Discretization of homogeneous systems preserving Lyapunov stability

Tesis que presenta:
Ana Cristina Silva Loredo

Para obtener el grado de :
Doctora en Control y Sistemas Dinámicos

Directores de Tesis:

Dr. Tonámetl Sánchez Ramírez
Dr. David Antonio Lizárraga Navarro

Instituto Potosino de Investigación Científica y Tecnológica

División de Control y Sistemas Dinámicos

San Luis Potosí, México, 2025

Créditos institucionales

Esta tesis fue elaborada en la División de Control y Sistemas Dinámicos del Instituto Potosino de Investigación Científca y Tecnológica, A.C., bajo la codirección del Dr. Tonámetl Sánchez Ramírez y el Dr. David A. Lizárraga Navarro.

Durante la realización del trabajo la autora recibió una beca académica del Consejo Nacional de Ciencia y Tecnología de México (No. de registro: 633357/336957) y del Instituto Potosino de Investigación Científca y Tecnológica, A. C.

Constancia de aprobación de la tesis

La tesis **Discretization of homogeneous systems preserving Lyapunov stability**, presentada para obtener el Grado de **Doctora en Control y Sistemas Dinámicos**, fue elaborada por **Ana Cristina Silva Loredo** y aprobada el 02 de Diciembre de 2025 por los suscritos, designados por el Colegio de Profesores de la División de la División de Control y Sistemas Dinámicos del Instituto Potosino de Investigación Científca y Tecnológica, A.C.

Dr. Tonámetl Sánchez Ramírez
(Codirector de la tesis)

Dr. David A. Lizárraga Navarro
(Codirector de la tesis)

Dr. Arturo Zavala Río
(Sinodal)

Dr. Adrián René Ramírez López
(Sinodal)

Dr. Jaime Alberto Moreno Pérez
(Sinodal externo)

Agradecimientos

A todas las personas e instituciones que caminaron conmigo en este proceso, quiero expresarles mi más sincera gratitud.

En primer lugar, deseo agradecer profundamente a mis directores de tesis, el Dr. Tonámetl y el Dr. David, y también al Dr. Arturo Zavala. Gracias por creer en mí incluso en los momentos en los que yo misma dudaba. Por la confianza que depositaron en mi trabajo, su inagotable paciencia y por haberme guiado con su vasto conocimiento durante estos años. Su rigor científico ha sido un ejemplo fundamental para mi formación.

Agradezco a mis sinodales, el Dr. Jaime y el Dr. Adrián, por sus valiosas preguntas y observaciones, que enriquecieron este trabajo. Asimismo, al IPICYT, por brindarme los espacios y recursos necesarios. Extiendo este agradecimiento a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación, por la beca otorgada, esencial para mi dedicación exclusiva a este proyecto. También agradezco a la fundación Sofía Kovalevskaia, por el apoyo otorgado para la culminación de este proceso.

A mis colegas y amigos gracias por los cafés, las discusiones científicas y las risas que hicieron los días largos más llevaderos, por escuchar mis ideas, mis quejas y mis lamentos, por animarme cuando el agotamiento parecía vencerme y por recordarme que incluso los caminos más complejos se recorren mejor acompañada.

No podría cerrar estos agradecimientos sin mencionar a mi familia. Gracias por ser mi refugio seguro, por el apoyo invaluable y la comprensión que me han dado para que pudiera cumplir esta meta. Por empujarme a seguir cuando sentía mis ganas flaquear. A mi esposo, gracias por tu comprensión infinita, por acompañarme en noches interminables frente a la computadora, y por darme fuerza cuando sentía que este proceso no tenía fin.

Este doctorado fue largo, arduo y muchas veces abrumador, pero no lo recorrí sola. A todas las personas que, con un gesto, una conversación o un abrazo, ayudaron a que este sueño siguiera vivo, les doy las gracias con el corazón. A todas y todos ustedes: gracias por no dejar que me perdiera en el camino.

Acronyms and notation

Acronyms

- **CS**: Continuous Solution.
- **IVP**: Initial Value Problem.
- **LFB**: Lyapunov-Function-Based.
- **LLS**: Lyapunov Level Set.
- **ODE**: Ordinary Differential Equation.
- **RK**: Runge-Kutta methods (e.g., RK4).

Notation

- $\mathbb{R}$: The set of real numbers.
- $\mathbb{R}_{\geq 0}$: The set of non-negative real numbers $[0, \infty)$.
- $\mathbb{R}_{>0}$: The set of positive real numbers $(0, \infty)$.
- $\mathbb{N}$: The set of natural numbers $\{0, 1, 2, \dots\}$.
- $\mathbb{R}^n$: n-dimensional Euclidean space.
- $T\mathbb{R}^n$: The tangent bundle of $\mathbb{R}^n$.
- $GL(n, \mathbb{R})$: The general linear group of degree n over $\mathbb{R}$.
- $\mathcal{B}$: Unit ball in $\mathbb{R}^n$.
- $S_r(\rho)$: Homogeneous sphere of radius ρ.
- C^p: Class of functions with continuous derivatives up to order p.
- $\mathcal{K}$: Class of strictly increasing functions $\gamma : [0, a) \to [0, \infty)$ with $\gamma(0) = 0$.
- $\delta_\lambda(x)$: Standard dilation of x with scale factor λ.

- $\Delta_\lambda^r(x)$: Weighted dilation of x with weights r and scale factor λ.
- $\Phi_{x_0}^f(\cdot)$: Integral curve (flow) of vector field f starting at x_0.
- $\phi(t, x_0)$: Solution of the continuous-time system starting at x_0.
- Ψ_f: Discrete-time map approximating the flow.
- V: Lyapunov function.
- W: Function defined as $W(x) = -\frac{\partial V}{\partial x} f(x)$.
- $\mathcal{M}$: General discretization scheme.
- r: Vector of weights $r = (r_1, \ldots, r_n)$.
- μ: Degree of homogeneity of the vector field f.
- m: Degree of homogeneity of the Lyapunov function V.
- h: Discretization step size.
- σ: Order of consistency of a discretization method.
- x_k: Discrete-time approximation of the solution at time t_k.
- e_k: Global error at step k.
- T_{k+1}: Local truncation error.
- $\|\cdot\|$: Euclidean norm.
- $\|\cdot\|_r$: Weighted homogeneous norm (r-homogeneous norm).
- $\|\cdot\|_d$: Canonical homogeneous norm.
- $\lfloor x \rfloor$: Integer part of x.
- $\lceil x \rfloor^\alpha$: Signed power function, defined as $\text{sign}(x)|x|^\alpha$.
- $[e, f]$: Lie bracket of vector fields e and f.
- $L_e V$: Lie derivative of function V along vector field e.
- $co(S)$: Convex closure of a set S.
- I_n: Identity matrix of size $n \times n$.
- $diag(\cdot)$: Diagonal matrix constructed from a vector.
- δ_j^i: Kronecker delta (equal to 1 if $i = j$, 0 otherwise).

