realiza una nueva serie de experimentos numéricos, esta vez comparando el método de gradiente discreto con su extensión donde se observa que si bien las soluciones obtenidas por ambos métodos son cualitativamente similares, la extensión minimiza el error local. Se utilizó matlab tanto para implementar los métodos propuestos como para realizar las simulaciones en las que se valida su funcionamiento comparando con otros métodos conocidos.

El trabajo está compuesto por seis capítulos. Este primer capítulo recoge la introducción general de la tesis así como algunas definiciones y resultados importantes relacionados con la teoría de sistemas dinámicos y el análisis numérico. En el segundo capítulo se introducen algunos conceptos fundamentales relacionados con la integración geométrica, así como algunos métodos de interés. En el tercer capítulo se introducen conceptos relacionados con los gradientes discretos, ya que se ha demostrado la capacidad de estos métodos para conservar algunas propiedades cualitativas de los sistemas dinámicos. Ya en el cuarto capítulo, luego de introducir las redes de Hopfield como caso de estudio puntualizando su importancia en diversas aplicaciones, se discretizan mediante un método de gradiente discreto. También en este capítulo se realiza un análisis del orden del método propuesto. Posteriormente, a partir de este método y utilizando algunas de las técnicas de integración geométrica vistas en el capítulo dos, se propone una extensión de alto orden de este. El capítulo 5 por su parte, se dedica a la validación de los métodos propuestos, mostrando el resultado de las simulaciones diseñadas en orden de mostrar su funcionamiento y ventajas respecto a otros métodos conocidos. Finalmente en el sexto capítulo se dan las conclusiones del trabajo y se dejan claras posibles líneas de investigación para trabajos futuros.

1.2. Sistemas dinámicos

Esta sección está dedicada al estudio de las ecuaciones diferenciales ordinarias vistas como un caso particular de los sistemas dinámicos. El capítulo comienza con un breve recuento sobre el surgimiento y desarrollo de las ecuaciones diferenciales. Luego se listan una serie de definiciones y resultados que aparecen en la literatura especializada, prestando atención especial a los sistemas de ecuaciones diferenciales ordinarias. Teniendo en cuenta que en la práctica muchas veces es imposible obtener la solución analítica de los sistemas de ecuaciones diferenciales, se introducen los conceptos necesarios para el análisis de la estabilidad de estos, tanto para el caso lineal, como para el no lineal. Se introducen además los conceptos de flujo y función de Lyapunov de una ecuación diferencial ordinaria. Finalmente, se establece la correspondencia entre los sistemas de gradiente y los sistemas con función de Lyapunov.

1.2.1. Introducción a las ecuaciones diferenciales

El surgimiento de las ecuaciones diferenciales tiene lugar conjuntamente con el desarrollo del cálculo diferencial e integral. Fue Leibniz el primero en adjudicarle el término de *ecuación diferencial* a la ecuación en donde la incógnita es una función y aparece junto a sus derivadas. A pesar de esto, no es hasta los siglos XVIII y XIX que las ecuaciones diferenciales llegan a considerarse como un elemento fundamental de las aplicaciones de la matemática a otras ciencias. Con el desarrollo de las ciencias en estos siglos se amplían grandemente el número de problemas de diferentes ramas que eran modelables y resolubles usando ecuaciones diferenciales.

Se denominan ecuaciones diferenciales a aquellas ecuaciones cuyas incógnitas son funciones de una o varias variables y de sus derivadas. En otras palabras, una ecuación diferencial es una ecuación que relaciona los diferenciales de una o varias funciones, mediante una ley dada. En el caso en que la incógnita de una ecuación diferencial sea función de una variable se está en presencia de una ecuación diferencial ordinaria (EDO) y si es función de varias variables la ecuación diferencial será una ecuación en derivadas parciales (EDP).

Los sistemas de EDO pueden contener explícitamente o no la variable independiente t . Usualmente a esta variable se le asigna un significado físico, el tiempo. En el caso en que dicha variable no aparezca de manera explícita, el sistema de EDO se dice autónomo, cuya formulación general viene dada por:

$$\frac{dy_i}{dt} = f_i(y_1, \dots, y_n), \quad i = 1, \dots, n \quad (1.1)$$

Las funciones $f_i(y_1, \dots, y_n)$ están definidas en un cierto conjunto abierto Ω del espacio real de dimensión n , mientras que las variables (y_1, \dots, y_n) denotan las coordenadas de un punto de dicho espacio. Este conjunto Ω se conoce como *espacio de fases* del sistema. Esta formulación también puede escribirse en forma vectorial:

$$\frac{d\mathbf{y}}{dt} = F(\mathbf{y}) = \begin{pmatrix} f_1(\mathbf{y}) \\ \vdots \\ f_n(\mathbf{y}) \end{pmatrix}, \quad \mathbf{y} = (y_1, \dots, y_n) \quad (1.2)$$

Hay que mencionar que el trabajo con sistemas autónomos no introduce pérdida de generalidad. En efecto, dado un sistema no autónomo, bastaría introducir una nueva variable $y_{n+1} = t$ y considerar que las funciones f_i dependen de vectores (y_1, \dots, y_{n+1}) de un espacio de fases aumentado, introduciendo así al modelo la dependencia del tiempo t .

Una solución del sistema (1.1) es una función $\mathbf{y} = \varphi(t)$ definida sobre el intervalo $[t_1, t_2]$ tal que al sustituirla en las ecuaciones del sistema en el lugar que ocupa la variable \mathbf{y} proporciona una identidad en todo el intervalo, o sea:

$$\frac{d\varphi}{dt} \equiv f(\varphi(t)) \quad \forall t \in [t_1, t_2]$$

Estrechamente vinculado a este concepto está el de órbita: específicamente se tiene que una órbita será la curva en el espacio de fases Ω dada por la imagen de una solución φ extendida a todo el intervalo $[t_1, t_2]$.

El estudio del conjunto de soluciones de una ecuación diferencial se puede abordar analizando el comportamiento de determinadas soluciones particulares, a saber,

las que cumplen la siguiente condición:

$$\varphi_i(t_1) = \varphi_i(t_2), \quad i = 1, \dots, n \quad \text{para} \quad t_1 \neq t_2 \quad (1.3)$$

Existen dos tipos de soluciones de esta clase:

1. La solución estacionaria o posición de equilibrio.
2. La solución periódica.

Haciendo uso del teorema de existencia y unicidad se supone que $y_i = \varphi_i(t)$ con $i = 1, \dots, n$ es una solución del sistema (1.1) que cumple con la condición (1.3). En este caso es posible prolongar la solución $y_i = \varphi_i(t)$ sobre todo el intervalo $(-\infty, +\infty)$, cumpliéndose además uno de los siguientes casos:

1. Para todos los valores de t se cumple que $\varphi_i(t) = y_i^*$, $i = 1, \dots, n$ donde (y_1^*, \dots, y_n^*) es un punto del conjunto abierto Ω . A la solución $y_i = \varphi_i(t)$ y al punto (y_1^*, \dots, y_n^*) se les denomina posición de *equilibrio* del sistema (1.1).
2. Existe un número positivo ω tal que para t arbitrario se cumple que:

$$\varphi_i(t + \omega) = \varphi_i(t), \quad i = 1, \dots, n$$

y para $|t_1 - t_2| < \omega$ existe al menos un índice i para el cual se cumple la desigualdad $\varphi_i(t_1) \neq \varphi_i(t_2)$. En este caso la solución se denomina *periódica* de periodo ω . Una solución periódica describe una trayectoria *cerrada*.

A continuación se enuncia una proposición que brinda una caracterización de la definición de posición de equilibrio.

Proposición 1.1 (Posición de equilibrio). : Dado el sistema (1.1) es necesario y suficiente para que el punto $(y_1^*, \dots, y_n^*) \in \Omega$ sea una posición de equilibrio del sistema que $f_i(y_1^*, \dots, y_n^*) = 0$ con $i = 1, \dots, n$.

En la literatura se pueden encontrar una gran variedad de métodos diseñados para la obtención de soluciones analíticas tanto de ecuaciones como de sistemas de EDO.

A pesar de su gran funcionalidad, en muchos casos estos métodos resultan ineficaces a la hora de obtener la solución analítica de la ecuación o el sistema de EDO según sea el caso, principalmente cuando se está modelando un problema real. Es entonces cuando entra a jugar su papel la teoría de estabilidad de las EDO. Haciendo uso de los resultados existentes en esta línea es posible realizar un estudio cualitativo del comportamiento de las soluciones del sistema alrededor de los puntos de equilibrio. De esta manera, si bien es cierto que no se tiene la expresión analítica de la solución exacta de la ecuación o el sistema de EDO en cuestión, al menos se pueden sacar conclusiones sobre su comportamiento en una cierta vecindad de los puntos de equilibrio.

1.2.2. Estabilidad de sistemas de ecuaciones diferenciales lineales

Dado un sistema de EDO como (1.1), el análisis de su estabilidad remite intuitivamente a la idea de que a una desviación pequeña de los valores iniciales corresponde una variación también pequeña de la solución, incluso en un intervalo de tiempo infinitamente grande. Para establecer la definición rigurosa, se introduce explícitamente en la notación la dependencia de los valores iniciales, de forma que $\varphi(t, \xi)$ será la solución del sistema (1.1) correspondiente a los valores iniciales ξ , es decir, se verifica $\varphi(0, \xi) = \xi$.

Definición 1.1 (Estabilidad según Lyapunov). Se dice que la posición de equilibrio $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$ del sistema (1.1) es estable según Lyapunov si cumple:

1. Existe $\rho > 0$ suficientemente pequeño tal que para $|\xi - \mathbf{y}^*| < \rho$, la solución $\varphi(t, \xi)$ del sistema (1.1) está determinada para todo $t > 0$.
2. Para todo $\varepsilon > 0$ existe $\delta < \rho$ tal que si $|\xi - \mathbf{y}^*| < \delta$ entonces $|\varphi(t, \xi) - \mathbf{y}^*| < \varepsilon$ para todo $t > 0$.

En otras palabras, dado el sistema (1.1), una posición de equilibrio \mathbf{y}^* es estable según Lyapunov, si toda solución del sistema que parte en el instante de tiempo $t = 0$ de un punto suficientemente próximo al punto de equilibrio \mathbf{y}^* , se mantiene en su curso ulterior en un entorno de este punto tan pequeño como se desee. Cuando un

punto de equilibrio no es estable, se denomina *inestable*. Otro concepto fundamental es el de estabilidad asintótica, como se refleja en la siguiente definición.

Definición 1.2 (Estabilidad asintótica). Sea $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$ una posición de equilibrio del sistema (1.1), se dice que \mathbf{y}^* es asintóticamente estable si, además de cumplir con las condiciones de estabilidad, se cumple para $\sigma < \rho$ suficientemente pequeño:

$$|\xi - \mathbf{y}^*| < \sigma \implies \lim_{t \rightarrow +\infty} |\varphi(t, \xi) - \mathbf{y}^*| = 0$$

Resumiendo, un punto de equilibrio \mathbf{y}^* es estable si todas las soluciones cercanas a él se mantienen cerca. Por su parte, se dice que es asintóticamente estable si las soluciones cercanas a él no solo se mantienen cerca, sino que tienden a él. No obstante, las trayectorias de las soluciones alrededor de un punto de equilibrio estable o asintóticamente estable pueden ser *cualitativamente* diferentes. En el siguiente epígrafe se recogen los comportamientos más comunes.

1.2.3. Retrato de fases de sistemas de ecuaciones diferenciales lineales con matriz de coeficientes constantes

En esta sección se ilustran algunos de los posibles comportamientos de las soluciones en el entorno de un punto de equilibrio. En cada caso, se clasifican los tipos de trayectorias que se pueden encontrar teniendo en cuenta su estabilidad. Por simplicidad se restringe el análisis a dimensión dos pero es preciso mencionar que en dimensiones superiores pueden aparecer comportamientos más complejos.

Sea el sistema:

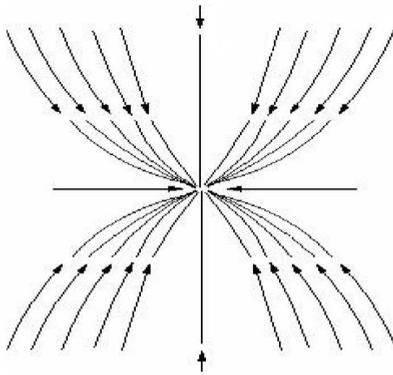
$$\frac{d\mathbf{y}}{dt} = A\mathbf{y} \tag{1.4}$$

donde $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ con a_{ij} constantes. La solución general del sistema será $\mathbf{y}(t) = c_1 v_1 e^{\lambda_1 t} + c_2 v_2 e^{\lambda_2 t}$ siendo:

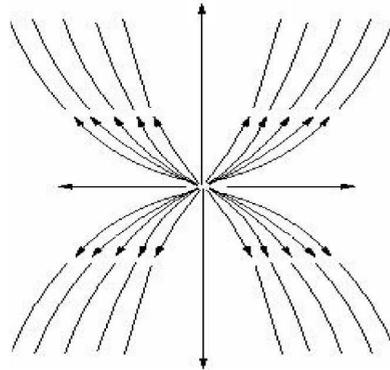
- λ_1, λ_2 los valores propios de la matriz A .

- v_1, v_2 los vectores propios correspondientes a los valores propios λ_1, λ_2 respectivamente.
- c_1, c_2 , constantes de integración, que se obtendrían de la condición inicial.

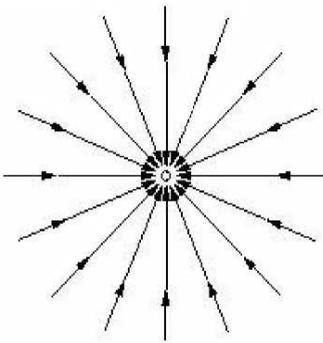
Bajo estas condiciones se pueden analizar tres casos principales, con distintas variantes:



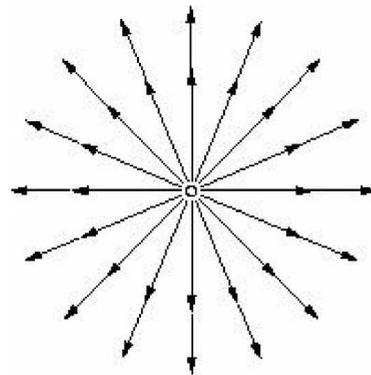
a) Nodo estable.



b) Nodo inestable.

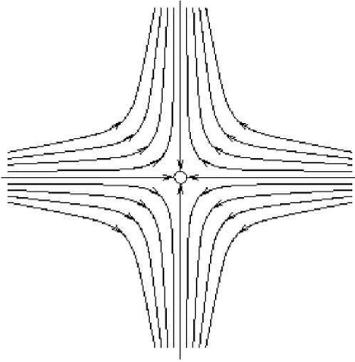


c) Nodo degenerado estable.

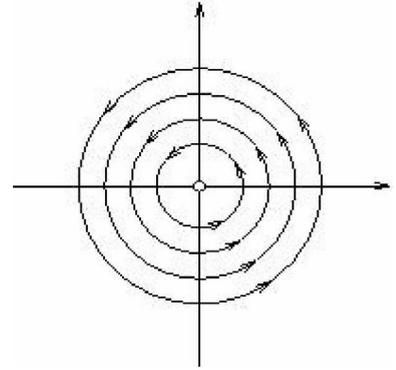


d) Nodo degenerado inestable.

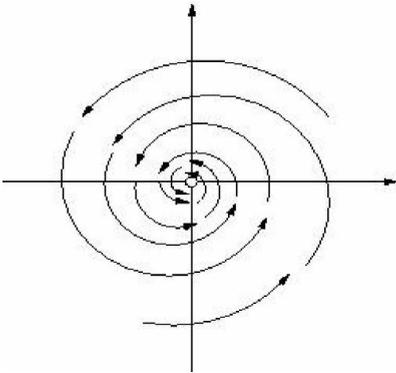
Figura 1.1: Distintos tipos de nodos.



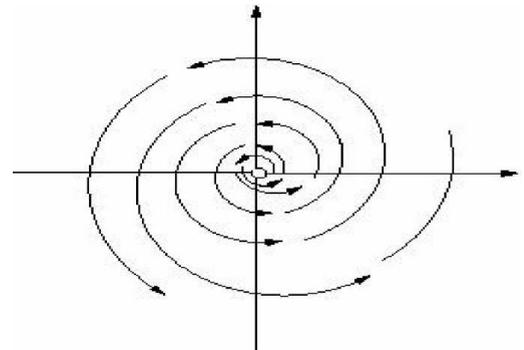
a) Punto de ensilladura.



b) Centro.



c) Foco estable.



d) Foco inestable.

Figura 1.2: Puntos de equilibrio con autovalores de signos diferentes o complejos.

1. Los valores propios de A son reales, no nulos y de igual signo. En este caso, el punto de equilibrio es un *nodo*, pudiendo distinguirse los cuatro comportamientos de la figura 1.1.

a) $\lambda_1 \neq \lambda_2 \Rightarrow$ Se obtienen trayectorias de fase curvilíneas que se acercan o se alejan del origen según sea el signo de los valores propios.

1) $\lambda_1 < \lambda_2 < 0 \Rightarrow$ Todas las trayectorias se aproximan al origen de

coordenadas, como se observa en el caso a) de la figura 1.1, por lo que se tiene un nodo estable.

2) $\lambda_1 > \lambda_2 > 0 \Rightarrow$ Todas las trayectorias se alejan del origen de coordenadas, véase caso b) de la figura 1.1, dando lugar a un nodo inestable.

b) $\lambda_1 = \lambda_2 = \lambda \Rightarrow$ Se obtienen trayectorias de fase rectas que se acercan o se alejan del origen según sea el signo del valor propio.

1) $\lambda < 0 \Rightarrow$ Si la matriz del sistema es diagonalizable, se obtiene un nodo estable. En caso de que la matriz no sea diagonalizable entonces se obtiene un nodo degenerado estable, como en el caso c) de la figura 1.1.

2) $\lambda > 0 \Rightarrow$ Si la matriz del sistema es diagonalizable, se obtiene un nodo inestable. En caso de que la matriz no sea diagonalizable entonces se obtiene un nodo degenerado inestable, véase caso d) de la figura 1.1.

2. Los valores propios de la matriz A son reales, no nulos y de signos diferentes. Se obtienen trayectorias curvilíneas que se aproximan al origen de coordenadas en la dirección del autovector con autovalor negativo y se alejan del origen en la dirección del autovector con autovalor positivo. Por ejemplo, véase el gráfico a) de la figura 1.2. En este caso, el punto de equilibrio se denomina *punto de ensilladura*.

3. Los valores propios de la matriz A son complejos conjugados $\lambda_{1,2} = a \pm ib$.

a) $a = 0 \Rightarrow$ Las trayectorias de fase, salvo la posición de equilibrio $(0, 0)$, son cerradas. Este caso, que se denomina *centro*, está ilustrado por el gráfico b) de la figura 1.2.

b) $a \neq 0 \Rightarrow$ Las soluciones serán oscilantes pero no periódicas. Las trayectorias se acercan o se alejan del origen de coordenadas en forma de espiral dependiendo del signo de a . En este caso, el punto de equilibrio se denomina *foco*.

- 1) $a < 0 \Rightarrow$ Las trayectorias se acercan al origen en forma de espiral logarítmica, como se ve en el caso c) de la figura 1.2.
- 2) $a > 0 \Rightarrow$ Las trayectorias se alejan del origen en forma de espiral logarítmica, mostrando un ejemplo en el caso d) de la figura 1.2.

De la observación de todos estos casos, se obtiene un criterio general: *un punto de equilibrio es estable si todos los valores propios de la matriz de coeficientes A tienen parte real negativa.*

1.2.4. Análisis de la estabilidad por linealización

En esta sección se recogen algunos resultados referentes al análisis de la estabilidad de sistemas de ecuaciones diferenciales no lineales. De manera general los problemas de la práctica se pueden modelar utilizando sistemas de ecuaciones diferenciales no lineales para los cuales es posible indagar sobre el comportamiento de sus soluciones de manera global. Sin embargo, conocer lo que ocurre alrededor de un punto de equilibrio resulta bastante complicado. Sería por lo tanto conveniente poder obtener alguna información del sistema no lineal a partir del lineal alrededor de un punto de equilibrio. De esta manera se podría realizar un análisis del comportamiento de las soluciones del sistema no lineal a partir de su linealización, al menos localmente.

Dado un sistema de EDO no lineal como en (1.2), la linealización alrededor del punto de equilibrio $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$ está determinada por los términos lineales del desarrollo en serie de Taylor de la función vectorial F , es decir:

$$\frac{d\mathbf{y}}{dt} = F(\mathbf{y}^*) + J_y F(\mathbf{y}^*) (\mathbf{y} - \mathbf{y}^*)$$

donde $J_y F(\mathbf{y}^*)$ es la matriz jacobiana de la función F evaluada en el punto de equi-

librio \mathbf{y}^* . Por ejemplo en el caso $n = 2$, si $F(\mathbf{y}) = (f_1(y_1, y_2), f_2(y_1, y_2))$, se tiene:

$$J_y F(\mathbf{y}^*) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{pmatrix}_{\mathbf{y}=\mathbf{y}^*} \quad (1.5)$$

Obsérvese que siempre es posible simplificar la notación utilizada. Bastaría con trasladar la posición de equilibrio \mathbf{y}^* al origen de coordenadas mediante el cambio de variable $Y = \mathbf{y} - \mathbf{y}^*$. Teniendo en cuenta que $F(\mathbf{y}^*) = 0$, ya que \mathbf{y}^* es un punto de equilibrio, se tiene entonces que $\frac{dY}{dt} = J_y F(\mathbf{y}^*) Y$, siendo $J_y F(\mathbf{y}^*)$ una matriz constante.

Luego de linealizar el sistema sería útil contar con alguna herramienta que permita relacionar el aspecto de fases alrededor de un punto de equilibrio de un sistema no lineal, con el aspecto de fases de la linealización correspondiente. Un resultado clásico al respecto es el teorema de Hartmann [14, 33] que se enuncia a continuación. Este teorema garantiza que el comportamiento de las trayectorias, en una vecindad del punto de equilibrio del sistema no lineal, coincide en lo esencial con el comportamiento de las trayectorias en el entorno del origen para el sistema lineal, siempre que los valores propios del jacobiano de la linealización no se anulen.

Teorema 1.1.

Dado un sistema autónomo de EDO como en (1.2), se asume que \mathbf{y}^ es un punto de equilibrio de dicho sistema. Sea $J_y F(\mathbf{y}^*)$ la matriz jacobiana de F evaluada en \mathbf{y}^* y sean λ_i , $i = 1, \dots, n$ sus respectivos valores propios. Entonces, la estabilidad del punto de equilibrio puede ser determinada de la siguiente manera:*

- *Si el sistema linealizado es estrictamente estable, es decir, si todos los valores propios de $J_y F(\mathbf{y}^*)$ se encuentran estrictamente en el semiplano izquierdo del plano complejo, entonces el punto de equilibrio \mathbf{y}^* es asintóticamente estable para el sistema no lineal original.*
- *Si el sistema linealizado es inestable, es decir, si al menos uno de los valores*

proprios de $J_y F(\mathbf{y}^)$ se encuentra estrictamente en el semiplano derecho del plano complejo, entonces el punto de equilibrio \mathbf{y}^* es inestable para el sistema no lineal original.*

- *Si el sistema linealizado es marginalmente estable, es decir, si todos los valores propios de $J_y F(\mathbf{y}^*)$ se encuentran en el semiplano izquierdo, pero al menos uno de ellos está en el eje iy , entonces no se puede concluir nada acerca de la aproximación lineal. El punto de equilibrio \mathbf{y}^* puede ser estable, asintóticamente estable, o inestable para el sistema no lineal original.*

Resumiendo, este teorema garantiza que el comportamiento de las trayectorias alrededor de un punto de equilibrio es cualitativamente similar en el sistema no lineal y en su linealización, salvo alrededor de los puntos de equilibrio no hiperbólicos, o sea, a los que le correspondan valores propios imaginarios puros.

1.2.5. Formalización de los sistemas dinámicos generales

Los conceptos vistos hasta el momento sobre EDO pueden integrarse en el marco de la teoría general de los sistemas dinámicos. Un sistema dinámico, intuitivamente, es el formalismo que describe el comportamiento de todos los elementos de un determinado espacio con el transcurso del tiempo. Este espacio puede ser, por ejemplo, el espacio de estados de un determinado sistema físico. Para ilustrar este ejemplo se puede considerar como sistema físico el movimiento de los planetas bajo la influencia de sus respectivas fuerzas gravitacionales. Sería entonces muy interesante poder analizar el comportamiento de los planetas con el transcurso del tiempo. En concreto, se debería poder predecir si en algún instante de tiempo t pudiese ocurrir una colisión entre estos, o si el sistema continuará evolucionando continuamente cualquiera sea el valor de t . Se procede entonces a definir formalmente qué es un sistema dinámico.

Definición 1.3 (Sistema dinámico). Un sistema dinámico es una aplicación ϕ de clase C^1 tal que $G \times \Omega \xrightarrow{\phi} \Omega$, donde Ω es un conjunto abierto del espacio euclídeo \mathbb{R}^n y G es un semigrupo. Denotando $\phi(t, y) = \phi_t(y)$ se tiene que la aplicación $\phi_t : \Omega \rightarrow \Omega$ satisface:

(1) $\phi_0 : \Omega \rightarrow \Omega$ es la identidad, es decir: $\phi_0(\mathbf{y}) = \mathbf{y}$.

(2) $\phi_t \circ \phi_s = \phi_{t+s}$ para cualesquiera $t, s \in G$.

En el caso en que G fuese un grupo, se dice que se trata de un sistema dinámico inversible. Nótese que, en este caso, la definición implica que la aplicación $\phi_t : \Omega \rightarrow \Omega$ sea C^1 para cada t y que tenga inversa C^1 , la cual se denota por ϕ_{-t} . En otras palabras, la aplicación ϕ_t será un difeomorfismo. Siempre que $G = \mathbb{N}_0$ o $G = \mathbb{Z}$ se está en presencia de un sistema dinámico discreto. En cambio cuando $G = \mathbb{R}^+$ o $G = \mathbb{R}$ el sistema dinámico es continuo. La aplicación ϕ se denomina *flujo* del sistema dinámico.

De manera general, se tiene que la solución de una EDO es *el flujo de un sistema dinámico inversible*. A continuación se muestra un ejemplo que ilustra esta afirmación. Sea A la matriz de una aplicación lineal en el espacio vectorial \mathbb{R}^n y sea ϕ la aplicación $\mathbb{R} \times \mathbb{R}^n \xrightarrow{\phi} \mathbb{R}^n$ definida por $\phi(t, \mathbf{y}) = e^{At}\mathbf{y}$. Entonces:

$$\begin{aligned} \phi : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \mathbf{y} &\longrightarrow \phi_t(\mathbf{y}) \end{aligned}$$

se representa por $\phi_t = e^{At}$. Puede comprobarse fácilmente que se cumplen las condiciones de la definición anterior.

(1) $\phi_0(\mathbf{y}) = e^0(\mathbf{y}) = \mathbf{y}$, por tanto ϕ_0 es el operador identidad.

(2) $\phi_t \circ \phi_s(\mathbf{y}) = e^{At}e^{As}(\mathbf{y}) = e^{A(t+s)}(\mathbf{y}) = \phi_{t+s}(\mathbf{y})$.

Existe una relación obvia entre el sistema dinámico definido anteriormente y la ecuación diferencial lineal de matriz constante $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$, ya que la solución de la ecuación diferencial es de la forma $\mathbf{y} = e^{At}\mathbf{y}_0 = \phi_t(\mathbf{y}_0)$.

Generalizando, se tiene que un sistema dinámico en un espacio vectorial caracteriza una ecuación diferencial o, de forma equivalente, un campo vectorial. En efecto, siendo el espacio vectorial E de dimensión finita y S un conjunto abierto de E , dado un sistema dinámico ϕ_t en S es posible construir un campo vectorial en S ,

$F : S \rightarrow E$ dado por:

$$F(\mathbf{y}) = \left. \frac{d\phi_t(\mathbf{y})}{dt} \right|_{t=0}; \quad \mathbf{y} \in S, \quad F(\mathbf{y}) \in E \quad (1.6)$$

Lo anteriormente planteado se puede reescribir en términos más convencionales. Si $\phi_t : S \rightarrow S$ es un sistema dinámico con $\mathbf{y} \in S$ tal que $\mathbf{y}(t) = \phi_t(\mathbf{y})$ y $F : S \rightarrow E$ está definida como en (1.6), se obtiene:

$$\frac{d\mathbf{y}}{dt} = F(\mathbf{y}) \quad (1.7)$$

siendo $\mathbf{y}(t) = \phi_t(\mathbf{y})$ la curva solución de (1.7) sujeta a las condiciones iniciales $\mathbf{y}(0) = \mathbf{y}_0$. De esta manera se obtiene una ecuación diferencial autónoma. Por este motivo, cuando se busca en la literatura información sobre sistemas dinámicos continuos aparece como uno de los ejemplos más ilustrativos el flujo de una ecuación diferencial autónoma.

1.2.6. Función de Lyapunov

El análisis de la estabilidad de los sistemas no lineales se puede realizar mediante el estudio de los valores propios de la linealización, como se ha visto anteriormente. Sin embargo este procedimiento resulta a menudo intratable por su complejidad. Otro de los caminos a seguir para el análisis de la estabilidad de un sistema dinámico es haciendo uso de la función de Lyapunov. El método de Lyapunov resulta muy efectivo, aunque tiene el inconveniente de que no existe ningún procedimiento constructivo para encontrar dicha función. Sin embargo, en ocasiones, la elección de la función de Lyapunov resulta natural, por ejemplo en el caso de sistemas mecánicos o eléctricos, donde la energía se considera a menudo como función de Lyapunov del sistema. De manera general, es usual denotar la función de Lyapunov por la letra V , la cual se define de la siguiente manera.

Definición 1.4 (Función de Lyapunov). Sea $\mathbf{y}^* \in \Omega$ un punto de equilibrio del sistema (1.2). Sea $V : U \rightarrow \mathbb{R}$ una función continua definida en una vecindad $U \subset \Omega$, diferenciable en $U - \mathbf{y}^*$. Se dice que V es función de Lyapunov en una vecindad de

\mathbf{y}^* si cumple:

- a) $V(\mathbf{y}^*) = 0, V(\mathbf{y}) > 0$ si $\mathbf{y} \neq \mathbf{y}^*$.
- b) $\frac{dV}{dt} \leq 0$ en $U - \mathbf{y}^*$.

Obsérvese que, utilizando la regla de la cadena, esta definición es equivalente a $\frac{dV}{dt} = \nabla V(\mathbf{y}) \cdot \mathbf{f}(\mathbf{y}) \leq 0$ con V acotada inferiormente. De esta condición se deduce que V será decreciente a lo largo de las trayectorias.

Uno de los objetivos del análisis cualitativo de las EDO es demostrar la estabilidad de un punto de equilibrio sin calcular la solución analítica de la ecuación diferencial. El siguiente teorema muestra que la existencia de la función de Lyapunov en una vecindad de un punto de equilibrio caracteriza la estabilidad de este.

Teorema 1.2 (Teorema 4.1 en [32]).

Sea \mathbf{y}^* un punto de equilibrio del sistema (1.2) y V una función de Lyapunov de este en la vecindad Ω de \mathbf{y}^* . Entonces, \mathbf{y}^* es asintóticamente estable.

Resumiendo, para concluir la estabilidad asintótica de un determinado punto de equilibrio de un sistema de EDO no lineal, basta con demostrar que se puede definir una función de Lyapunov en una vecindad de este punto.

1.2.7. Formulación de gradiente para sistemas con función de Lyapunov

En esta sección se pretende establecer la correspondencia entre los sistemas de gradiente y los sistemas con función de Lyapunov. A simple vista no resulta aparente, pero se puede establecer una correspondencia entre ellos que será útil en próximos capítulos.

Un *sistema de gradiente* definido en un conjunto abierto $\Omega \subset \mathbb{R}^n$ es un sistema dinámico de la forma:

$$\frac{d\mathbf{y}}{dt} = -\nabla V(\mathbf{y}) \tag{1.8}$$

donde $V : U \rightarrow \mathbb{R}$ es una función de clase C^2 y $\nabla V(\mathbf{y}) = \left(\frac{\partial V}{\partial y_1}, \dots, \frac{\partial V}{\partial y_n} \right)$ es el vector gradiente de V definido tal que $\nabla V(\mathbf{y}) : U \rightarrow \mathbb{R}^n$. Ahora bien, considerando

$\dot{V} : U \rightarrow \mathbb{R}^n$ como la derivada de la función V a lo largo de las trayectorias de (1.8), para cada punto $\mathbf{y}^* = \mathbf{y}(t^*)$ se tiene:

$$\dot{V}(\mathbf{y}^*) = \left. \frac{dV(\mathbf{y}(t))}{dt} \right|_{t=t^*}$$

Teorema 1.3.

Se tiene que $\dot{V}(\mathbf{y}) \leq 0$ para todo $\mathbf{y} \in U$. Asimismo, $\dot{V}(\mathbf{y}^*) = 0$ si y solo si \mathbf{y}^* es un punto de equilibrio de (1.8).

Corolario 1.1.

Sea \mathbf{y}^* un mínimo aislado de V . Entonces \mathbf{y}^* es un punto de equilibrio asintóticamente estable del sistema de gradiente (1.8).

Nótese que teniendo en cuenta lo visto en la sección anterior referente a la función de Lyapunov de un sistema de EDO, se puede establecer la analogía y concluir que la función V en este caso, no es más que la función de Lyapunov del sistema (1.8).

Por su parte los sistemas con función de Lyapunov se pueden escribir de la forma:

$$\frac{d\mathbf{y}}{dt} = L(\mathbf{y})\nabla V(\mathbf{y})$$

siendo $L(\mathbf{y})$ una matriz simétrica y definida negativa. Esta formulación se conoce como la forma de *gradiente lineal del sistema* [39] y se puede considerar como una generalización de un sistema de gradiente. En efecto, la ecuación (1.8) se obtiene como un caso particular del gradiente lineal tomando $L(\mathbf{y})$ como la matriz identidad cambiada de signo. De aquí, se concluye la equivalencia entre los sistemas de gradiente y los sistemas con función de Lyapunov por lo que, en lo sucesivo, se hará referencia a estos últimos como sistemas de gradiente.

En próximos capítulos, la conservación de la estabilidad del sistema bajo discretización, será el centro de atención. Es por esto que a continuación se recuerdan algunos resultados fundamentales del análisis numérico.

1.3. Métodos numéricos para ecuaciones diferenciales ordinarias

En esta sección se hace una breve introducción a los métodos numéricos diseñados para la resolución de ecuaciones diferenciales ordinarias. Se introducen los conceptos de estabilidad, orden, consistencia y convergencia de un método numérico. Se parte de su formulación general para luego diferenciar los métodos explícitos de los implícitos, puntualizando las ventajas y desventajas en cada caso. Se diferencian además los métodos de paso simple y los multipaso poniendo ejemplos de cada uno.

1.3.1. Aproximación numérica y sus propiedades

En este apartado se resumen algunas de las definiciones y resultados más significativos del análisis numérico, específicamente de los métodos diseñados para la resolución de ecuaciones diferenciales ordinarias [19], permitiendo su referencia en el resto del trabajo. Por otra parte, se habla de la necesidad de diseñar métodos numéricos que conserven las propiedades cualitativas del sistema continuo, tales como el flujo, la estabilidad o la función de Lyapunov en caso de que la posea, entre otras. Centrándose en la incapacidad de los métodos numéricos usuales para conservar estas propiedades, se introducen los conocidos como métodos de integración geométrica. En el diseño de estos métodos se centrará el trabajo en próximos capítulos.

De manera general se le llama *método numérico* [46] a un algoritmo que permite obtener la solución de un problema matemático mediante el uso de un número finito de operaciones aritméticas, aunque sea con carácter aproximado. Durante el estudio de los métodos numéricos siempre se pone en consideración la mejor aproximación de la solución calculada, así como el menor coste computacional.