- G_d: Generator matrix of a dilation group.
- S_c: Level set of a function (typically Lyapunov), defined as $\{x : V(x) = c\}$.

Resumen

En esta tesis se estudia el problema de discretización de sistemas homogéneos. En particular, se propone una mejora a un esquema de discretización diseñado para preservar algunas propiedades importantes de los sistemas homogéneos asintóticamente estables. Como resultado, se obtiene una familia de esquemas de discretización que preserva las mismas propiedades de los sistemas homogéneos, pero que puede llegar a tener un costo computacional menor. El esquema del cual se partió preserva una función de Lyapunov asociada al sistema, así como la estabilidad del origen y el tipo de convergencia de las soluciones. En cuanto a sus propiedades numéricas, ya se había demostrado que es consistente de orden uno, pero no se había probado su convergencia. La familia de esquemas de discretización que proponemos sigue preservando la estabilidad del origen, una función de Lyapunov y el tipo de convergencia de las soluciones. Se demuestra que cada miembro de esta familia es numéricamente convergente y consistente de algún orden, que depende del orden de consistencia de un método de discretización auxiliar utilizado en dos etapas del esquema, como por ejemplo: los métodos de Euler, el método del punto medio, la familia de Runge-Kutta, entre otros. El orden de consistencia de los esquemas que proponemos impacta el costo computacional de su implementación con respecto al esquema inicial: a mayor orden de consistencia, mayor será el tamaño del paso de discretización requerido para alcanzar una cierta precisión en las soluciones aproximadas. En contraste con otros esquemas de discretización, que pueden conducir a inconsistencias como divergencia de las soluciones en lugar de convergencia, o a un comportamiento oscilatorio (chattering) producido por condiciones iniciales grandes o pequeñas, o cuando el campo vectorial es discontinuo en el origen, la familia propuesta evita estos problemas porque preserva las características esenciales de los sistemas homogéneos, proyectando la dinámica del sistema sobre un conjunto de nivel definido a partir de una función de Lyapunov asociada al sistema. Además, mientras que los métodos de discretización diseñados para sistemas homogéneos existentes en la literatura típicamente imponen restricciones estrictas sobre el tamaño máximo del paso de discretización para garantizar sus propiedades, los esquemas introducidos aquí no presentan tales limitaciones, siempre que el método auxiliar utilizado satisfaga ciertas condiciones de regularidad.

Abstract

In this thesis, the problem of discretizing homogeneous systems is studied. In particular, one discretization scheme, designed for asymptotically stable homogeneous systems, is improved. The result is a family of discretization schemes that preserves the essential characteristics of homogeneous systems. The original scheme preserves a Lyapunov function associated with the system, as well as the stability of the origin and the type of convergence of the solutions. Regarding its numerical properties, it was previously shown to be consistent of order one, but its convergence was not proven. The family of discretization schemes we propose still preserves the stability of the origin, a Lyapunov function and the type of convergence of the solutions. We prove that each member of this family is numerically convergent and consistent of some order, depending on that of an auxiliary discretization method used at two stages of the scheme (e.g. Euler, midpoint, Runge-Kutta family, etc.). The consistency order of the schemes we propose impacts the computational cost of their implementation: a higher consistency order allows for a larger discretization step size to achieve a certain accuracy in the approximate solutions. In contrast to other discretization schemes that may lead to inconsistencies -such as solution divergence instead of convergence, or chattering with large/small initial conditions or when the vector field is discontinuous at the origin- the proposed family avoids these issues because it projects the system's dynamics onto a level set of an associated Lyapunov function, thus preserving the decreasing nature of the solutions. Moreover, while existing discretization methods for homogeneous systems typically impose strict limitations on the maximum step size for which their properties are guaranteed, the schemes introduced here do not have such constraints, provided that the used auxiliary method satisfies certain regularity conditions.

Contents

1 Introduction **1**

2 Preliminaries **9**
2.1 Stability for nonlinear dynamical systems 10
2.2 Homogeneous systems 11
2.3 Numerical methods for discretization of dynamical systems 15
2.3.1 Single step methods 18
2.3.2 Linear Multistep methods 21
2.3.3 Example of discretization procedure 21

3 Discretization methods for r-homogeneous systems **23**
3.1 Improved Euler's scheme with a state-dependent discretization step . . . 23
3.1.1 Explicit Euler scheme 24
3.1.2 Implicit Euler scheme 25
3.2 Consistent discrezitation for r-homogeneous systems 30
3.3 Projection methods preserving Lyapunov functions 34
3.3.1 Projection schemes 35
3.3.2 Symmetric projection methods 36

4 A Lyapunov-Function-Based family of discretization schemes for homogeneous systems **39**
4.1 Order one LFB method with explicit Euler 39
4.2 LFB family of discretization schemes 48
4.2.1 Implementations 59

5 Conclusions and Future Work **69**

Bibliography **75**

List of Figures

1.1 Graphical representation of a solution of the discretization of (1.3). . . . 3
1.2 Graphic representation of a solution of the discretization of (1.7). 5

2.1 Graphic representation of one discretization step for the explicit Euler's method. 19
2.2 Graphic representation of one discretization step for the midpoint method. 20

4.1 Comparison between approximate solutions of (4.24) with explicit Euler's method (left) and with the method of (Sanchez et al., 2020) (right). . . . 47
4.2 Comparison between approximate solutions of (4.61) with explicit Euler's method (left) and with the method of (Sanchez et al., 2020) (right). . . . 48
4.3 Implementation 1. Norm of z_k for different values of h, using three methods to compute (4.39). Numerical verification of Assumption 4. 61
4.4 Implementation 1. Norm of the states of the LFB discretization of (4.59) for different sizes of h. The Matlbab ODE45 approximation is taken as reference. 62
4.5 Implementation 1. Global errors up to k for the LFB discretization of (4.59) for different values of h. 63
4.6 Implementation 2. $\log_{10}$ of the norm of the states of the discretization of (4.61) with LFB Midpint method (left) and Explicit Euler's method (right). 64
4.7 Implementation 2. $\log_{10}$ of the norm of the states of the LFB discretization of (4.61) with MidPoint method for different initial conditions. 65
4.8 Implementation 3. Norm of z_k using two methods to compute (4.39). Numerical verification of Assumption 4. 66
4.9 Implementation 3. Norm of the states of the discretization of (4.63), approximating (4.9) and (4.4) with two LFB methods. 66
4.10 Implementation 4. Discretization of (4.64) for explicit Euler's, LFB Heun and LFB RK4. 68

List of Tables

4.1 Maximum global errors. 60

1

Introduction

Dynamical systems are mathematical frameworks used to model the evolution of artificial and natural processes over time, which can be treated either as a continuous variable or as a sequence of discrete instants. When modeled in continuous time, dynamical systems are typically described by differential equations, with each solution being a continuous-time function and representing the time evolution of one variable. In contrast, discrete-time dynamical systems are formulated using difference equations, where solutions are given by sequences that specify the values of one variable at some time instants. A fundamental purpose of dynamical systems is estimation of future states based on current or past information. However, explicit analytical solutions are available only for a limited class of systems, such as linear systems or certain low-dimensional nonlinear systems with simple structural characteristics. For the vast majority of nonlinear systems, obtaining closed-form solutions is generally infeasible. To address this challenge, discretization schemes, which are mathematical and computational tools that transform continuous-time systems into discrete-time representations, are employed. These schemes convert systems of differential equations into systems of difference equations, thereby enabling the approximation of continuous-time solutions through discrete-time sequences. Discretization plays a critical role in applications such as control engineering, where it facilitates both the simulation of dynamical behavior and the digital implementation of control algorithms.

Let us consider the continuous-time dynamical system

$$\dot{x} = f(x), \tag{1.1}$$

where $f : \mathbb{R} \longrightarrow T\mathbb{R}$ is a vector field that may be discontinuous at the origin. Conditions for existence and uniqueness of solutions in forward time are assumed to be satisfied for (1.1). Broadly speaking, a discretization scheme $\mathcal{M}$ allows one to compute a discrete-time system in the form

$$x_{k+1} = \Psi\left(t_k, \cdots, h\right), \tag{1.2}$$

where the middle dots in the arguments of Ψ in (1.2) indicate a number and type of arguments of Ψ that may vary depending on the involved discretization scheme, for example if Ψ depends on values prior to x_{k+1}, but does not depend on x_{k+1}, the method is said to be explicit, and is said to be implicit otherwise. The subindex k in x_k indicates that the k-th element of the solution $(x_k)_{k\in\mathbb{N}}$ of (1.2) approximates the continuous solution x of (1.1) at the time instant t_k, i.e., $x_k \approx x(t_k)$. The discrete sampling times are usually defined as $t_k = t_0 + kh$, with h denoting the discretization step that may be changed for each implementation of $\mathcal{M}$, nonetheless, certain schemes allow the discretization step size to vary throughout the course of a single execution. As the discretization step approaches zero, (1.2) should approximate (1.1) more closely.

The field of numerical methods for approximating solutions to continuous-time dynamical systems is well-established and has evolved through a variety of approaches and perspectives (Ascher & Petzold, 1998; Hairer et al., 1993; Lambert, 1991; Stuart & Humphries, 1996). A wide range of discretization techniques exists for nonlinear systems, including general-purpose methods—such as those detailed in (Hairer et al., 1993)—as well as specialized methods tailored to particular classes of systems, such as discontinuous systems, as discussed in (Acary & Brogliato, 2008). According to numerical analysis, any practically useful discretization scheme must satisfy two fundamental properties: consistency and convergence. These properties are closely related to the local and global truncation errors, respectively, and they ensure that the discrete system accurately approximates the continuous system as the time step becomes sufficiently small. Local truncation error measures the error introduced in a single step, ignoring errors from prior steps, while global truncation error accounts for the accumulation of errors over multiple steps. Convergence ensures that, as the discretization step approaches zero, the discrete solution of (1.2) approaches the continuous solution of (1.1). Consistency guarantees that reducing the step size leads to a reduction in the local error. Convergence implies consistency, but the converse is not satisfied. The order of consistency is a number that indicates how large the discretization step can be set to ensure that the discrete solution approximates the continuous solution with a certain desired precision : the lower the order of consistency of a method, the smaller the size of the discretization step needed to achieve the same precision as with a method whose order of consistency is higher. To illustrate this fact, let us consider, for example, the system

$$\dot{x} = x. \tag{1.3}$$

The continuous-time solution of (1.3), for initial condition x_0, is given by

$$x(t) = x_0 e^t. \tag{1.4}$$

Figure 1.1 shows the results obtained from the discretization of (1.3), using two convergent and consistent discretization schemes, with different consistency orders:

- **Explicit Euler's method.** (consistency order 1)

$$x_{k+1} = x_k + h x_k \tag{1.5}$$

- **Midpoint method** (consistency order 2)