De la teoría de ecuaciones diferenciales, revisada en la sección anterior, se sabe

que el problema de Cauchy, o de valores iniciales:

$$\begin{aligned}\frac{dy}{dt} &= F(t, \mathbf{y}), \quad t \geq t_0 \\ \mathbf{y}(t_0) &= \mathbf{y}_0\end{aligned}\tag{1.9}$$

posee una única solución si $F : [t_0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función suficientemente suave. Esta suavidad puede caracterizarse de varias maneras por ejemplo, bastaría con que la función F fuese de Lipschitz, o sea:

Definición 1.5. Se dice que $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es de Lipschitz si existe una constante $K > 0$ tal que:

$$\|F(\mathbf{x}) - F(\mathbf{y})\| \leq K \|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

En los casos en los que no es posible obtener la expresión analítica de la solución exacta mediante el uso de los métodos de resolución de EDO, se procede a obtener una aproximación numérica de la solución. Se sabe entonces que un método numérico para el problema (1.9) viene dado por la ecuación:

$$\mathbf{y}_{k+1} = \Psi_{F, h_k}(\mathbf{y}_k, \dots, \mathbf{y}_0)\tag{1.10}$$

donde se asume que existe la función Ψ_{F, h_k} y está bien definida. Por su parte, \mathbf{y}_k es una aproximación del verdadero valor de la solución $\mathbf{y}(t_k)$, siendo h_k el incremento del tiempo entre instantes sucesivos $h_k = t_{k+1} - t_k$. Para simplificar, en lo adelante se considera h_k constante y se escribe solo h . Al hablar de métodos numéricos implícitos en próximas secciones, se verá que la función $\Psi_{F, h}$ puede estar definida en forma implícita.

Cuando se trabaja con métodos numéricos hay tres conceptos que no se pueden dejar de mencionar: convergencia, estabilidad y consistencia [44, 35, 36]. El análisis de estas propiedades, ayudará a caracterizar correctamente el método numérico con el que se esté trabajando. Es por esto que se recuerdan a continuación algunas de las definiciones más utilizadas en la literatura especializada sobre este tema.

La idea de **convergencia** está relacionada con el *error global*, es decir, con la diferencia entre el verdadero valor de la solución y la aproximación numérica calculada

tras cierto número de pasos.

Definición 1.6. Un método numérico se dice convergente si para cada EDO como en la ecuación (1.9) con F función de Lipschitz y todo $t^* > 0$ se cumple que:

$$\lim_{h \rightarrow 0^+} \max_{k=0,1,\dots, \lfloor t^*/h \rfloor} \|\mathbf{y}_k - \mathbf{y}(t_k)\| = 0$$

donde $\lfloor \alpha \rfloor \in \mathbb{Z}$ es la parte entera de $\alpha \in \mathbb{R}$, siendo \mathbf{y}_k la aproximación del verdadero valor de la solución \mathbf{y} en t_k .

En otras palabras se dice que el método numérico es convergente si para cada función de Lipschitz, la solución numérica de la EDO tiende a la solución real tomando el tamaño de paso cada vez más pequeño sobre un intervalo de tiempo determinado. Es válido señalar que la convergencia no es solo otra propiedad deseada sino una condición fundamental, ya que un método numérico que no converja, en la práctica resulta inútil.

Definición 1.7. Se dice que un método numérico, como en la ecuación (1.10) para la EDO (1.9) es de orden p si:

$$\mathbf{y}(t_{k+1}) - \Psi_{F,h}(\mathbf{y}(t_k), \dots, \mathbf{y}(t_0)) = O(h^{p+1})$$

para todo $k = 0, 1, \dots$

Una definición alternativa es que un método numérico es de orden p si recupera exactamente cada solución polinómica de grado p o menos. El orden de un método numérico proporciona información acerca de su comportamiento local. En otras palabras, al avanzar de t_k a t_{k+1} con h suficientemente pequeño se incurre en un error de orden $O(h^{p+1})$.

Otra propiedad deseada cuando se trabaja con métodos numéricos es la consistencia. Se sabe que así como la convergencia está estrechamente relacionada con el error global del método numérico, la **consistencia** se relaciona con el *error local*. Grosso modo, se puede decir que es el error tras un solo paso del método y refleja en qué medida los verdaderos valores de la solución satisfacen las ecuaciones del método numérico.

Definición 1.8. Un método numérico como en (1.10) es consistente si:

$$\Psi_{F,0}(\mathbf{y}) = F(\mathbf{y}) \quad (1.11)$$

Obsérvese que un método numérico tal que (1.11) se satisface es un método numérico al menos de orden uno.

Por su parte el concepto de **estabilidad** está estrechamente relacionado con la manera en la que se propagan los errores en el algoritmo al calcular sucesivamente los valores aproximados de la solución. Existen varios conceptos de estabilidad citados en la literatura [18]. A continuación se discuten los conceptos de 0-estabilidad y estabilidad absoluta o A-estabilidad.

La idea de 0-estabilidad es acotar la propagación del error. Ese error se refiere al cometido al estimar la solución en un instante de tiempo determinado. Si un método es 0-estable, entonces el efecto de este error sobre las aproximaciones sucesivas se puede acotar uniformemente para tamaños de paso suficientemente pequeños. La definición rigurosa de 0-estabilidad sería:

Definición 1.9. Un método numérico como en (1.10) es 0-estable si Ψ cumple la condición de Lipschitz, con λ constante de Lipschitz respecto de la variable \mathbf{y} para $h \leq h_0$, o sea:

$$\|\Psi_{F,h}(\mathbf{y}) - \Psi_{F,h}(\mathbf{y}^*)\| \leq \lambda \|\mathbf{y} - \mathbf{y}^*\|$$

En cambio, se dice que un método numérico es *absolutamente estable* o *A-estable* si para un tamaño de paso fijo la evolución de la solución discretizada permanece acotada en el infinito. Esta propiedad se relaciona con el comportamiento asintótico de la solución. Es, por tanto, un concepto antagónico a la 0-estabilidad, en la que para un intervalo de integración fijo, la solución permanece acotada cuando el tamaño de paso es suficientemente pequeño.

Para realizar el análisis de la A-estabilidad de un método numérico se trabaja con el problema de Cauchy lineal y escalar:

$$\begin{aligned} y'(t) &= \lambda y(t), \quad t > 0 \\ y(0) &= 1 \end{aligned} \quad (1.12)$$

con $\lambda \in \mathbb{C}$, cuya solución viene dada por $y(t) = e^{\lambda t}$. Nótese que $\lim_{x \rightarrow \infty} |y(t)| = 0$ si la parte real de λ es negativa.

Definición 1.10. Un método numérico que aproxime (1.12) es *A-estable* si:

$$|y_n| \rightarrow 0 \quad \text{cuando} \quad t_n \rightarrow +\infty$$

siendo h el paso de discretización. La solución numérica y_n de (1.12) evidentemente depende de h y de λ . La *región de estabilidad absoluta* de un método numérico es el subconjunto del plano complejo:

$$\mathcal{A} = \left\{ \hat{h} = h\lambda \in \mathbb{C} : |y_n| \rightarrow 0 \quad \text{cuando} \quad t_n \rightarrow +\infty \right\}$$

Luego \mathcal{A} es el conjunto de todos los valores de $h\lambda$ para los cuales las soluciones obtenidas a través del método numérico tienden a cero cuando t_n tiende a infinito.

Si bien es cierto que cada una de estas propiedades tiene su valor por separado, si existiese un resultado que las vinculara sería muy útil en la práctica. De esta manera no sería necesario comprobar que se cumplen cada una de las definiciones antes mencionadas para hacer un estudio detallado del método numérico. Afortunadamente este resultado existe y se considera uno de los más importantes de la teoría de análisis numérico, se le conoce como *Teorema de equivalencia de Lax* [50]. Este teorema relaciona los conceptos de convergencia, estabilidad y consistencia de un método numérico de la siguiente manera.

Teorema 1.4.

Un método numérico es convergente si y solo si es consistente y estable.

También está demostrado en la literatura que si un determinado método numérico es 0-estable y consistente de orden p , entonces es convergente de orden p y su error global viene dado por $\mathbf{y}(t_k) - \mathbf{y}_k = O(h^p)$. Obsérvese la reducción en el orden del error, frente a la definición 1.7 para el error local.

Hasta aquí el resumen de las propiedades generales de un método numérico. En el próximo apartado se clasifican los métodos numéricos en explícitos, implícitos,

de paso simple y multipaso o paso múltiple, poniendo ejemplos en cada caso. Vale señalar que no todas las clasificaciones son excluyentes o sea, es posible hablar de un método explícito de paso simple, de un método implícito de paso múltiple o viceversa.

1.3.2. Método de Euler

El primero de todos los métodos numéricos diseñados para integrar una EDO fue introducido por Euler en 1768. El método de Euler [10], como se conoce hasta la actualidad, está considerado como la piedra angular de la evolución del análisis numérico de las ecuaciones diferenciales, a pesar de ser sencillo y elemental computacionalmente.

La idea detrás de este método es relativamente sencilla. Supóngase que se quiere integrar una EDO como en (1.9) y solo se conoce el valor de la solución exacta en un instante de tiempo t_0 . Se plantea la cuestión de si sería entonces posible aproximar el valor de la solución en un nuevo instante de tiempo. La respuesta a esta pregunta es afirmativa y el enfoque más elemental sería utilizar una interpolante lineal constante. En otras palabras, estimar el valor de la solución exacta $y(t)$ haciendo la siguiente aproximación:

$$F(y(t)) \approx F(y(t_0)) \quad \forall t \in [t_0, t_0 + h]$$

siendo $h > 0$ suficientemente pequeño. Integrando en (1.9) se obtiene:

$$y(t) = y(t_0) + \int_{t_0}^t F(y(\tau)) d\tau \approx y_0 + (t - t_0) F(y_0) \quad (1.13)$$

dado que $F(y) = \frac{dy}{dt}$. Luego, si se considera la sucesión $t_0, t_1 = t_0 + h, t_2 = t_0 + 2h, \dots$ donde $h > 0$ es el tamaño de paso, se denota por y_k a la estimación numérica de la solución exacta $y(t_k)$ siendo $k = 0, 1, \dots$ y se tiene en cuenta la ecuación (1.13) se obtiene:

$$y_{k+1} = y_k + h F(y_k), \quad k = 0, 1, \dots \quad (1.14)$$

A este esquema se le conoce como método de Euler. Nótese que F no depende de y_{k+1}

y que el término \mathbf{y}_{k+1} está despejado. Esto equivale a decir que el método obtenido es explícito y de paso simple, como se verá más adelante. Por otra parte, se sabe que el método de Euler es convergente y de primer orden. Finalmente, haciendo uso del Teorema de Lax se puede concluir que este método es estable y consistente.

1.3.3. Formulación implícita y explícita de métodos numéricos

Un método implícito es aquel en el que el término \mathbf{y}_{k+1} aparece implícitamente en la ecuación que describe al método. De manera general un método implícito se puede expresar de la siguiente manera:

$$\Phi_{F,h}(\mathbf{y}_{k+1}, \mathbf{y}_k, \dots, \mathbf{y}_0) = 0 \quad (1.15)$$

Por ejemplo, para obtener el *método de Euler implícito* se considera un esquema de diferencias para la aproximación de la primera derivada por retroceso en el instante $\mathbf{y} = \mathbf{y}_{k+1}$, o sea:

$$\frac{d\mathbf{y}(t_{k+1})}{dt} \approx \frac{\mathbf{y}(t_{k+1}) - \mathbf{y}(t_k)}{h}$$

Teniendo en cuenta que:

$$\frac{d\mathbf{y}(t_{k+1})}{dt} = F(\mathbf{y}_{k+1}) \quad \Rightarrow \quad \frac{\mathbf{y}(t_{k+1}) - \mathbf{y}(t_k)}{h} = F(\mathbf{y}_{k+1})$$

con lo que finalmente se obtiene el *método de Euler implícito*:

$$\mathbf{y}(t_{k+1}) = \mathbf{y}(t_k) + h F(\mathbf{y}_{k+1})$$

Nótese que al aplicar el método de Euler implícito habría que resolver una ecuación no lineal en cada paso y esto conlleva a que el coste computacional sea mayor. Esto sucede de manera general con todos los métodos implícitos, constituyendo una desventaja con respecto a los métodos explícitos. En cambio, los métodos implícitos tienen propiedades favorables desde el punto de vista de la estabilidad numérica y juegan un papel fundamental en la resolución de ecuaciones algebraico-diferenciales y ecuaciones *stiff* [13].

Por su parte, un método explícito es aquel en el que la variable y_{k+1} aparece despejada. En otras palabras, la función Φ de la ecuación (1.15) puede escribirse como sigue:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h \Psi_{F,h}(\mathbf{y}_k, \dots, \mathbf{y}_0) \quad (1.16)$$

con lo que el cálculo de \mathbf{y}_{k+1} no depende de sí mismo, sino de las aproximaciones precedentes.

1.3.4. Métodos de paso simple o multipaso

Existen una gran variedad de métodos numéricos en los que se obtiene el valor de la solución numérica en un tiempo futuro usando información relativa a la historia reciente de la solución. El cálculo de la aproximación en estos casos no se limita a utilizar información sobre el estado presente de la misma. En otras palabras, se puede decir que para calcular el término \mathbf{y}_{k+1} se utiliza no solo el término \mathbf{y}_k como en el método de Euler, sino que además se necesitan los términos $\mathbf{y}_{k-1}, \dots, \mathbf{y}_{k-r+1}$ siendo $r \geq 1$ el número de pasos del método con el que se esté trabajando. Estos métodos son los conocidos como *métodos multipaso*. Nótese que esta descripción es consistente con la formulación dada en la ecuación (1.10). En cambio, si para calcular \mathbf{y}_{k+1} basta con utilizar el término \mathbf{y}_k , se dice que se está en presencia de un método de paso simple.

En general un método lineal multipaso con r pasos, se puede describir utilizando $2r+2$ constantes $\alpha_0, \dots, \alpha_r, \beta_0, \dots, \beta_r$ que son los coeficientes de los polinomios que describen el método:

$$\alpha_r \mathbf{y}_{k+1} + \dots + \alpha_0 \mathbf{y}_{k-r+1} = h [\beta_r F(\mathbf{y}_{k+1}) + \dots + \beta_0 F(\mathbf{y}_{k-r+1})] \quad (1.17)$$

Nótese que si $\beta_r = 0$ el método será explícito y en caso contrario será implícito, ya que para efectuar un paso del método es necesario hallar todas las componentes de \mathbf{y}_{k+1} resolviendo un sistema de ecuaciones. La formulación general de los métodos multipaso implícitos y explícitos coinciden con las ecuaciones (1.15) y (1.16) respectivamente.

Ahora bien, si la función $\Phi_{F,h}$ de (1.15) es lineal, entonces la ecuación (1.15) representa un método lineal multipaso. Existen varios métodos lineales múltipaso, entre los que se pueden citar los métodos de Adams explícitos e implícitos.

1.3.5. Métodos de Runge-Kutta

Una segunda generalización para la mejora del método de Euler son los métodos de Runge-Kutta. Al igual que con los métodos multipaso, existen formulaciones explícitas e implícitas de estos métodos. A continuación se comenta muy brevemente uno de los enfoques a seguir para la obtención de este tipo de métodos.

La idea de los métodos descritos en 1895 por Runge y elaborados más ampliamente por Kutta para la resolución de un problema de Cauchy de primer orden, como el dado en la ecuación (1.9), consiste en calcular la nueva ordenada adicionando a la anterior un incremento. Este incremento debe coincidir con el desarrollo de Taylor de la solución exacta hasta el término de la derivada de orden s evaluada en $t_k + h$, pero solo utilizando la primera derivada de F . Esta idea es similar a la de los métodos obtenidos a partir del desarrollo de Taylor, pero en estos se trabaja con la evaluación de la función y sus derivadas superiores. El incremento que caracteriza un método de Runge-Kutta se obtendrá como combinación lineal de valores obtenidos al evaluar F en s puntos del subintervalo formado por dos instantes de tiempo sucesivos. En su forma más general, un método de Runge-Kutta se define de la siguiente manera:

Definición 1.11. Sean a_{ij} , b_i , $i = 1, \dots, s$ números reales y sea $c_i = \sum_{j=1}^s a_{ij}$. Un método de Runge-Kutta viene dado por:

$$y_{k+1} = y_k + \Delta y_k = y_k + h \sum_{i=1}^s b_i K_i \tag{1.18}$$

$$K_i = F \left(t_k + c_i h, y_k + h \sum_{j=1}^s a_{ij} K_j \right), \quad i = 1, 2, \dots, s$$

siendo s el número de etapas del método.

Como se deduce de la definición, el método está completamente caracterizado

por los coeficientes a_{ij} , b_i y c_i . Por definición, $c_1 = 0$, mientras que $0 < c_i \leq 1$ para $i > 1$. Estos coeficientes usualmente se encuentran agrupados en el conocido como *tableau* de Butcher [10]:

$$\begin{array}{c|cccc}
 c_1 & a_{11} & a_{12} & \dots & a_{1s} \\
 c_2 & a_{21} & a_{22} & \dots & a_{2s} \\
 \vdots & \vdots & & \ddots & \vdots \\
 c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\
 \hline
 & b_1 & b_2 & \dots & b_s
 \end{array}
 \quad \text{o bien, en forma matricial:} \quad
 \begin{array}{c|c}
 c & A \\
 \hline
 & b^T
 \end{array}$$

donde $A = (a_{ij}) \in \mathbb{R}^{s \times s}$, $b = (b_1, b_2, \dots, b_s)^T \in \mathbb{R}^s$ y $c = (c_1, c_2, \dots, c_s)^T \in \mathbb{R}^s$. Por su parte, s denota el número de etapas del método. Si los elementos a_{ij} son nulos para $j \geq i$ con $i = 1, 2, \dots, s$ entonces cada K_i puede ser explícitamente calculado en términos de cada uno de los $i - 1$ coeficientes K_1, \dots, K_{i-1} . En este caso, el método de Runge-Kutta es explícito. En caso contrario, se está en presencia de un método de Runge-Kutta implícito, en el que para calcular el coeficiente K_i es necesario resolver un sistema no lineal con s ecuaciones.

Ahora bien, los coeficientes a_{ij} , b_i y c_i se determinan bajo la condición de que el valor aproximado y_{k+1} calculado según (1.18) coincida con el que se obtendría evaluando el desarrollo en serie de Taylor hasta el término de orden s :

$$y_{k+1} = y_k + \sum_{i=1}^s \frac{h^i}{i!} y^{(i)} \Big|_{t_k} \tag{1.19}$$

lo cual equivale a exigir que el incremento de Runge coincida con el incremento de Taylor, o sea:

$$h \sum_{i=1}^s b_i K_i = \sum_{i=1}^s \frac{h^i}{i!} y^{(i)} \Big|_{t_k} \tag{1.20}$$

Comparando las expresiones (1.18) y (1.19) se observa que, en la primera, el incremento se construye como combinación lineal de las funciones K_i , es decir, de la función F evaluada en s puntos cuyas abscisas se encuentran en el intervalo $[t_k, t_{k+1}]$; por el contrario, en la segunda, el incremento se construye como combinación lineal de las s primeras derivadas de y evaluadas en $t = t_k$. Es decir, la idea fundamental

de los métodos de Runge-Kutta consiste en contraponer dos formas de construir el incremento Δy_k :

1. **Taylor:** Evaluación de s funciones (las derivadas) en un solo punto t_k .
2. **Runge-Kutta:** Evaluación de una función (F) en s puntos del intervalo $[t_k, t_{k+1}]$.

Desde el punto de vista computacional, resulta más eficiente evaluar una sola función en s puntos, que hallar las derivadas superiores de una determinada función para luego evaluarlas en un único punto, como se requiere en el caso de los métodos diseñados a partir del desarrollo en serie de Taylor. De ahí la vigencia de los métodos de Runge-Kutta más de un siglo después de ser propuestos.

1.4. Conclusiones

Este capítulo además de introducir el tema de la tesis, consta de otras dos grandes secciones. La primera se ha dedicado al estudio de los sistemas dinámicos y a las ecuaciones diferenciales ordinarias, como casos particulares de estos. Se ha hecho énfasis en el análisis de la estabilidad, tanto de los sistemas lineales como de los no lineales. En el caso de los sistemas no lineales, se recordaron resultados básicos de la teoría de estabilidad de ecuaciones diferenciales, que garantizan que el estudio de la estabilidad de un sistema no lineal puede realizarse a través de su linealización alrededor de los puntos de equilibrio del sistema. Como método alternativo se introduce el concepto de función de Lyapunov, puntualizando que su existencia garantiza la estabilidad del sistema dado, al menos en una vecindad del punto de equilibrio en cuestión. Finalmente, se establece la correspondencia entre los sistemas de gradiente y los sistemas con función de Lyapunov.

En la otra sección se recogen algunos de los resultados más relevantes de la teoría de análisis numérico. Algunos de los conceptos y definiciones que se recogen en este capítulo se utilizarán más adelante, por lo que el objetivo principal del capítulo es sentar las bases para que la tesis sea autocontenido. Se mencionan diferentes clasificaciones de los métodos numéricos: métodos explícitos, implícitos, paso simple y

multipaso, ilustrando con ejemplos cada caso. Asimismo, se hace un resumen de sus principales propiedades: orden de convergencia, estabilidad y consistencia.

Es válido señalar que, si bien los métodos numéricos clásicos brindan buenas aproximaciones de las soluciones en la mayoría de los casos, la discretización no conserva las propiedades cualitativas del sistema continuo, siendo este el punto de partida para el diseño de métodos de integración geométrica, a los que se les dedica el próximo capítulo.

Capítulo 2

Integración geométrica

Resumen del capítulo

En este capítulo se introducen los métodos de integración geométrica. Estos métodos están diseñados con el objetivo de conservar las propiedades cualitativas del sistema continuo bajo discretización, cualquiera sea el tamaño de paso con el que se trabaje. En cambio, los métodos numéricos clásicos vistos en el capítulo anterior solo las conservan para tamaños de paso suficientemente pequeños. Tras hacer una breve introducción a los métodos de integración numérica geométrica, se define el adjunto de un método numérico que, junto con la composición de métodos numéricos, constituye una vía alternativa para el incremento del orden de un determinado método a la vez que se conservan determinadas propiedades cualitativas del sistema continuo. Finalmente, se analizan los métodos de proyección diseñados con el objetivo de conservar la función de Lyapunov de un sistema dado. Comienza aquí la contribución de la presente tesis, pues los métodos de proyección se implementan sobre diversos sistemas y, del estudio detallado de sus limitaciones, se deduce el interés de los métodos de gradiente discreto que se introducen en el próximo capítulo.

2.1. Integración numérica geométrica

Se sabe que las ecuaciones diferenciales juegan un papel fundamental en la aplicación de la matemática a otras ciencias, por lo que el esfuerzo dedicado a la tarea de resolverlas o integrarlas está ampliamente justificado. Como se vio en el capítulo anterior, en ocasiones los métodos elementales de integración son incapaces de resolverlas. Es entonces cuando entra a jugar su papel el estudio cualitativo de las soluciones alrededor de un punto de equilibrio del sistema, ya que de esta manera se obtiene al menos una noción del comportamiento de sus soluciones. Por otra parte, esta incapacidad motivó la introducción de técnicas de mayor potencia y generalidad, entre las que se pueden citar el trabajo con series y los métodos numéricos. Por ejemplo, los métodos numéricos multipaso y Runge-Kutta citados en las secciones precedentes, son herramientas útiles para la resolución aproximada de sistemas de ecuaciones diferenciales. Pero incluso los métodos numéricos más sofisticados fallan cuando se trata de reproducir la dinámica del modelo continuo. En otras palabras, si bien es cierto que brindan una aproximación suficientemente cercana a la solución exacta, en muchas ocasiones la solución obtenida no conserva las propiedades cualitativas del sistema continuo. Por ejemplo, son incapaces de conservar las propiedades geométricas del flujo de la ecuación diferencial, primeras integrales, puntos de equilibrio, o la estabilidad, entre otras propiedades del sistema continuo. Un ejemplo clásico de esta incapacidad se obtiene al integrar el sistema de ecuaciones diferenciales que describe el movimiento de los planetas del sistema solar alrededor del sol, conocido como problema de Kepler [2]. *La primera ley de Kepler* garantiza que los planetas del sistema solar se mueven formando una órbita elíptica con el sol en uno de sus focos. Sin embargo, al integrar el sistema de ecuaciones diferenciales que describe esta dinámica utilizando la regla de Euler, la solución aproximada obtenida no describe una trayectoria elíptica por lo que no respeta la dinámica del sistema continuo. Existen una gran variedad de problemas en los que, al discretizar el sistema continuo, la solución aproximada obtenida no presenta las mismas propiedades que la solución exacta, propiedades que pueden ser conocidas incluso sin tener la expresión analítica de la solución. Con vistas a solucionar este problema, surge una nueva

línea de investigación, la cual se conoce como integración geométrica [17].

Como se menciona en el capítulo anterior, la integración geométrica establece un vínculo entre el análisis de sistemas dinámicos y el diseño de métodos numéricos, de manera tal que al discretizar el sistema continuo mediante el método numérico diseñado se conserven sus propiedades cualitativas. Se diseñan métodos con el objetivo de conservar las propiedades del flujo de la ecuación diferencial, por ejemplo: métodos simplécticos para sistemas Hamiltonianos (métodos que conserven la energía o la forma simpléctica del sistema), métodos que conservan primeras integrales y métodos numéricos en variedades, métodos que conservan estructuras como el volumen o invariantes lineales y cuadráticas, métodos que conservan la estabilidad incluyendo puntos de equilibrio y función de Lyapunov del sistema, entre otros. La motivación de la creación y desarrollo de algoritmos que preserven estas propiedades en diferentes problemas, viene dada independientemente en cada una de las áreas de investigación ya que el interés por conservar una determinada propiedad del sistema continuo está estrechamente relacionada con las aplicaciones de dicho sistema. Por ejemplo, se sabe que las soluciones de un sistema Hamiltoniano viven en una variedad en la que una determinada función, que a menudo corresponde a la noción física de energía, se mantiene constante. Por tanto, cualquier aproximación numérica para la que esta función deje de ser constante, no sería una buena aproximación para dicho sistema. El problema sería entonces el diseño de métodos numéricos, de manera tal que la trayectoria discreta del método tenga las mismas propiedades que el sistema continuo.

Frente a los numerosos métodos que se han propuesto en el ámbito de los sistemas Hamiltonianos, la implementación de métodos para sistemas de gradiente es relativamente escasa. Desde un punto de vista físico, las ecuaciones de Hamilton describen sistemas conservativos, mientras que los sistemas de gradiente, a través de la existencia de la función de Lyapunov, introducen la noción de sistema disipativo. En esta dirección se trabaja en próximos capítulos, centrando la atención en el diseño de métodos numéricos que conserven la estabilidad del sistema y con esta, su función de Lyapunov.

2.2. Adjunto de un método numérico

Como se vio en secciones anteriores, una de las propiedades fundamentales de un método numérico es su orden de convergencia. Teniendo en cuenta la relación directa que existe entre el orden de un método numérico y su error local, sería conveniente, una vez diseñado un método numérico, ser capaces de incrementar su orden, ya que de esta manera se obtendría una aproximación de la solución más cercana a la realidad. Con este propósito, se dedica esta sección al estudio del adjunto de un método numérico ya que, como se verá más adelante, componiendo un determinado método numérico con su adjunto, se obtiene un método con las mismas propiedades que el inicial pero de al menos un orden mayor.

En esta sección se hablará de los métodos numéricos simétricos, de los que se sabe poseen propiedades favorables. La clave para comprender la *simetría* es la definición de adjunto de un método numérico, la cual viene dada por:

Definición 2.1. El adjunto Φ_h^* del método Φ_h es el mapa inverso del método original con tamaño de paso inverso $-h$, o sea:

$$\Phi_h^* = \Phi_{-h}^{-1}$$

En otras palabras, $y_{k+1} = \Phi_h^*(y_k)$ está definido implícitamente por $\Phi_{-h}(y_{k+1}) = y_k$. Por otra parte, un método en el que $\Phi_h^* = \Phi_h$ se llama *simétrico*. Resumiendo, un método numérico es simétrico si coincide con su adjunto.

Como se comentó en el capítulo anterior, se sabe que el flujo de una ecuación diferencial autónoma es simétrico, o sea, coincide con su inverso cambiando de signo la escala de tiempo. En cambio, de manera general los métodos numéricos de paso simple no presentan esta propiedad. Es posible citar varios ejemplos que lo demuestran, de los cuales, el más sencillo es la regla de Euler:

- El método de Euler implícito es el adjunto del método de Euler explícito.
- La regla del punto medio implícita es simétrica, por lo que coincide con su adjunto.

- La regla trapezoidal y el método de Störmer-Verlet también son simétricas.

En cuanto a las propiedades de los métodos adjuntos, se tiene:

- $(\Phi_h^*)^* = \Phi_h$.
- $(\Phi_h \circ \Psi_h)^* = \Psi_h^* \circ \Phi_h^*$.

El teorema que se enuncia a continuación muestra la relación que existe entre el orden de un método numérico y el de su adjunto.

Teorema 2.1.

Sea φ_t el flujo exacto de una ecuación diferencial autónoma y sea Φ_h un método de paso simple de orden p tal que:

$$\Phi_h(y_0) = \varphi_h(y_0) + C(y_0)h^{p+1} + O(h^{p+2}) \quad (2.1)$$

Entonces, el método adjunto Φ_h^* es también de orden p y se cumple que:

$$\Phi_h^*(y_0) = \varphi_h(y_0) + (-1)^p C(y_0)h^{p+1} + O(h^{p+2}) \quad (2.2)$$

En particular, si el método es simétrico, su orden mínimo será par [17].

2.3. Métodos de composición

Los métodos de composición son aquellos que se obtienen al considerar la acción consecutiva de varios métodos numéricos. Se fracciona el tamaño de paso de manera tal que en cada uno de los subpasos actúa uno de los métodos numéricos que forman la composición. En los subpasos pueden actuar métodos iguales o diferentes y eventualmente la composición puede considerarse como la acción consecutiva del mismo método numérico. Existen varias variantes de esta idea: la composición de varios métodos de Runge-Kutta considerando el mismo tamaño de paso [17], la composición cíclica de métodos multipaso [49] o la composición de métodos de Runge-Kutta de bajo orden [28]. En próximos capítulos se utilizarán los métodos de composición con

el objetivo de aumentar el orden de convergencia a la vez que se conservan ciertas propiedades cualitativas del método original. Con este objetivo, a continuación se considera la composición de un método básico de paso fijo con diferentes tamaños de paso y eventualmente la composición con su adjunto, idea que ha sido desarrollada previamente en [52, 58, 40].

Definición 2.2. Sea Φ_h un método numérico de paso simple y $\gamma_1, \gamma_2, \dots, \gamma_s$ números reales. Entonces se define su composición con tamaños de paso $\gamma_1 h, \gamma_2 h, \dots, \gamma_s h$ de manera tal que:

$$\Psi_h = \Phi_{\gamma_s h} \circ \dots \circ \Phi_{\gamma_1 h} \quad (2.3)$$

siendo Ψ_h el método de composición correspondiente.

El teorema que se enuncia a continuación muestra que la elección de los escalares $\gamma_1 h, \gamma_2 h, \dots, \gamma_s h$ influye en el orden del método de composición obtenido. Específicamente, se tiene:

Teorema 2.2.

Sea Φ_h un método de paso simple de orden p . Si:

$$\begin{aligned} \gamma_1 + \dots + \gamma_s &= 1 \\ \gamma_1^{p+1} + \dots + \gamma_s^{p+1} &= 0 \end{aligned} \quad (2.4)$$

entonces el método de composición (2.3) es al menos de orden $p + 1$.

Ahora bien, si en lugar de componer un mismo método numérico con diferentes tamaños de paso, se compusiera un método con su adjunto la composición dada por la ecuación (2.3) debería ser reemplazada por la ecuación:

$$\Psi_h = \Phi_{\alpha_s h} \circ \Phi_{\beta_s h}^* \circ \dots \circ \Phi_{\alpha_1 h} \circ \Phi_{\beta_1 h}^* \quad (2.5)$$

con lo que la condición (2.4) para la obtención de un método de orden al menos $p + 1$

se convertiría en:

$$\begin{aligned}\beta_1 + \alpha_1 + \dots + \beta_s + \alpha_s &= 1 \\ (-1)^p \beta_1^{p+1} + \alpha_1^{p+1} \dots + (-1)^p \beta_s^{p+1} + \alpha_s^{p+1} &= 0\end{aligned}\tag{2.6}$$

En particular, la solución $\alpha_1 = \beta_1 = \frac{1}{2}$ para $p = s = 1$ convierte todo método de paso simple de primer orden en un método simétrico de segundo orden, tal que:

$$\Psi_h = \Phi_{h/2} \circ \Phi_{h/2}^*\tag{2.7}$$

Por ejemplo, si Φ_h es el método de Euler explícito, entonces Ψ_h en (2.7) define el método de punto medio implícito. En próximos capítulos se utilizará esta técnica para incrementar el orden de un método de integración geométrica de gradiente discreto, siendo éste uno de los principales aportes de la tesis.

2.4. Método de proyección que conserva la función de Lyapunov

En esta sección se trabaja con un método de proyección diseñado específicamente con el objetivo de conservar la función de Lyapunov del sistema continuo. Los resultados que se discuten a continuación fueron publicados con anterioridad en [12].

2.4.1. Definición del método de proyección

Para diseñar el método de proyección, los autores proponen el uso de un método de Runge-Kutta de s etapas explícito al que denotan por φ_h . Este método es de orden p y está definido por los coeficientes a_{ij} , b_i . Otra característica requerida de φ_h es que esté provisto de salida densa, la cual sea de orden \bar{p} con \bar{s} etapas. Usualmente, $\bar{p} \geq p - 1$, con $\bar{s} \geq s$. Específicamente, la salida del método φ_h para un instante

intermedio $t_n \leq t_n + \theta h \leq t_n + h$ está definida por las ecuaciones:

$$y_{n+\theta} = y_n + h \sum_{i=1}^{\bar{s}} b_i(\theta) g_i, \quad \theta \in [0, 1] \quad (2.8)$$

$$g_i = f \left(y_n + h \sum_{j=1}^{i-1} a_{ij} g_j \right), \quad i = 1, \dots, \bar{s} \quad (2.9)$$

Considérese un problema general de valores iniciales:

$$\begin{aligned} \frac{dy}{dt} &= F(t, y), \quad t \geq t_0 \\ y(t_0) &= y_0 \end{aligned} \quad (2.10)$$

Al aplicar el método φ_h , se calcula una aproximación y_{n+1} de la solución de (2.10) en el tiempo $t_{n+1} = t_n + h$ a partir de los valores de (t_n, y_n) , siguiendo el esquema que se detalla a continuación:

Paso 1: Después de calcular las etapas internas g_i , se calcula $\tilde{y}_{n+1} = \varphi_h(y_n)$, aproximación dada por el método de Runge-Kutta básico:

$$\tilde{y}_{n+1} = y_n + h \sum_{i=1}^s b_i g_i \quad (2.11)$$

Paso 2: Se calcula una aproximación V_{n+1} de la función de Lyapunov:

$$V_{n+1} = V(y_n) + h \sum_{i=1}^m b_i^* \alpha(X_{n+c_i^*}) \quad (2.12)$$

siendo b_i^* y c_i^* , $i = 1, \dots, m$ los coeficientes y los nodos de la fórmula de cuadratura Gaussiana en $[0, 1]$ y α el flujo de la ecuación diferencial (2.10). Los valores $y_{n+c_i^*} \approx y(t_n + c_i^* h)$, $i = 1, \dots, m$ se calculan según (2.8) y corresponden al método de Runge-Kutta con salida densa.

El sentido de la ecuación (2.12) resulta evidente al considerar que:

$$\frac{dV(y)}{dt} = \nabla V(y) \cdot F(y) \equiv \alpha(y)$$

Entonces:

$$\int_{t_n}^{t_{n+1}} \alpha(y(t)) dt = V(y_{n+1}) - V(y_n) \quad (2.13)$$

Esta integral se puede reescribir en el intervalo $[-1, 1]$, para esto se considera el siguiente cambio de variable:

$$s = \frac{2t - t_{n+1} - t_n}{t_{n+1} - t_n} = \frac{2t - t_n - h - t_n}{h} = 2\left(\frac{t - t_n}{h}\right) - 1 \Rightarrow t - t_n = \frac{h}{2}(s + 1)$$

Con lo que finalmente se tiene que:

$$t = t_n + \frac{h}{2}(s + 1)$$

Resulta evidente que si $t \in [t_n, t_{n+1}]$ entonces $s \in [-1, 1]$. Por otra parte, se tiene que $dt = \frac{h}{2} ds$, con lo que la siguiente cadena de igualdades es válida:

$$\int_{t_n}^{t_{n+1}} \alpha(y(t)) dt = \int_{-1}^1 \alpha\left(y\left(t_n + \frac{h}{2}(s + 1)\right)\right) \frac{h}{2} ds = h \int_0^1 \alpha(y(t_n + zh)) dz$$

La última igualdad se cumple al tomar $z = \frac{s + 1}{2}$, ya que entonces $dz = \frac{ds}{2}$.