$$x_{k+1} \quad = \quad x_k + h\left(x_k + \frac{h}{2}x_k\right) \tag{1.6}$$

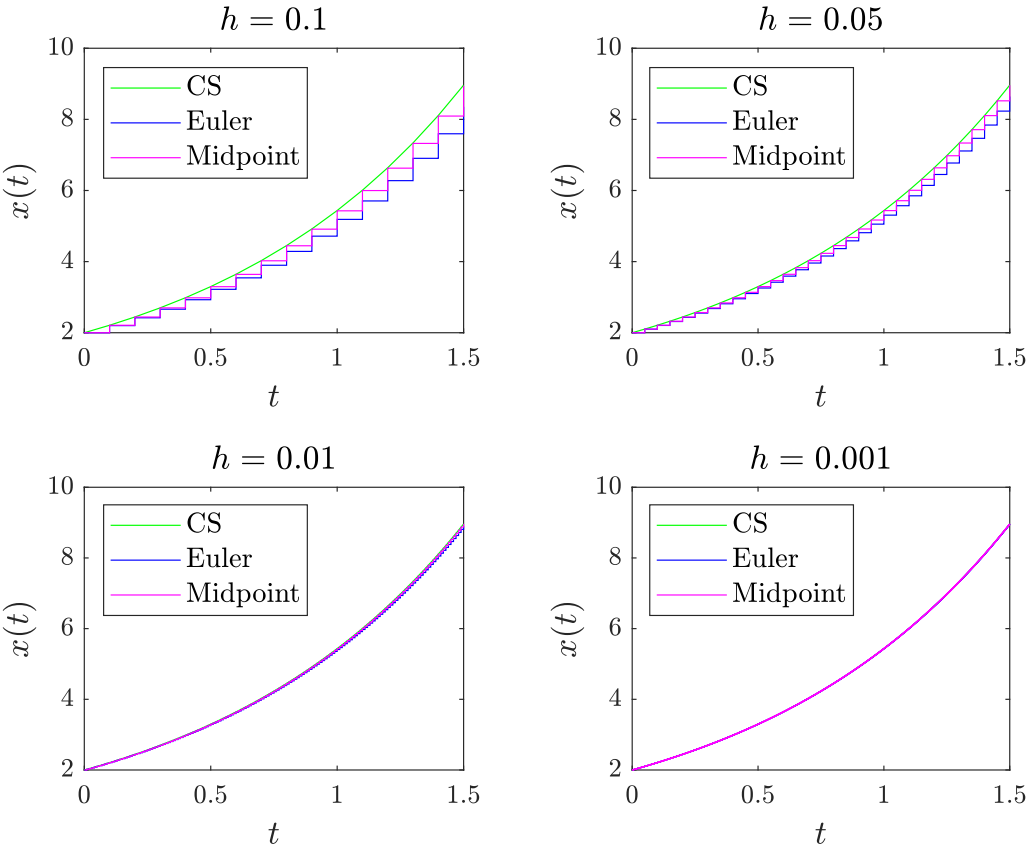

Figure 1.1: Graphical representation of a solution of the discretization of (1.3).

It can be appreciated how the midpoint method achieves higher precision with a

larger discretization step compared to the Euler method.

There exists several well-established discretization methods, including for example the Runge-Kutta and Adams families, that have been carefully developed, with their numerical convergence, consistency order, and other essential properties thoroughly validated by their developers. These methods are designed to ensure that the discrete approximations they produce remain computationally precise when the discretization step tends to zero. However, these schemes are so general that they do not guarantee that particular qualitative properties of each type of system, such as stability or instability, are preserved. Discretization schemes are generally not employed for the analysis of qualitative properties of dynamical systems—such as stability or the nature of convergence of solutions—as these characteristics are typically established through formal mathematical techniques prior to the discretization process. Instead, such schemes are primarily utilized to investigate quantitative aspects of the system, including, for instance, the rate of convergence or divergence of solutions. Therefore, when used to discretize dynamical systems, they are generally expected to possess additional properties that assess how closely their approximations align with the true solutions of the continuous systems.

In general, the well-known methods (Euler, Runge-Kutta, Adams family, etc.) do not guarantee the preservation of particular properties of the systems they approximate. Let us consider for example the system

$$\begin{aligned} \dot{x}_1 &= -2\lceil x_1 \rfloor^{\frac{3}{2}} + x_2, \\ \dot{x}_2 &= -\lceil x_1 \rfloor^2, \end{aligned} \tag{1.7}$$

where $\lfloor y \rceil^\alpha = \text{sign}(y)|y|^\alpha$. It can be proved that the origin of $\mathbb{R}^2$ is an asymptotically stable equilibrium point for (1.7). Moreover, solutions of (1.7) converge to any neighborhood of the origin in finite time. If one simulates the approximation of the solution of (1.7), using the explicit Euler method, for initial condition $[10^5, 0]$ and discretization step $h = 0.1$ one obtains the solution illustrated in Figure 1.2. Note that the solution in Figure 1.2 does not converge to the origin, which contradicts the asymptotic stability of the origin for (1.7). This occurs because the explicit Euler method only preserves the stability of equilibrium points for sufficiently small discretization steps. There exist discretization schemes in the literature specifically designed to preserve certain properties of particular systems, for instance, methods with A-stability that preserve the asymptotic stability of linear systems, or geometric integration approaches that aim at preserving first integrals (e.g., energy or momentum) in Hamiltonian systems (Hairer, Lubich, & Wanner, 2006; Marsden, 1992; Sanz-Serna, 1992; Stuart & Humphries, 1996).

One challenge when discretizing Lyapunov-stable nonlinear dynamical systems is ensuring the preservation of Lyapunov stability. This challenge can be effectively addressed if the Lyapunov function that holds for the continuous-time system should also serve as a valid Lyapunov function for the discrete-time approximation. Several approaches have

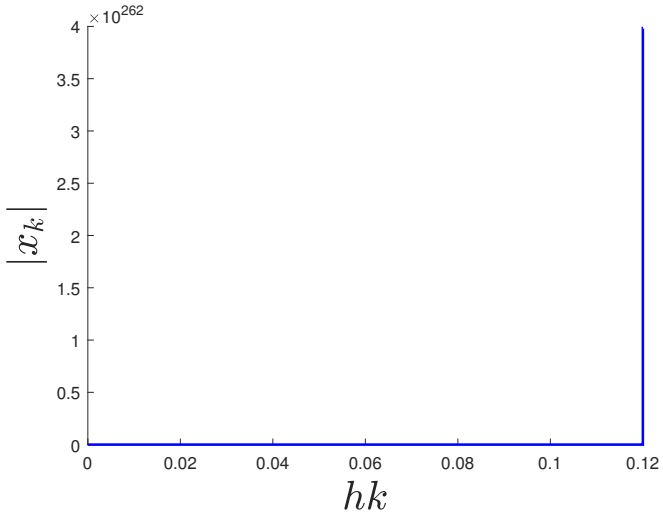

Figure 1.2: Graphic representation of a solution of the discretization of (1.7).

been explored to develop methods that maintain Lyapunov stability, notable examples include: restricting the discretization step size in certain Runge-Kutta methods for gradient systems (Hairer & Lubich, 2014), using discrete-gradient methods for gradient systems (McLachlan et al., 1999), and applying projection methods (M. P. Calvo et al., 2010; Grimm & Quispel, 2005).

Homogeneous systems form a significant class of dynamical systems, encompassing all linear systems as well as a subset of nonlinear systems that share many key characteristics with linear ones, for example, the scalability of solutions. In addition to these similarities, homogeneous nonlinear systems exhibit unique features that are particularly valuable for the analysis and design of control systems. Research on this type of systems includes: nilpotent and homogeneous approximation (Andrieu et al., 2008; Hermes, 1986), stability and stabilization (Grüne, 2000; Hahn, 1967; Hermes, 1991; Hong et al., 1999; Kawski, 1990; Mendoza-Avila et al., 2023; Orlov, 2004; Polyakov et al., 2016; Zubov, 1964), Lyapunov functions (Nakamura et al., 2002), (Moreno & Osorio, 2012), (Sanchez & Moreno, 2019), (Efimov et al., 2018), robustness (Bernuau et al., 2013), (Rosier, 1992), (Sepulchre & Aeyels, 1996), sliding modes (Cruz-Zavala & Moreno, 2017; Levant, 2005), observer design (Qian & Lin, 2006), differentiators (Moreno, 2022), generalized homogeneity (Polyakov, 2020) and gain design (Sanchez et al., 2023).

Some of the interesting properties of homogeneous systems are intrinsic robustness and *scaling*, the property whereby trajectories of a homogeneous system are obtained by scaling solutions that evolve in arbitrarily small neighborhoods of the origin, which turns out to be quite helpful to deduce the *global* qualitative nature of the solutions from their *local* behavior. Unfortunately, many standard discretization methods are not well-suited to the specific nature of homogeneous systems. As a result, these methods

often provide suboptimal approximations, failing to leverage the important features of homogeneity. A notable example arises when the system's solutions exhibit finite-time or nearly-fixed time convergence (sometimes referred to as rational convergence). In such cases, discretization methods that do not align with the system's homogeneity can produce approximations that suffer from numerical chattering phenomena (Levant, 2010), (Acary & Brogliato, 2010), or even result in finite escape times (Efimov et al., 2017; Levant, 2010), as is the case with system (1.7).

Distinct strategies have been developed to discretize homogeneous systems, including sliding-mode controllers and observers, that feature improved properties with respect to the well known schemes, e.g., implicit discretization of sliding mode controllers in (Drakunov & Utkin, 1989) and (Utkin, 1994), discretization of an arbitrary-order robust exact differentiator (Koch & Reichhartinger, 2018), and discrete-time realization of sliding-mode algorithms—via implicit discretization of differential inclusions—as discussed in (Huber et al., 2016) and (Acary et al., 2012). In (Polyakov et al., 2019, 2022), a discretization approach was developed for asymptotically stable homogeneous systems, based on the fact that every homogeneous stable system is topologically equivalent to a quadratically stable one. This method was shown to preserve the convergence type of the continuous-time system's solutions for finite-time and nearly-fixed time convergence. Although the methods proposed in (Polyakov et al., 2019) preserve the type of convergence, they are implicit and depend on a generally implicitly defined homogeneous norm, moreover, their proven order of consistency is one. In (Efimov et al., 2019), Euler-like schemes with a dynamical scaling of the discretization step were proposed to discretize homogeneous systems. The implicit and the explicit methods in (Efimov et al., 2019) are not proven to preserve the type of convergence of the solutions nor the Lyapunov function, and their proven order of consistency is one. Other methods that preserve Lyapunov functions also include the method based on discrete-gradients proposed in (McLachlan et al., 1999), which is restricted to gradient systems; and the projection method (M. P. Calvo et al., 2010), which is quite involved since it requires Gaussian quadratures, dense outputs of Runge–Kutta methods, solving implicit equations to compute the projections, and it imposes restrictions on the discretization step; moreover, none of these two methods guarantees the preservation of the convergence rates from the continuous-time system.

In (Sanchez et al., 2020), a Lyapunov-function-based ("LFB") discretization method was introduced for asymptotically stable homogeneous systems. This method preserves both Lyapunov stability and the type of convergence of the system's trajectories. It can be understood as a two-step projection technique: first, the explicit Euler method is applied to compute a discrete approximation of the system's dynamics, *projected* onto a level set of the Lyapunov function. Then, information from the Lyapunov function is used to *lift* the approximation back to the original state space. The method was shown to be numerically consistent of first order in (Sanchez et al., 2020). A natural question that arises from this work is whether the performance of the approximation scheme in

(Sanchez et al., 2020), such as its precision and order of consistency, can be improved by using more precise or higher-order auxiliary methods to approximate the projected dynamics, instead of the explicit Euler method. Since the order of the auxiliary method is expected to influence the order of the overall discretization scheme, the ideal objective would be to develop a higher-order discretization scheme that continues to preserve the Lyapunov function and the convergence type of the solutions.

In this thesis, the discretization problem of weighted homogeneous systems, preserving some features related to homogeneity, such as stability and the type of convergence of the solutions, is studied. The advantages and disadvantages of the methods that were proposed in (Polyakov et al., 2019), (Efimov et al., 2019), (M. Calvo et al., 2006), and (Sanchez et al., 2020) are studied. The main contribution of the present work is the proposal of an improvement to the discretization scheme in (Sanchez et al., 2020). This improvement generalizes the one from (Sanchez et al., 2020) in the sense of allowing one to replace Euler's method by more general discretization schemes, of arbitrary order $\sigma \in \mathbb{N}$, to approximate the projected dynamics; thus obtaining a family of discretization schemes whose consistency order σ can be selected as large as desired, as long as an auxiliary method with the same consistency order is used to approximate the projected dynamics. Numerical convergence and consistency of order σ, for any discretization step value, are proved for this family of schemes. It is also proved that all the methods in this family preserve a Lyapunov function for the original system and the convergence type of the solutions.

The thesis is organized as follows: Chapter 1 provides a brief introduction to the research. Chapter 2 presents essential preliminaries on stability, homogeneous systems, and numerical discretization methods. Chapter 3 discusses the discretization schemes for homogeneous systems introduced in (Polyakov et al., 2019), (Efimov et al., 2019) and (M. P. Calvo et al., 2010). Chapter 4 is dedicated to the LFB family of discretization schemes, which constitutes the main contribution of this work. Here, the method proposed in (Sanchez et al., 2020) is reviewed, the principal results of the research are proved, and several implementation results are included. Finally, Chapter 5 offers concluding remarks and outlines potential directions for future research.