Denotando:

$$\beta(s) = \frac{h}{2} \alpha\left(y\left(t_n + \frac{h}{2}(s + 1)\right)\right)$$

se tiene:

$$\int_{-1}^1 \alpha\left(y\left(t_n + \frac{h}{2}(s + 1)\right)\right) \frac{h}{2} ds = \int_{-1}^1 \beta(s) ds \approx \sum_{i=1}^m b_i^* \beta(s_i)$$

y considerando $s_i = 2c_i^* - 1$ entonces:

$$h \int_0^1 \alpha(y(t_n + zh)) dz \approx \sum_{i=1}^m b_i^* \beta(2c_i^* - 1) = \sum_{i=1}^m b_i^* \frac{h}{2} \alpha(y(t_n + hc_i^*))$$

Introduciendo esta aproximación en (2.13), se tiene:

$$V(y_{n+1}) - V(y_n) = h \int_0^1 \alpha(y(t_n + zh)) dz \approx \frac{h}{2} \sum_{i=1}^m b_i^* \alpha(y(t_n + hc_i^*))$$

Por tanto, las aproximaciones $y(t_n + hc_i^*) \approx y_{n+c_i^*}$ se calculan utilizando el método de Runge-Kutta con salida densa. Finalmente se tiene que:

$$V(y_{n+1}) \approx V(y_n) + \frac{h}{2} \sum_{i=1}^m b_i^* \alpha(y_{n+c_i^*}) \equiv V_{n+1}$$

Paso 3: Se calcula y_{n+1} , que es la proyección de \widetilde{y}_{n+1} sobre la variedad diferenciable de dimensión $(n - 1)$ dada por:

$$\mathcal{M} = \{y : V(y) = V_{n+1}\} \tag{2.14}$$

Existen varias técnicas de proyección que se pueden utilizar en este paso del algoritmo. En este caso se utilizará una de las más conocidas, la proyección ortogonal. Esta técnica permite calcular la proyección de y_{n+1} , como:

$$y_{n+1} = \widetilde{y}_{n+1} + \lambda_n w_n \tag{2.15}$$

donde $w_n = w_n(y_n, h) \in \mathbb{R}^n$ define la dirección de proyección y $\lambda_n = \lambda_n(y_n, h) \in \mathbb{R}$ se escoge de manera tal que $y_{n+1} \in \mathcal{M}$. La proyección estándar calcula y_{n+1} resolviendo el problema de optimización:

$$\min \|y_{n+1} - \widetilde{y}_{n+1}\|$$

$$\text{sujeto a: } V(y_{n+1}) = V_{n+1}$$

Si la norma considerada es la euclidiana, las condiciones necesarias de Lagrange, después de aproximar $\nabla V(\tilde{y}_{n+1}) \approx \nabla V(y_{n+1})$ conducen a la resolución del sistema:

$$y_{n+1} = \tilde{y}_{n+1} + \lambda_n \nabla V(\tilde{y}_{n+1}) \quad \text{con} \quad V(y_{n+1}) = V_{n+1} \quad (2.16)$$

Se dice entonces que y_{n+1} es la proyección ortogonal de \tilde{y}_{n+1} sobre \mathcal{M} , aunque es válido señalar que se pueden escoger otras direcciones de proyección.

Obsérvese que, una vez escogida la dirección de proyección w_n , el valor obtenido equivale a resolver en λ la ecuación:

$$V(\varphi_h(y_n) + \lambda w_n) - V_{n+1} = 0 \quad (2.17)$$

Nótese que esta ecuación es no lineal para λ y que existen varias maneras de resolver una ecuación de este tipo. Por ejemplo, se podría utilizar el método de Newton, aunque el tipo de ecuación que se obtenga en cada caso, puede sugerir la utilización de otras vías de solución. Es estrictamente necesario que la ecuación (2.17) tenga solución ya que esto garantiza que $y_{n+1} \in \mathcal{M}$. El siguiente teorema garantiza la existencia de la solución bajo ciertas condiciones.

Teorema 2.3.

Se considera el método de proyección descrito en los pasos 1, 2 y 3, el cual conserva la función de Lyapunov V del sistema de ecuaciones diferenciales (2.10). Si además se tiene que $\nabla V(y_n)^T \cdot w_n(y_n, 0) \neq 0$ entonces:

- (i) *Existe $h^* > 0$ tal que $V(\varphi_h(y_n) + \lambda w_n) - V_{n+1} = 0$ define una única función $\lambda_n = \lambda(y_n, h)$ para cualquier $h \in [0, h^*]$.*
- (ii) *El orden del método de proyección es $\hat{p} \geq \min\{p, \bar{p} + 1, 2m\}$*

La prueba de este teorema se encuentra en [12]. Una vez realizado un análisis exhaustivo de este método teóricamente, resulta necesario para una mejor comprensión ilustrar el funcionamiento práctico del mismo. Es por esto que en la próxima sección se escoge un sistema de EDO sencillo y se integra mediante este método,

pretendiendo de esta manera darle sentido a cada uno de los pasos del algoritmo que lo definen.

2.4.2. Ilustración del funcionamiento del método

Con el objetivo de ilustrar el funcionamiento del método propuesto anteriormente, se introduce un ejemplo sencillo y ampliamente estudiado. El objetivo es discretizar este sistema mediante el método de proyección que se introdujo en la sección anterior. El sistema en estudio es:

$$\begin{aligned}\frac{dy}{dt} &= -ay, \quad a > 0 \\ y(0) &= 1\end{aligned}\tag{2.18}$$

De la teoría de ecuaciones diferenciales ordinarias se sabe que este problema de Cauchy tiene solución analítica, la cual viene dada por $y(t) = e^{-at}$, con lo que siempre que $a > 0$ el sistema será estable. También se puede afirmar que $V = \frac{1}{2}y^2$ es función de Lyapunov de este sistema, ya que:

$$\alpha(y) = \nabla V \cdot F = -ay^2 < 0 \quad \text{para todo } a > 0$$

Nótese que $F = -ay$ es la parte derecha de la ecuación diferencial y que el gradiente $\nabla V = y$. Esta observación muestra que el sistema en estudio posee función de Lyapunov.

A continuación se procede a aplicar el método de proyección, para lo que se siguen los 3 pasos del algoritmo que lo definen.

Paso 1: Calcular $\tilde{y}_{n+1} = \varphi_h(y_n)$ siendo φ_h un método de Runge-Kutta de orden p . Para mostrar el funcionamiento del método y a la vez simplificar el trabajo algebraico, basta con tomar φ_h como la regla de Euler, considerándola como un Runge-Kutta de primer orden, resultando:

$$\tilde{y}_{n+1} = y_n + hF(y_n) = y_n - ah y_n$$

con lo que:

$$\tilde{y}_{n+1} = (1 - ah) y_n \quad (2.19)$$

Luego de calcular \tilde{y}_{n+1} , se procede al cálculo de la aproximación de la función de Lyapunov V_{n+1} . Para esto se trabaja con el mismo método del paso anterior, pero considerando su salida densa.

Paso 2: Según el método se tiene que calcular V_{n+1} de manera tal que:

$$V_{n+1} = V(y_n) + h \sum_{i=1}^m b_i^* \alpha(y_{n+c_i^*}) \quad (2.20)$$

siendo b_i^* , c_i^* , $i = 1, \dots, m$ los coeficientes y nodos de la cuadratura Gaussiana en $[0, 1]$ respectivamente. Por su parte los valores $y_{n+c_i^*} \simeq y(t_n + c_i^* h)$ se calculan utilizando la salida densa del método utilizado en el paso anterior.

Ahora bien, siendo α el flujo de la ecuación (2.18) y V su función de Lyapunov, se tiene:

$$\int_{t_n}^{t_n+h} \alpha(y) dt = V(y(t_{n+1})) - V(y_n)$$

de manera que haciendo el cambio de variable $\sigma = \frac{t - t_n}{h}$ se obtiene:

$$\int_{t_n}^{t_n+h} \alpha(y) dt = h \int_0^1 \alpha(y(t_n + \sigma h)) d\sigma$$

pudiendo ahora hacer corresponder σ a los nodos de la cuadratura Gaussiana. En particular, dado que solo se pretende alcanzar primer orden, se trabaja con $\sigma = \frac{1}{2}$ y su coeficiente asociado $b^* = 1$. Entonces, aplicando la fórmula de la cuadratura en $[0, 1]$:

$$h \int_0^1 \alpha(y(t_n + \sigma h)) d\sigma \simeq \frac{h}{2} \alpha\left(y\left(t_n + \frac{h}{2}\right)\right)$$

para luego aplicar la regla de Euler en $t_n + \frac{h}{2}$:

$$\frac{h}{2} \alpha \left(y \left(t_n + \frac{h}{2} \right) \right) \approx \frac{h}{2} \alpha \left(y_n + \frac{h}{2} F(y_n) \right) = \frac{h}{2} \alpha \left(y_n - a \frac{h}{2} y_n \right) = \frac{h}{2} \alpha \left(\left(1 - a \frac{h}{2} \right) y_n \right)$$

y teniendo en cuenta que $\alpha(y) = -ay^2$ se obtiene:

$$\frac{h}{2} \alpha \left(\left(1 - a \frac{h}{2} \right) y_n \right) = -a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 y_n^2$$

Sustituyendo en (2.20):

$$V_{n+1} = V(y_n) - a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 y_n^2 = \frac{1}{2} y_n^2 - a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 y_n^2$$

con lo que, finalmente, se tiene:

$$V_{n+1} = \left[\frac{1}{2} - a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 \right] y_n^2 \quad (2.21)$$

Luego de calcular las aproximaciones tanto de la solución como de la función de Lyapunov, a las que se denotan como \tilde{y}_{n+1} y V_{n+1} respectivamente, solo faltaría obtener y_{n+1} que, como se vio anteriormente, es la proyección de \tilde{y}_{n+1} sobre la variedad \mathcal{M} . Para esto basta aplicar el tercer paso del método.

Paso 3: En este paso se calcula y_{n+1} como la proyección de \tilde{y}_{n+1} sobre la variedad:

$$\mathcal{M} = \{y : V(y_n) = V_{n+1}\}$$

Según la ecuación (2.16), el objetivo es calcular λ de manera tal que la siguiente ecuación tenga sentido:

$$0 = V(\tilde{y}_{n+1} + \lambda \nabla V(\tilde{y}_{n+1})) - V_{n+1} = V((1 + \lambda)\tilde{y}_{n+1}) - V_{n+1}$$

la segunda igualdad se satisface ya que en este problema, $\nabla V(y) = y$. Entonces,

sustituyendo las expresiones de \tilde{y}_{n+1} , V_{n+1} , dadas en (2.19) y (2.21), se tiene:

$$V((1 + \lambda)(1 - ah)y_n) - \left[\frac{1}{2} - a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 \right] y_n^2 = 0$$

$$\frac{1}{2}(1 + \lambda)^2(1 - ah)^2 y_n^2 - \left[\frac{1}{2} - a \frac{h}{2} \left(1 - a \frac{h}{2} \right)^2 \right] y_n^2 = 0$$

que, tras la correspondiente manipulación algebraica, resulta:

$$(1 + \lambda)^2(1 - ah)^2 = 1 - ah \left(1 - a \frac{h}{2} \right)^2$$

$$(1 + \lambda)^2 = \frac{1 - ah \left(1 - a \frac{h}{2} \right)^2}{(1 - ah)^2}$$

$$1 + \lambda = \frac{\sqrt{1 - ah \left(1 - a \frac{h}{2} \right)^2}}{1 - ah}$$

obteniéndose:

$$\lambda = \frac{\sqrt{1 - ah + (ah)^2 - \frac{(ah)^3}{4}}}{1 - ah} - 1 \quad (2.22)$$

Sustituyendo el valor de λ y \tilde{y}_{n+1} en la ecuación (2.16), se tiene:

$$y_{n+1} = (1 - ah)y_n + \left(\frac{\sqrt{1 - ah + (ah)^2 - \frac{(ah)^3}{4}}}{1 - ah} - 1 \right) (1 - ah)y_n$$

con lo que finalmente se obtiene el método de proyección buscado:

$$y_{n+1} = \sqrt{1 - ah + (ah)^2 - \frac{(ah)^3}{4}} y_n \quad (2.23)$$

Nótese que el método obtenido es explícito y que, teniendo en cuenta la manera en

la que está diseñado, permite conservar la función de Lyapunov del sistema continuo. Sin embargo, se puede apreciar en este ejemplo que no respeta de manera general la dinámica del sistema continuo. Una de las circunstancias que pueden dar lugar a un comportamiento incorrecto es la aparición de puntos espurios, o sea, soluciones del sistema discreto que no lo son del continuo. Por ejemplo, en los casos en los que el término $ah = 2$, el radical que aparece en la ecuación del método se hace uno, con lo que $y_{n+1} = y_n$, o sea, el método no avanza, por lo que se entiende que ha encontrado un punto de equilibrio. Esta solución es incorrecta, dado que el único punto de equilibrio del sistema continuo es el origen de coordenadas.

Por otra parte, cuando se enunció el método, se puntualizó la necesidad de que existiera un λ de manera tal que la ecuación (2.16) tenga sentido. Sin embargo, al observar la expresión de λ obtenida en este ejemplo, se puede concluir que su existencia depende del signo que tome el polinomio de tercer grado en términos de ah que aparece en el numerador de (2.22). Ahora bien, considérese dicho polinomio:

$$f(r) = 1 - r + r^2 - \frac{r^3}{4}$$

con $r = ah$. Este polinomio tiene dos raíces complejas y una real, estando esta última localizada aproximadamente en $r \approx 7,5$. Haciendo un análisis del signo de $f(r)$ alrededor de este punto, se puede concluir que $f(r) > 0$ cuando $0 < r < 7,5$. Luego, en este intervalo, la ecuación que define a λ tendría solución y el método de proyección funcionaría correctamente, salvo por la aparición de puntos espurios. Por otra parte siempre que $r > 7,5$, $f(r) < 0$ con lo que la ecuación (2.22) no tendría solución y por ende no se podría aplicar el método de proyección.

A modo de conclusión se puede decir que este ejemplo no solo sirve para ilustrar el funcionamiento del método de proyección, sino que muestra la complejidad computacional que encierran cada uno de sus pasos. También, se evidencia la incapacidad del método para encontrar solución cualquiera sea el tamaño de paso de discretización. Nótese que en el análisis anterior, se concluye que el buen funcionamiento del método depende del valor de ah , por lo que el tamaño de paso juega un papel fundamental. En próximos capítulos se compara la eficiencia de este método

con el de un método de gradiente discreto. La elección crítica del tamaño de paso será un punto clave de la comparación.

2.5. Conclusiones

En este capítulo se ha centrado la atención en los métodos de integración numérica geométrica. Se vio cómo la composición de un método con su adjunto es una estrategia plausible a la hora de incrementar el orden de un determinado método numérico, a la vez que se conservan las propiedades del sistema continuo en cuestión.

También en este capítulo se han discutido aspectos claves del funcionamiento de un método de proyección diseñado para la conservación de la función de Lyapunov de un determinado sistema de EDO. Este método fue propuesto en [12] y los autores demuestran que, siguiendo un determinado algoritmo de trabajo, se puede diseñar un método de proyección para el que existirá un h de manera tal que se conserve la función de Lyapunov del sistema continuo. Hay que señalar que el proceso de obtención de este valor del paso de discretización no es constructivo, por lo que en la práctica, para obtener un funcionamiento correcto del método, podría ser necesario escoger un tamaño de paso extremadamente pequeño. También en este capítulo se ilustra el funcionamiento del método aplicado a un sistema de EDO escalar simple, evidenciándose algunas limitaciones del mismo.

Capítulo 3

Gradientes discretos y discretización de sistemas de gradiente lineal

Resumen del capítulo

En este capítulo se describen los métodos numéricos no convencionales basados en gradientes discretos, mostrando los trabajos previos encontrados en la literatura. Asimismo, se introducen los resultados fundacionales de la tesis, al construir una familia de estos métodos. En la sección 3.1 se introduce el concepto de gradiente discreto, así como algunas de las definiciones y resultados más importantes relacionados con estos. En la sección 3.2 se enumeran algunos ejemplos de gradientes discretos, mientras que en la sección 3.3 se repasa la formulación de un método de integración de gradiente discreto. Se puntualiza el hecho de que, en principio, los métodos de gradiente discreto son métodos numéricos implícitos. Estos métodos están diseñados de manera tal que conserven las propiedades cualitativas de los sistemas dinámicos, tales como el flujo, la energía, la estabilidad, entre otras, cualquiera sea el tamaño de paso con el que se discretice el sistema continuo. Es válido señalar que este trabajo se centra específicamente en la conservación de la función de Lyapunov y por ende de la estabilidad del sistema. Comenzando con

las aportaciones novedosas del capítulo, en la sección 3.4 se describe la metodología de construcción de métodos de gradiente discreto, haciendo énfasis en la elección de los parámetros. Con el objetivo de mostrar de manera práctica el funcionamiento de dicha metodología de trabajo, se seleccionan tres sistemas escalares simples, para los que se diseñan métodos de gradiente discreto. A la vista de los resultados obtenidos con los ejemplos sencillos, se establece como gradiente discreto para el resto del trabajo el del incremento de las coordenadas. Es entonces cuando se obtiene uno de los resultados más importantes del capítulo: *con una elección adecuada de los parámetros, el método de gradiente discreto obtenido puede reescribirse de forma explícita para el caso de sistemas dinámicos con función de Lyapunov multilineal*. Finalmente se hace el estudio del orden de los métodos obtenidos en el caso general y en los casos particulares de los sistemas escalares escogidos como ejemplos ilustrativos.

3.1. Formulaciones de gradiente para sistemas dinámicos y gradientes discretos.

El diseño de métodos de gradiente discreto tiene su fundamento en una idea relativamente sencilla. Supóngase que se tiene una EDO general:

$$\begin{aligned}\frac{dy}{dt} &= F(t, y), \quad t \geq t_0 \\ y(t_0) &= y_0\end{aligned}\tag{3.1}$$

de la cual se sabe que posee una función de Lyapunov, con las características descritas en capítulos anteriores, a la que se denota por la letra V . La idea, entonces, consiste en sustituir la derivada de V por un incremento finito. Incidentalmente, se hace notar que dicha función podría ser una primera integral del sistema, pero teniendo en cuenta que el objetivo de la tesis es el diseño de métodos de integración que conserven la función de Lyapunov del sistema bajo discretización, en lo que resta, todos los resultados que se muestran sobre métodos de gradiente discreto se reducen al caso de sistemas con función de Lyapunov.

Se puede afirmar que toda EDO como en (3.1) para la cual se conoce su función de Lyapunov puede ser reescrita en la forma de *gradiente lineal* [39]:

$$\frac{dy}{dt} = L(y) \nabla V(y) \quad (3.2)$$

siendo ∇V el gradiente de la función de Lyapunov del sistema continuo y L una matriz simétrica y definida negativa. Es válido señalar que tanto L como V son funciones continuamente diferenciables, aunque se asume tanta diferenciable como sea necesaria para ambas funciones en el resto del trabajo. Además se sabe que esta descomposición no es única y que las diferentes formas de escribir $L(y)$ pueden considerarse como diferentes estructuras métricas [8]. Esta última afirmación podría dar lugar a una interesante línea de investigación, aunque en estos momentos queda fuera del alcance de este trabajo. Por su parte, los gradientes discretos pueden utilizarse para construir aproximaciones discretas de la ecuación diferencial, las cuales preservan funciones de Lyapunov exactamente.

Definición 3.1. Sea V una función diferenciable. Se dice entonces que $\bar{\nabla}V$ es un gradiente discreto de V si es una función continua y cumple que:

$$\begin{aligned} \bar{\nabla}V(y, z) \cdot (z - y) &= V(z) - V(y) \\ \bar{\nabla}V(y, y) &= \nabla V(y) \end{aligned} \quad (3.3)$$

siendo $y = y(k)$, $z = y(k + 1)$, notación que se mantiene en el resto del trabajo.

A continuación se enuncia una proposición que brinda una caracterización de la definición de gradiente discreto.

Proposición 3.1. La función $\bar{\nabla}V$ es un gradiente discreto si es continua y cumple:

$$\bar{\nabla}V(y, z) = \frac{V(z) - V(y)}{|z - y|^2} (z - y) + \omega(y, z), \quad (y \neq z)$$

donde $\omega(\mathbf{y}, \mathbf{z})$ es una función vectorial tal que:

$$\begin{cases} \omega(\mathbf{y}, \mathbf{z})(\mathbf{z} - \mathbf{y}) = 0 & \text{si } \mathbf{y} \neq \mathbf{z} \\ \lim_{\mathbf{z} \rightarrow \mathbf{y}} (\omega(\mathbf{y}, \mathbf{z}) - \pi_{(\mathbf{z}-\mathbf{y})}^\perp \nabla V(\mathbf{y})) = 0 \end{cases}$$

siendo $\pi_{(\mathbf{z}-\mathbf{y})}^\perp$ la proyección de la componente perpendicular a $\mathbf{z} - \mathbf{y}$.

Nótese que en dimensión uno existe un único gradiente discreto y viene dado por:

$$\bar{\nabla} V(\mathbf{y}, \mathbf{z}) = \frac{V(\mathbf{z}) - V(\mathbf{y})}{\mathbf{z} - \mathbf{y}} \quad (3.4)$$

En cambio en dimensiones más altas existe una gran diversidad de gradientes discretos. Una definición análoga para campos vectoriales permite estudiar el análogo discreto a la regla de la cadena.

Definición 3.2. Sea F un campo vectorial diferenciable. Entonces $\bar{D}F$ es una derivada discreta si es continua y cumple:

$$\bar{D}F(\mathbf{y}, \mathbf{z}) \cdot (\mathbf{z} - \mathbf{y}) = F(\mathbf{z}) - F(\mathbf{y})$$

$$\bar{D}F(\mathbf{y}, \mathbf{y}) = DF(\mathbf{y})$$

De la definición se deduce que $\bar{D}F$ es una derivada discreta si y solo si cada una de sus filas es un gradiente discreto para la coordenada correspondiente, es decir, el gradiente discreto es una derivada discreta unidimensional.

Proposición 3.2. (Regla de la Cadena). Sean F y G campos vectoriales, $\bar{D}F$ y $\bar{D}G$ sus correspondientes derivadas discretas. Entonces, $\bar{D}F(G(\mathbf{y}), G(\mathbf{z})) \bar{D}G(\mathbf{y}, \mathbf{z})$ es una derivada discreta para $F \circ G$.

En el siguiente epígrafe se listan algunos de los gradientes discretos de alto orden más utilizados en la literatura.

3.2. Algunos gradientes discretos.

Como se menciona anteriormente, existe una amplia gama de gradientes discretos de alto orden en la literatura, por ejemplo, el gradiente discreto del valor medio, el gradiente discreto del punto medio y el gradiente discreto del incremento de las coordenadas.

Ejemplo 3.1. El gradiente discreto del valor medio [21] viene dado por:

$$\bar{\nabla}_1 V(\mathbf{y}, \mathbf{z}) = \int_0^1 \nabla V((1 - \xi)\mathbf{y} + \xi\mathbf{z}) d\xi, \quad (\mathbf{y} \neq \mathbf{z})$$

donde $\bar{\nabla}_1 V(\mathbf{y}, \mathbf{z})$ es el promedio del gradiente de V en $[\mathbf{y}, \mathbf{z}]$ siendo $[\mathbf{y}, \mathbf{z}]$ el segmento que une \mathbf{y} y \mathbf{z} .

Ejemplo 3.2. El gradiente discreto del punto medio introducido por Gonzalez [15] se define como:

$$\bar{\nabla}_2 V(\mathbf{y}, \mathbf{z}) = \frac{V(\mathbf{z}) - V(\mathbf{y})}{|\mathbf{z} - \mathbf{y}|^2} (\mathbf{z} - \mathbf{y}) + \pi_{(\mathbf{z} - \mathbf{y})^\perp} \nabla V\left(\frac{\mathbf{z} + \mathbf{y}}{2}\right), \quad (\mathbf{y} \neq \mathbf{z})$$

Ambos gradientes discretos son exactos en el punto medio de $[\mathbf{y}, \mathbf{z}]$ cuando ∇V es lineal. En otras palabras, se tiene que $\nabla V = \bar{\nabla}_1 V = \bar{\nabla}_2 V$ en el punto medio del intervalo $[\mathbf{y}, \mathbf{z}]$ siempre que ∇V sea lineal.

Ejemplo 3.3. Otro de los gradientes más utilizados en la literatura es el gradiente discreto del incremento de las coordenadas. Este gradiente discreto está asociado con la trayectoria lineal a trozos que une \mathbf{y} con \mathbf{z} , siendo cada trozo paralelo a uno de los ejes coordenados. Este hecho contrasta con los gradientes discretos anteriormente reseñados, que se basan intuitivamente en una trayectoria a lo largo del segmento $[\mathbf{y}, \mathbf{z}]$. Este gradiente discreto se denota por $\bar{\nabla}_3 V$ para $\mathbf{y} \in \mathbb{R}^n$ tal que $\mathbf{y} = (y_1, y_2, \dots, y_n)$.

Asumiendo como orden de las coordenadas y_1, y_2, \dots, y_n se tiene:

$$\bar{\nabla}_3 V(\mathbf{y}, \mathbf{z}) = \begin{pmatrix} \frac{V(z_1, y_2, \dots, y_n) - V(y_1, y_2, \dots, y_n)}{z_1 - y_1} \\ \frac{V(z_1, z_2, y_3, \dots, y_n) - V(z_1, y_2, y_3, \dots, y_n)}{z_2 - y_2} \\ \vdots \\ \frac{V(z_1, \dots, z_{n-2}, z_{n-1}, y_n) - V(z_1, \dots, z_{n-2}, y_{n-1}, y_n)}{z_{n-1} - y_{n-1}} \\ \frac{V(z_1, \dots, z_{n-2}, z_{n-1}, z_n) - V(z_1, \dots, z_{n-2}, z_{n-1}, y_n)}{z_n - y_n} \end{pmatrix} \quad (3.5)$$

El gradiente discreto del incremento de las coordenadas se puede ver como una aproximación de primer orden del gradiente de V en el punto medio del intervalo $[\mathbf{y}, \mathbf{z}]$.

3.3. Método de integración

Para la EDO escrita en forma de gradiente lineal como en (3.2) se consideran métodos de integración de la forma:

$$\frac{\mathbf{z} - \mathbf{y}}{h} = \tilde{L}(\mathbf{y}, \mathbf{z}, h) \bar{\nabla} V(\mathbf{y}, \mathbf{z}) \quad (3.6)$$

Como se sugiere anteriormente, $\mathbf{y}(k)$ tiene como objetivo proporcionar una aproximación adecuada de $\mathbf{y}(t)$ en $t = kh$. Nótese que el método dado por la ecuación (3.6) es implícito, al menos en principio, ya que \mathbf{z} aparece en la parte derecha de la ecuación que describe el método. A continuación se describen los requisitos y el significado de los parámetros del método, $\tilde{L}, \bar{\nabla} V$.

La matriz $\tilde{L}(\mathbf{y}, \mathbf{z}, h)$ de funciones continuamente diferenciables representa una especie de discretización de $L(\mathbf{y})$, mientras que $\bar{\nabla} V(\mathbf{y}, \mathbf{z})$ es un *gradiente discreto*. Se ha demostrado en [39] que, siendo $\bar{\nabla} V$ un *gradiente discreto*, el método dado por la

ecuación (3.6) es consistente siempre que:

$$\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y}) \quad (3.7)$$

con lo que respetando esta condición y la de ser $\bar{\nabla}V$ gradiente discreto, la elección de los parámetros \tilde{L} y $\bar{\nabla}V$ puede hacerse libremente.

Por otra parte se tiene que el método (3.6) conserva V si \tilde{L} es definida negativa, o sea, la función V evaluada en la trayectoria discreta de la solución también decrece, en el sentido de la siguiente definición.

Definición 3.3. Sea la función $V(\mathbf{y})$, se dice que V es una función de Lyapunov del método numérico dado por la ecuación (3.6), si $V(\mathbf{z}) < V(\mathbf{y})$ para todo \mathbf{y} , salvo cuando $\mathbf{z} = \mathbf{y}$. En este último caso, se cumple la igualdad $V(\mathbf{z}) = V(\mathbf{y})$.

Proposición 3.3. Si $\tilde{L}(\mathbf{y}, \mathbf{z}, h)$ es definida negativa para todo $\mathbf{y}, \mathbf{z}, h$, entonces V es una función de Lyapunov de la aplicación (3.6).

Esta proposición garantiza que los métodos de integración de gradiente discreto conservan la función de Lyapunov del sistema continuo, cualquiera sea el tamaño de paso h que se utilice en la discretización.

Demostración 3.1. Para probar que V es función de Lyapunov, basta con demostrar que decrece a lo largo de las trayectorias, o sea:

$$V(\mathbf{z}) - V(\mathbf{y}) < 0$$

En efecto, se tiene:

$$\begin{aligned} V(\mathbf{z}) - V(\mathbf{y}) &= \bar{\nabla}V(\mathbf{z}, \mathbf{y})(\mathbf{z} - \mathbf{y}) = \bar{\nabla}V(\mathbf{z}, \mathbf{y}) \left(h \tilde{L}(\mathbf{y}, \mathbf{z}, h) \bar{\nabla}V(\mathbf{z}, \mathbf{y}) \right) \\ &= h \bar{\nabla}V(\mathbf{z}, \mathbf{y})^T \tilde{L}(\mathbf{y}, \mathbf{z}, h) \bar{\nabla}V(\mathbf{z}, \mathbf{y}) \end{aligned}$$

donde la primera igualdad resulta de la definición (3.3), y la segunda, de la ecuación (3.6). Ahora bien, denotando $\mathbf{v} = \bar{\nabla}V(\mathbf{z}, \mathbf{y})$ se tiene:

$$V(\mathbf{z}) - V(\mathbf{y}) = h \mathbf{v}^T \tilde{L} \mathbf{v} < 0 \quad \text{si } \mathbf{v} \neq 0$$

siempre que \tilde{L} sea definida negativa.

3.4. Formulación de un método de gradiente discreto

En esta sección se brinda el mecanismo de trabajo que se utiliza cuando se pretende discretizar un sistema de EDO, como el dado por la ecuación (3.1), con el enfoque del gradiente discreto. Para ello, se propone el siguiente esquema:

1. Reescribir el sistema en la forma de *gradiente lineal* dada por la ecuación (3.2), con L continuamente diferenciable.
2. Escoger los parámetros del método, \tilde{L} y $\bar{\nabla}V$, de manera tal que se cumpla la condición de consistencia (3.7) y la definición de gradiente discreto.
3. Finalmente, discretizar el sistema continuo, considerando un método de integración con la siguiente estructura:

$$\frac{z - y}{h} = \tilde{L}(y, z, h) \bar{\nabla}V(y, z) \quad (3.8)$$

en el que se asume, como se mencionó anteriormente, que \tilde{L} y V son continuamente diferenciables.

3.4.1. Métodos de gradiente discreto para sistemas escalares simples

Para mostrar prácticamente el mecanismo de obtención de un método de gradiente discreto, se consideran algunos ejemplos de ecuaciones escalares simples, tomados de [42]. Se escogieron estos sistemas por dos razones fundamentalmente. En primer lugar, son sistemas para los que es relativamente sencillo encontrar una función de Lyapunov. Es válido señalar que la elección de la función de Lyapunov del sistema es un punto sensible, ya que en muchos casos se puede afirmar que un determinado sistema de EDO tiene función de Lyapunov, pero no es posible determinar la ecuación que la define, ni existe una metodología de trabajo infalible establecida con este propósito. La segunda razón para la elección de estos sistemas es que presentan al

menos un punto fijo asintóticamente estable. De la teoría de estabilidad de sistemas de EDO, se conoce que siendo y^* un punto fijo asintóticamente estable del sistema, entonces $V = (y - y^*)^2$ es función de Lyapunov del sistema en cuestión, al menos localmente. Haciendo uso de estos ejemplos, se muestra además cómo una elección adecuada de la matriz \tilde{L} permite reescribir el método de gradiente en forma explícita.

Primeramente, se considera el problema de valores iniciales:

$$P1 : \quad \frac{dy}{dt} = -ay, \quad y(0) = y_0 \quad (3.9)$$

con $a > 0$. De la teoría de ecuaciones diferenciales ordinarias se sabe que este problema de Cauchy tiene solución analítica, la cual viene dada por:

$$y(t) = y_0 e^{-at} \quad (3.10)$$

por lo que siempre que $a > 0$ este sistema será asintóticamente estable. Esto permite proponer como candidata a función de Lyapunov del sistema a la función $V = \frac{1}{2}y^2$. Para demostrar que efectivamente V es función de Lyapunov de $P1$, basta con comprobar que se cumplen las condiciones de la definición dada en capítulos anteriores. Estas son triviales, salvo la cuestión crítica de que V decrezca a lo largo de las trayectorias. Se tiene, entonces:

$$\frac{dV}{dt} = \nabla V \cdot f = y(-ay) = -ay^2 < 0$$

que se cumple para todo y , siempre que $a > 0$. En resumen, efectivamente $V = \frac{1}{2}y^2$ puede considerarse como función de Lyapunov del sistema (3.9).

Luego de escoger la función de Lyapunov del sistema, el primer paso para obtener un método de gradiente discreto es reescribirlo en la forma de gradiente lineal (3.2). Se puede escribir:

$$\frac{dy}{dt} = -ay = L(y) \nabla(V) \quad (3.11)$$

teniendo en cuenta que $\nabla V = y$ y tomando $L = -a$. Obsérvese además que con esta elección de L se garantiza que $L < 0$.

El segundo paso para la obtención del método de gradiente discreto es la selección de los parámetros del método $\tilde{L}(y, z, h)$ y $\bar{\nabla}V(y, z)$. Los parámetros se escogen de manera tal que cumplan la condición de consistencia, $\tilde{L}(y, y, 0) = L(y)$ y la definición de gradiente discreto, $\bar{\nabla}V(y, y) = \nabla V(y)$. En este caso, la elección del gradiente es evidente, ya que se está trabajando en dimensión uno, en la que existe un único gradiente discreto, cuya definición está dada en (3.4). Resulta, entonces:

$$\bar{\nabla}V(y, z) = \frac{z^2 - y^2}{2(z - y)} = \frac{z + y}{2} \quad (3.12)$$

Para la elección de $\tilde{L}(y, z, h)$ hay una gran variedad de opciones. En este caso se hace la elección que resulta más natural, $\tilde{L} = L$, teniendo en cuenta que L es constante.

Finalmente sustituyendo la elección de los parámetros en (3.8), se tiene:

$$z = y + h\tilde{L}(y, z)\bar{\nabla}V(y, z) = y - \frac{ah(z + y)}{2} = y - \frac{ah}{2}z - \frac{ah}{2}y \quad (3.13)$$

de manera que puede despejarse z :

$$z = \frac{1 - \frac{ah}{2}}{1 + \frac{ah}{2}}y \quad (3.14)$$

obteniéndose así un método de gradiente discreto, que además de poseer todas las características que tienen estos métodos por construcción, también es explícito.

Como segundo ejemplo, se considera el sistema de EDO con condiciones iniciales:

$$P2 : \quad \frac{dy}{dt} = -ay^3, \quad y(0) = y_0 \quad (3.15)$$

con $a > 0$. En este caso también es posible obtener la solución analítica del sistema y viene dada por la expresión:

$$y(t) = \frac{y_0}{\sqrt{2at y_0^2 + 1}} \quad (3.16)$$

Existen varias formas de comprobar que el punto de equilibrio $y^* = 0$ es asintóticamente estable. El jacobiano de la EDO dada por la ecuación (3.15) es negativo en $y^* = 0$, o bien el límite $t \rightarrow \infty$ de la solución exacta dada por la ecuación (3.16) es nulo. En este caso, también se propone como función de Lyapunov a $V = \frac{1}{2}y^2$. Realizando un análisis similar al efectuado en el ejemplo anterior se demuestra que, en efecto, V es función de Lyapunov del sistema. Además, esta propiedad se cumple globalmente, ya que:

$$\frac{dV}{dt} = \nabla V \cdot f = y(-ay^3) = -ay^4 < 0$$

para cualquier $y \neq 0$. Dada la definición de V y tomando $L(y) = -ay^2$ se reescribe el sistema (3.15) en la forma de gradiente lineal, tal que:

$$\frac{dy}{dt} = -ay^3 = -ay^2(y) = L(y)\nabla V \quad (3.17)$$

Al igual que en el ejemplo anterior, se deben escoger los parámetros del método, $\tilde{L}(y, z, h)$ y $\bar{\nabla}V(y, z)$ de manera tal que cumplan las condiciones que garantizan la consistencia del método de gradiente discreto que se obtendrá a continuación. En este ejemplo, al igual que en el caso anterior, existe una única posibilidad para la elección del gradiente discreto, ya que se está trabajando también en dimensión uno. Teniendo en cuenta que se escogió V como en P1, el gradiente discreto vendrá dado por la misma expresión, o sea:

$$\bar{\nabla}V(y, z) = \frac{z+y}{2} \quad (3.18)$$

También haciendo la elección trivial, $\tilde{L} = L$ se tiene:

$$z = y + h\tilde{L}(y, z)\bar{\nabla}V(y, z) = y - \frac{ahy^2}{2}(z+y) = y - \frac{ahy^2z}{2} - \frac{ahy^3}{2} \quad (3.19)$$

de manera que en este caso también es posible despejar z :

$$z = \frac{1 - \frac{ah}{2}y^2}{1 + \frac{ah}{2}y^2}y \quad (3.20)$$

obteniéndose así un método de gradiente discreto que, al igual que en el ejemplo anterior, es explícito.

Como tercer y último ejemplo, se considera el problema de valores iniciales, conocido en la literatura como ecuación logística:

$$P3 : \quad \frac{dy}{dt} = ay(1 - y), \quad y(0) = y_0 \quad (3.21)$$

con $y > 0$, $a > 0$. Al igual que en los ejemplos anteriores, es posible obtener la solución analítica de este sistema y viene dada por:

$$y(t) = \frac{1}{1 + \left(\frac{1}{y_0} - 1\right)e^{-at}} \quad (3.22)$$

En este caso, el sistema presenta dos puntos de equilibrio, $y^* = 0$, $y^* = 1$. Razonando de manera similar a como se hizo anteriormente, se puede demostrar que $y^* = 1$ es asintóticamente estable. Por tanto, una buena candidata a función de Lyapunov en este caso, sería la función $V = \frac{1}{2}(1 - y)^2$. Se verifica fácilmente que cumple las condiciones para ser función de Lyapunov del sistema para este punto de equilibrio, ya que:

$$\frac{dV}{dt} = \nabla V \cdot f = -(1 - y)ay(1 - y) = -ay(1 - y)^2 < 0 \quad (3.23)$$

siempre que $y > 0$, $y \neq 1$, $a > 0$. Es válido señalar que, a diferencia de los casos anteriores, en este caso la condición para que V sea función de Lyapunov del sistema solo se satisface localmente. La importancia de este factor radica en que el método numérico puede dar lugar a una trayectoria que salga del dominio en el que la V seleccionada es función de Lyapunov. Ahora bien, V es función de Lyapunov del

sistema (3.21) en una vecindad de $y^* = 1$, la cual es válida para cualquier valor inicial $y_0 > 0$. Entonces, la EDO puede reescribirse en la forma de gradiente lineal, definiendo:

$$L(y) = -ay \quad (3.24)$$

de manera que:

$$\frac{dy}{dt} = ay(1-y) = L(y) \nabla V \quad (3.25)$$

con $L(y)$ definida negativa, como se requiere. La formulación general del gradiente discreto en este caso también coincide con el dado en la ecuación (3.4) ya que es de dimensión uno, con lo que:

$$\bar{\nabla}V(y, z) = \frac{(1-z)^2 - (1-y)^2}{2(z-y)} = \frac{-2+z+y}{2} \quad (3.26)$$

Respecto a la elección de $\tilde{L}(y, z, h)$ también es este caso se asume $\tilde{L} = L$. La construcción del método de gradiente discreto es análoga a los casos anteriores. Sustituyendo entonces los parámetros escogidos en la ecuación (3.8) se obtiene el método de gradiente discreto:

$$\begin{aligned} z &= y + h\tilde{L}(y, z)\bar{\nabla}V(y, z) \\ &= y - ah y \left(\frac{-2+z+y}{2} \right) \\ &= y + ah y - \frac{ah}{2} y z - \frac{ah}{2} y^2 \end{aligned}$$

Despejando z , se tiene:

$$z = \frac{\left(1 + ah - \frac{ah}{2}y\right)y}{1 + \frac{ah}{2}y} \quad (3.27)$$

Como se observa, en este caso el método de gradiente discreto obtenido también es explícito, a pesar de que la definición de estos métodos es implícita. A modo de conclusión, se puede decir que para sistemas de dimensión uno con función de Lyapunov cuadrática, siempre que se escoja $\tilde{L} = L$, el método de gradiente discreto obtenido puede reescribirse en forma explícita.

3.4.2. Formulación de un método de gradiente discreto como un método numérico explícito

Como se vio en el epígrafe anterior, bajo ciertas condiciones de los parámetros, es posible obtener métodos de integración de gradiente discreto explícitos. En esta sección se formaliza este resultado mostrando una familia de sistemas para los que se puede escribir el método de gradiente discreto en forma explícita.

A partir de este momento, se trabajará con el gradiente discreto del incremento de las coordenadas. Es válido señalar que, en lo sucesivo, se prescindirá del subíndice que aparece en la definición inicial de este gradiente. Dicho esto y teniendo en cuenta la definición del gradiente discreto del incremento de las coordenadas (3.5), se puede reescribir su coordenada i -ésima de la siguiente manera:

$$(\bar{\nabla}V)_i(\mathbf{y}, \mathbf{z}) = \frac{V(z_1, \dots, z_i, y_{i+1}, \dots, y_n) - V(z_1, \dots, z_{i-1}, y_i, y_{i+1}, \dots, y_n)}{z_i - y_i} \quad (3.28)$$

Considerando sistemas de EDO cuya función de Lyapunov sea multilineal, entonces V se puede reescribir según:

$$V(y_1, \dots, y_n) = A_i(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n)y_i + B_i(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n) \quad (3.29)$$

para cada componente i , siendo A_i y B_i funciones escalares. Nótese que se dan las igualdades $A_i = \frac{\partial V}{\partial y_i}$, $\frac{\partial A_i}{\partial y_i} = 0$, $B_i = V|_{y_i=0}$. Entonces, la i -ésima componente del incremento de la función de Lyapunov V , es decir, el numerador de (3.28), puede ser calculado de la siguiente manera:

$$\begin{aligned} \Delta_i V &= V(z_1, \dots, z_{i-1}, z_i, y_{i+1}, \dots, y_n) - V(z_1, \dots, z_{i-1}, y_i, y_{i+1}, \dots, y_n) \\ &= A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n)z_i + B_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) - \\ &\quad - A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n)y_i - B_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) \\ &= A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n)(z_i - y_i) \end{aligned}$$

Sustituyendo esta expresión en (3.28), se tiene que la i -ésima componente del gra-

diente discreto del incremento de las coordenadas viene dada por:

$$\left(\overline{\nabla}V\right)_i(\mathbf{y}, \mathbf{z}) = A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) \quad (3.30)$$

siempre que $y_i \neq z_i$. Ahora bien, por brevedad se escribe:

$$A_i(\mathbf{y}, \mathbf{z}) = A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n)$$

enfaticando que $A_i(\mathbf{y}, \mathbf{z})$ no depende de z_j para $j \geq i$.

El siguiente paso para el diseño de un método de gradiente discreto es la elección de $\widetilde{L}(\mathbf{y}, \mathbf{z}, h)$, donde el único requerimiento es la condición de consistencia (3.7). Se asume ahora que es posible escoger $\widetilde{L}(\mathbf{y}, \mathbf{z}, h)$ como una matriz diagonal donde cada componente \widetilde{L}_{ii} dependa linealmente de z_i , es decir que se pueda escribir como:

$$\widetilde{L}_{ii}(\mathbf{y}, \mathbf{z}, h) = M_{ii}(\mathbf{y}) + N_{ii}(\mathbf{y}) z_i \quad (3.31)$$

siendo M, N funciones matriciales diagonales. Entonces, al sustituir la matriz \widetilde{L} definida en (3.31) y el gradiente del incremento de las coordenadas en (3.8), se tiene:

$$\begin{aligned} z_i - y_i &= h \widetilde{L}_{ii}(\mathbf{y}, \mathbf{z}, h) \left(\overline{\nabla}V\right)_i(\mathbf{y}, \mathbf{z}) = h [M_{ii}(\mathbf{y}) + N_{ii}(\mathbf{y}) z_i] A_i(\mathbf{y}, \mathbf{z}) \\ &= h M_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z}) + h N_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z}) z_i \end{aligned}$$

con lo que se puede despejar z_i :

$$\begin{aligned} z_i - h N_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z}) z_i &= y_i + h M_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z}) \\ z_i (1 - h N_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z})) &= y_i + h M_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z}) \end{aligned}$$

para, finalmente, obtener la expresión explícita:

$$z_i = \frac{y_i + h M_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z})}{1 - h N_{ii}(\mathbf{y}) A_i(\mathbf{y}, \mathbf{z})} \quad (3.32)$$

Aunque la parte derecha de la expresión formalmente depende del vector \mathbf{z} , en la práctica \mathbf{z} puede calcularse sin tener que resolver una ecuación no lineal, ya que

solo aparece en (3.32) a través del gradiente discreto $\bar{\nabla}V(\mathbf{y}, \mathbf{z})$. Nótese que en la ecuación (3.30) la componente i -ésima del gradiente discreto solo depende de las componentes z_j de \mathbf{z} con $j < i$. Por tanto, al calcular z_i , todas las variables en la ecuación (3.32) ya han sido calculadas. En resumen, el método definido por la ecuación (3.30) es una expresión explícita que puede ser calculada secuencialmente. A modo de conclusión se puede decir que, cuando V es multilineal, entonces se puede obtener un método numérico explícito con el gradiente discreto del incremento de las coordenadas, escogiendo los elementos de la diagonal de \tilde{L} como funciones lineales de la variable z_i .

3.5. Análisis del orden del método

En esta sección se estudia el orden del método numérico obtenido en la sección anterior mediante el procedimiento sistemático habitual [19]: el desarrollo en serie de Taylor alrededor de $h = 0$, tanto de la solución exacta del sistema de ecuaciones diferenciales, como de la solución aproximada obtenida por el método numérico. Es válido señalar que el método de gradiente discreto es consistente por construcción [39], con lo que es al menos de orden uno. Luego el objetivo principal de esta sección es determinar si este método numérico eventualmente podría alcanzar orden $p > 1$ y calcular la constante de error del método. En otras palabras, se considera, por una parte, la solución exacta $\mathbf{y}(t)$ de la ecuación diferencial \mathbf{y} , por otra, la solución aproximada \mathbf{y} provista por el método numérico. Asumiendo que no se comete ningún error al calcular la solución en t , el error local de un método de orden p con constante de error C puede ser acotado por la expresión $\|\mathbf{y}(t+h) - \mathbf{z}\| \leq Ch^{p+1}$, donde \mathbf{z} es el resultado de un paso h del método numérico. A continuación, se muestra este análisis en detalle, para luego aplicarlo a sistemas escalares simples, específicamente a los sistemas $P1$, $P2$ y $P3$, con los que se trabajó anteriormente.

3.5.1. Análisis del orden del método de gradiente discreto

Para analizar el orden de un método numérico, primeramente se realiza el desarrollo en serie de Taylor de la solución exacta $\mathbf{y}(t)$, para luego compararlo con el desarrollo de la aproximación numérica. En general el método de gradiente discreto es de orden uno, ya que los términos de segundo orden no se anulan.

Se sabe que un sistema de EDO puede reescribirse en la forma de gradiente lineal cuando posee función de Lyapunov, o sea:

$$\frac{d\mathbf{y}}{dt} = f(\mathbf{y}) = L(\mathbf{y}) \nabla V(\mathbf{y}) \quad (3.33)$$

Por claridad de expresión, se omitirá la dependencia de t de $\mathbf{y}(t)$ siempre que no cree confusión. Por tanto, el desarrollo en serie de Taylor de la solución exacta de la EDO viene dado por:

$$\begin{aligned} \mathbf{y}(t+h) &= \mathbf{y} + h \frac{d\mathbf{y}}{dt} + \frac{h^2}{2} \frac{d^2\mathbf{y}}{dt^2} + O(h^3) \\ &= \mathbf{y} + h f(\mathbf{y}) + \frac{h^2}{2} J^c(\mathbf{y}) f(\mathbf{y}) + O(h^3) \\ &= \mathbf{y} + h L(\mathbf{y}) \nabla V(\mathbf{y}) + \frac{h^2}{2} J^c(\mathbf{y}) f(\mathbf{y}) + O(h^3) \end{aligned} \quad (3.34)$$

siendo $J^c(\mathbf{y})$ la matriz Jacobiana de f para \mathbf{y} :

$$J^c = \frac{\partial f}{\partial \mathbf{y}} = \left[\frac{\partial f_i}{\partial y_j} \right]_{ij} = \left[\frac{\partial (L(\mathbf{y}) \nabla V(\mathbf{y}))_i}{\partial y_j} \right]_{ij} \quad i, j = 1, \dots, n \quad (3.35)$$

A continuación se considera la solución \mathbf{z} del método numérico de gradiente discreto, como en la ecuación (3.32). Se supone, para simplificar, que la aproximación \tilde{L} no depende del tamaño de paso h . Se puede entonces definir la función F :

$$F(\mathbf{z}, h) = \mathbf{z} - \mathbf{y} - h \tilde{L}(\mathbf{y}, \mathbf{z}) \cdot \bar{\nabla} V(\mathbf{y}, \mathbf{z}) = \mathbf{0} \quad (3.36)$$

La función $F(\mathbf{z}, h)$ define implícitamente $\mathbf{z} = g(h)$. Por tanto, el desarrollo en serie

de Taylor de la solución numérica en $h = 0$ viene dado por:

$$z = y + h \left. \frac{dg}{dh} \right|_{h=0} + \frac{h^2}{2} \left. \frac{d^2g}{dh^2} \right|_{h=0} + O(h^3) \quad (3.37)$$

ya que $g(0) = y$. Entonces, la derivada $\frac{dg}{dh}$ se obtiene con ayuda del teorema de la función implícita:

$$\frac{dF}{dh} = \frac{\partial F}{\partial z} \cdot \frac{dz}{dh} + \frac{\partial F}{\partial h} = \frac{\partial F}{\partial z} \cdot \frac{dg}{dh} + \frac{\partial F}{\partial h} = 0 \quad (3.38)$$

tal que:

$$\frac{dg}{dh} = - \left(\frac{\partial F}{\partial z} \right)^{-1} \cdot \frac{\partial F}{\partial h} \quad (3.39)$$

Teniendo en cuenta la expresión de F dada en la ecuación (3.36), se tiene, por una parte, que:

$$\frac{\partial F}{\partial z} = I_n - h \frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial z} \quad (3.40)$$

donde $\frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial z}$ es la matriz jacobiana, o sea, una matriz cuadrada en la que el elemento ij viene dado por $\frac{\partial(\tilde{L} \cdot \bar{\nabla} V)_i}{\partial z_j}$. Por otra parte:

$$\frac{\partial F}{\partial h} = -\tilde{L} \cdot \bar{\nabla} V \quad (3.41)$$

Sustituyendo (3.40) y (3.41) en (3.39) se obtiene la derivada de $g(h)$:

$$\frac{dg}{dh} = \left(I_n - h \frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial z} \right)^{-1} \cdot (\tilde{L} \cdot \bar{\nabla} V) \quad (3.42)$$

Nótese que para $h = 0$ la ecuación (3.39) se reduce a:

$$\left. \frac{dg}{dh} \right|_{h=0} = L(y) \cdot \nabla V(y) = f(y) \quad (3.43)$$

ya que se cumplen las igualdades $z = y$, $\tilde{L}(y, y, 0) = L(y)$. Ahora bien, el siguiente paso sería obtener la derivada de segundo orden de $g(h)$, lo que da lugar a calcular el diferencial de la ecuación (3.38) definiendo la función $G(z, h)$:

$$G(z, h) = \frac{dF}{dh} = \frac{\partial F}{\partial z} \cdot \frac{dg}{dh} + \frac{\partial F}{\partial h} = \left(I_n - h \frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial z} \right) \cdot \frac{dg}{dh} - (\tilde{L} \cdot \bar{\nabla} V) = 0 \quad (3.44)$$

El teorema de la función implícita garantiza que:

$$\frac{dG}{dh} = \frac{\partial G}{\partial z} \cdot \frac{dz}{dh} + \frac{\partial G}{\partial h} = \frac{\partial G}{\partial z} \cdot \frac{dg}{dh} + \frac{\partial G}{\partial h} = 0 \quad (3.45)$$

donde se utiliza la definición implícita de $z = g(h)$. Diferenciando la definición de $G(z, h)$ en la ecuación (3.44) se obtiene, por una parte:

$$\frac{\partial G}{\partial z} = -h \mathbf{H} - J^d \quad (3.46)$$

siendo \mathbf{H} la matriz cuya componente ij es $\sum_{k=1}^n \frac{\partial^2(\tilde{L} \cdot \bar{\nabla} V)}{\partial z_j \partial z_k} \frac{dg_k}{dh}$ y J^d el jacobiano de $\tilde{L} \cdot \bar{\nabla} V$, o sea:

$$J^d = \frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial z} = \left[\frac{\partial(\tilde{L} \cdot \bar{\nabla} V)_i}{\partial z_j} \right]_{ij} \quad i, j = 1, \dots, n \quad (3.47)$$

Por otra parte, se tiene que:

$$\frac{\partial G}{\partial h} = -J^d \cdot \frac{dg}{dh} + (I_n - h J^d) \cdot \frac{d^2 g}{dh^2} \quad (3.48)$$

Como se está trabajando con los desarrollos en serie de Taylor alrededor de $h = 0$, entonces:

$$\left. \frac{\partial G}{\partial z} \right|_{h=0} = -J^d \Big|_{h=0} \quad \left. \frac{\partial G}{\partial h} \right|_{h=0} = -J^d \Big|_{h=0} \cdot \frac{dg}{dh} + \frac{d^2 g}{dh^2} \quad (3.49)$$

Sustituyendo entonces en la ecuación (3.45) se obtiene la derivada de segundo orden:

$$\left. \frac{d^2 g}{dh^2} \right|_{h=0} = 2 J^d|_{h=0} \cdot \frac{dg}{dh} = 2 J^d|_{h=0} \cdot f(\mathbf{y}) \quad (3.50)$$

donde la última igualdad resulta del cálculo previo de la primera derivada, como en la ecuación (3.43). Finalmente, sustituyendo en la ecuación (3.37) resulta:

$$\mathbf{z} = \mathbf{y} + h f(\mathbf{y}) + h^2 J^d|_{h=0} f(\mathbf{y}) + O(h^3) \quad (3.51)$$

Comparando esta ecuación con el desarrollo de la solución exacta que aparece en la ecuación (3.34), se obtiene que la constante de error de los términos de segundo orden es:

$$C_{GD} = \left(\frac{1}{2} J^c - J^d|_{h=0} \right) f(\mathbf{y}) \quad (3.52)$$

por lo que la condición $J^c = 2 J^d|_{h=0}$ garantizaría que el método de gradiente discreto obtenido es de segundo orden.

A modo de comparación, se calcula la constante de error correspondiente a la regla de Euler explícita definida por:

$$\mathbf{z} = \mathbf{y} + h f(\mathbf{y}) = \mathbf{y} + h L(\mathbf{y}) \cdot \nabla V(\mathbf{y}) \quad (3.53)$$

la cual puede ser comparada con el desarrollo de la solución exacta dada en (3.34), siendo la constante de error en este caso:

$$C_E = \frac{1}{2} J^c f(\mathbf{y}) \quad (3.54)$$

Más adelante se muestra, en algunos casos particulares, cómo este valor es muchas veces mayor que la constante de error del método de gradiente discreto propuesto para las aplicaciones examinadas, aunque ambos métodos sean de primer orden.

3.5.2. Análisis del orden del método, para sistemas escalares simples

Se ha mostrado en [39] que los métodos de gradiente discreto, además de ser métodos numéricos cuya formulación general es implícita, son métodos de primer orden. En este epígrafe se pretende particularizar los resultados obtenidos en el epígrafe anterior, en cuanto al análisis del orden del método de gradiente discreto obtenido, al caso particular de los ejemplos $P1$, $P2$ y $P3$. También en estos casos se compara con la regla de Euler. Ahora bien, se sabe que:

$$C_{GD} = \left(\frac{1}{2} J^c - J^d \right) f(y)$$

siendo J^c el jacobiano discreto dado en (3.35) y J^d el jacobiano continuo cuya expresión viene dada en (3.47) y $f(y)$ la parte derecha de la EDO en cuestión.

Trabajando con $P1$ se tiene que $f(y) = -ay$, $J^c = -a$, $J^d = -\frac{a}{2}$ con lo que $C_{GD} = 0$. Esto demuestra que el método de gradiente obtenido en este caso es al menos de segundo orden, hecho que se corrobora al demostrar que coincide con la regla trapezoidal. En efecto, para este último método, se tiene que:

$$z = y + \frac{h}{2}[f(z) + f(y)] = y + \frac{h}{2}[-az - ay]$$

de manera que, despejando z , resulta:

$$z = \frac{1 - \frac{ah}{2}}{1 - \frac{ah}{2}} y$$

que es el método de gradiente discreto obtenido en (3.14). Resumiendo, en este caso se obtiene un método de gradiente discreto que, además de ser explícito, es de segundo orden. Por otra parte se sabe que la constante de error asociada al término de segundo orden para el método de Euler se calcula según la ecuación (3.54). Para este ejemplo sería $C_E = \frac{a^2 y}{2}$, con lo que se puede concluir que si se considera el error en valor absoluto, el error producido por el método de Euler siempre será mayor que

el producido por el método de gradiente discreto. Por supuesto, todo el análisis se refiere a h suficientemente pequeño de forma que domine el término $O(h^2)$.

Trabajando con $P2$, se tiene que $f(y) = -ay^3$, $J^c = -3ay^2$, $J^d = -\frac{ay^2}{2}$ con lo que al sustituir en la expresión de C_{GD} se tiene:

$$C_{GD} = a^2y^5$$

Para calcular la constante de error asociada a la regla de Euler hay que tener en cuenta que para este sistema toma la forma:

$$z = y - ay^3$$

Por su parte, de acuerdo a la ecuación (3.54) se tiene:

$$C_E = \frac{3}{2}a^2y^5$$

con lo que se puede concluir que en efecto el método de gradiente discreto es de primer orden. Ahora bien, hay que señalar que cualquiera sea el valor de y , la constante de error asociada a los términos de segundo orden para el método de gradiente discreto va a ser menor que la asociada a la regla de Euler. En otras palabras:

$$C_{GD} < C_E, \quad \text{para todo } y \in \mathbb{R}$$

Este resultado basta para garantizar que la precisión del método de gradiente discreto será mayor que la de la regla de Euler o, en términos de error local, el método de gradiente discreto tiene menor error local.

Haciendo un análisis similar para $P3$ se tiene que:

$$f(y) = ay(1-y), \quad J^c = a(1-2y), \quad J^d = -\frac{ay}{2}$$

con lo que al sustituir en la expresión de C_{GD} se obtiene la siguiente expresión:

$$C_{GD} = \frac{a^2 y}{2} (1 - y)^2$$

En este caso, la regla de Euler viene dada por la expresión:

$$z = y + a y (1 - y)$$

Teniendo en cuenta la expresión de la constante de error asociada a los términos de segundo orden, se tiene:

$$C_E = \frac{a^2 y}{2} (1 - 2y)(1 - y)$$

En este caso también resulta fácil concluir que el método de gradiente discreto es de primer orden, al igual que la regla de Euler. Sin embargo, a diferencia del caso anterior, C_{GD} no siempre va a ser menor que C_E . Resulta relativamente sencillo determinar los valores de y para los cuales $C_{GD} < C_E$, o sea, $C_{GD} - C_E < 0$, ya que sustituyendo las expresiones calculadas anteriormente para C_{GD} y C_E se tiene:

$$(1 - y)^2 - (1 - 2y)(1 - y) < 0 \Rightarrow y(1 - y) < 0$$

Esto se cumple para $y < 0$ ó $y > 1$ por lo que en estos intervalos la precisión del método de gradiente discreto será mayor que la de la regla de Euler. No obstante, hay que señalar que el objetivo fundamental del trabajo es la preservación de las propiedades cualitativas del sistema, por lo que la búsqueda de un error extremadamente pequeño tiene un interés limitado.

3.6. Experimentos numéricos

A continuación se muestran gráficamente los resultados de algunas simulaciones realizadas al integrar numéricamente los problemas $P1$, $P2$ y $P3$. En todos los casos se trabaja con $a = 100$.

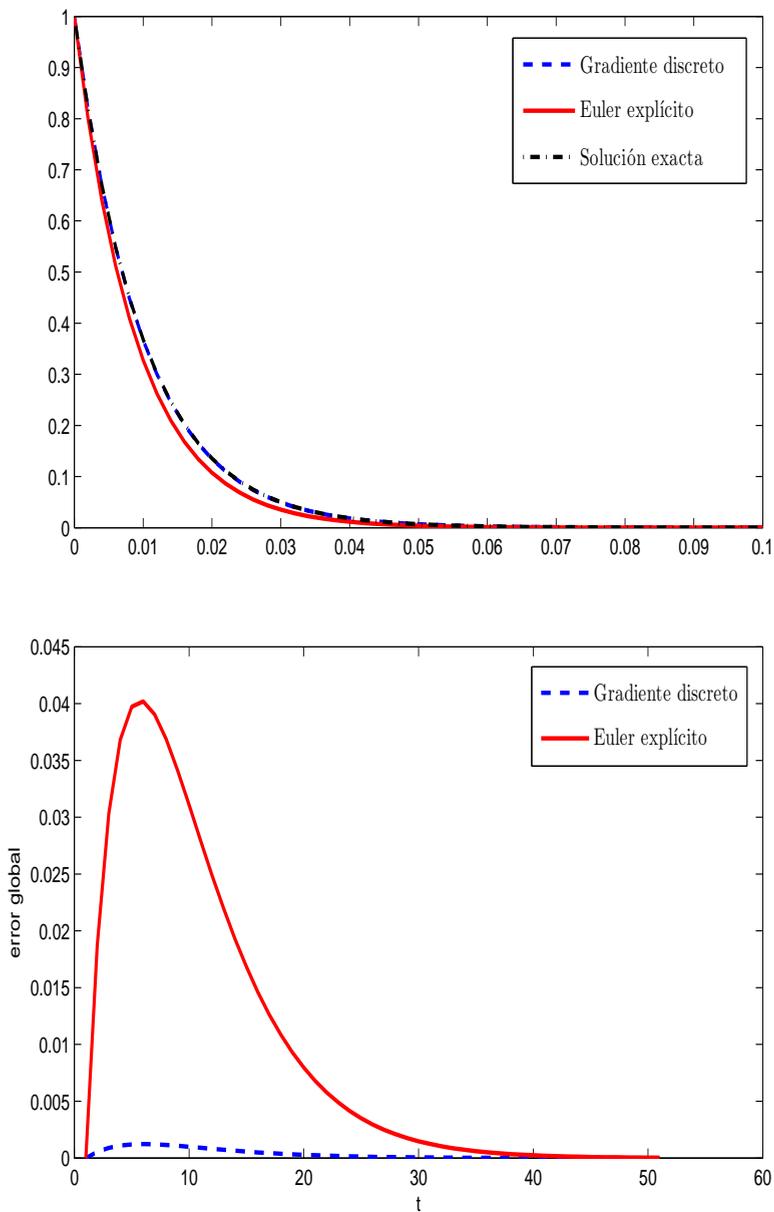


Figura 3.1: Trayectoria de las soluciones (figura superior) y error global (figura inferior), para el problema $P1$, $h = 0,002$.

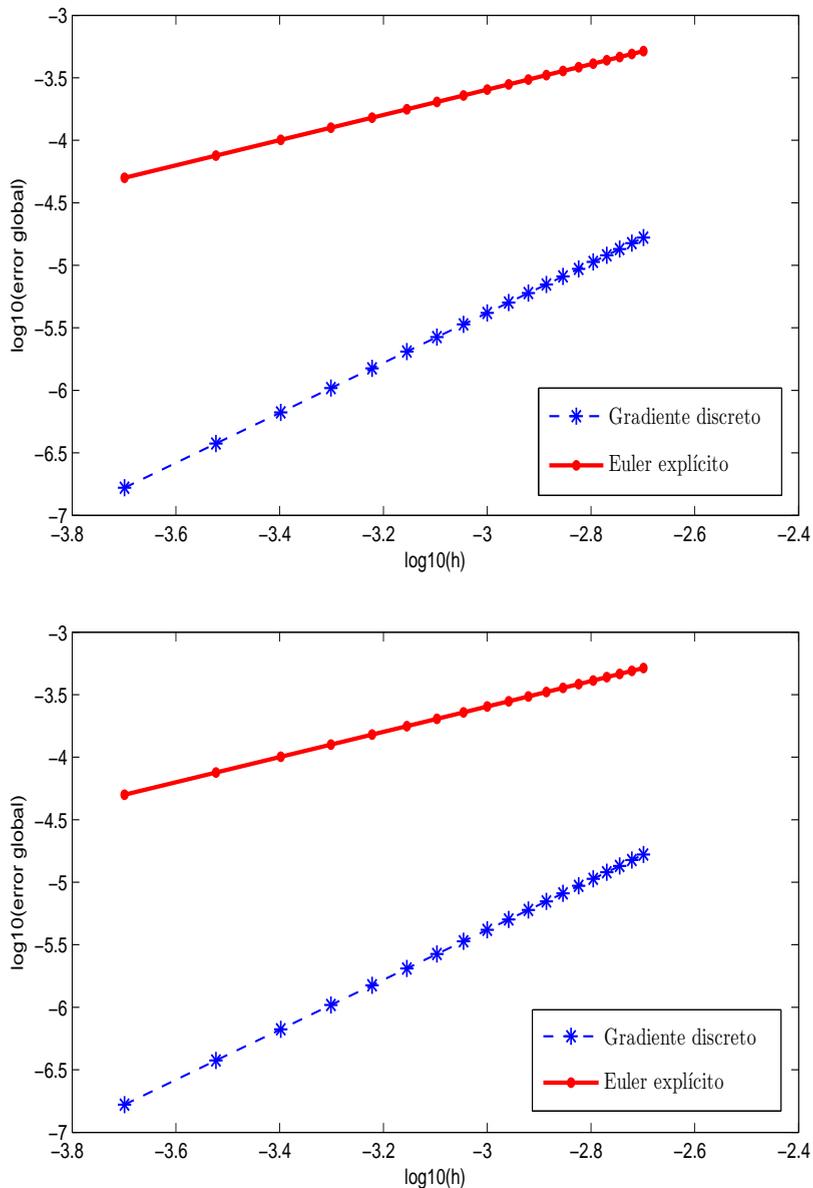


Figura 3.2: Relación entre error global y tamaño de paso en el problema $P1$, $y_0 = 1$ (figura superior), $y_0 = -1$ (figura inferior).

Para el problema $P1$, se trabajó primeramente con $y_0 = 1$, $h = 0,002$, mostrando los resultados de las simulaciones en la figura 3.1. En esta figura se observan dos gráficas, la primera de las cuales muestra las trayectorias de las soluciones obtenidas para ambos métodos, o sea, para el método de gradiente discreto propuesto y para la regla de Euler, así como la solución exacta de este problema que fue calculada analíticamente en el epígrafe 3.4. Por su parte, la segunda gráfica representa el error *global* a lo largo de la trayectoria, que se obtuvo restando la trayectoria obtenida mediante el método definido en la ecuación (3.14) luego de k pasos, menos la solución exacta dada por la ecuación (3.10) en $t = kh$. Se observa que tanto el método de gradiente como la regla de Euler brindan una solución aproximada cercana a la solución exacta, aunque se puede apreciar que la solución obtenida por el método de gradiente discreto es prácticamente la misma que la exacta, mientras que la de la regla de Euler difiere un poco de esta última. Se observa además que el error global del método de Euler es mucho mayor que el del método de gradiente discreto. Por su parte, la figura 3.2 muestra los resultados de un segundo experimento. Esta vez se trabajó con valores del tamaño de paso entre $h = 2 \cdot 10^{-4}$ y $h = 2 \cdot 10^{-3}$, con el objetivo de mostrar cómo variaba el error global de los métodos numéricos respecto del tamaño de paso h . La primera gráfica corresponde al punto inicial $y_0 = 1$ y la segunda a $y_0 = -1$, observándose en ambos casos que la dependencia del error global respecto de h para la regla de Euler es lineal, como se esperaba ya que es un método numérico de primer orden. No puede dejar de señalarse que la escala de ambos ejes en la figura 3.2 es logarítmica. En cambio, para el método de gradiente discreto, a pesar de que la dependencia apreciada también es lineal, la recta que muestra esta dependencia tiene pendiente dos, hecho que valida el resultado obtenido en la sección anterior cuando se demostró que en este caso el método de gradiente discreto es de segundo orden. Asimismo, se muestra que en ambos casos el error para el método de Euler es superior. Por otra parte, se observa que si se aumenta el tamaño de paso, por ejemplo, considerando $h = 0,02$ el método de Euler produce resultados completamente incorrectos. Uno de los comportamientos observados con mayor frecuencia, si bien las figuras correspondientes no resultaban particularmente ilustrativas y se han omitido, es que las soluciones producidas por el método de Euler entran en ciclos y

comienzan a oscilar.

Para el problema $P2$, en primer lugar, se definió $y_0 = 1$, $h = 0,003$, dando lugar a los resultados que se muestran en la figura 3.3. Aquí, la primera gráfica muestra las trayectorias de las soluciones obtenidas tanto para el método de gradiente discreto propuesto, como para la regla de Euler explícita. Esta figura también muestra la trayectoria de la solución exacta calculada analíticamente con anterioridad. Por su parte, la segunda gráfica representa el *error global* a lo largo de la trayectoria, obtenido de manera similar al caso anterior, pero esta vez restando la trayectoria obtenida por el método dado en la ecuación (3.20) luego de k pasos y la solución exacta dada por la ecuación (3.16) en $t = kh$. Se observa que tanto el método de gradiente como la regla de Euler brindan una solución aproximada cercana a la solución exacta, pero el error global del método de Euler es superior al del método de gradiente discreto. En un segundo experimento, mostrado en la figura 3.4, se trabajó con h en el intervalo $h \in (10^{-4}, 10^{-3})$, con el objetivo de mostrar la influencia del tamaño de paso en el error global de los métodos numéricos. La primera gráfica corresponde al punto inicial $y_0 = 1$ y la segunda a $y_0 = -1$. Obsérvese que en ambos casos la dependencia del error global respecto de h es lineal, como se esperaba, ya que al igual que en el ejemplo anterior ambos métodos son de primer orden. Asimismo, se muestra que en ambos casos el error para el método de Euler es superior. Ahora bien, si se aumenta el tamaño de paso, por ejemplo, considerando $h = 0,03$ el método de Euler produce resultados completamente incorrectos y el error crece de forma no acotada hasta sobrepasar la capacidad del ordenador.

Un conjunto similar de experimentos se realizó para el problema $P3$ definido por la ecuación logística. En la figura 3.5 se muestran las trayectorias de las soluciones y los errores para $y_0 = 1,5$ y $h = 0,004$. Para el análisis de la influencia del tamaño de paso sobre el error global, mostrado en la figura 3.6, se utilizó primero $y_0 = 1,5$ y luego $y_0 = 0,5$ para mostrar distintos comportamientos cualitativos, teniendo en cuenta que en este caso el punto de equilibrio estable es $y = 1$. En ambos casos se aprecia la capacidad del método de gradiente discreto, no solo para brindar una buena aproximación de la solución del sistema, sino también para conservar la función de Lyapunov del sistema continuo.