2

Preliminaries

An ordinary differential equation (ODE) is an equation that involves some ordinary derivatives (as opposed to partial derivatives) of a function. The order η of an ODE is determined by the highest order derivative it contains. Every ODE of order $\eta \in \mathbb{N}$ may be written in the form

$$f\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta)}\right) = 0, \tag{2.1}$$

where $y : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is a solution of (2.1) and $y^{(i)}$ denotes the i-th derivative of y. ODE (2.1) is said to be *autonomous* if f does not depend explicitly on t, and *non-autonomous* otherwise. A system of ODEs of order $\eta \in \mathbb{N}$ may be written in the form (2.1), with the understanding that $y : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n$ and that f actually represents a set of ODEs, i.e,

$$f\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta)}\right) = \begin{pmatrix} f_1\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta)}\right) \\ f_2\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta)}\right) \\ \vdots \\ f_n\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta)}\right) \end{pmatrix}. \tag{2.2}$$

Every order η ODE in the form

$$y^{(\eta)} = f\left(t, y, y^{(1)}, y^{(2)}, \ldots, y^{(\eta-1)}\right) \tag{2.3}$$

is equivalent to a first order system of ODEs by means of a recursive change of variables. Let us consider for example

$$\frac{d^3y}{dt^3} = f\left(t, y, \frac{dy}{dt}, \frac{d^2y}{dt^2}\right). \tag{2.4}$$

Let us define $x_1 = y$, $x_2 = \frac{dy}{dt}$ and $x_3 = \frac{dx_2}{dt}$. Variables x_1 to x_3 and y satisfy the first order ODE system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3, \\ \dot{x}_3 &= f\left(t, x_1, x_2, x_3\right). \end{aligned}$$

A dynamical system is an ODE system for which y describes the evolution over time of a state. Dynamical autonomous systems, for which f does not depend explicitly on t, are commonly represented as first-order ODE systems in the form

$$\dot{x} = f\left(x\right), \tag{2.5}$$

where $x : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \longrightarrow T\mathbb{R}^n$, where $T\mathbb{R}^n$ denotes the tangent bundle of $\mathbb{R}^n$, and for every $x \in \mathbb{R}^n$ the fiber $T_x\mathbb{R}^n$ is homoemorphic to $\mathbb{R}^n$. Throughout this thesis, we will consider initial value problems (IVP for short) given by systems of the form (2.5), under a set of assumptions that will be made explicit later on, together with initial conditions

$$x(t_0) = x_0, \tag{2.6}$$

with $x_0 \in \mathbb{R}^n$.

2.1 Stability for nonlinear dynamical systems

Hereafter in this document, we will assume that System (2.5) satisfies conditions for the existence and uniqueness of solutions in forward time. A point $\bar{x} \in \mathbb{R}^n$ is said to be an *equilibrium point* of (2.5) if $f\left(\bar{x}\right) = 0$. Intuitively, equilibrium points are considered stable if solutions may be made to remain arbitrarily close to the equilibrium solution by considering initial conditions sufficiently close to the equilibrium; if not, they are said to be unstable.

Definition 1 (Stability in the sense of Lyapunov, (Lyapunov, 1992))**.** *For* (2.5)*, the origin is said to be :*

- ***Stable*** *if and only if for every $\epsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that for every $t_0 \in \mathbb{R}_{>0}$, if $\|x(t_0)\| < \delta$, then $\|x(t)\| < \epsilon$, for every $t \geq t_0$.*
- ***Unstable*** *if it is not stable.*
- ***Asymptotically stable*** *if and only if it is stable and δ can be choosen such that if $\|x(t_0)\| < \delta$, then $\lim_{t\to\infty} x(t) = 0$.*

A class C^1, positive definite function $V : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, with D a neighborhood of $\bar{x}$, is said to be a *weak Lyapunov function* for (2.5) if the derivative od V along the trajectories of (2.5), denoted $\dot{V}$, satisfies $\dot{V}(x) \leq 0$ for every $x \in D$ and it is said to be

a *strict Lyapunov function* if $\dot{V}(x) < 0$ for every $x \in D \setminus \{0\}$. Stability of equilibrium points is usually characterized in terms of the existence of Lyapunov functions.

Theorem 1 (Stability characterized by Lyapunov functions ((Lyapunov, 1992))). *For* (2.5)*, the origin is said to be:*

- ***Stable*** *if there exists a weak Lyapunov function for* (2.5)*.*
- ***Asymptotically stable*** *if there exists a strict Lyapunov function for* (2.5)*.*
- ***Globally asymptotically stable*** *(resp.* ***globally stable****) if it is asymptotically stable (resp. stable,* $D = \mathbb{R}^n$ *and* V *is radially unbounded.*

Let $c \in \mathbb{R}_{\geq 0}$. The set

$$S_c = \{x \in \mathbb{R}^n : V(x) = c\} \tag{2.7}$$

will be called a *Lyapunov level set* (LLS for short). When the origin exhibits asymptotic stability, the time derivative of a strict Lyapunov function V is negative definite, which implies that solutions starting on a LLS S_c move towards inner Lyapunov level sets, characterized by smaller values of c. Consequently, as c diminishes, the LLS S_c progressively approaches the origin, leading the trajectory to do the same.

By means of the chain rule, the derivative of any Lyapunov function V along the trajectories of (2.5) is given by

$$\dot{V} = -W(x) \tag{2.8}$$

where

$$W(x) = -\frac{\partial V}{\partial x}(x) f(x). \tag{2.9}$$

2.2 Homogeneous systems

Linear systems are a specific and restricted category of dynamical systems that exhibit notable characteristics, such as the ability to be solved explicitly. When an equilibrium point of a linear system is asymptotically stable, solutions will always approach it in an exponential manner, and a variety of linear algebra techniques can be applied to analyze them. These traits are not generally present in nonlinear systems, as they lack closed-form solutions and their behavior near equilibrium points can be more unpredictable. Homogeneous systems, a set containing all the linear and some nonlinear systems, act as a link between linear and nonlinear dynamics, possessing some properties similar to the ones of linear systems that simplify their study, such as solution scalability, which allows local properties to be extended to global ones, and the fact that solutions to these systems will approach equilibrium points only in one of three distinct manners.

Definition 2 (Standard dilation (Polyakov, 2020)). *The* ***standard dilation*** *of* $x \in \mathbb{R}^n$*, with a scale factor* $\lambda \in \mathbb{R}_{>0}$*, is defined by*

$$\delta_\lambda(x) = \lambda x. \tag{2.10}$$

Definition 3 (Homogeneity (Rosier, 1992))**.** *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is said to be* ***homogeneous of degree*** *$\nu \in \mathbb{R}$ if and only if it satisfies*

$$f\left(\delta_\lambda(x)\right) = \lambda^\nu f(x), \tag{2.11}$$

for every $\lambda \in \mathbb{R}_{>0}$ and every $x \in \mathbb{R}^n$.

Definition 4 (Weighted dilation (Rosier & Bacciotti, 2005))**.** *The* ***weighted dilation*** *$\Delta^r_\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, with* ***weights vector*** *$r \in \mathbb{R}^n_{>0}$ and* ***scale factor*** *$\lambda \in \mathbb{R}_{>0}$, is defined for every $x \in \mathbb{R}^n$ by*

$$\Delta^r_\lambda(x) = \left(D_{\lambda,r}\right) x. \tag{2.12}$$

Weighted dilations scale each component of a vector $x \in \mathbb{R}^n$, by a uniform dilation, which permits the application of distinct scaling factors for each component.