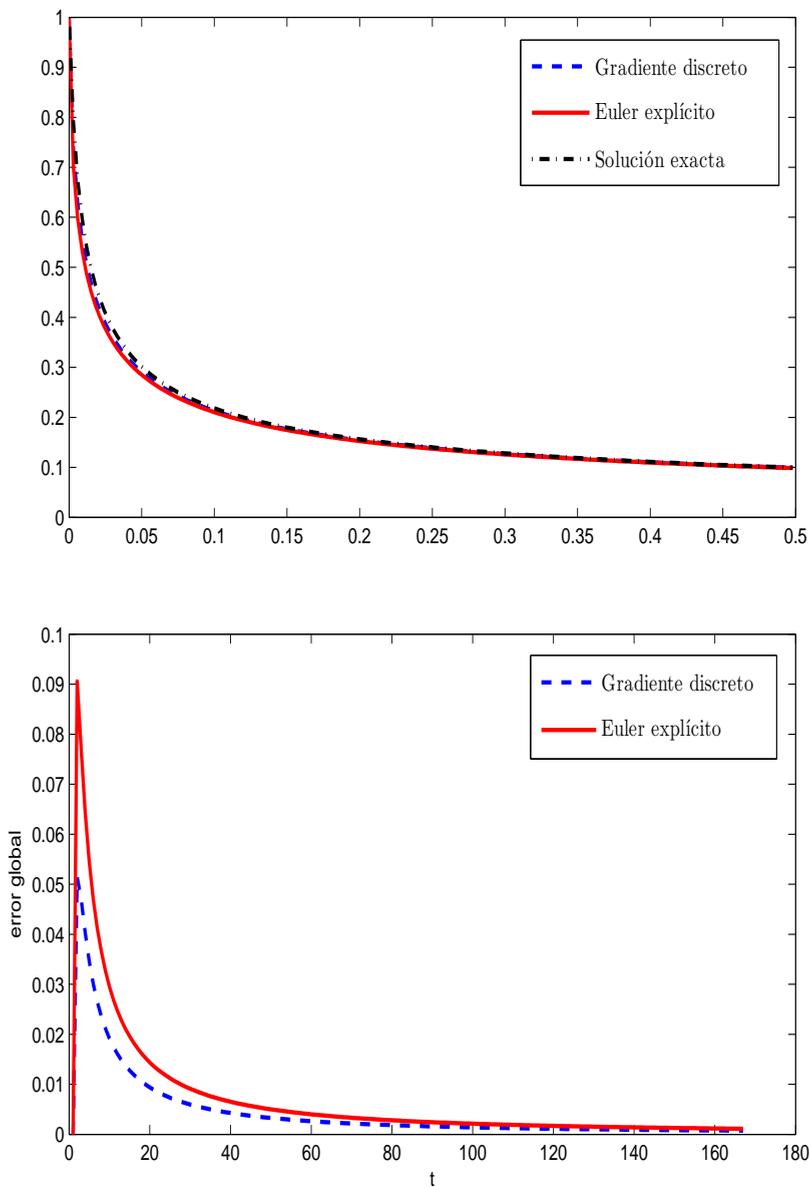


Figura 3.3: Trayectoria de las soluciones (figura superior) y error global (figura inferior), para el problema $P2$, $h = 0,003$.

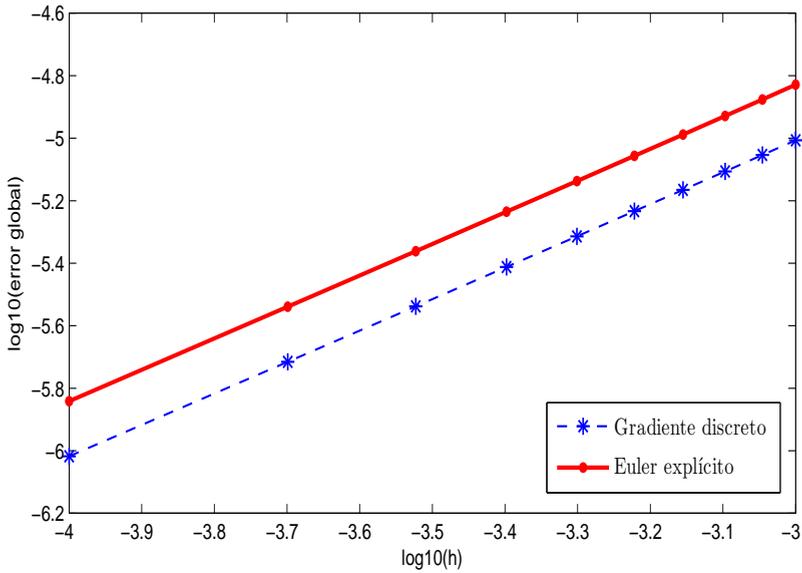
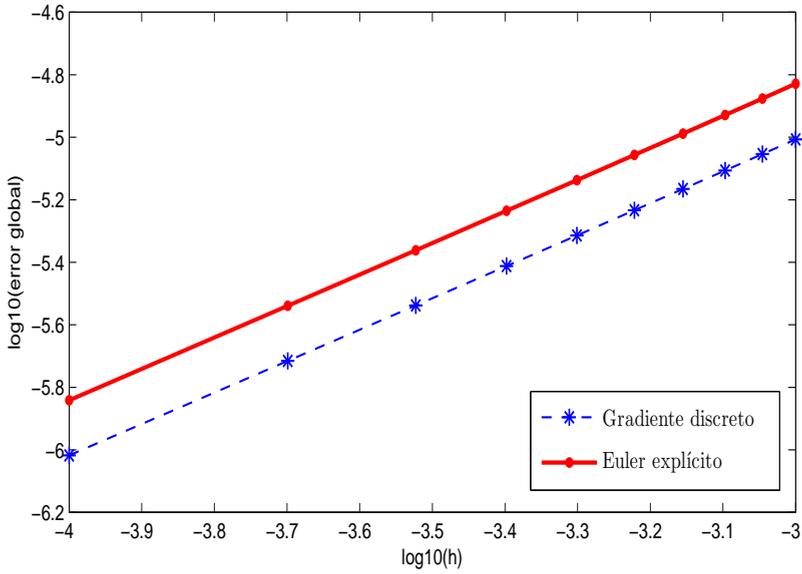


Figura 3.4: Relación entre error global y tamaño de paso en el problema $P2$, $y_0 = 1$ (figura superior), $y_0 = -1$ (figura inferior).

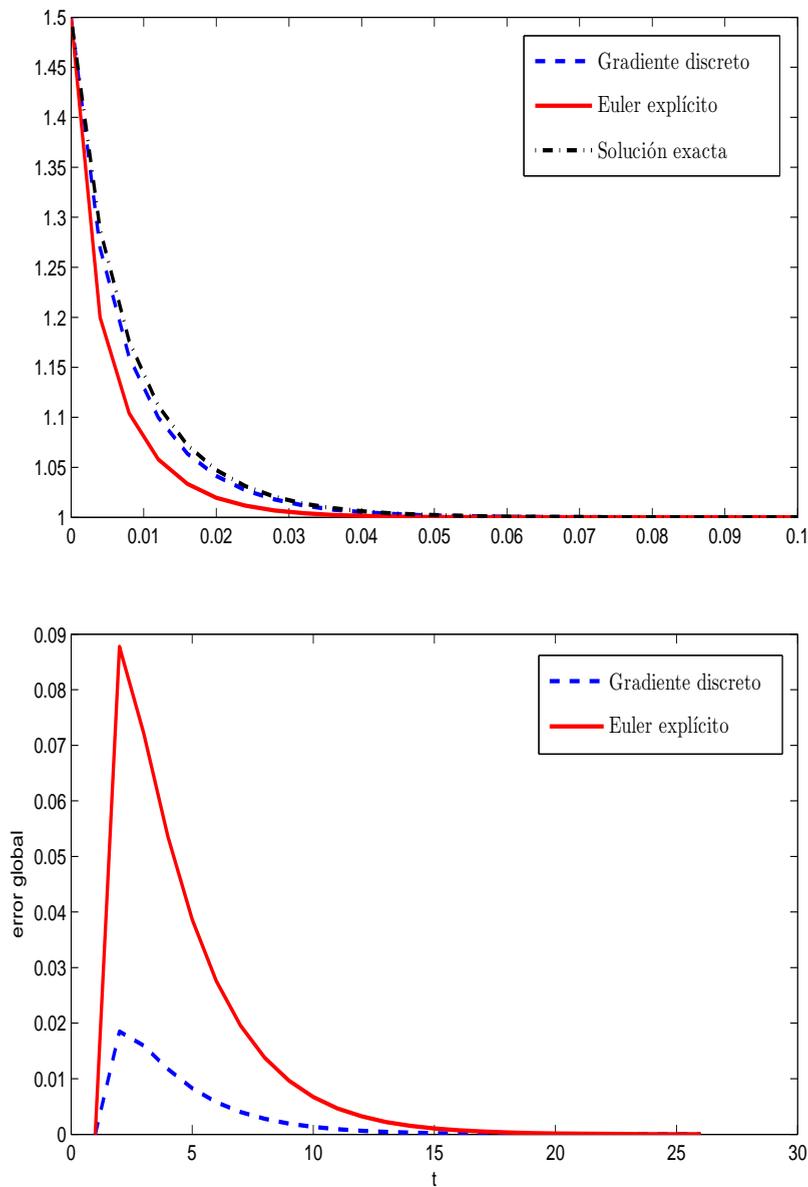


Figura 3.5: Trayectoria de las soluciones (figura superior) y error global (figura inferior), para el problema $P3$, $h = 0,004$.

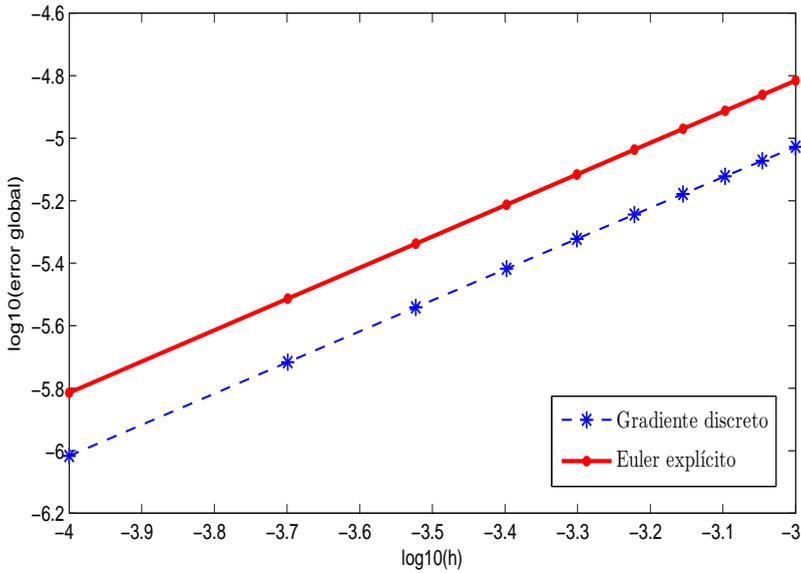
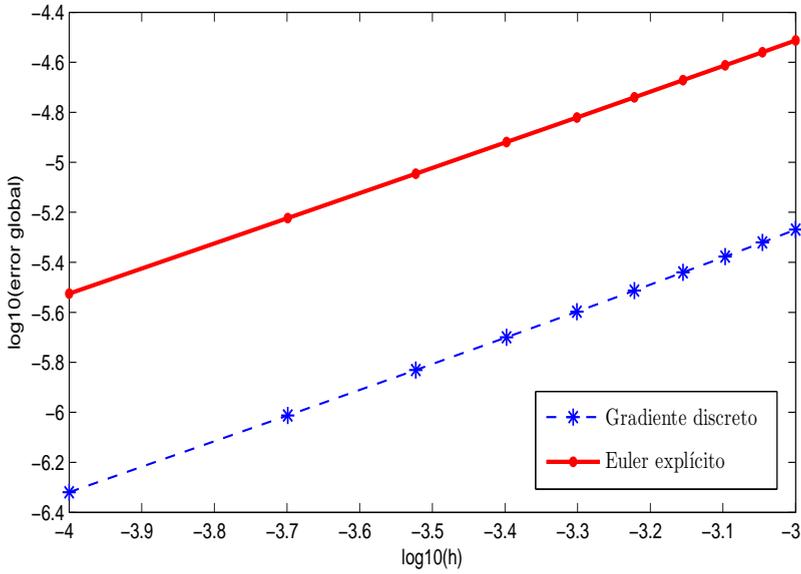


Figura 3.6: Relación entre error global y tamaño de paso, para el método de gradiente discreto y la regla de Euler, ambos aplicados a la ecuación diferencial logística. $y_0 = 1,5$ (figura superior), $y_0 = 0,5$ (figura inferior).

A modo de conclusión, se puede afirmar que los métodos de gradiente discreto, a pesar de ser implícitos, se pueden reescribir de manera explícita si se consideran condiciones especiales de los parámetros. También se establece que, a pesar de ser en principio métodos de primer orden, existen condiciones que garantizan que puedan obtenerse métodos de gradiente discreto de orden superior. Incluso en los casos en los que se obtengan métodos de gradiente discreto explícitos de primer orden, los métodos de gradiente discreto suelen aproximar mejor la solución del sistema, pero sobre todo respetan el comportamiento cualitativo del sistema original, ya que conservan la función de Lyapunov del sistema continuo cualquiera sea el tamaño de paso con el que se trabaje. En próximos capítulos se generalizan estos resultados trabajando con sistemas de alta dimensión.

3.7. Conclusiones

En el inicio del capítulo se mostraron algunos preliminares de la teoría de los gradientes discretos y el diseño de métodos de integración basados en estos. Se centra la atención en los métodos de gradiente discreto diseñados específicamente para la conservación de la función de Lyapunov del sistema continuo, dando lugar posteriormente a la exposición de las principales aportaciones del capítulo. Se describe la metodología de construcción de métodos de gradiente discreto, haciendo énfasis en la elección de los parámetros. Se ilustra el funcionamiento de esta metodología, diseñando métodos de gradiente discreto para tres sistemas escalares simples. A continuación, se centra la atención en el gradiente discreto del incremento de las coordenadas y se formalizan los resultados obtenidos, trabajando con una familia de sistemas de EDO, cuya función de Lyapunov es multilínea. De esta manera, se obtiene el resultado más importante del capítulo: *con una elección adecuada de los parámetros, el método de gradiente discreto obtenido puede reescribirse de forma explícita, para el caso de sistemas dinámicos con función de Lyapunov multilínea*. Finalmente se realiza un estudio del orden de los métodos obtenidos en el caso general y en los casos particulares de los sistemas escalares simples. Los resultados se ilustran mediante una serie de experimentos numéricos, que corroboraron la capacidad de los métodos

obtenidos para conservar la función de Lyapunov del sistema continuo, a la vez que reducen, en la mayoría de los casos, el error global, en comparación con la regla de Euler.

Capítulo 4

Discretización de las redes de Hopfield con gradiente discreto

Resumen del capítulo

En este capítulo se exponen algunos de los resultados más importantes de la tesis. Se propone el uso de un método de gradiente discreto para la integración de sistemas de ecuaciones diferenciales ordinarias de manera tal que se preserven la función de Lyapunov del sistema al tiempo que las demás propiedades favorables del sistema continuo. Como campo de aplicación se utilizan las redes de Hopfield con y sin autopesos. Se muestra cómo en ambos casos es posible reescribir el método de gradiente discreto explícitamente teniendo en cuenta la forma de la función de Lyapunov del sistema y la elección de los parámetros del método. Posteriormente se hace el análisis del orden del método propuesto y se propone su extensión de alto orden. Específicamente se propone la composición del método de gradiente discreto con su adjunto dando lugar de esta manera a un método de segundo orden que además de conservar la función de Lyapunov del sistema continuo puede escribirse explícitamente al igual que en el caso anterior.

4.1. Las redes de Hopfield

En este capítulo se propone el uso de un método de integración diseñado de manera tal que, al discretizar un sistema de EDO con función de Lyapunov, ésta se conserve conjuntamente con el resto de propiedades favorables que posea el sistema. El capítulo sigue la siguiente estructura: en esta sección se introducen las redes de Hopfield como campo de aplicación; en la sección 4.2 se propone el diseño de un método de gradiente discreto para discretizar las redes de Hopfield de dimensión dos, con el objetivo de conservar la función de Lyapunov del sistema cualquiera sea el tamaño de paso con el que se trabaje; a continuación, en la sección 4.3 se generaliza el resultado de la sección anterior trabajando con el modelo de Hopfield de dimensión n ; en la sección 4.4 se hace el análisis del orden del método de gradiente discreto propuesto; ya en la sección 4.5, se considera la presencia de autopesos en el modelo, discretizando también en este caso el sistema utilizando un método de gradiente discreto; finalmente, en la sección 4.6 se propone una generalización del método propuesto a alto orden, componiendo el método de gradiente discreto obtenido inicialmente con su adjunto.

Como se dijo anteriormente, se utilizarán como campo de aplicación las redes de Hopfield con y sin autopesos. Este sistema es un modelo recurrente inspirado en las redes neuronales biológicas [22]. Desde el punto de vista dinámico, las redes de Hopfield son un sistema estable, ya que poseen función de Lyapunov. Precisamente esta peculiaridad de ser sistemas con función de Lyapunov es la que ha sido utilizada para resolver problemas de optimización combinatoria [26, 53] y estimación de parámetros [7, 56]. Por tanto, se desea que al discretizar el sistema, la red discreta también posea función de Lyapunov, en otras palabras, que se conserve la función de Lyapunov del sistema continuo bajo discretización. Trabajos previos [4] muestran que la discretización usual utilizada en el marco de la redes neuronales no conserva la función de Lyapunov del sistema, ya que aparecen soluciones periódicas. Resulta natural, por tanto, la búsqueda de un método de integración alternativo que conserve la función de Lyapunov a la vez que integre el sistema. Por otra parte, las redes de Hopfield también pueden considerarse como una generalización del modelo de Ising

de la mecánica estadística [34], y su simulación por ordenador permitiría investigar propiedades de tales modelos, siempre que la discretización preserve sus características dinámicas. Por último, desde un punto de vista estrictamente matemático, la generalización a alto orden de las redes de Hopfield [43, 30] permite definir sistemas de EDO de dimensión arbitraria, que constituyen un interesante banco de pruebas de métodos numéricos. En estos sistemas, la conservación de las propiedades cualitativas cobra más importancia que la minimización del error local en la obtención de aproximaciones de las soluciones. Por todas estas razones, resultan las redes de Hopfield un modelo atractivo para mostrar el funcionamiento de los métodos de gradiente discreto. Como se mencionó en el capítulo 3, éstos métodos están diseñados específicamente para la conservación de la función de Lyapunov del sistema continuo, cualquiera sea el tamaño de paso con el que se trabaje, a la vez que brindan una buena aproximación de las soluciones.

Una red neuronal de Hopfield se puede describir de manera general como un sistema dinámico continuo, compuesto por n elementos de proceso o neuronas. En cada instante de tiempo la variable y_i , $i = 1, \dots, n$ refleja la salida de la neurona i -ésima. Por otra parte, desde el punto de vista dinámico, las redes de Hopfield son sistemas estables, pues poseen función de Lyapunov, que usualmente se denota con la letra V . A esta función también se le conoce como función de energía del sistema. La consideración de la función de Lyapunov como energía permite ver, desde el punto de vista físico, este modelo como un sistema disipativo, ya que como se vio en el capítulo 1, una de las principales características de una función de Lyapunov es que decrece a lo largo de las trayectorias. En la literatura aparecen varias formulaciones para las redes de Hopfield, pero en este trabajo se centra la atención en la conocida como formulación de Abe [1, 55, 31].

La formulación de Abe de las redes de Hopfield está definida por el siguiente

sistema de ecuaciones n -dimensional:

$$\begin{aligned} \frac{du_i}{dt} &= \sum_{j=1}^n w_{ij} y_j - b_i \\ y_i &= \tanh\left(\frac{u_i}{\beta}\right) \end{aligned} \quad (4.1)$$

donde u_i puede interpretarse biológicamente como el potencial interno de la neurona i , $\beta > 0$ es un parámetro, cuya función es regular la pendiente de la tangente hiperbólica y n es el número de neuronas. Por su parte, los elementos w_{ij} con $j = 1, \dots, n$ representan el peso de la conexión de la neurona j con la i y los b_i con $i = 1, \dots, n$ son los valores del sesgo de la neurona i . Tanto los valores de w_{ij} como los de b_i son constantes. A partir de este momento, se asume $\beta = 1$ sin pérdida de generalidad. Esta suposición solo lleva a un cambio en la escala en el tiempo. Obsérvese que la ecuación (4.1) no es exactamente un sistema de EDO debido a las variables internas u_i . Estrictamente hablando, es una ecuación Algebraico-Diferencial [9] donde la parte diferencial es lineal.

A partir de ahora se asumirá que la matriz de autopesos es simétrica y con diagonal nula, o sea $w_{ij} = w_{ji}$ y $w_{ii} = 0$. Se demuestra [57] que cuando se utilizan las redes de Hopfield como método de optimización, esta suposición no conduce a pérdida de generalidad.

Desde el punto de vista dinámico, se puede demostrar que las redes de Hopfield dadas en la ecuación (4.1) poseen puntos de equilibrio estables. Este hecho se deriva de la existencia de una función de Lyapunov ya que, la existencia de tal función, al menos en una vecindad de un punto de equilibrio, caracteriza su estabilidad [32]. Ahora bien, se sabe que la función de Lyapunov del sistema viene dada por:

$$V(\mathbf{y}) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} y_i y_j + \sum_{i=1}^n b_i y_i = -\frac{1}{2} \mathbf{y}^T \mathbf{W} \mathbf{y} + \mathbf{b}^T \mathbf{y} \quad (4.2)$$

de manera que resulta fácil establecer que este modelo es un sistema de gradiente, ya

que:

$$\frac{du_i}{dt} = \sum_{j=1}^n w_{ij} y_j - b_i = -\frac{\partial V}{\partial y_i}$$

Esta igualdad es un punto clave a la hora de demostrar que V es, en efecto, función de Lyapunov del sistema, pues su derivada temporal resulta entonces:

$$\frac{dV}{dt} = \sum_{i=1}^n \frac{\partial V}{\partial y_i} \frac{dy_i}{dt} = -\sum_{i=1}^n \left(\frac{\partial V}{\partial y_i} \right)^2 (1 - \tanh^2(u_i)) = -\sum_{i=1}^n \left(\frac{\partial V}{\partial y_i} \right)^2 (1 - y_i^2) \leq 0 \quad (4.3)$$

La última desigualdad resulta al observar que las variables de estado permanecen acotadas por $|y_i| < 1$, ya que se obtienen como resultado de la función tangente hiperbólica. Consecuentemente, todo punto fijo es también un mínimo local de V y por ende un punto de equilibrio estable.

Con el objetivo de exponer resultados generales, tanto para el sistema dinámico que define la red continua como para su discretización, se reformula la ecuación (4.1) en términos de las variables y_i . De esta manera se eliminan las variables u_i que se habían considerado variables internas del sistema. Se tiene entonces que:

$$\begin{aligned} \frac{dy_i}{dt} &= \frac{dy_i}{du_i} \frac{du_i}{dt} = (1 - \tanh^2(u_i)) \left(\sum_{j=1}^n w_{ij} y_j - b_i \right) \\ &= (1 - y_i^2) \left(\sum_{j=1}^n w_{ij} y_j - b_i \right) = -(1 - y_i^2) \left(\frac{\partial V}{\partial y_i} \right) \end{aligned}$$

resultando finalmente la que se considera como definición de sistema de Hopfield:

$$\frac{dy_i}{dt} = -(1 - y_i^2) \left(\frac{\partial V}{\partial y_i} \right) \quad (4.4)$$

Este sistema puede también escribirse en forma matricial:

$$\frac{d\mathbf{y}}{dt} = L(\mathbf{y}) \nabla V(\mathbf{y}) \quad (4.5)$$

siendo:

$$L(\mathbf{y}) = \begin{pmatrix} 1 - y_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 - y_n^2 \end{pmatrix}$$

mientras que $\nabla V(\mathbf{y})$ es el gradiente de la función de Lyapunov definida en la ecuación (4.2). Nótese que la ecuación (4.5) coincide con la forma de gradiente lineal del sistema que se vio en el capítulo 3.

Una propiedad singular de este sistema es que las trayectorias permanecen dentro del hipercubo $(-1, 1)^n$, siempre que el valor inicial también se encuentre dentro de este. En otras palabras, el hipercubo es invariante, lo cual es obvio de acuerdo a la ecuación (4.1), ya que la tangente hiperbólica toma valores entre $(-1, 1)$. Sin embargo, si un método numérico estándar fuese aplicado a la ecuación (4.4), la discretización podría destruir esta propiedad.

En cuanto a la dinámica del sistema, debido a la existencia de la función de Lyapunov, el único posible comportamiento a largo plazo es que las trayectorias se acerquen asintóticamente a un punto que pertenezca al conjunto de puntos fijos del sistema. En particular, no existirían soluciones periódicas. De la ecuación (4.4), se puede concluir que los puntos de equilibrio del sistema, ver capítulo 1, pueden clasificarse en: vértices del hipercubo, o sea, los \mathbf{y} tal que $|y_i| = 1$; y puntos de equilibrio interiores, que son aquellos en los que $\sum_{j=1}^n w_{ij} y_j - b_i = 0$. Nótese que, al asumir la ausencia de autopesos, como en el caso que nos ocupa, la función de Lyapunov V es lineal en cada una de sus variables, ya que $\frac{\partial V}{\partial y_i} = - \sum_{j=1}^n w_{ij} y_j - b_i$ no depende de y_i . A la hora de caracterizar la estabilidad del sistema, se distinguen dos casos fundamentales, según se considere la ausencia de autopesos o no. Por una parte, cuando los autopesos se anulan, puede demostrarse [5] que los únicos puntos fijos estables de la red de Hopfield continua, dada por la ecuación (4.1), son los vértices del hipercubo. Por otra parte, el análisis del caso en el que los autopesos no se anulan resulta mucho más complicado. Resumiendo, tanto la existencia de la función de Lyapunov como la invarianza del hipercubo son características críticas para el comportamiento

favorable de las redes de Hopfield. Es por esto que el objetivo fundamental de este trabajo es diseñar o proponer métodos numéricos que, por construcción, preserven la función de Lyapunov que presenta la red continua, independientemente del tamaño de paso con que se trabaje.

4.2. Aplicación del gradiente discreto a las redes de Hopfield

En esta sección se procede al diseño de un método de gradiente discreto específicamente construido para conservar la función de Lyapunov del sistema que define las redes de Hopfield. Existen algunos trabajos que han abordado ya el problema desde este punto de vista, entre los que se puede citar la tesis doctoral del profesor Miguel Atencia Ruiz [3]. En estos trabajos, a partir de resultados previos sobre el sistema continuo, se formula la red discreta como un método numérico no estándar que resuelve la ecuación continua, dependiendo del tamaño de paso de discretización. Desde el punto de vista numérico, dicha red discreta es una aproximación de primer orden de la continua. Asimismo, el análisis dinámico muestra que se preservan los puntos fijos del sistema, así como su estabilidad o inestabilidad. Sin embargo, en general, este método no estándar no garantiza la conservación de la función de Lyapunov, lo que se demuestra mediante la construcción de soluciones periódicas. En resumen, la destrucción de la estructura de gradiente, dificulta la capacidad del sistema como algoritmo de optimización.

A la vista de los resultados anteriores, surge la necesidad de diseñar un método numérico que permita discretizar la red continua a la vez que conserve las propiedades cualitativas de ésta, específicamente su función de Lyapunov. Para esto se centra la atención en los métodos de gradiente discreto y se sigue el esquema descrito para el diseño de éstos en el capítulo anterior, o sea:

1. Reescribir el sistema (4.1) en la forma de gradiente lineal $\frac{dy}{dt} = L(y) \nabla V(y)$ definiendo explícitamente la expresión de la matriz $L(y)$ y del gradiente de la función $V(y)$.

2. Discretizar el sistema usando el método:

$$\frac{z - \mathbf{y}}{h} = \tilde{L}(\mathbf{y}, z, h) \bar{\nabla} V(\mathbf{y}, z) \quad (4.6)$$

siendo $\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y})$, $\bar{\nabla} V(\mathbf{y}, \mathbf{y}) = \nabla V(\mathbf{y})$.

3. Sustituir las expresiones de $\tilde{L}(\mathbf{y}, z, h)$ y $\bar{\nabla} V(\mathbf{y}, z)$ en (4.6) para obtener la expresión del método numérico resultante.

Con vistas a ilustrar la metodología de obtención de un método de gradiente discreto, se escoge como ejemplo ilustrativo el sistema de Hopfield en dimensión dos asumiendo la ausencia de autopesos. Como se dijo anteriormente, se escoge el gradiente discreto del incremento de las coordenadas $\bar{\nabla} V$.

Teniendo en cuenta la formulación general de las redes de Hopfield dada en la ecuación (4.1), en dimensión dos el sistema adquiere la siguiente forma:

$$\begin{aligned} \frac{du_1}{dt} &= w_{12} y_2 - b_1, & y_1 &= \tanh u_1 \\ \frac{du_2}{dt} &= w_{21} y_1 - b_2, & y_2 &= \tanh u_2 \end{aligned} \quad (4.7)$$

El primer paso para la obtención de un método de gradiente discreto para este sistema, sería reescribirlo en la forma de gradiente lineal.

4.2.1. Formulación del sistema en forma de gradiente lineal

Dado que se trabaja con $\mathbf{y} \in \mathbb{R}^2$, se usará la notación¹ $\mathbf{y} = (y_1, y_2)$. Ahora bien, teniendo en cuenta la expresión de V dada en la ecuación (4.2) para el sistema (4.7)

¹ Cuando no haya confusión se mantiene la notación habitual de usar subíndices para las distintas componentes de un vector.

y asumiendo $w_{ij} = w_{ji}$ para $i \neq j$ y $w_{ii} = 0$, entonces:

$$\begin{aligned} V &= -\frac{1}{2} (w_{12} y_1 y_2) + b_1 y_1 - \frac{1}{2} (w_{21} y_2 y_1) + b_2 y_2 \\ &= -\frac{1}{2} (2 w_{12} y_1 y_2) + b_1 y_1 + b_2 y_2 \\ &= -w_{12} y_1 y_2 + b_1 y_1 + b_2 y_2 \end{aligned}$$

con lo que:

$$\nabla V = \left(\frac{\partial V}{\partial y_1}, \frac{\partial V}{\partial y_2} \right) = (-w_{12} y_2 + b_1, -w_{21} y_1 + b_2)$$

Como se vio en la sección anterior, se puede reescribir el sistema eliminando las variables internas u_i , quedando:

$$\begin{aligned} \frac{dy_1}{dt} &= (1 - y_1^2) \left(-\frac{\partial V}{\partial y_1} \right) = (1 - y_1^2) (w_{12} y_2 - b_1) \\ \frac{dy_2}{dt} &= (1 - y_2^2) \left(-\frac{\partial V}{\partial y_2} \right) = (1 - y_2^2) (w_{21} y_1 - b_2) \end{aligned}$$

Finalmente, en forma matricial, puede escribirse:

$$\begin{aligned} \begin{pmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \end{pmatrix} &= \begin{pmatrix} 1 - y_1^2 & 0 \\ 0 & 1 - y_2^2 \end{pmatrix} \begin{pmatrix} w_{12} y_2 - b_1 \\ w_{21} y_1 - b_2 \end{pmatrix} \\ \frac{d\mathbf{y}}{dt} &= L(\mathbf{y}) \nabla V(\mathbf{y}) \end{aligned}$$

donde:

$$L(\mathbf{y}) = -\begin{pmatrix} 1 - y_1^2 & 0 \\ 0 & 1 - y_2^2 \end{pmatrix}, \quad \nabla V(\mathbf{y}) = \begin{pmatrix} -w_{12} y_2 + b_1 \\ -w_{21} y_1 + b_2 \end{pmatrix}$$

Teniendo en cuenta que $L(\mathbf{y})$ es simétrica y definida negativa, se cumple:

$$\frac{dV}{dt} = (\nabla V)^T L(\mathbf{y}) \nabla V \leq 0$$

ya que:

$$\begin{aligned} (\nabla V)^T L(\mathbf{y}) \nabla V &= -(-w_{12}y_2 + b_1, -w_{21}y_1 + b_2) \begin{pmatrix} 1 - y_1^2 & 0 \\ 0 & 1 - y_2^2 \end{pmatrix} \begin{pmatrix} -w_{12}y_2 + b_1 \\ -w_{21}y_1 + b_2 \end{pmatrix} \\ &= -\left[(-w_{12}y_2 + b_1)^2(1 - y_1^2) + (-w_{21}y_1 + b_2)^2(1 - y_2^2)\right] \leq 0 \end{aligned}$$

Esto implica que efectivamente $V(\mathbf{y})$ es una función de Lyapunov para el sistema (4.7). Luego de reescribir el sistema en la forma de gradiente lineal, corresponde elegir los parámetros del método de gradiente discreto.

4.2.2. Definición de los parámetros del método

Una vez reescrito el sistema (4.7) en la forma de gradiente lineal, se seleccionan los parámetros del método de gradiente discreto mediante el que se discretizará la red continua. Los parámetros se escogen de manera tal que se cumpla $\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y})$, $\bar{\nabla}V(\mathbf{y}, \mathbf{y}) = \nabla V(\mathbf{y})$. En este caso se elige:

$$\tilde{L}(\mathbf{y}, \mathbf{z}, h) = - \begin{pmatrix} 1 - y_1z_1 & 0 \\ 0 & 1 - y_2z_2 \end{pmatrix}$$

Con esta elección de \tilde{L} es fácil comprobar que se cumple la condición de consistencia, o sea, $\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y})$. Por otra parte, la expresión para el gradiente discreto del incremento de las coordenadas sería:

$$\bar{\nabla}V(\mathbf{y}, \mathbf{z}) = \begin{pmatrix} \frac{V(z_1, y_2) - V(y_1, y_2)}{z_1 - y_1} \\ \frac{V(z_1, z_2) - V(z_1, y_2)}{z_2 - y_2} \end{pmatrix}$$

y teniendo en cuenta la definición de V resulta:

$$\begin{aligned}\bar{\nabla}V(\mathbf{y}, \mathbf{z}) &= \begin{pmatrix} \frac{(-w_{12} z_1 y_2 + b_1 z_1 + b_2 y_2) - (-w_{12} y_1 y_2 + b_1 y_1 + b_2 y_2)}{z_1 - y_1} \\ \frac{(-w_{12} z_1 z_2 + b_1 z_1 + b_2 z_2) - (-w_{12} z_1 y_2 + b_1 z_1 + b_2 y_2)}{z_2 - y_2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{-w_{12} y_2 (z_1 - y_1) + b_1 (z_1 - y_1)}{z_1 - y_1} \\ \frac{-w_{12} z_1 (z_2 - y_2) + b_2 (z_2 - y_2)}{z_2 - y_2} \end{pmatrix} = \begin{pmatrix} -w_{12} y_2 + b_1 \\ -w_{12} z_1 + b_2 \end{pmatrix}\end{aligned}$$

Finalmente se tiene que:

$$\bar{\nabla}V(\mathbf{y}, \mathbf{z}) = \begin{pmatrix} -w_{12} y_2 + b_1 \\ -w_{12} z_1 + b_2 \end{pmatrix} \quad (4.8)$$

con lo que $\bar{\nabla}V(\mathbf{y}, \mathbf{y}) = \begin{pmatrix} -w_{12} y_2 + b_1 \\ -w_{12} y_1 + b_2 \end{pmatrix} = \nabla V(\mathbf{y})$ como se requería.

Una vez elegidos los parámetros del método y luego de comprobar que cumplen con la definición de gradiente discreto y la condición de consistencia, solo queda sustituir en la ecuación general de los métodos de gradiente discreto y obtener la fórmula de discretización.