Definition 5 (Weighted homogeneity ((Rosier & Bacciotti, 2005)))**.** *Given a vector of weights $r \in \mathbb{R}^n_{>0}$:*

- *A function $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be r-* ***homogeneous of degree*** *$m \in \mathbb{R}$ if and only if it satisfies*

  $$V\left(\Delta^r_\lambda(x)\right) = \lambda^m V(x), \tag{2.13}$$

  for every $\lambda \in \mathbb{R}_{>0}$, and every $x \in \mathbb{R}^n$.

- *A vector field $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is said to be r-* ***homogeneous of degree*** *$\mu \in \mathbb{R}$ if and only if it satisfies*

  $$f\left(\Delta^r_\lambda(x)\right) = \lambda^\mu \Delta^r_\lambda\left(f(x)\right), \tag{2.14}$$

  for every $\lambda \in \mathbb{R}_{>0}$, and every $x \in \mathbb{R}^n$.

- *System* (2.5) *is said to be r-homogeneous of degree μ if its associated vector field f is r-* ***homogeneous of degree*** *μ.*

There exists one vector field associated with every family of weighted dilations Δ^r_λ, called the *generalized vector field* e associated with Δ^r_λ, which is given by

$$e = \sum_{j=1}^{n} r_i x_i \frac{\partial}{\partial x_i}. \tag{2.15}$$

Weighted homogeneity of functions and vector fields was defined in terms of the generalized vector field e associated with Δ^r_λ by (Kawski, 1995) as follows:

- V is r-homogeneous of degree m if and only if

  $$L_e V = mV \tag{2.16}$$

- f is r-homogeneous of degree μ if and only if

$$L_e f = \mu f, \tag{2.17}$$

where $L_e(\cdot)$ denotes the Lie derivative of tensor field in the direction of e. Since $L_e V = e \cdot V$ and $L_e f = [e, f]$, it follows that:

- V is r-homogeneous of degree m if and only if

$$e \cdot V = mV \tag{2.18}$$

- f is r-homogeneous of degree μ if and only if

$$[e, f] = \mu f, \tag{2.19}$$

where $[\cdot, \cdot]$ denotes the Lie bracket of vector fields.

For asymptotically stable r-homogeneous systems, the existence of an homogeneous, class C^p, Lyapunov function is not only a sufficient condition to determine the stability of an equilibrium point, but it is also a necessary condition.

Theorem 2 ((Rosier, 1992)). *Let the origin be a locally asymptotically stable equilibrium point of* (2.5) *and let the vector field f associated with* (2.5) *be continuous on $\mathbb{R}^n$ and r-homogeneous of degree $\mu \in \mathbb{R}$ for some $r \in \mathbb{R}^n$. Then for any $p \in \mathbb{N}$ and any $m > p \cdot \max_i\{r_i\}$, there exists a strict Lyapunov function V of* (2.5)*, which is r-homogeneous of degree m and of class C^p. As a direct consequence, the function W, given by* (2.8)*, is r-homogeneous of degree $m + \mu$.*

In (Rosier, 1992) it was proved that

$$\dot{V}(x) \leq -\alpha V^{\frac{m+\mu}{m}}(x),$$

for every $x \in \mathbb{R}^n$, where α is given by

$$\alpha = \min_{x \in S_1} W(x). \tag{2.20}$$

Unlike linear systems, where if the equilibrium point is asymptotically stable, the solutions converge to it exponentially, in the case of r-homogeneous systems, convergence of solutions towards the origin can be achieved in 3 different ways, which depend on the degree of homogeneity of the system.

Definition 6 (Classification of global asymptotic stability (see e.g. (Nekhoroshikh et al., 2020), (Kumar & Efimov, 2023))). *A globally asymptotically stable equilibrium point of System* (2.5) *is said to be:*

- ***Globally exponentially stable*** *with respect to an r-homogeneous norm $\|\cdot\|_r$, if there exist constants $M \geq 1$ and $\omega > 0$ such that, for every $x_0 \in \mathbb{R}^n$, the solution $x(\cdot)$ starting at $x(0) = x_0$ satisfies*

$$\|x(t)\|_r \leq M\|x_0\|_r e^{-\omega t}, \qquad \forall t \in \mathbb{R}_{>0}. \tag{2.21}$$

- ***Globally finite-time stable*** *if there exists a positive definite function $T : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0}$, called the settling-time function, such that every solution $x(\cdot)$, starting at $x_0 \in \mathbb{R}^n$, satisfies*

$$\begin{aligned} \text{if } x_0 \neq 0 \text{ then } x(t) &\neq 0, \qquad \forall t \in [0, T(x_0)) \\ x(T(x_0)) &= 0. \end{aligned}$$

- ***Globally nearly-fixed-time stable*** *if for any $\rho \in \mathbb{R}_{>0}$ there exists $\overline{T} \in \mathbb{R}_{>0}$, called the transition time, such that for every $x_0 \in \mathbb{R}^n$, the solution $x(\cdot)$ starting at $x(0) = x_0$ satisfies*

$$\|x(t)\| \leq \rho, \qquad \forall t \geq \overline{T}. \tag{2.22}$$

If the origin is globally nearly-fixed-time stable, then, for any ball centered at the origin, all solutions of (2.5) are guaranteed to belong to that ball after a time that may depend on the radius of the ball, but not on the initial conditions. Once such a ball is specified, the convergence time bound is fixed, hence the name of this property. In this work, the origin is assumed to be an equilibrium point because any other equilibrium point may be mapped to the origin using an appropriate translation. The stability of the origin is characterized in terms of the bounds on a Lyapunov function, as expressed in the following lemma (see, e.g., (Nakamura et al., 2002; Rosier & Bacciotti, 2005)).