4.2.3. Discretización mediante gradiente discreto

La siguiente fase del proceso de construcción del método, tras escoger sus parámetros, es propiamente la discretización del sistema continuo dado por la ecuación (4.7), según:

$$\mathbf{z} = \mathbf{y} + h \tilde{L}(\mathbf{y}, \mathbf{z}, h) \bar{\nabla}V(\mathbf{y}, \mathbf{z}) \quad (4.9)$$

o sea:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - h \begin{pmatrix} 1 - y_1 z_1 & 0 \\ 0 & 1 - y_2 z_2 \end{pmatrix} \begin{pmatrix} -w_{12} y_2 + b_1 \\ -w_{12} z_1 + b_2 \end{pmatrix}$$

de manera que queda:

$$z_1 = y_1 + h [(-1 + y_1 z_1) (-w_{12} y_2 + b_1)]$$

$$z_2 = y_2 + h [(-1 + y_2 z_2) (-w_{12} z_1 + b_2)]$$

En este caso es posible despejar z_1 y z_2 , con lo que se puede expresar el método en forma explícita, o sea:

$$\begin{aligned} z_1 &= y_1 + h [(w_{12} y_2 - b_1) - y_1 z_1 (w_{12} y_2 - b_1)] \\ z_1 [1 + h y_1 (w_{12} y_2 - b_1)] &= y_1 + h (w_{12} y_2 - b_1) \\ z_1 &= \frac{y_1 + h (w_{12} y_2 - b_1)}{1 + h y_1 (w_{12} y_2 - b_1)} \end{aligned}$$

De manera similar se despeja z_2 y se obtiene:

$$z_2 = \frac{y_2 + h (w_{12} z_1 - b_2)}{1 - h y_2 (w_{12} z_1 - b_2)}$$

Resumiendo, se obtiene un método de gradiente discreto explícito para las redes de Hopfield de dimensión dos, cuya formulación viene dada por:

$$z_1 = \frac{y_1 + h (w_{12} y_2 - b_1)}{1 + h y_1 (w_{12} y_2 - b_1)}, \quad z_2 = \frac{y_2 + h (w_{12} z_1 - b_2)}{1 + h y_2 (w_{12} z_1 - b_2)}$$

Este método fue diseñado para que, cualquiera que sea el paso de discretización, se conserve la función de Lyapunov del sistema continuo. En la siguiente sección se generalizan los resultados aplicándolos a las redes de Hopfield de dimensión n .

4.3. Generalización del gradiente discreto al sistema de Hopfield de dimensión n

Con vistas a generalizar los resultados obtenidos en la sección anterior, se considera $\mathbf{y} \in \mathbb{R}^n$, $n \in \mathbb{N}$. Entonces, se tiene que las redes de Hopfield vienen dadas por el sistema:

$$\begin{aligned} \frac{du_i}{dt} &= \sum_{j=1}^n w_{ij} y_j - b_i \\ y_i &= \tanh(u_i) \end{aligned} \tag{4.10}$$

cuya función de Lyapunov resulta ser:

$$V(\mathbf{y}) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} y_i y_j + \sum_{i=1}^n b_i y_i$$

Obsérvese que asumir la ausencia de autopesos es equivalente a decir que V es multilínea, por lo que responde a la descomposición dada en la ecuación (3.29) del capítulo 3.

Por otra parte, se sabe que es posible reescribir este sistema en términos de las variables y_i , coincidiendo esta formulación con la de gradiente lineal dada en (4.5). En concreto, se tiene:

$$\frac{dy_i}{dt} = (1 - y_i^2) \left(\sum_{j=1}^n w_{ij} y_j - b_i \right)$$

Esta formulación es conveniente, para posteriormente pasar a seleccionar los parámetros del método. La correcta elección de los parámetros permitirá reescribir el método numérico de gradiente discreto en forma explícita. Los parámetros escogidos en este caso son:

$$\tilde{L}(\mathbf{y}, \mathbf{z}, h) = - \begin{pmatrix} 1 - y_1 z_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 - y_n z_n \end{pmatrix}$$

Es fácil comprobar que $\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y})$, ya que cuando $h = 0$, se cumple $\mathbf{y} = \mathbf{z}$. Se

considera ahora el gradiente discreto del incremento de las coordenadas:

$$\bar{\nabla}V(\mathbf{y}, \mathbf{z}) = \begin{pmatrix} \frac{V(z_1, y_2, \dots, y_n) - V(y_1, y_2, \dots, y_n)}{z_1 - y_1} \\ \frac{V(z_1, z_2, y_3, \dots, y_n) - V(z_1, y_2, \dots, y_n)}{z_2 - y_2} \\ \vdots \\ \frac{V(z_1, \dots, z_{i-1}, z_i, y_{i+1}, \dots, y_n) - V(z_1, \dots, z_{i-1}, y_i, y_{i+1}, \dots, y_n)}{z_i - y_i} \\ \vdots \\ \frac{V(z_1, \dots, z_n) - V(z_1, \dots, z_{n-1}, y_n)}{z_n - y_n} \end{pmatrix}$$

Como se comentó en el capítulo 3, cuando la función de Lyapunov es multilinear, se puede simplificar la notación escribiendo el gradiente discreto propuesto como:

$$(\bar{\nabla}V)_i(\mathbf{y}, \mathbf{z}) = A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) \tag{4.11}$$

siempre que $y_i \neq z_i$, donde $A_i = \frac{\partial V}{\partial y_i}$. En ocasiones, por brevedad se escribirá:

$$A_i(\mathbf{y}, \mathbf{z}) = A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n)$$

pero es importante tener en cuenta que $A_i(\mathbf{y}, \mathbf{z})$ no depende de z_j cuando $j \geq i$. Utilizando esta notación, se puede reescribir el gradiente discreto del incremento de las coordenadas de la siguiente manera:

$$\bar{\nabla}V(\mathbf{y}, \mathbf{z}) = \begin{pmatrix} A_1(y_2, \dots, y_n) \\ A_2(z_1, y_3, \dots, y_n) \\ \vdots \\ A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) \\ \vdots \\ A_n(z_1, \dots, z_{n-1}) \end{pmatrix}$$

Nótese que en este caso también resulta sencillo comprobar que, para $h = 0$

$$(\bar{\nabla}V)_i(\mathbf{y}, \mathbf{y}) = (\nabla V)_i(\mathbf{y})$$

Luego de escoger los parámetros del método y comprobar que cumplen las condiciones bajo las cuales se garantiza la consistencia del método numérico, se procede a la discretización del sistema (4.10) como en (4.9). Teniendo en cuenta la elección de los parámetros, se puede garantizar la obtención de un método de gradiente discreto explícito. Esto ocurre específicamente debido a que V es multilineal, $\tilde{L}(\mathbf{y}, \mathbf{z}, h)$ es lineal respecto a la variable \mathbf{z} y el gradiente discreto seleccionado es el incremento de las coordenadas. Finalmente, el método obtenido es:

$$z_i = \frac{y_i + h(A_i - b_i)}{1 + h y_i (A_i - b_i)}, \quad A_i = \sum_{j < i} w_{ij} z_j + \sum_{j > i} w_{ij} y_j \quad (4.12)$$

Nótese que, a pesar de que a simple vista parece que el método obtenido es implícito, hay que tener en cuenta que $A_i(\mathbf{y}, \mathbf{z})$ no depende de z_j cuando $j \geq i$. Entonces, es evidente que cada componente puede ser calculada utilizando solamente las componentes anteriores a esta, por lo que el método de gradiente discreto obtenido es explícito en la práctica. De esta manera, es posible aprovechar todas las ventajas de los métodos explícitos sobre los implícitos, sobre todo en cuanto a eficiencia a la hora de implementarlos.

4.4. Análisis del orden de la aplicación al sistema de Hopfield

En esta sección se analiza el método de gradiente discreto propuesto, analizando su orden en el caso de la integración de las redes de Hopfield n -dimensionales. Se sabe que el orden del método numérico obtenido puede ser estudiado siguiendo el procedimiento estándar, en otras palabras, comparando el desarrollo en serie de Taylor alrededor de $h = 0$ de la solución exacta del sistema de ecuaciones diferenciales y de la solución aproximada que brinda el método numérico. Nótese que el método de

gradiente discreto es consistente por construcción, por lo que es al menos de orden uno y por ende el error luego de un paso viene dado por $\mathbf{y}(t+h) - \mathbf{z} = C_{GD} h^2 + O(h^3)$, siendo C_{GD} la constante de error del método. En el capítulo 3, se vio que la constante de error depende de J^d y J^c según la ecuación que se reproduce a continuación por comodidad de referencia:

$$C_{GD} = \left(\frac{1}{2} J^c - J^d \Big|_{h=0} \right) f(\mathbf{y}) \quad (4.13)$$

siendo $J^d = \frac{\partial(\tilde{L} \cdot \bar{\nabla} V)}{\partial \mathbf{z}}$, $J^c = \frac{\partial(L \cdot \nabla V)}{\partial \mathbf{y}}$ los jacobianos del método numérico y de la ecuación continua, respectivamente, mientras que $f(\mathbf{y})$ es la parte derecha de la EDO.

Como se mencionó anteriormente, la simple aplicación del método de gradiente discreto a sistemas de EDO lineales y escalares conduce a la obtención de la regla trapezoidal, que como se sabe del análisis numérico es un método de segundo orden, por lo que es ventajoso en comparación con la regla de Euler. A continuación se analizará si esta ventaja se mantiene cuando se aplica este método a sistemas más complejos. Nótese que, teniendo en cuenta la elección de \tilde{L} y $\bar{\nabla} V$, la i -ésima coordenada del término $\tilde{L} \cdot \bar{\nabla} V$ para las redes de Hopfield viene dada por:

$$(\tilde{L} \cdot \bar{\nabla} V)_i = (1 - y_i z_i) A_i(\mathbf{y}, \mathbf{z}) \quad (4.14)$$

donde $A_i(\mathbf{y}, \mathbf{z})$ se define como en (4.11). Ahora bien, la matriz jacobiana del sistema continuo J^d se define según:

$$J_{ii}^d = -y_i A_i(\mathbf{y}, \mathbf{z}) \qquad J_{ij}^d = (1 - y_i z_i) \frac{\partial A_i}{\partial z_j} = (1 - y_i z_i) w_{ij}$$

ya que $\frac{\partial A_i}{\partial z_i} = 0$ si $i \neq j$. Teniendo en cuenta que cuando $h = 0$, se cumplen $\mathbf{z} = \mathbf{y}$,

$A_i(\mathbf{y}, \mathbf{y}) = \frac{\partial V}{\partial y_i}$ resulta:

$$J_{ii}^d \Big|_{h=0} = -y_i \frac{\partial V}{\partial y_i} \qquad J_{ij}^d \Big|_{h=0} = (1 - y_i^2) w_{ij} \quad (4.15)$$

Por su parte, teniendo en cuenta la expresión de J^c , la definición de la matriz L y del gradiente discreto ∇V se obtiene:

$$J_{ii}^c = -2y_i \frac{\partial V}{\partial y_i} \qquad J_{ij}^c = (1 - y_i^2) w_{ij} \quad \text{si } i \neq j \qquad (4.16)$$

Calculando la expresión de la constante de error C_{GD} del método de gradiente discreto, se observa que solo los términos de la diagonal se anulan, ya que habría que calcular la diferencia $\frac{1}{2} J^c - J^d \Big|_{h=0}$. Este resultado basta para concluir que, de manera general, el método de gradiente discreto obtenido es solo de primer orden. No obstante, si se escogen los parámetros del método de manera tal que $J^c = 2 J^d \Big|_{h=0}$ se obtendrá un método de gradiente discreto al menos de segundo orden. Por otra parte, se recuerda que la constante de error del método de Euler viene dada por:

$$C_E = \frac{1}{2} J^c f(y)$$

Si se compara este valor con el de la constante de error del método de gradiente discreto, que viene dada por:

$$C_{GD} = \left(\frac{1}{2} J^c - J^d \Big|_{h=0} \right) f(y)$$

lo que sugiere que $C_E > C_{GD}$, a pesar de ser ambos métodos de primer orden. Es decir, el método de gradiente discreto brindaría una mejor aproximación de las soluciones, a la vez que conserva la función de Lyapunov del sistema. En el próximo capítulo, mediante simulaciones numéricas se respalda esta suposición.

El análisis que se ha realizado en esta sección no es fácilmente generalizable para la obtención de condiciones superiores a orden dos. El establecimiento de una teoría sistemática del orden de métodos de gradiente discreto, similar a la desarrollada para los métodos de Runge-Kutta, es una dirección prometedora para futuras investigaciones que, sin embargo, se sale de los objetivos trazados para este trabajo.

4.5. Método de gradiente discreto para el sistema de Hopfield con autopesos

En los epígrafes anteriores se trabajó con la formulación de Abe de las redes neuronales de Hopfield considerando que no existen autopesos, o sea, asumiendo $w_{ii} = 0$. Como se dijo anteriormente, en este caso los puntos de equilibrio estables del sistema se encuentran en los vértices del hipercubo de definición. Este modelo encuentra una aplicación significativa como método de optimización cuando la función objetivo es multilineal. En este epígrafe se asume la presencia de autopesos en el modelo y se diseña un método de gradiente discreto para discretizarlo. En este caso, los puntos de equilibrio del sistema continuo se encuentran en el interior del hipercubo. Una variante del modelo resultante se ha aplicado a la construcción de un método de estimación de parámetros, por lo que es también un interesante caso de estudio.

Se recuerda que la formulación de Abe viene dada por:

$$\frac{du_i}{dt} = \sum_{j=1}^n w_{ij} y_j - b_i$$

$$y_i = \tanh\left(\frac{u_i}{\beta}\right)$$

y que este sistema posee función de Lyapunov, dada por:

$$V(y) = -\frac{1}{2} \sum_i \sum_j w_{ij} y_i y_j + \sum_{i=1}^n b_i y_i$$

Como en este caso se está asumiendo la presencia de autopesos en el modelo, o sea, $w_{ii} \neq 0$, la función de Lyapunov no sería multilineal por lo que no es posible reescribirla como en el ejemplo anterior, en el que se asumió la ausencia de autopesos.

En cambio, la siguiente descomposición es válida:

$$V = - \sum_i \left(\sum_{j>i} w_{ij} y_i y_j + \frac{1}{2} w_{ii} y_i^2 \right) + \sum_i b_i y_i$$

También en este caso puede reescribirse el sistema en términos de las variables y_i , resultando:

$$\frac{dy_i}{dt} = (1 - y_i^2) \left(\sum_j w_{ij} y_j - b_i \right)$$

Luego de reescribir el sistema en la forma de gradiente lineal y siguiendo la metodología de trabajo para el diseño de métodos de gradiente discreto, dada al inicio de este capítulo, el siguiente paso sería escoger los parámetros del método, a saber: el gradiente discreto $\bar{\nabla}V$ y la matriz \tilde{L} . Como se vio anteriormente, estos parámetros se escogen de manera tal que se cumplan las dos condiciones $\bar{\nabla}V(\mathbf{y}, \mathbf{y}) = \nabla V(\mathbf{y})$, $\tilde{L}(\mathbf{y}, \mathbf{y}, 0) = L(\mathbf{y})$, garantizando de esta manera la obtención de un método consistente. En este caso también se trabaja con el gradiente discreto del incremento de las coordenadas. Se sabe que de manera general la coordenada i -ésima de este gradiente discreto viene dada por:

$$(\bar{\nabla}V)_i = \frac{V(z_1, \dots, z_{i-1}, z_i, y_{i+1}, \dots, y_n) - V(z_1, \dots, z_{i-1}, y_i, y_{i+1}, \dots, y_n)}{z_i - y_i}$$

y teniendo en cuenta la descomposición de V en este caso, se tiene:

$$\begin{aligned} (\bar{\nabla}V)_i &= \frac{A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) z_i + B_i - \frac{1}{2} w_{ii} z_i^2}{z_i - y_i} \\ &\quad - \frac{\left(A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) y_i + B_i - \frac{1}{2} w_{ii} y_i^2 \right)}{z_i - y_i} \\ &= \frac{\left(A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) - \frac{1}{2} w_{ii} (z_i + y_i) \right) (z_i - y_i)}{(z_i - y_i)} \end{aligned}$$

De esta manera, finalmente se tiene que la i -ésima coordenada del gradiente discreto del incremento de las coordenadas, se define según:

$$(\bar{\nabla}V)_i = A_i(z_1, \dots, z_{i-1}, y_{i+1}, \dots, y_n) - \frac{1}{2} w_{ii} (z_i + y_i) \quad (4.17)$$

donde:

$$A_i = \sum_{j < i} w_{ij} z_j + \sum_{j > i} w_{ij} y_j$$

igual que en el caso sin autopesos, aunque cabe señalar que en este caso no coincide con el gradiente de la función de Lyapunov.

Por su parte, en el modelo en que se consideraba que todos los autopesos w_{ii} se anulaban, se asumió:

$$\tilde{L} = -\text{diag}(1 - y_i z_i)$$

Si se toma ahora este valor para \tilde{L} , se habrían seleccionado exactamente los mismos parámetros que para el modelo de Hopfield sin autopesos. Se trabajará provisionalmente con esta definición, con el único objetivo de ilustrar la importancia que tiene una correcta selección de los parámetros. En efecto, a continuación se podrá apreciar cómo en este caso esta selección no es adecuada, lo que dará lugar a un cambio en la elección de \tilde{L} .

Sustituyendo los parámetros escogidos en la ecuación (4.6) que define el método de integración, se tiene:

$$\begin{aligned} \frac{z_i - y_i}{h} &= -(1 - y_i z_i) \left(A_i - \frac{1}{2} w_{ii} (z_i + y_i) \right) = -(1 - y_i z_i) \left[A_i - \frac{1}{2} w_{ii} z_i - \frac{1}{2} w_{ii} y_i \right] \\ z_i &= y_i - h(1 - y_i z_i) \left(A_i - \frac{1}{2} w_{ii} z_i - \frac{1}{2} w_{ii} y_i \right) \\ &= y_i - h \left(A_i - \frac{1}{2} w_{ii} z_i - \frac{1}{2} w_{ii} y_i - A_i y_i z_i + \frac{1}{2} w_{ii} y_i z_i^2 + \frac{1}{2} w_{ii} y_i^2 z_i \right) \\ &= y_i - h \left[\left(\frac{1}{2} w_{ii} y_i \right) z_i^2 + \left(\frac{1}{2} w_{ii} y_i^2 - \frac{1}{2} w_{ii} - A_i y_i \right) z_i + \left(A_i - \frac{1}{2} w_{ii} y_i \right) \right] \end{aligned}$$

Finalmente, agrupando por potencias de z_i , resulta:

$$\left(-\frac{h}{2} w_{ii} y_i \right) z_i^2 + \left[h \left(-\frac{1}{2} w_{ii} y_i^2 + \frac{1}{2} w_{ii} + A_i y_i \right) - 1 \right] z_i + \left[h \left(-A_i + \frac{1}{2} w_{ii} y_i \right) + y_i \right] = 0 \quad (4.18)$$

con lo que si se denota:

$$\begin{aligned}\alpha &= -\frac{h}{2} w_{ii} y_i \\ \beta &= \frac{h}{2} w_{ii} + h A_i y_i - \frac{h}{2} w_{ii} y_i^2 - 1 \\ \gamma &= \frac{h}{2} w_{ii} y_i - h A_i + y_i\end{aligned}\tag{4.19}$$

la ecuación (4.18) no es más que una ecuación de segundo grado en términos de z_i .

A continuación se hace un análisis cualitativo para decidir cuál de las dos soluciones de la ecuación de segundo grado obtenida anteriormente, definirá el método de gradiente discreto que se está diseñando. Está claro que, por consistencia, cuando $h = 0$ debe cumplirse que $z = y$. Si se considera que la ecuación (4.19) define los coeficientes $\alpha(h)$, $\beta(h)$, $\gamma(h)$ como funciones de h , cuando $h = 0$ se tendrán los valores $\alpha(0) = 0$, $\beta(0) = -1$, $\gamma(0) = y_i$. Se considera ahora la ecuación $\alpha(h)z^2 + \beta(h)z + \gamma(h) = 0$ y se comienza el análisis con la solución que viene dada por:

$$z_i = \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}\tag{4.20}$$

Analizando el límite cuando $h \rightarrow 0$, se tiene:

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} &= \lim_{h \rightarrow 0} \left(\frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} \cdot \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}} \right) \\ &= \lim_{h \rightarrow 0} \frac{\beta^2 - (\beta^2 - 4\alpha\gamma)}{2\alpha(-\beta - \sqrt{\beta^2 - 4\alpha\gamma})} \\ &= \lim_{h \rightarrow 0} \frac{2\gamma}{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}} \\ &= \lim_{h \rightarrow 0} \frac{2\gamma}{-\beta - |\beta|} = \infty\end{aligned}$$

ya que $\beta(0) = -1$. La no existencia del límite es un criterio suficiente para desechar esta solución, dado que se asume que el método debe estar definido por una función continua.

Haciendo un análisis similar con la otra solución, se obtiene:

$$\lim_{h \rightarrow 0} \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha} = \lim_{h \rightarrow 0} \frac{2\gamma}{-\beta + |\beta|} = \frac{2\gamma}{-2\beta} = \frac{2y_i}{2} = y_i$$

De esta manera se concluye que $z = y$ cuando $h = 0$, que coincide con la condición intuitiva de consistencia. Así, se puede afirmar que el método de gradiente discreto estaría definido por la ecuación $z_i = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$ con α, β, γ como en la ecuación (4.19).

Luego de efectuar algunos experimentos numéricos con vistas a validar el funcionamiento del método obtenido, se puede arribar a la conclusión de que este no es adecuado en varios aspectos:

1. Habría que chequear en cada paso el signo de β , ya que cuando $\beta > 0$ la discretización correcta es la que viene dada por la ecuación:

$$z_i = \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

2. A pesar de conservar la función de Lyapunov del sistema continuo, la discretización no respeta el dominio de definición de este, ya que las soluciones se salen del hipercubo $(-1, 1)$.
3. Otro inconveniente de esta discretización es que siempre que $\beta^2 - 4\alpha\gamma < 0$, el método no encuentra solución.

Teniendo en cuenta el mal funcionamiento del método obtenido, se propone una elección diferente de los parámetros, con el objetivo de mejorar su eficiencia. Se continúa trabajando con el mismo gradiente discreto, o sea:

$$(\bar{\nabla}V)_i = A_i - \frac{1}{2} w_{ii} (z_i + y_i)$$

pero se toma la elección más sencilla para la matriz \tilde{L} , es decir, $\tilde{L} = L = -\text{diag}(1 - y_i^2)$.

Entonces, el método queda de la siguiente forma:

$$\begin{aligned} \frac{z_i - y_i}{h} &= - (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} (z_i + y_i) \right) \\ &= - (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} z_i - \frac{1}{2} w_{ii} y_i \right) \\ &= - (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} y_i \right) + (1 - y_i^2) \left(\frac{1}{2} w_{ii} z_i \right) \end{aligned}$$

con lo que:

$$\begin{aligned} z_i &= y_i - h (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} y_i \right) + h (1 - y_i^2) \left(\frac{1}{2} w_{ii} z_i \right) \\ z_i - \frac{h}{2} (1 - y_i^2) w_{ii} z_i &= y_i - h (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} y_i \right) \\ \left(1 - \frac{h}{2} (1 - y_i^2) w_{ii} \right) z_i &= y_i - h (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} y_i \right) \end{aligned}$$

Finalmente, se obtiene un método de gradiente discreto explícito que viene dado por el siguiente modelo iterativo:

$$z_i = \frac{y_i - h (1 - y_i^2) \left(A_i - \frac{1}{2} w_{ii} y_i \right)}{1 - \frac{h}{2} (1 - y_i^2) w_{ii}}$$

Este método de gradiente discreto permite discretizar las redes de Hopfield con autopesos a la vez que conserva la función de Lyapunov de la red continua. Obsérvese que en este caso también fue posible reescribir el método en forma explícita incluso cuando la función de Lyapunov no es multilineal. Por otra parte, haciendo un análisis del orden similar al realizado para el método obtenido con ausencia de autopesos, se concluye que este método es también de primer orden.

Teniendo en cuenta que en ambos casos se han obtenido métodos de gradiente discreto de primer orden, la propuesta de una generalización de alto orden de dichos métodos es una línea natural a seguir. Es en esta dirección en la que se encuentran los

resultados de la próxima sección.

4.6. Método de gradiente discreto adjunto

El objetivo de esta sección es obtener un método de gradiente discreto de orden superior. Específicamente, se propone la extensión de alto orden del método de gradiente discreto diseñado al inicio del capítulo, aplicado a las redes de Hopfield sin autopesos. Para esto, se considera la composición de este con su adjunto. De manera general, para obtener el adjunto de un método numérico se considera el inverso del método y se retrocede en el tiempo. En otras palabras, se efectúan los intercambios z_i por y_i , h por $-h$, tal como se explicó con más detalle en el capítulo 2. Teniendo esto en cuenta y considerando la expresión del método de gradiente discreto obtenido con anterioridad, cuya expresión viene dada en la ecuación (4.12), se tiene:

$$y_i = \frac{z_i - h A_i(y_1, \dots, y_{i-1}, z_{i+1}, \dots, z_n)}{1 - h z_i A_i(y_1, \dots, y_{i-1}, z_{i+1}, \dots, z_n)} \quad i = 1, \dots, n \quad (4.21)$$

Luego de despejar, es necesario invertir el orden de cálculo de las coordenadas para obtener un método explícito, el cual será el adjunto del método de gradiente discreto dado en la ecuación (4.12), o sea:

$$z_i = \frac{y_i - h A_i(y_1, \dots, y_{i-1}, z_{i+1}, \dots, z_n)}{1 - h y_i A_i(y_1, \dots, y_{i-1}, z_{i+1}, \dots, z_n)}, \quad i = n, \dots, 1 \quad (4.22)$$

Teniendo en cuenta que el objetivo de esta sección es obtener un método de gradiente discreto de orden superior, se propone la composición del método de gradiente discreto de primer orden con su adjunto. Como se dijo en el capítulo 2, la composición de métodos numéricos tiene muchas finalidades pero cuando se trata de incrementar el orden de convergencia, una de las vías a seguir es la composición:

$$\Psi_h = \Phi_{h/2} \circ \Phi_{h/2}^* \quad (4.23)$$

que convierte todo método Φ consistente de paso simple y orden uno en un método Ψ simétrico de segundo orden [17]. En particular, siendo Φ_h el método de gradiente

discreto propuesto en la ecuación (4.12) y Φ_h^* su adjunto, cuya expresión viene dada en la ecuación (4.22), entonces se puede afirmar que Ψ_h será un método numérico explícito, simétrico y de segundo orden. Resulta relativamente sencillo comprobar que además este método conserva la función de Lyapunov del sistema continuo. Se tiene que:

$$\Phi_{h/2}(\mathbf{y}_n) = \tilde{\mathbf{y}}_{n+1} \Rightarrow V(\tilde{\mathbf{y}}_{n+1}) < V(\mathbf{y}_n)$$

ya que Φ_h es un método de gradiente discreto, siendo V la función de Lyapunov del sistema. Por otra parte:

$$\Psi_h(\mathbf{y}_n) = \Phi_{h/2}^*(\Phi_{h/2}(\mathbf{y}_n)) = \Phi_{h/2}^*(\tilde{\mathbf{y}}_{n+1}) = \mathbf{y}_{n+1}$$

y en este caso también se cumple que $V(\mathbf{y}_{n+1}) < V(\tilde{\mathbf{y}}_{n+1})$ ya que Φ_h^* es también un método de gradiente discreto, con lo que se concluye la validez de la siguiente cadena de desigualdades:

$$V(\mathbf{y}_{n+1}) < V(\tilde{\mathbf{y}}_{n+1}) < V(\mathbf{y}_n)$$

En resumen, el método numérico propuesto, además de ser explícito, simétrico y de orden dos, conserva la función de Lyapunov del sistema.

4.7. Conclusiones

En este capítulo se recogen algunos de los resultados más importantes de la tesis. En primer lugar, se introducen las redes de Hopfield como campo de aplicación de las técnicas de integración que se proponen posteriormente. Se diferencian dos casos concretos, según se asuma la ausencia o la presencia de autopesos en el modelo. A continuación, se diseña un método de gradiente discreto de manera tal que se conserve la función de Lyapunov bajo discretización, cualquiera sea el tamaño de paso. Se muestra, además, que una elección adecuada de los parámetros del método de gradiente discreto permite reescribirlo explícitamente, a pesar de ser estos métodos implícitos en principio. Este resultado se obtiene tanto para el modelo sin autopesos como, posteriormente, para el que incluye autopesos y, por tanto, cuya función

de Lyapunov no es multilineal. Se realiza además, en este capítulo, el estudio del orden del método propuesto, obteniéndose condiciones de segundo orden sobre los parámetros. Finalmente, se propone una extensión de alto orden del método obtenido, al discretizar el sistema de Hopfield considerando la ausencia de autopesos y se demuestra su capacidad para conservar la función de Lyapunov del sistema continuo.

En el próximo capítulo, se muestran simulaciones en las que se ilustra el funcionamiento del método de gradiente discreto obtenido para la discretización de las redes de Hopfield sin autopesos. Se muestran también en las simulaciones las mejoras que introduce su extensión de segundo orden ya que, si bien es cierto que las soluciones obtenidas son similares cualitativamente, el método de segundo orden minimiza el error local cometido en la discretización, a la vez que conserva la función de Lyapunov del sistema, respetando a su vez el dominio de definición de la red continua.

Capítulo 5

Experimentos numéricos

Resumen del capítulo

En este capítulo se muestran un grupo de simulaciones numéricas con el objetivo de respaldar los resultados teóricos de capítulos anteriores respecto al orden, convergencia y exactitud de los métodos de gradiente discreto propuestos. Los experimentos se diseñaron para la discretización de las redes de Hopfield asumiendo la ausencia de autopesos. En primer lugar, se compara el método de primer orden diseñado con la regla de Euler, algunos esquemas de Runge-Kutta y el método de proyección presentado en capítulos anteriores. En todos los casos, se aprecia el buen comportamiento del método de gradiente discreto, tanto en la precisión con la que aproxima la solución del sistema, como en su capacidad para conservar la función de Lyapunov del sistema continuo cualquiera sea el tamaño de paso de discretización. Este buen comportamiento persiste incluso al aumentar el tamaño de paso, hasta el punto de que el resto de los métodos muestran comportamientos erróneos desde el punto de vista dinámico. A continuación, se diseñan algunos experimentos con objeto de mostrar las mejoras obtenidas al proponer la extensión de alto orden del método de gradiente discreto, mediante la composición con el método adjunto.

5.1. Método de gradiente discreto

En el capítulo anterior se propone un método de gradiente discreto para la discretización de las redes de Hopfield que, como ya se ha dicho, vienen dadas por la siguiente ecuación algebraico-diferencial:

$$\frac{du_i}{dt} = \sum_{j=1}^n w_{ij} y_j - b_i$$

$$y_i = \tanh u_i$$

El método de gradiente discreto obtenido cuando se asume la ausencia de autopesos en el modelo viene dado por:

$$z_i = \frac{y_i + h(A_i - b_i)}{1 + h y_i (A_i - b_i)}, \quad A_i = \sum_{j < i} w_{ij} z_j + \sum_{j > i} w_{ij} y_j$$

que, como se vio en el capítulo anterior, es un método explícito ya que todas las componentes z_j con $j < i$ ya han sido calculadas cuando se afronta el cómputo de z_i .

En este capítulo se muestra el resultado de algunos experimentos numéricos realizados con el fin de evaluar el funcionamiento de este método. De la teoría del análisis numérico se sabe que cualquier método numérico puede proporcionar una aproximación correcta de las soluciones de un determinado sistema de ecuaciones diferenciales, si se trabaja con un tamaño de paso suficientemente pequeño. Es por esto que, con objeto de mostrar las ventajas de trabajar con este método de gradiente discreto, se diseñan casos de prueba en los que usualmente los métodos tradicionales fallan al obtener una aproximación adecuada de las soluciones o al conservar la dinámica del sistema continuo. En la sección 4.1 vimos que las redes de Hopfield, desde el punto de vista dinámico, constituyen un sistema estable, ya que es un sistema con función de Lyapunov. Por otra parte, teniendo en cuenta la ecuación que define este modelo, se sabe que su dominio de definición se encuentra dentro del hipercubo $(-1, 1)^n$ y que al asumir la ausencia de autopesos, o sea $w_{ii} = 0$, los puntos de equilibrio estables coinciden con los vértices del hipercubo, puntos en los que $|y_i| = 1$. Puede decirse, por

consiguiente, que tanto la existencia de la función de Lyapunov como su dominio de definición son dos características fundamentales de este sistema, teniendo en cuenta, entre otras aplicaciones, la resolución de problemas de optimización. Por tanto, el objetivo de los experimentos será mostrar la superioridad del método de gradiente discreto, en lo tocante a que, al discretizar el sistema, se continúen conservando estas propiedades.

Para diseñar los experimentos numéricos, se tuvo en cuenta que en una vecindad de un punto crítico, el comportamiento de las soluciones puede ser cualitativamente diferente, dependiendo de la discretización y el tamaño de paso con el que se trabaje, por lo que tomar el punto inicial cerca de un punto crítico sería un interesante caso de estudio. Por tanto, se fuerza el sistema a tener un punto de equilibrio inestable y_c , el cual se define como un punto arbitrario en el interior del hipercubo $(-1, 1)^n$, estableciendo convenientemente la relación entre los parámetros, es decir: $b = W \cdot y_c$. Nótese que el punto de ensilladura que presenta el sistema pertenece a la frontera entre las cuencas de atracción asociadas a diferentes puntos de equilibrio estables. Finalmente, se escoge el punto inicial y_0 en una vecindad del punto crítico y_c , de manera tal que pequeñas perturbaciones conduzcan a comportamientos dinámicos cualitativamente diferentes.

Con respecto a la matriz de pesos W , inicialmente, se tomaron los coeficientes de la matriz W aleatorios, pero de manera tal que la matriz fuese simétrica y con diagonal nula. Los experimentos realizados con esta definición de los parámetros no mostraron ningún comportamiento relevante, o sea, no se aprecia ninguna ventaja del método de gradiente discreto sobre los métodos tradicionales. En el siguiente conjunto de experimentos, se definen los coeficientes de la matriz W diferentes en distintos órdenes de magnitud, de manera tal que el sistema sea moderadamente *stiff*. Fue entonces cuando comenzaron a ser visibles las ventajas de la utilización del método de gradiente discreto propuesto. Los experimentos numéricos se realizaron con cuatro enfoques diferentes, que se explicaran detalladamente en próximas secciones. La idea general de cada enfoque sería:

1. *Experimentos cualitativos*: comparación con la regla de Euler, con el objetivo de mostrar la factibilidad, al menos en principio, del método propuesto. Para

esto se utilizan como ejemplo de prueba las redes de Hopfield con $n = 3$.

2. *Experimentos cuantitativos*: comparación con la regla de Euler, confirmando los resultados teóricos obtenidos en la sección 4.4 referentes al orden y precisión del método propuesto.
3. *Comparación práctica* sistemática con métodos de Runge-Kutta comerciales [54]: se implementan los métodos sobre sistemas de alta dimensión, y los resultados apoyan el uso práctico del método propuesto.
4. Comparación con el método de *proyección* [12]: el objetivo es mostrar las ventajas del método de gradiente discreto propuesto frente a métodos que también han sido diseñados expresamente para la integración de sistemas de gradiente.

5.1.1. Enfoque cualitativo. Comparación con el método de Euler

El primero de los experimentos está diseñado desde el punto de vista cualitativo, siendo el objetivo mostrar el funcionamiento del método de gradiente discreto, comparando su funcionalidad con la del método de Euler explícito. La elección del método de Euler resultó ser la más natural, teniendo en cuenta que el método de gradiente discreto es también explícito y de primer orden. Para los experimentos se considera el sistema de Hopfield con $n = 3$. La matriz de los autopesos se escoge simétrica y con diagonal nula y, en concreto, los valores exactos para este experimento son:

$$W = \begin{pmatrix} 0 & 100 & 1 \\ 100 & 0 & 100 \\ 1 & 100 & 0 \end{pmatrix}$$

Por su parte, el punto inicial se escoge aleatoriamente pero en una vecindad reducida del punto crítico, que en este caso se fija en $y_c = (0,4 \ -0,2 \ 0,1)^T$, mientras que b viene dada por $b = W \cdot y_c$, garantizando así que y_c es un punto de equilibrio. Primeramente, se trabaja con un tamaño de paso pequeño, específicamente $h = 0,001$, como grupo de control. En este caso el comportamiento dinámico de ambos métodos es similar. Ambos conservan la función de Lyapunov del sistema continuo, aunque se observa

que el error local en el caso del método de Euler es levemente superior al del método de gradiente discreto. Dicho sea de paso, no es de extrañar que el método de Euler se comporte correctamente desde el punto de vista dinámico, pues se ha demostrado que este método conserva la función de Lyapunov cuando se considera un tamaño de paso suficientemente pequeño [51]. Las trayectorias obtenidas se muestran en las figuras 5.1 y 5.2.

Sin embargo, cuando se aumenta el tamaño de paso a $h = 0,005$, el resultado de las simulaciones es radicalmente diferente. Las trayectorias para este caso se muestran en la figura 5.3, donde se observa que la segunda componente entra en una solución periódica cuando se integra el sistema utilizando el método de Euler. En cambio, la solución obtenida a través del método de gradiente discreto continúa conservando un comportamiento correcto: las tres componentes de la solución convergen a uno de los puntos de equilibrio estables del sistema. Respecto a la función de Lyapunov, cuya trayectoria se muestra en la figura 5.4, es evidente que la regla de Euler es incapaz de conservarla y este hecho se corrobora al aparecer puntos en los que el incremento de V es visible.

El comportamiento del error local también se muestra en la figura 5.4, observándose un comportamiento totalmente erróneo para la regla de Euler, a pesar de estar trabajando con un tamaño de paso pequeño. Por otra parte se realizó una batería completa de experimentos con diferentes tamaños de paso. Comenzando con $h = 0,004$, aparecen puntos donde la regla de Euler deja de conservar la función de Lyapunov, aunque la solución sigue convergiendo al punto de equilibrio estable. Para $h = 0,0045$, se obtiene una solución de período dos, mientras que la trayectoria que se muestra en la figura 5.4 para $h = 0,005$ parece ser de período cuatro. Aunque el análisis dinámico de las soluciones discretas obtenidas no constituya el objeto del presente trabajo, cabe comentar que la aparente existencia de una bifurcación que dobla el período, cuando el tamaño de paso se considera como un parámetro de bifurcación, sugiere una posible ruta al caos si el tamaño de paso se incrementara suficientemente. El comportamiento cualitativo obtenido (la preservación de la función de Lyapunov por el método del gradiente discreto y la aparición de soluciones periódicas para el método de Euler) aparece de forma consistente para una amplia

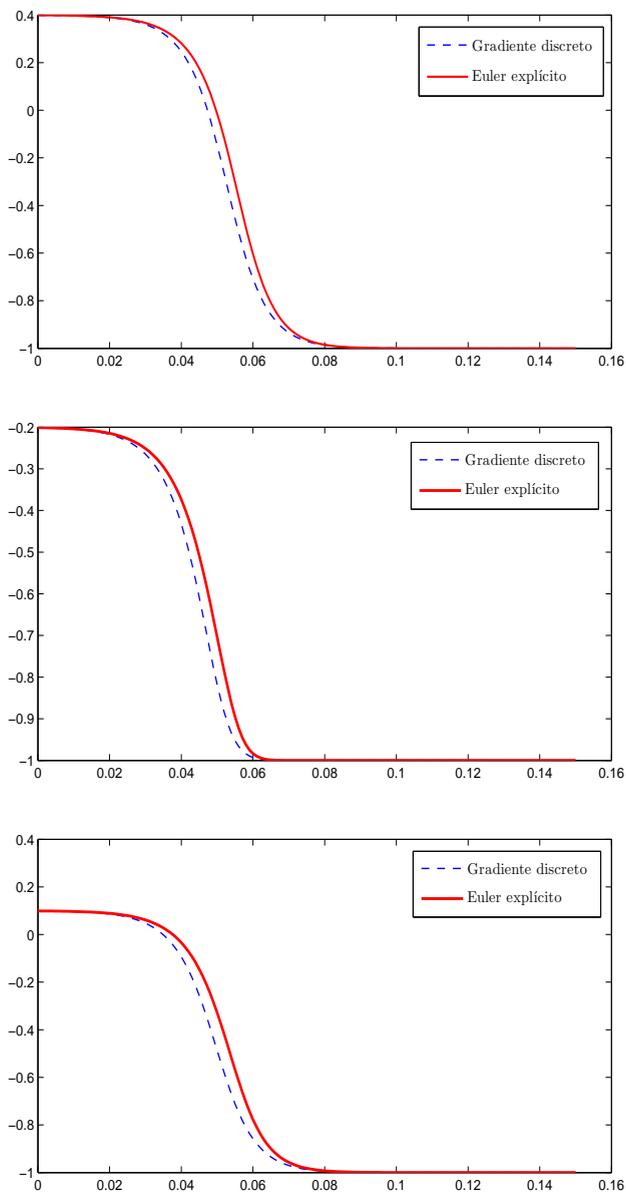


Figura 5.1: Trayectoria de las tres componentes de la solución de la red de Hopfield de dimensión $n = 3$ para $h = 0,001$.

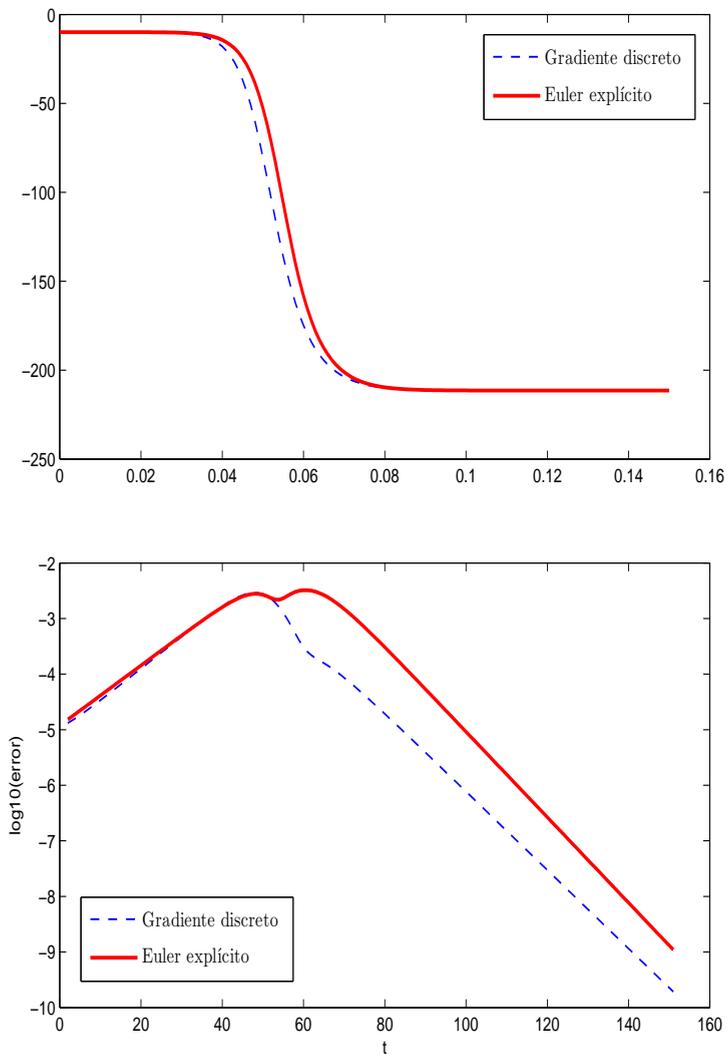


Figura 5.2: Función de Lyapunov V y error local de la red de Hopfield de dimensión $n = 3$ para $h = 0,001$.

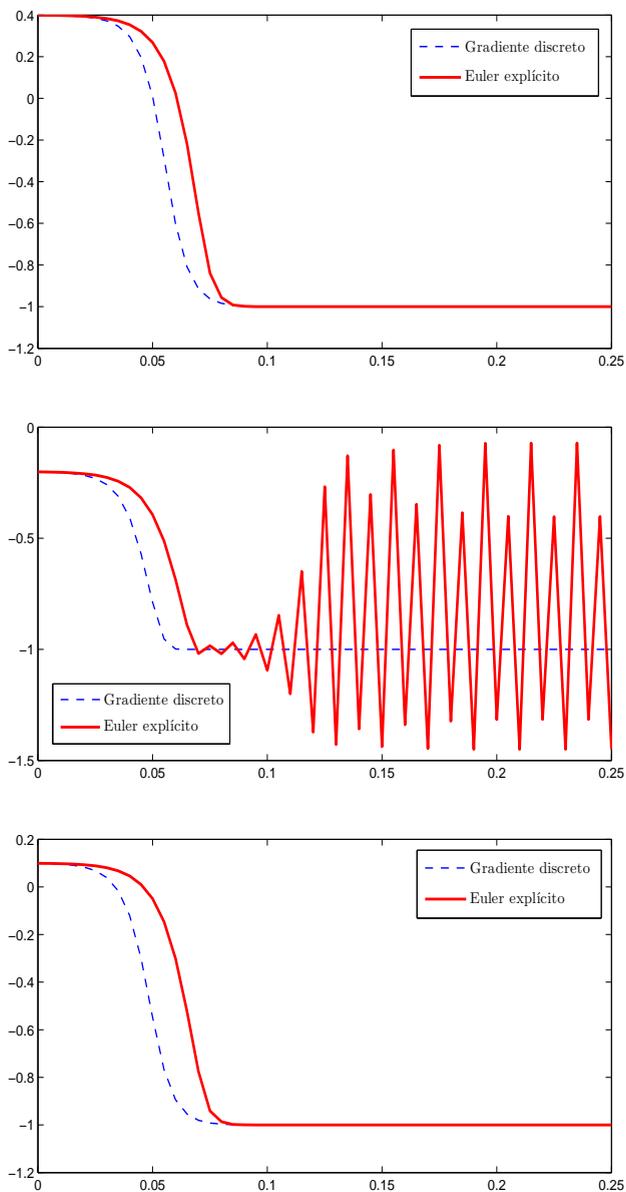


Figura 5.3: Trayectoria de las tres componentes de la solución de la red de Hopfield de dimensión $n = 3$ para $h = 0,005$.

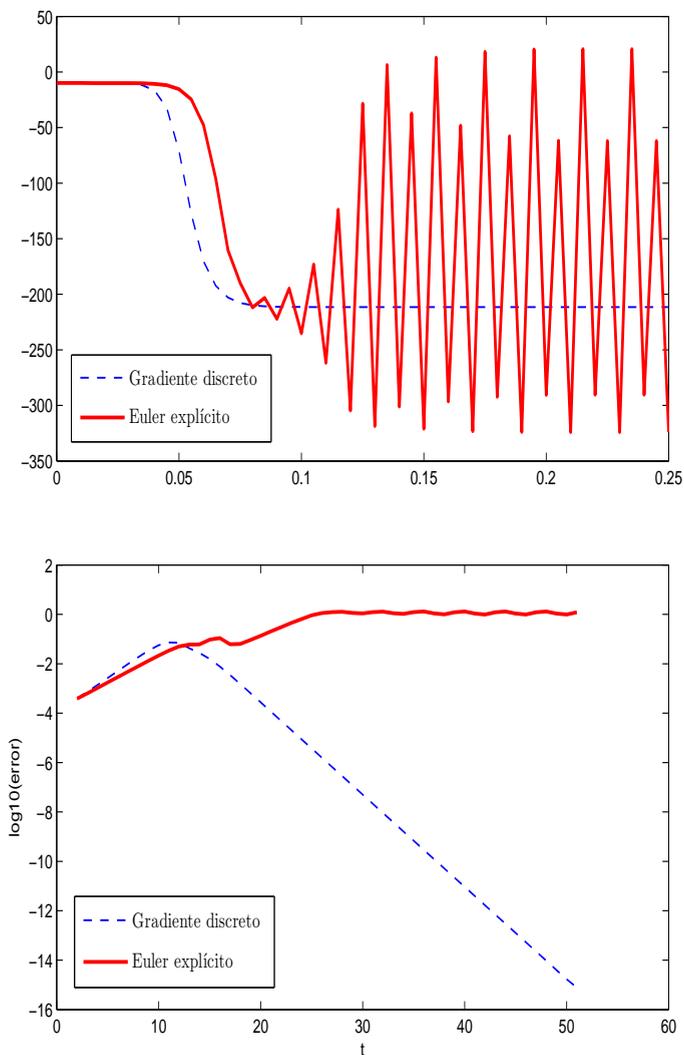


Figura 5.4: Función de Lyapunov V y error local de la red de Hopfield de dimensión $n = 3$ para $h = 0,005$.

gama de valores de los coeficientes W , b , así como de los valores iniciales y_0 .

5.1.2. Enfoque cuantitativo: análisis del error

El segundo grupo de experimentos se llevó a cabo con los mismos parámetros que en el caso anterior, pero ahora la elección de los tamaños de paso estaba guiada por un enfoque cuantitativo. A diferencia de los casos simples del capítulo 3, no se dispone de la expresión analítica de la solución exacta del sistema. Por tanto, no puede concluirse *a priori* cuál de los dos métodos está brindando una mejor aproximación de esta solución. En orden a tener una idea más clara al respecto, se discretiza considerando un tamaño de paso extremadamente pequeño, exactamente $h = 0,0001$, y las trayectorias son consideradas como la solución *exacta* del sistema, lo que se confirma teniendo en cuenta que las soluciones obtenidas por ambos métodos son muy similares. Entonces, se discretiza el sistema utilizando tanto el método de Euler como el de gradiente discreto utilizando $h = 0,001$ y se comparan con la solución considerada exacta, mostrando el resultado en la figura 5.5. Ambas trayectorias, cuando se considera $h = 0,001$, se alejan de la trayectoria exacta, pero la perturbación ocurrida para el método de Euler es considerablemente mayor. También en la figura 5.5 se muestra una vista ampliada de las trayectorias obtenidas, para que la comparación pueda realizarse con mayor claridad. Este proceso de degradación más lenta del método de gradiente discreto, cuando el tamaño de paso se aumenta, sugiere que los métodos diseñados para preservar el comportamiento cualitativo del sistema original también poseen propiedades favorables desde el punto de vista cuantitativo.

Por otra parte, con el fin de estimar la relación e influencia del tamaño de paso de discretización respecto al error local producido, tanto por el método de Euler como por el método de gradiente discreto, se realizan otro grupo de simulaciones pero esta vez considerando el sistema con $n = 30$. El resto del contexto experimental se mantiene, es decir, el punto crítico se escoge aleatoriamente dentro del hipercubo de definición y el punto inicial en una vecindad de éste, mientras que los elementos de la matriz de autopesos W se escogen aleatoriamente respetando su simetría y la ausencia de autopesos. El error se estima con el procedimiento usual [18], que se resume en

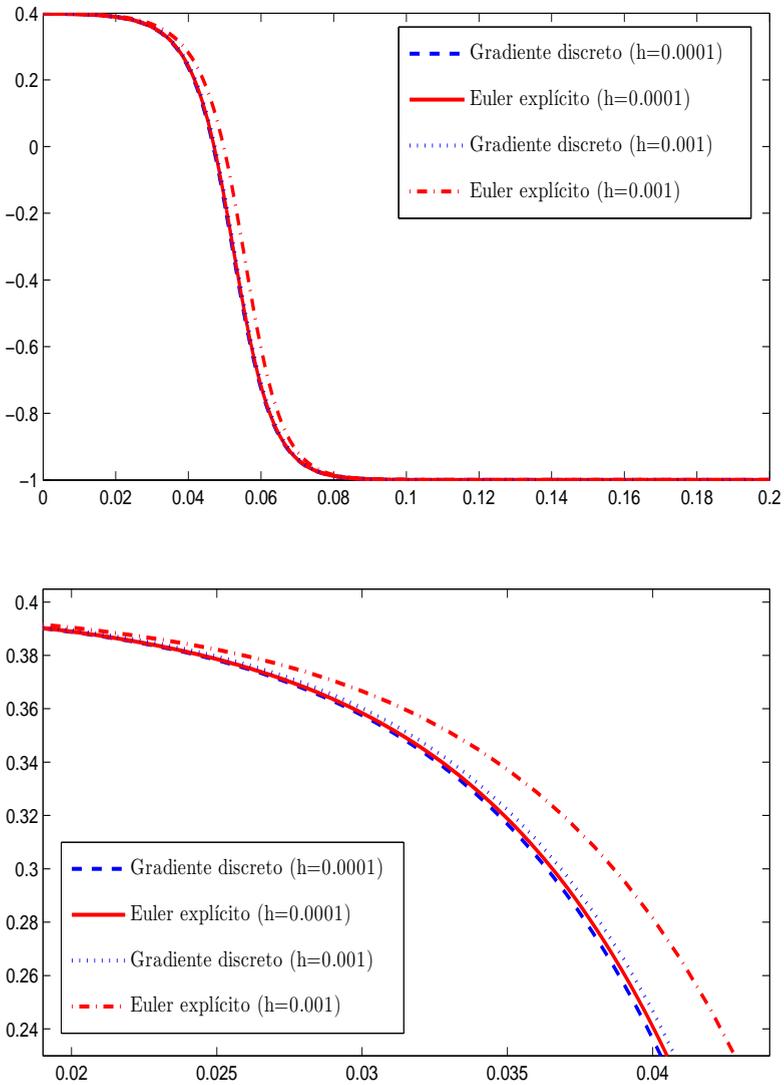


Figura 5.5: Primera componente de la trayectoria de la solución de la red de Hopfield de dimensión $n = 3$ calculada por ambos métodos con diferentes tamaños de paso (figura superior). Vista ampliada de la misma trayectoria (figura inferior).

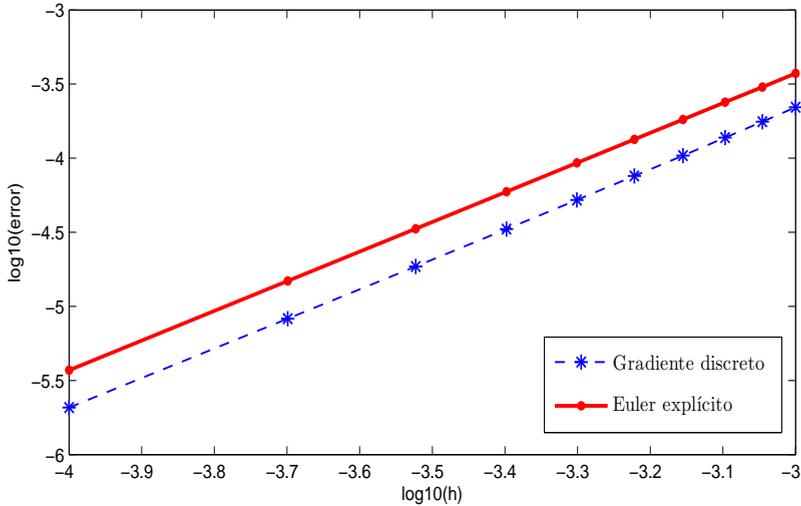


Figura 5.6: Error local de la red de Hopfield de dimensión $n = 30$ para ambos métodos con diferentes tamaños de paso.

restar la solución obtenida mediante un método de orden n de la proporcionada por un método de orden al menos $n + 1$. En este caso, se analizan el método de Euler y el método de gradiente discreto, que son ambos de orden uno. Por tanto, se comparan las soluciones obtenidas con la que resulta al discretizar el sistema utilizando un método de Runge-Kutta de segundo orden. De esta manera, se obtiene una aproximación del error local, tanto para el método de Euler como para el de gradiente discreto. Los resultados se muestran, en escala logarítmica, en la figura 5.6 evidenciando que el error local producido por el método de gradiente discreto es menor que el que se obtiene al discretizar el sistema mediante el método de Euler. En efecto, se aprecia que la constante de error del método de gradiente discreto es un orden de magnitud menor que la de la regla de Euler, tal como se había demostrado en la sección 3.5. Además, el hecho de que la dependencia entre el error local y el tamaño de paso sea lineal en la escala logarítmica, corrobora el resultado obtenido también en la sección 3.5: *el método de gradiente discreto propuesto es de primer orden.*

Una vez validado el funcionamiento y superioridad del método de gradiente discreto propuesto respecto a la regla de Euler, tanto en su capacidad para conservar la función de Lyapunov del sistema continuo como al brindar una mejor aproximación de las soluciones, se pone a prueba su eficiencia comparando con algunos métodos de Runge-Kutta disponibles en la biblioteca de *Matlab* [54].

5.1.3. Comparación con métodos de Runge-Kutta

Es bien sabida la eficiencia de los esquemas de Runge-Kutta en la resolución de sistemas de ecuaciones diferenciales. Existe una gran variedad de estos esquemas, incluyendo formulaciones explícitas, implícitas, de paso fijo, de paso variable, etc., como se vio en la sección 1.3.5. Puede decirse que las implementaciones comerciales de los métodos de Runge-Kutta constituyen el estándar a batir por nuevas propuestas, y por esta razón parece conveniente utilizarlos como referencia de comparación frente al método de gradiente discreto aquí expuesto.

Ahora bien, el método de gradiente discreto es un método explícito de paso fijo, para el que no se ha diseñado ninguna estrategia adaptativa de paso variable como mecanismo de control de error. Por su parte, los métodos de Runge-Kutta poseen sofisticados métodos de análisis de control de error, orden y adaptación del tamaño de paso, entre otros. Surge entonces de manera natural la cuestión de cómo realizar una comparación equitativa entre ambos métodos, que permita validar la eficiencia del método de gradiente discreto. Siguiendo la documentación de *Matlab*, la cual aconseja no manipular directamente el tamaño de paso, se configuran los esquemas de Runge-Kutta escogidos para el experimento fijando la tolerancia absoluta (*tol*) máxima permitida. Esta se puede considerar comparable al error local de un método de paso fijo. Entonces, el tamaño de paso del método de gradiente discreto se ajusta como $h = \sqrt{\frac{tol}{105}}$, ya que de la figura 5.6 se infiere que el error local es aproximadamente:

$$y(t+h) - z = 105 h^2 \quad (5.1)$$

En resumen, la idea es realizar la comparación, en términos de coste y fiabilidad, entre métodos que cumplen los requisitos de exactitud impuestos, enfoque que está

Tabla 5.1: Comparación del método de gradiente discreto con varios métodos de Runge-Kutta con $t \in [0, 0,5]$. DG: método de gradiente discreto; *ode45*, *ode23*, *ode23s*: métodos de *Matlab*.

Tolerancia	Método	Tiempo (seg)	máx ΔV	% pasos $\Delta V > 0$
10^{-2}	DG	0.03	$-1.85 \cdot 10^{-4}$	0.00 %
	<i>ode45</i>	0.41	7.81	31.75 %
	<i>ode23</i>	0.26	9.93	38.50 %
	<i>ode23s</i>	0.47	$1.44 \cdot 10^{-2}$	6.45 %
10^{-3}	DG	0.08	$-4.25 \cdot 10^{-7}$	0.00 %
	<i>ode45</i>	0.41	$6.54 \cdot 10^{-1}$	24.82 %
	<i>ode23</i>	0.26	$9.94 \cdot 10^{-1}$	26.09 %
	<i>ode23s</i>	0.53	$5.97 \cdot 10^{-4}$	3.38 %
10^{-4}	DG	0.24	$-1.14 \cdot 10^{-7}$	0.00 %
	<i>ode45</i>	0.42	$6.59 \cdot 10^{-1}$	24.82 %
	<i>ode23</i>	0.28	$9.94 \cdot 10^{-1}$	25.00 %
	<i>ode23s</i>	0.60	$1.78 \cdot 10^{-3}$	2.44 %
10^{-5}	DG	1.00	$-3.37 \cdot 10^{-8}$	0.00 %
	<i>ode45</i>	0.41	$6.59 \cdot 10^{-1}$	24.82 %
	<i>ode23</i>	0.27	$9.94 \cdot 10^{-1}$	24.87 %
	<i>ode23s</i>	0.58	$1.34 \cdot 10^{-3}$	2.22 %

sustentado en las principales referencias consultadas sobre el tema [27].

Los métodos de Runge-Kutta (ver [48] y las referencias que contiene) seleccionados para comparar con el método de gradiente discreto propuesto fueron:

- ***ode45***: esquema explícito de Runge-Kutta de órdenes 4 y 5, diseñado por Dormand y Prince.
- ***ode23***: esquema explícito de Runge-Kutta de órdenes 2 y 3, diseñado por Bogacki y Shampine. Puede ser más eficiente que el *ode45* para tolerancias de error bajas y en presencia de problemas moderadamente *stiff*.
- ***ode23s***: se basa en una fórmula modificada del método de Rosenbrock de orden 2 para ecuaciones *stiff*.

En estos experimentos también se trabaja con el sistema de Hopfield de dimensión $n = 30$, y la matriz de los autopesos se obtiene aleatoriamente para luego multiplicar uno de sus elementos por una constante positiva suficientemente grande, de manera tal que el jacobiano del sistema tenga valores propios diferentes. De esta manera se garantiza que el sistema sea moderadamente *stiff*. Los resultados de las simulaciones se muestran en la tabla 5.1, que recoge el valor de la tolerancia utilizado en cada caso, el tiempo en segundos que demora cada método en converger, el máximo valor que alcanza el incremento de la función de Lyapunov ($\max \Delta V$) y el porcentaje de pasos en los que el incremento de la función de Lyapunov es positivo ($\% \text{ pasos } \Delta V > 0$). Este último valor refleja la tasa de ocasiones en las que cada método deja de conservar la función de Lyapunov del sistema continuo, ya que al ser V decreciente a lo largo de las trayectorias, por ser función de Lyapunov del sistema, su incremento ΔV tendría que ser negativo en todo momento. Teniendo todo esto en cuenta, se puede concluir que el indicador fundamental de fiabilidad de los métodos es el incremento máximo de la función de Lyapunov $\max \Delta V$, y cualquier valor positivo indicaría la destrucción de la estabilidad del sistema.

Se puede apreciar en los resultados que, para todos los valores de la tolerancia analizados, los métodos de Runge-Kutta fallan a la hora de conservar la función de Lyapunov del sistema continuo, incluso cuando $tol = 10^{-5}$, lo que implica que el paso de discretización es del orden de 10^{-4} , según la relación establecida entre ellos. En otras palabras, incluso para tamaños de paso suficientemente pequeños, los métodos de Runge-Kutta no brindan una solución cualitativamente correcta, o sea, las soluciones obtenidas no conservan la función de Lyapunov del sistema. Sin embargo, se puede diferenciar el comportamiento del método *ode23s* del resto. Por ejemplo, cuando la tolerancia toma el valor $tol = 10^{-2}$, los métodos explícitos *ode45* y *ode23* fallan a la hora de conservar la función de Lyapunov, ya que se aprecian considerables incrementos positivos de V . Por su parte, el método *ode23s* falla por un margen muy pequeño. No es de extrañar que este método funcione mejor que los esquemas explícitos, ya que está diseñado para la resolución de ecuaciones *stiff*, y se conocen las ventajas en este sentido de los métodos implícitos. De la teoría del análisis numérico se sabe que los métodos explícitos usualmente no funcionan correctamente al discre-

tizar sistemas *stiff*, ya que cualquiera que sea la condición inicial con la que se trabaje, se ven obligados a disminuir excesivamente el tamaño de paso en orden a obtener la solución del sistema. Por otra parte, cuando se observa el porcentaje de pasos en los que la función de Lyapunov aumenta ($\% \text{ pasos } \Delta V > 0$), el método *ode23s* falla con una frecuencia limitada, pero significativa, mientras que los métodos explícitos fallan aproximadamente un tercio de los pasos. En cambio, el método de gradiente discreto obtiene la solución del sistema, a la vez que conserva la función de Lyapunov del sistema continuo, cualquiera sea el tamaño de paso de discretización. Es sorprendente el hecho de que, al imponer valores de tolerancia más estrictos, los resultados se mantienen. Por ejemplo, para $tol = 10^{-5}$, todos los métodos de Runge-Kutta analizados fallan en conservar la función de Lyapunov. También se puede concluir de los experimentos que para cualquier valor de la tolerancia comprendido en el rango analizado, el método de gradiente discreto es al menos un orden de magnitud más rápido que los métodos de Runge-Kutta. De esta manera queda justificada la elección del método de gradiente discreto sobre los esquemas de Runge-Kutta analizados, a pesar de ser un método de primer orden. En la tabla 5.1 se muestran los resultados para un determinado juego de valores de los parámetros, pero es válido señalar que los resultados se repiten de forma consistente para una amplia gama de valores de W , b , y_0 , al igual que en los experimentos realizados con la regla de Euler.

5.1.4. Comparación con el método de proyección

Finalmente, se compara el método propuesto con el método de proyección que se introduce en la sección 2.4. Como se vio anteriormente, este método también está diseñado con el objetivo de conservar la función de Lyapunov del sistema continuo. Como método de Runge-Kutta básico para la construcción del método de proyección, se utilizó el esquema de Fehlberg 4(5) considerando la salida densa del método de orden 5.

Los experimentos muestran que, tomando valores pequeños del tamaño de paso, por ejemplo, valores entre $h = 10^{-5}$ y $h = 10^{-3}$, el método de gradiente discreto y el método de proyección se comportan correctamente. En otras palabras, no solo

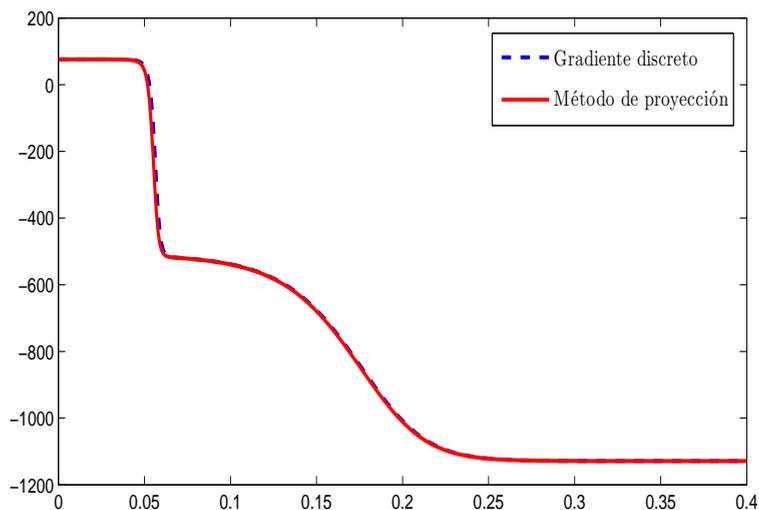
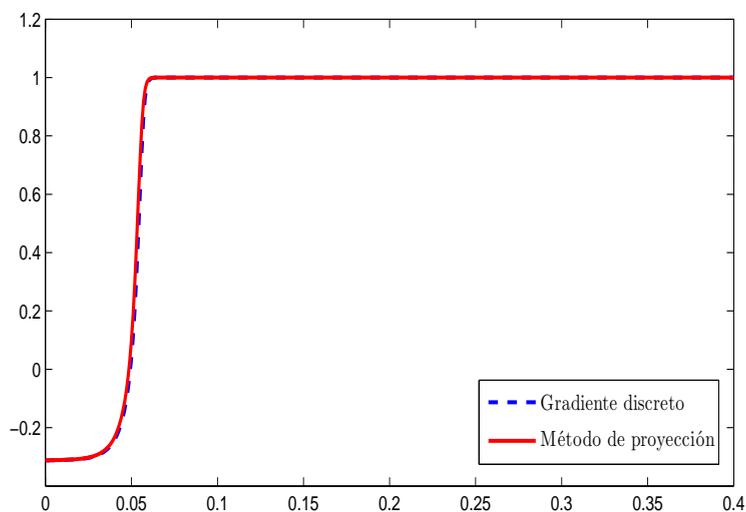


Figura 5.7: Componente y_2 de la solución numérica obtenida por el método de gradiente discreto y el método de proyección para $h = 10^{-3}$ (figura superior). Función de Lyapunov V obtenida por ambos métodos también para $h = 10^{-3}$ (figura inferior).

obtienen soluciones cualitativamente similares, sino que dichas soluciones conservan la función de Lyapunov del sistema continuo, como se muestra en la figura 5.7. A pesar de tener un comportamiento similar, es válido señalar que en estos experimentos, el coste computacional del método de proyección es significativamente superior al del método de gradiente discreto. No obstante esta observación, hay que aclarar que una comparación rigurosa en términos de tiempo de CPU requeriría una cuidadosa programación con una mejora del rendimiento. Sin embargo, el descubrimiento más relevante aparece cuando, al aumentar el tamaño de paso, por ejemplo hasta $h = 10^{-2}$, el método de proyección no encuentra solución. Por su parte, el método de gradiente continúa funcionando satisfactoriamente, incluso para tamaños de paso aún mayores, como se muestra en la tabla 5.1.

Para interpretar la razón por la que el método de proyección fracasa completamente en proporcionar una solución del sistema, se realizó un análisis detallado del algoritmo. El punto crítico es la resolución de la ecuación no lineal que define la proyección en cada paso de discretización. En la sección 2.4 se vio que en la tercera etapa del algoritmo que define el método de proyección, se proyecta la aproximación de la solución \tilde{y}_{n+1} calculada en el primer paso, sobre una variedad a la que se denota con la letra \mathcal{M} . Esta variedad está diseñada de manera tal que, mientras la solución se mantenga sobre ella, el método conservará la función de Lyapunov del sistema. Ahora bien, para obtener la proyección de la solución sobre \mathcal{M} , es necesario obtener un valor de λ que satisfaga la siguiente ecuación:

$$V(\tilde{y}_{n+1} + \lambda \nabla g(\tilde{y}_{n+1})) - V_{n+1} = 0$$

Para solucionar esta ecuación podría utilizarse cualquier método iterativo, por ejemplo el método de Newton. En este caso se recurrió a la implementación que proporciona la función *fzero* de *Matlab*. El análisis de los experimentos mostró que este método puede fallar en tres puntos:

1. En los experimentos realizados se observa que es necesario ajustar cuidadosamente el valor inicial de esta ecuación, ya que el algoritmo no converge para algunos valores, por ejemplo con la elección obvia $\lambda_0 = 0$.

2. En algunos casos no fue posible encontrar un valor inicial que convergiera a una proyección válida.
3. En otras ocasiones, el método iterativo produce un desbordamiento tras calcular algunas aproximaciones, sin llegar a converger.

Si se tiene en cuenta que resolver esta ecuación no lineal equivale a proyectar sobre la variedad \mathcal{M} , y que es precisamente esta proyección lo que garantiza que se conserve la función de Lyapunov, la falta de convergencia del método de solución se vislumbra como un gran inconveniente del método de proyección.

A modo de resumen, se puede decir que las dos últimas situaciones del párrafo anterior sugieren que, en ocasiones, la variedad \mathcal{M} podría no existir, impidiendo el buen funcionamiento de este método. Esto no contradice los resultados expuestos en [12], donde se demuestra que existe un h suficientemente pequeño para el cual la proyección será válida. Sin embargo, el resultado teórico, al ser no constructivo, no evita que, en la práctica, pueda fijarse un tamaño de paso para el que no exista la proyección. Este hecho evidencia una deficiencia del método de proyección respecto al de gradiente discreto, ya que este último encuentra solución a la vez que conserva la función de Lyapunov del sistema continuo, cualquiera sea el tamaño de paso escogido.

5.2. Método de composición

Esta sección está dedicada a la validación de la extensión de alto orden del método de gradiente discreto propuesto. Como se vio en el capítulo 3, los métodos de composición usualmente se utilizan con el objetivo de aumentar el orden de convergencia de un determinado método numérico, a la vez que se conservan ciertas propiedades cualitativas del método original. Recordemos que en el capítulo anterior se propone un método de gradiente discreto para la integración de las redes de Hopfield, asumiendo la ausencia de autopesos. El método de gradiente discreto obtenido para la discretización de este sistema se define según la ecuación (4.12), es de primer orden y está diseñado de manera tal que al discretizar el sistema se conserva su función de

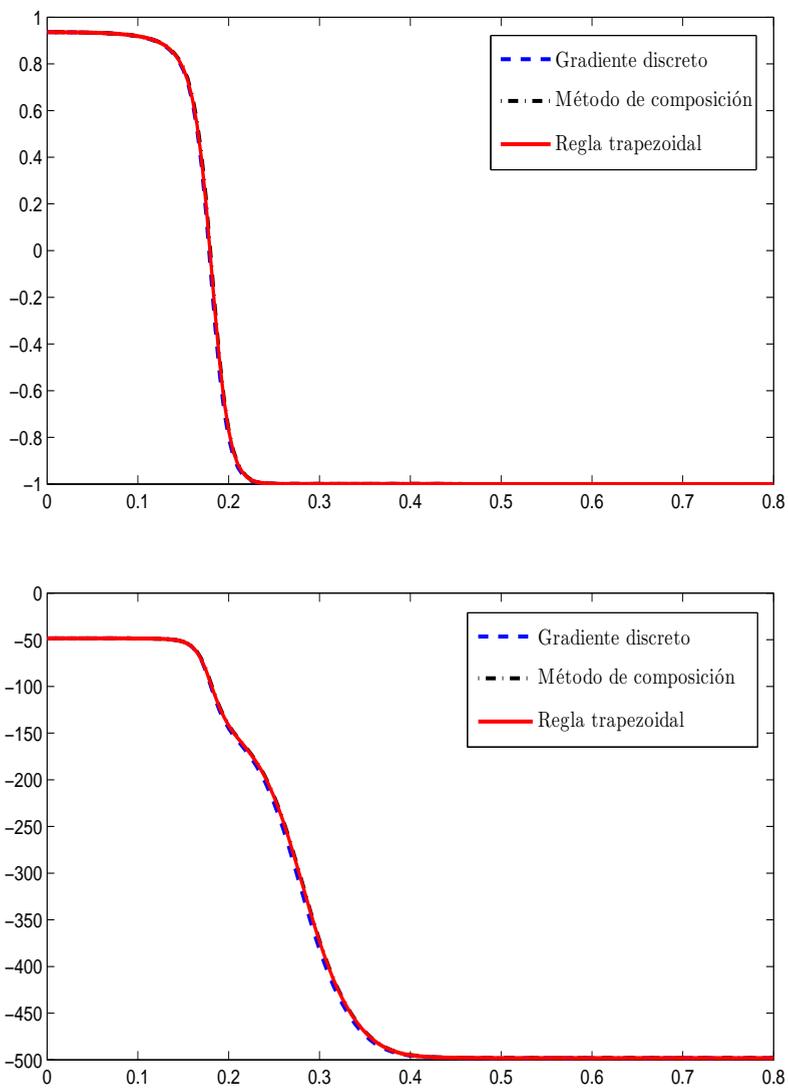


Figura 5.8: Aproximación numérica de una de las componentes de la solución obtenida por cada método para $h = 5 \cdot 10^{-3}$ (figura superior). Función de Lyapunov V obtenida por cada método también para $h = 5 \cdot 10^{-3}$ (figura inferior).

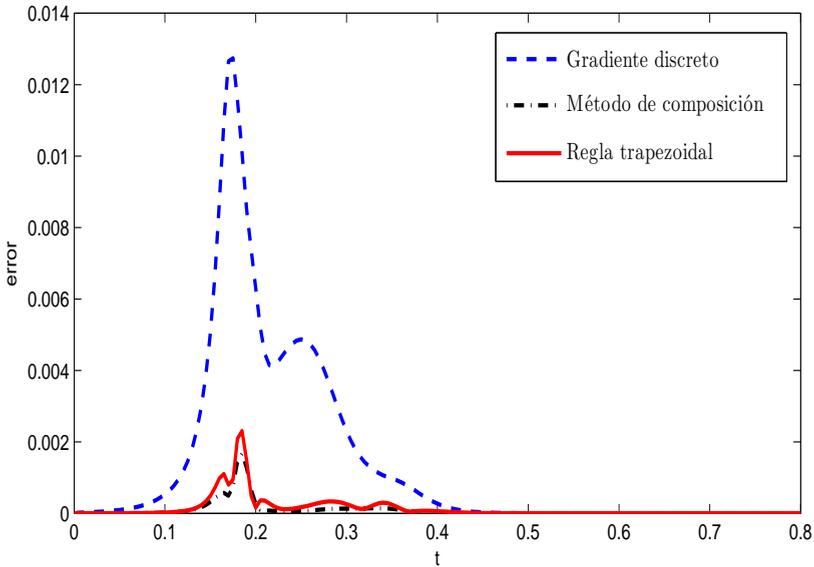


Figura 5.9: Error local para cada uno de los métodos, al calcular la solución aproximada para $h = 5 \cdot 10^{-3}$.

Lyapunov, cualquiera sea el tamaño de paso.

A partir de este método, y tomando en consideración que la composición de un método de paso simple de primer orden con su adjunto da lugar a un método simétrico de segundo orden, se propuso en el capítulo 4 un método de composición, que es el que se estudia en esta sección. Específicamente se considera el método:

$$\Psi_h = \Phi_{h/2} \circ \Phi_{h/2}^*$$

siendo Φ_h el método de gradiente discreto propuesto en la ecuación (4.12) y Φ_h^* su adjunto, cuya expresión viene dada en la ecuación (4.22). Se puede afirmar entonces que este es también un método de gradiente discreto, por lo que conserva la función de Lyapunov bajo discretización, es explícito y de segundo orden.

Los experimentos numéricos se diseñaron con dos enfoques diferentes:

1. *Cualitativamente*: comparación con el método de gradiente discreto y con la regla trapezoidal, para probar la validez, en principio, de la propuesta.
2. *Cuantitativamente*: comparación con el método de gradiente discreto de primer orden, con el objetivo de mostrar las mejoras que aporta esta generalización.

En ambos enfoques se utilizan como caso de estudio las redes de Hopfield, asumiendo la ausencia de autopesos y la dimensión $n = 30$. El resto de los parámetros, entiéndase W , b , y_0 , se escogen de manera similar a los casos mostrados en la sección anterior.

El primer grupo de simulaciones se diseñó como prueba de concepto, con el objetivo de mostrar el funcionamiento del método de composición propuesto. Para esto se compara con el método de gradiente discreto de primer orden y con la regla trapezoidal. La elección de la regla trapezoidal resultó ser la elección más natural, teniendo en cuenta que es un método explícito y de segundo orden, al igual que el método de composición. Los resultados se muestran en las figuras 5.8 y 5.9. En la primera figura se muestra la aproximación numérica de una de las componentes de la solución del sistema obtenida mediante cada uno de los métodos. Se muestra solo una componente porque la información obtenida cualquiera sea la componente analizada es similar. Los tres métodos muestran comportamientos cualitativos similares, tanto a la hora de obtener una aproximación de la solución válida, como al conservar la función de Lyapunov del sistema. La conservación de la función de Lyapunov se observa en la segunda gráfica de la figura. Todos los resultados mostrados hasta ahora, son el resultado de discretizar el sistema utilizando $h = 10^{-3}$ como paso de discretización. Cuando se incrementa levemente el tamaño de paso, por ejemplo a $h = 5 \cdot 10^{-3}$, la regla trapezoidal deja de mostrar un comportamiento correcto, ya que en la mayoría de los casos no es capaz de integrar el sistema. De esta manera, se sustenta una vez más la elección de los métodos de integración geométrica sobre otros métodos tradicionales, en este caso, la regla trapezoidal. Por su parte, en la figura 5.9 se muestra el comportamiento del error local respecto al tiempo, también para $h = 10^{-3}$. El resultado de este experimento permite afirmar que, incluso en los casos en los que tanto los métodos de gradiente discreto como la regla trapezoidal brindan soluciones cualitativamente similares a la vez que conservan la función de Lyapunov, es el método

de composición el que minimiza el error local en el proceso. Era de esperar que este método mejorara la precisión respecto al de gradiente discreto, ya que es de un orden superior, pero es incluso mejor que la regla trapezoidal, que es de su mismo orden de convergencia.

El segundo grupo de experimentos fue diseñado con el objetivo de comparar el funcionamiento de ambos métodos de gradiente discreto y está motivado, en cierta manera, teniendo en cuenta los resultados del experimento anterior. El objetivo es mostrar otras ventajas que sustenten la elección del método de composición sobre el de gradiente discreto de primer orden. La pregunta fundamental a responder es si la extensión de alto orden del método de gradiente discreto de primer orden, además de proveer mejoras en cuanto a su precisión y capacidad de conservación de la función de Lyapunov del sistema, funciona mejor a la hora de conservar otras propiedades cualitativas del sistema, como por ejemplo, su dominio de definición. En la primera gráfica de la figura 5.10 se muestra la aproximación numérica de una de las componentes de la solución obtenida mediante ambos métodos de gradiente discreto para $h = 10^{-2}$. Por su parte, en la segunda gráfica se observa el comportamiento de la función de Lyapunov V , también para $h = 10^{-2}$, evidenciando la capacidad de ambos métodos para conservar V , obteniéndose soluciones similares. No obstante, en la figura 5.11 se observa que el error local para el método de gradiente discreto es considerablemente mayor que el obtenido mediante el método de composición. A simple vista, estos experimentos no brindan ningún elemento nuevo a nuestro análisis pero nos sirven de punto de partida para el próximo grupo de simulaciones. A continuación, se aumenta el tamaño de paso a $h = 3 \cdot 10^{-2}$ y en la primera gráfica de la figura 5.12 se observa que se continúan obteniendo soluciones cualitativamente similares para ambos métodos. En cambio, si se observa el comportamiento de la función de Lyapunov en la segunda gráfica de esta figura, se aprecia un leve incremento de esta, para la solución obtenida mediante el método de gradiente discreto. En otras palabras, el método de gradiente discreto deja de conservar la función de Lyapunov. Dado que la teoría garantiza que un método de gradiente discreto conserva la función de Lyapunov, este comportamiento puede calificarse de sorprendente y merece un análisis detallado. La revisión del experimento confirmó que la solución

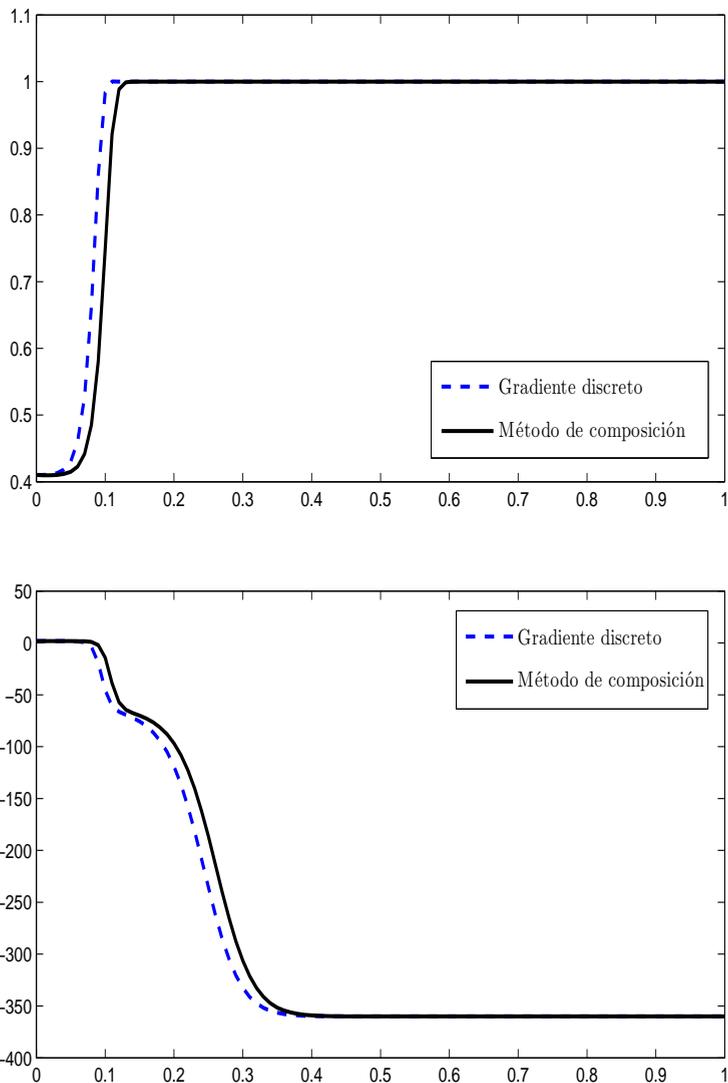


Figura 5.10: Aproximación numérica de una de las componentes de la solución obtenida mediante ambos métodos de gradiente discreto para $h = 10^{-2}$ (figura superior). Función de Lyapunov V obtenida mediante ambos métodos de gradiente discreto también para $h = 10^{-2}$ (figura inferior).

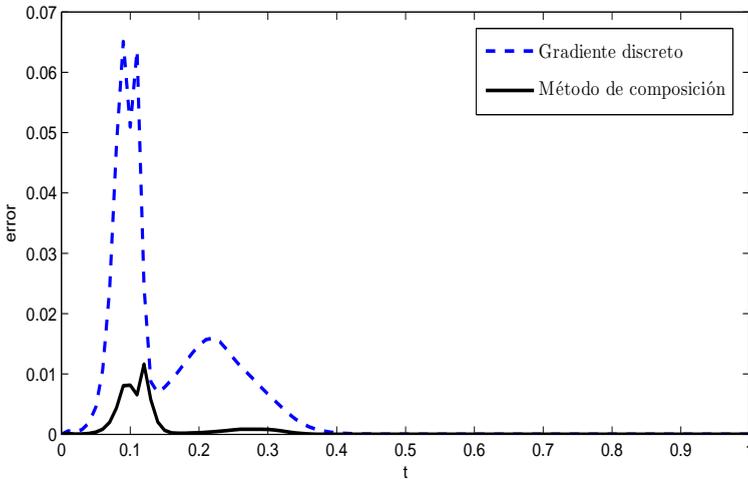


Figura 5.11: Error local para cada uno de los métodos, al calcular la solución aproximada para $h = 10^{-2}$.

obtenida por el método de gradiente discreto se sale del dominio de definición de la red continua, que como se sabe es el hipercubo $(-1, 1)$. Resumiendo, el fallo del método de gradiente no tiene que ver con la conservación de la función de Lyapunov del sistema, sino con la conservación de su dominio de definición. En cambio, el método de composición continúa brindando soluciones cualitativamente correctas, a la vez que minimiza el error local, como se puede observar en la figura 5.13. El establecimiento de una fundamentación teórica de estos resultados es un punto que merece investigación ulterior.

5.3. Conclusiones

Este capítulo se ha dedicado a validar el funcionamiento del método de gradiente discreto de primer orden propuesto y de su generalización a alto orden. La validación se realizó al compararlo con la regla de Euler, algunos esquemas comerciales de Runge-Kutta y el método de proyección propuesto en [12]. Este último método,

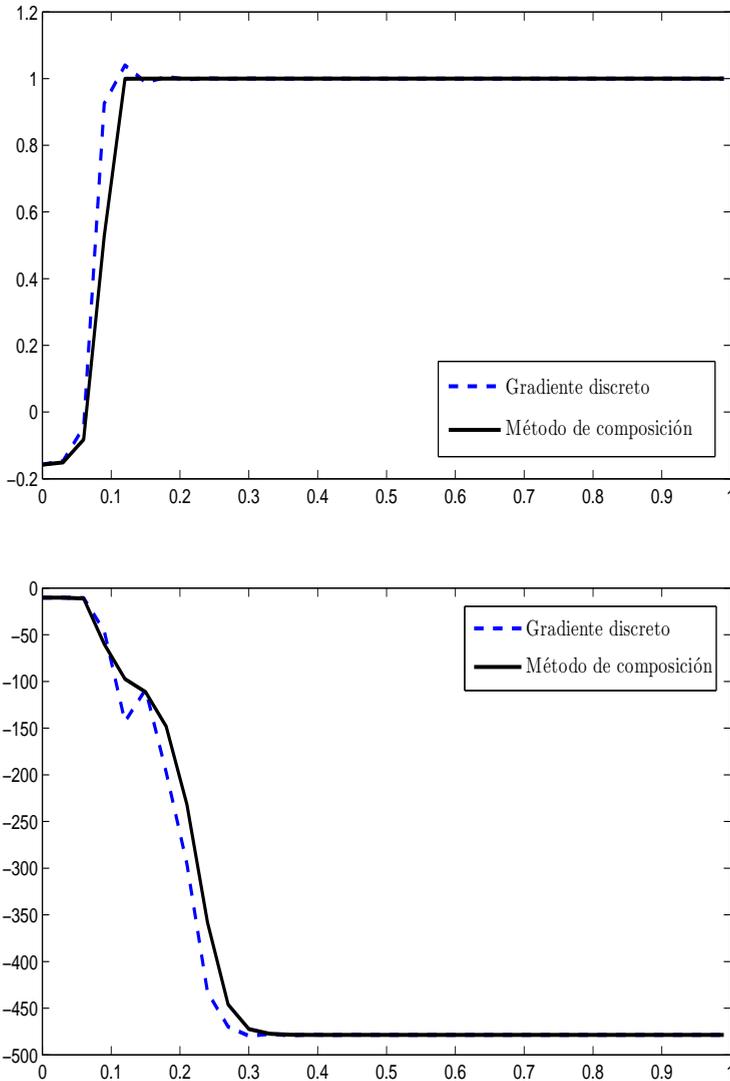


Figura 5.12: Aproximación numérica de una de las componentes de la solución obtenida mediante ambos métodos de gradiente discreto para $h = 3 \cdot 10^{-2}$ (figura superior). Función de Lyapunov V obtenida mediante ambos métodos de gradiente discreto también para $h = 3 \cdot 10^{-2}$ (figura inferior).

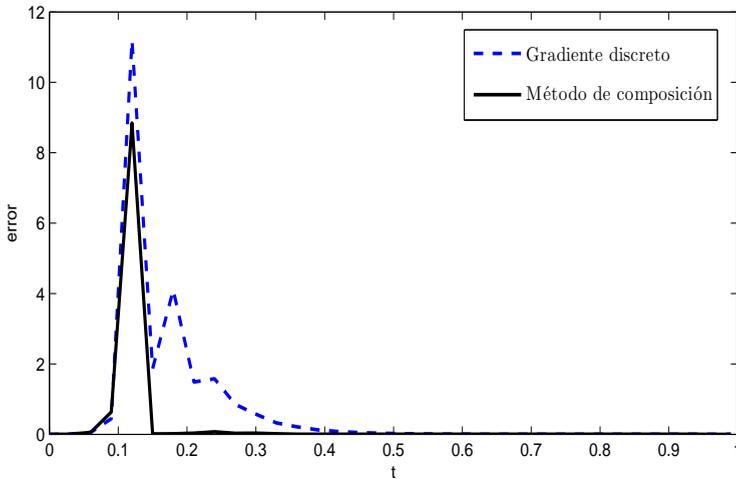


Figura 5.13: Error local para cada uno de los métodos, al calcular la solución aproximada para $h = 3 \cdot 10^{-2}$.

como se mostró en el capítulo 2, fue diseñado concretamente para garantizar la conservación de la función de Lyapunov del sistema continuo.

Al comparar con la regla de Euler, se observó que, para tamaños de paso suficientemente pequeños, el método de gradiente discreto funciona de manera similar a la regla de Euler. Sin embargo, al incrementar levemente el tamaño de paso, la regla de Euler muestra un comportamiento totalmente erróneo, apareciendo incluso soluciones periódicas, las cuales destruyen totalmente la estructura de gradiente del sistema continuo. Otra forma de ver este resultado es que la regla de Euler es incapaz de conservar la función de Lyapunov del sistema, siendo visibles puntos en los que se aprecia un aumento considerable. Por el contrario, el método de gradiente discreto continúa funcionando correctamente.

Por su parte, al comparar el funcionamiento del método de gradiente discreto con algunos esquemas comerciales de Runge-Kutta de órdenes superiores a él, se observa que, incluso para tamaños de paso pequeños como $h = 10^{-5}$, los tres métodos de Runge-Kutta son incapaces de conservar la función de Lyapunov del sistema

continuo. Incluso el método implícito *ode23s*, que está diseñado para la resolución de problemas *stiff*, muestra un incremento de la función de Lyapunov, siendo el método de gradiente discreto el único en mostrar un comportamiento cualitativamente correcto: conservación de la función de Lyapunov y del dominio de definición.

Por último, al comparar con el método de proyección se obtienen comportamientos similares de ambos métodos, siempre y cuando el paso de discretización sea menor o igual que 10^{-3} . Para tamaños de paso superiores a este valor, el método de proyección es incapaz de encontrar una solución del sistema que, a su vez, conserve la función de Lyapunov. Un análisis detallado de las posibles razones del fallo del método de proyección, pese a estar diseñado específicamente para garantizar la conservación de la función de Lyapunov, sugiere que en ocasiones no existe la variedad sobre la que debe proyectarse.

Finalmente, se valida el funcionamiento de la extensión de alto orden de este método, obtenida al componerlo con su adjunto. Este método, al que se denominó método de composición, se compara con el método de gradiente discreto de primer orden y con la regla trapezoidal. Los experimentos mostraron, por una parte, la incapacidad de la regla trapezoidal para integrar el sistema conservando a su vez la función de Lyapunov del sistema continuo. Por otra parte, se evidencian las mejoras que aporta esta generalización, en dos sentidos: el incremento del orden de convergencia, lo cual resultaba evidente, al ser este un método de segundo orden por construcción; y la capacidad para conservar las propiedades cualitativas del sistema continuo. Se observa que, incluso cuando el método de gradiente discreto falla en conservar el dominio de definición del sistema y por ende su función de Lyapunov, el método de composición continúa mostrando un comportamiento cualitativo correcto.

De esta manera, el conjunto de experimentos numéricos mostrados, pretende señalar la utilización, tanto del método de gradiente discreto como de su extensión, como elecciones válidas a la hora de discretizar sistemas de gradiente con función de Lyapunov.

Capítulo 6

Conclusiones

La presente tesis se ha desarrollado en el marco de la integración geométrica, prestando especial interés a la conservación, bajo discretización, de la función de Lyapunov de sistemas de gradiente. Se centra la atención en los métodos de gradiente discreto, trabajando principalmente con el gradiente discreto del incremento de las coordenadas.

En el primer capítulo, además de hacer una breve introducción al tema central de la tesis, se recopila un gran número de resultados de la literatura referentes al estudio de las ecuaciones diferenciales ordinarias (EDOs), vistas como un caso particular de los sistemas dinámicos y de la teoría de análisis numérico, con vistas a que la tesis fuese autocontenida. Se presta especial atención a los resultados referentes al análisis de la estabilidad, tanto de los sistemas lineales como de los no lineales, listando resultados que garantizan un comportamiento similar entre las órbitas de un sistema no lineal y su linealización bajo ciertas condiciones. Se introduce además el concepto de función de Lyapunov, mostrando su existencia como un método alternativo para el análisis de la estabilidad del sistema dado, al menos en un determinado entorno. Por otra parte, se establece la correspondencia entre los sistemas con función de Lyapunov y los sistemas de gradiente. Se recuerdan además algunas clasificaciones básicas de los métodos numéricos: explícitos, implícitos, paso simple y multipaso, así como un breve resumen de sus principales propiedades: consistencia, estabilidad y orden

de convergencia. Finalmente, se puntualiza la incapacidad de los métodos numéricos clásicos para conservar las propiedades cualitativas del sistema continuo, siendo este el punto de partida para el capítulo 2 de la tesis, que se dedica a los métodos de integración geométrica.

En el capítulo 2 se introduce la idea del diseño de métodos de integración numérica geométrica, centrando la atención en los métodos diseñados con el objetivo de conservar la función de Lyapunov del sistema continuo y con ella su estabilidad. Se muestra la composición de métodos numéricos como una vía alternativa para incrementar el orden de convergencia de un determinado método numérico, sin renunciar a que se conserven las propiedades del sistema que se pretende discretizar. También en este capítulo se realiza un estudio detallado de un método de proyección diseñado en [12] para la conservación de la función de Lyapunov de un determinado sistema de EDO, con vistas a comparar su funcionamiento con el de los métodos de gradiente discreto, que son el eje central del capítulo 3. Una primera contribución preliminar de la tesis, incluida en este capítulo, es el análisis de la implementación del método de proyección sobre un sistema escalar lineal, ilustrando la complejidad y las limitaciones del método incluso en esta ecuación trivial.

Teniendo en cuenta que la tesis propone el uso de métodos de gradiente discreto, se dedica el capítulo 3 a su estudio, mostrando inicialmente algunos aspectos básicos que proporcionan el fundamento al diseño de métodos de integración basados en gradientes discretos. Se dirige la atención principalmente a los métodos de gradiente discreto diseñados específicamente para la conservación de la función de Lyapunov del sistema continuo. Ya en este capítulo se muestran algunos de los resultados de la tesis:

1. Se ilustra la metodología de construcción de métodos de gradiente discreto, integrando tres sistemas escalares simples.
2. Se demuestra que, con una elección adecuada de los parámetros, el método de gradiente discreto obtenido puede reescribirse de forma explícita, para el caso de sistemas dinámicos con función de Lyapunov multilineal, siempre que se trabaje con el gradiente discreto del incremento de las coordenadas.

3. Se realiza un estudio detallado del orden de los métodos obtenidos en el caso general y en los casos particulares de los sistemas escalares simples.
4. Se validan los resultados obtenidos, realizando experimentos numéricos, que evidencian la capacidad de los métodos de gradiente discreto diseñados en cada caso para conservar la función de Lyapunov del sistema continuo, a la vez que minimizan, casi siempre, el error global.

El capítulo 4, por su parte, recoge algunos de los resultados más importantes de la tesis. En este capítulo se introduce como caso de estudio el sistema algebraico-diferencial conocido como redes de Hopfield. Este modelo se utiliza para mostrar el funcionamiento de los métodos propuestos, dando lugar a los principales aportes del capítulo, entre los que se enumeran:

1. Se diseña un método de gradiente discreto que conserva la función de Lyapunov del sistema bajo discretización, cualquiera sea el tamaño de paso.
2. Se muestra que, en este caso, es posible reescribir explícitamente el método de gradiente discreto obtenido realizando una elección adecuada de los parámetros.
3. Se obtienen condiciones sobre los parámetros que permitirían diseñar un método de gradiente discreto de segundo orden.
4. Se propone una extensión de alto orden del método obtenido, para la discretización del sistema de Hopfield considerando la ausencia de autopesos.

Los resultados teóricos obtenidos en este capítulo se validan mediante simulaciones numéricas implementadas en *Matlab* de los métodos diseñados en la tesis, siendo este el contenido del capítulo 5 de la tesis.

El capítulo 5 se dedica a la validación del funcionamiento del método de gradiente discreto de primer orden propuesto y de su generalización a alto orden. Los principales resultados son:

1. Se valida el funcionamiento del método de gradiente discreto de primer orden comparándolo con la regla de Euler, algunos esquemas comerciales de Runge-Kutta y el método de proyección propuesto en [12].
2. Se valida el funcionamiento de la extensión de alto orden de este método. En este caso, se compara con el método de gradiente discreto de primer orden y con la regla trapezoidal.
3. Teniendo en cuenta los resultados de las simulaciones en cada caso, se concluye que tanto el método de gradiente discreto como su extensión son elecciones válidas a la hora de discretizar sistemas de gradiente con función de Lyapunov.

Como se ha puntualizado en varias ocasiones durante el desarrollo de este trabajo, el diseño de métodos numéricos que conserven la función de Lyapunov dista de ser un camino poco fructífero. Por el contrario, queda mucho camino por recorrer y muchas direcciones que explorar. En la línea de esta investigación se podría profundizar en las siguientes direcciones:

1. Generalizar las condiciones bajo las cuales se obtienen métodos de gradiente discreto explícitos, a pesar de ser implícita la formulación general de estos.
2. Continuar trabajando en la búsqueda de condiciones que permitan incrementar el orden de los métodos de gradiente discreto. En esta dirección, el objetivo general sería establecer una teoría del orden que permita diseñar métodos de gradiente discreto de orden arbitrario, similar a la que se ha desarrollado para los métodos de Runge-Kutta y los métodos multipaso.
3. Respecto a los métodos de proyección, sería interesante proponer diferentes direcciones de proyección, a la vez que se imponen condiciones sobre el tamaño de paso para garantizar que el método esté bien definido y conserve la función de Lyapunov del sistema continuo.
4. Aplicar los resultados obtenidos a otros sistemas, más allá de las redes de Hopfield, con objeto de probar la generalidad de la metodología obtenida.

Bibliografía

- [1] S. Abe. Theories on the Hopfield Neural Networks. In *Proc. IEE International Joint Conference on Neural Networks*, volume I, pages 557–564, 1989.
- [2] V. I. Arnold. *Mathematical methods of classical mechanics*. Number 60 in Graduate texts in mathematics. Springer, New York, 2nd ed edition, 1997.
- [3] M. Atencia. *Estabilidad en Sistemas Neuronales Realimentados. Aplicación al Control*. PhD thesis, Departamento de Tecnología Electrónica, Universidad de Málaga, 2004.
- [4] M. Atencia, G. Joya, and F. Sandoval. Continuous-state hopfield dynamics based on implicit numerical methods. In D. J., editor, *Artificial Neural Networks ICANN 2002*, volume 2415 of *Lecture Notes in Computer Science*, pages 794–794. Springer Berlin / Heidelberg, 2002. 10.1007/3-540-46084-5_220.
- [5] M. Atencia, G. Joya, and F. Sandoval. Dynamical Analysis of Continuous Higher Order Hopfield Networks for Combinatorial Optimization. *Neural Computation*, 17(8):1802–1819, 2005.
- [6] M. Atencia, G. Joya, and F. Sandoval. Hopfield neural networks for parametric identification of dynamical systems. *Neural Processing Letters*, 21(2):143–152, 2005.
- [7] M. Atencia, G. Joya, and F. Sandoval. Hopfield Neural Networks for Parametric Identification of Dynamical Systems. *Neural Processing Letters*, 21(2):143–152, 2005.

- [8] T. Bárta, R. Chill, and E. Fašangová. Every ordinary differential equation with a strict Lyapunov function is a gradient system. *Monatshefte für Mathematik*, 166(1):57–72, 2011.
- [9] K. Brenan, S. Campbell, and L. Petzold. *Numerical solution of initial-value problems in differential-algebraic equations*. SIAM, 1995.
- [10] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley and Sons, Ltd, 2008.
- [11] M. Calvo, M. P. Laburta, J. I. Montijano, and L. Rández. On the preservation of Lyapunov functions by Runge-Kutta methods. In T. E. Simos, G. Psihoyios, and C. Tsitouras, editors, *International Conference on Numerical Analysis and Applied Mathematics*, pages 735–738, 1168, 2009. AIP.
- [12] M. Calvo, M. P. Laburta, J. I. Montijano, and L. Rández. Projection methods preserving lyapunov functions. *BIT Numerical Mathematics*, 50(2):223–241, Mar. 2010.
- [13] G. Dahlquist. A special stability problem for linear multistep methods. *BIT 3*, pages 27–43, 1963.
- [14] L. Edelstein-Keshet. *Mathematical Models in Biology*. Birkhauser, New York, 1988.
- [15] O. Gonzalez. Time integration and discrete hamiltonian systems. *Journal of Nonlinear Science*, 6(5):449–467, Sept. 1996.
- [16] E. Hairer and C. Lubich. Energy-diminishing integration of gradient systems. *IMA Journal of Numerical Analysis*, Oct. 2013.
- [17] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*. Springer, Berlin, 2002.
- [18] E. Hairer, S. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff problems*. Springer, 1987.

-
- [19] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations I: nonstiff problems*. Number 8 in Springer series in computational mathematics. Springer, Heidelberg ; London, 2nd rev. ed edition, 2009.
- [20] E. Hairer and G. Wanner. *Analysis by its history*. Springer, New York, 1997.
- [21] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25(1):35–61, 1983.
- [22] S. Haykin. *Neural Networks. A Comprehensive Foundation*. Macmillan College Publishing Company, 1994.
- [23] Y. Hernández-Solano, M. Atencia, and G. Joya. Integración de sistemas dinámicos mediante métodos numéricos que conserven la función de Lyapunov. In *Congreso de la Real Sociedad Matemática Española*, 2011.
- [24] J. Hertz, A. Krogh, and R. Palmer. *Introduction to the theory of neural computation*. Addison-Wesley, 1991.
- [25] J. Hoffman. *Dreams of calculus: perspectives on mathematics education*. Springer, Berlin ; New York, 2004.
- [26] J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA*, 81:3088–3092, 1984.
- [27] T. E. Hull, W. H. Enright, B. M. Fellen, and A. E. Sedgwick. Comparing numerical methods for ordinary differential equations. *SIAM Journal on Numerical Analysis*, 9(4):603–637, Dec. 1972.
- [28] A. Iserles. *Solving linear ordinary differential equations by exponentials of iterated commutators*, volume Numer. Math. 45. Springer-Verlag, Berlin, 1984.
- [29] A. Iserles. *A first course in the numerical analysis of differential equations*. Cambridge texts in applied mathematics. Cambridge University Press, Cambridge ; New York, 2nd ed edition, 2009.
-

- [30] G. Joya, M. Atencia, and F. Sandoval. Associating arbitrary-order energy functions to an artificial neural network. Implications concerning the resolution of optimization problems. *Neurocomputing*, 14:139–156, 1997.
- [31] G. Joya, M. Atencia, and F. Sandoval. Hopfield neural networks for optimization: Study of the different dynamics. *Neurocomputing*, 43(1-4):219–237, 2002.
- [32] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.
- [33] E. Kostelich and D. Arnbornster. *Introductory Differential Equations, from linearity to chaos*. Massachusetts, 1996.
- [34] M. Krasnytska, B. Berche, and Y. Holovatch. Phase transitions in the potts model on complex networks. *Condensed Matter Physics*, 16(2):23602, 2013.
- [35] J. López-Marcos and J. Sanz Serna. A definition of stability for nonlinear problems. In *Numerical treatment of differential equations*, pages 216–226. Teubner, 1988.
- [36] J. C. López-Marcos and J. M. Sanz-Serna. Stability and convergence in numerical analysis III: linear investigation of nonlinear stability. *IMA Journal of Numerical Analysis*, 8(1):71–84, 1988.
- [37] R. McLachlan, G. Quispel, and N. Robidoux. Unified approach to hamiltonian systems, poisson systems, gradient systems, and systems with lyapunov functions or first integrals. *Physical Review Letters*, 81(12):2399–2403, Sept. 1998.
- [38] R. McLachlan, G. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *Philos. Trans. of the Royal Society of London Series A*, 357(1754):1021–1045, 1999.
- [39] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1754):1021 –1045, Apr. 1999.

-
- [40] R. I. McLachlan and C. Scovel. Equivariant constrained symplectic integration. *J. Nonlinear. Sci.*, 5:233–256, 1995.
- [41] G. R. W. Quispel and G. S. Turner. Discrete gradient methods for solving ODEs numerically while preserving a first integral. *Journal of Physics A: Mathematical and General*, 29(13):L341–L349, July 1996.
- [42] J. Ramos and C. García-López. Piecewise-linearized methods for initial-value problems. *Applied Mathematics and Computation*, 82(2-3):273–302, Mar. 1997.
- [43] T. Samad and P. Harper. High-order Hopfield and Tank optimization networks. *Parallel Computing*, 16:287–292, 1990.
- [44] J. Sanz Serna. Stability and convergence in numerical analysis I: Linear problems-A simple, comprehensive account. In J. Hale and P. Martinez-Amores, editors, *Nonlinear Differential Equations*, pages 64–113. Pitman, 1985.
- [45] J. M. Sanz-Serna. Symplectic integrators for hamiltonian problems: an overview. *Acta Numerica*, 1:243, 1992.
- [46] J. M. Sanz-Serna. *Diez lecciones de Cálculo Numérico*. Universidad de Valladolid, Valladolid, 1998.
- [47] J. Schropp. Conserving first integrals under discretization with variable step size integration procedures. *Journal of Computational and Applied Mathematics*, 115(1-2):503–517, Mar. 2000.
- [48] L. F. Shampine and M. W. Reichelt. The MATLAB ODE suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, Jan. 1997.
- [49] H. J. Stetter. *Analysis of Discretization Methods for Ordinary Differential Equations*. Springer-Verlag, Berlin, 1973.

- [50] J. C. Strikwerda. *Finite difference schemes and partial differential equations*. Chapman and Hall, 2004.
- [51] A. Stuart and A. Humphries. *Dynamical systems and numerical analysis*. Cambridge University Press, 1996.
- [52] M. Suzuki. Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations. *Phys. Lett. A*, 146:319–323, 1990.
- [53] D. Tank and J. Hopfield. 'Neural' Computation of Decisions in Optimization Problems. *Biological Cybernetics*, 52:141–152, 1985.
- [54] The MathWorks. *MATLAB Function Reference*.
- [55] Y. Uesaka. Mathematical Aspects of Neuro-Dynamics for Combinatorial Optimization. In T. Kohonen, K. Makisara, O. Simula, and J. Kangas, editors, *Artificial Neural Networks (ICANN 1991)*, pages 1011–1014. Elsevier, 1991.
- [56] A. van den Bos. *Parameter estimation for scientists and engineers*. Wiley-Interscience, 2007.
- [57] M. Vidyasagar. Minimum-seeking properties of analog neural networks with multilinear objective functions. *IEEE Trans. on Automatic Control*, 40(8):1359–1375, 1995.
- [58] H. Yoshida. Construction of higher order symplectic integrators. *Phys. Lett. A*, 150:262–268, 1990.