Lemma 1. *Let (2.5) be an r-homogenous system, of degree $\mu \in \mathbb{R}$ for some $r \in \mathbb{R}^n_{>0}$, with an asymptotically stable equilibrium at the origin. Let $V : \mathbb{R}^n \longrightarrow \mathbb{R}_{>0}$ be a strict Lyapunov function for (2.5), being r-homogeneous of degree $m \in \mathbb{R}_{>0}$. Then, for every initial condition $x_0 \in \mathbb{R}^n$ and any solution $x(\cdot)$ of (2.5) starting at x_0, the following is satisfied for any $t \in \mathbb{R}_{>0}$, with α as given in (2.20):*

- *If $\mu = 0$, then the origin is exponentially stable with respect to the homogeneous norm given by $V^{\frac{1}{m}}$ and*

$$V(x(t)) \leq V(x_0)\exp(-\alpha t).$$

- *If $\mu > 0$ then the origin is nearly-fixed-time stable and*

$$V(x(t)) \leq \frac{V(x_0)}{\left(1 + \frac{\mu}{m}\alpha V^{\frac{\mu}{m}}(x_0)t\right)^{\frac{m}{\mu}}}.$$

For every $\rho > 0$ the transition time $\overline{T}$ is given by $\frac{m\overline{V}^{\frac{m}{\mu}}}{\alpha\mu}$, with $\overline{V} > \inf_{x \in B_\rho} V(x)$.

- *If $\mu < 0$, then the origin is finite-time stable and*

$$V(x(t)) \leq \begin{cases} \left(V^{\frac{-\mu}{m}}(x_0) - \frac{-\mu}{m}\alpha t\right)^{\frac{-m}{\mu}}, & t < \frac{-m}{\mu\alpha}V^{\frac{-\mu}{m}}(x_0) \\ 0, & t \geq \frac{-m}{\mu\alpha}V^{\frac{-\mu}{m}}(x_0) \end{cases} .$$

The settling time function T is upper bounded by $\frac{-m}{\alpha\mu}V^{\frac{-\mu}{m}}(x_0)$.

2.3 Numerical methods for discretization of dynamical systems

Only a very limited class of ODE systems can be solved using analytical methods. In general, for many differential equations arising in applications, it is impossible or impractical to have solution formulas. Numerical methods are mathematical techniques used for solving problems that cannot be solved, or are very difficult to solve, analytically. Some methods perform a series of calculations iteratively to obtain discrete-time sequences that approximate the continuous solutions of the ODE system. The accuracy and precision with which a numerical method approximates a solution depend on the properties and characteristics discussed in this section.

To simplify and clarify the explanation of the discretization process, let us denote by $f_{\mathcal{C}}$ the vector field f associated with system (2.5), where subindex $\mathcal{C}$ indicates that it is a continuous vector field. Broadly speaking, the solution to the problem of discretizing (2.5) consists of obtaining a discrete-time system

$$x_{k+1} = f_{\mathcal{D}}\left(x_0, \ldots, x_{k+1}, t_k\right), \tag{2.23}$$

where $f_{\mathcal{D}} : \Xi \longrightarrow \mathbb{R}^n$ is a discrete-time map that approximates the flow generated by the vector field $f_{\mathcal{C}}$ at certain instants of time, with the distance between those instants given by the *discretization step*. The domain Ξ may vary for each discretization scheme, since $f_{\mathcal{D}}$ may depend on values of x at previous time instants, on the discretization step size, or on the discretization instant, among others. Once the system is discretized, to obtain the solution of the IVP given by (2.23) and x_0, on a discrete set of time instants $t_0 < t_1 < \cdots < t_f$, one computes a sequence $(x_k)_{k\in K_h}$ from (2.23), for a fixed discretization step h, with $K_h = \{t_0, \ldots, N_h\}$ and $N_h = \lfloor\frac{t_f - t_0}{h}\rfloor$, where $\lfloor a\rfloor$ denotes the integer part of a. Thus $x_k \approx x(t_k)$ on a discrete set of time instants, usually defined to be evenly spaced as $t_k = t_0 + kh$, where $h \in \mathbb{R}_{>0}$ is called the discretization step. Since the discretization process is iterative, in practice, the computational calculations of both the discrete-time system and the sequence approximating the continuous solution are performed in tandem at each iteration. For this reason, several times throughout this document, mainly when we explain results from other references in an intuitive and non-rigorous way, we refer to $(x_k)_{k\in\mathbb{N}}$ as a *discrete-time solution* of (2.5), obtained with a given discretization scheme, in the understanding that $(x_k)_{k\in\mathbb{N}}$ is actually a solution of

the discretization (2.23), which approximates a solution of (2.5), at some discrete instants of times, for a given initial condition. We use the correct and complete terminology in critical parts that require it, as for example in theorems and when we present the main results of this work.

As mentioned for example in (Lambert, 1991), a general discretization method $\mathcal{M}$, for the IVP given by (2.5) and (2.6), can be written in the form

$$\begin{aligned} x_{k+s} &= h\psi_f(t_k, x_{k+s}, x_{k+s-1}, \ldots, x_k, h) - \sum_{j=0}^{s-1} \alpha_j x_{k+j} \\ &= \Psi_f\left(t_k, x_{k+s}, x_{k+s-1}, \ldots, x_k, h\right), \end{aligned} \tag{2.24}$$

where:

- s is the number of *steps* of the method. If $s = 1$ then the method is said to be *single-step*, otherwise is said to be a *multi-step* method.
- The index f in ψ_f indicates that ψ depends on $(t_k, x_{k+s}, x_{k+s-1}, \ldots, x_k, h)$ through the vector field $f_{\mathcal{C}}$, for example

$$\psi_f(x, y) = \frac{h}{4} f(x_k, y_k, x_{k-1}, y_{k-1}, h) - 12hf(x_{k-1}, y_{k-1}).$$

- Method (2.24) is said to be *explicit* if ψ_f does not depend on x_{k+s} and *implicit* otherwise.
- Method (2.24) is said to be *linear* if the multiplication $h\psi_f$ gives a linear combination of $f(x_0), \ldots, f(x_{k+1})$ and is said to be *nonlinear* otherwise.

The problem of discretizing (2.5), with a discretization scheme $\mathcal{M}$, is said to be *well posed* in the Hadamard sense (Hadamard, 1902) if and only if it satisfies the following requirements:

1. For every initial condition $x_k \in \mathbb{R}^n$ and every discretization step $h \in \mathbb{R}_{>0}$, there exists $x_{k+1} = \Psi_f(\cdots)$, for every $k \in K_h$.
2. For every initial condition $x_k \in \mathbb{R}^n$ and every discretization step $h \in \mathbb{R}_{>0}$, the scheme $\mathcal{M}$ produces a unique value $x_{k+1} = \Psi_f(\cdots)$, for every $k \in K_h$.
3. Small perturbations on x_k produces small perturbations on x_{k+1}.

In standard well-established references for numerical methods, e.g., (Iserles, 2009), (Eugene & Keller, 1994), (Epperson, 2002), (Lambert, 1991), (Süli & Mayers, 2003), it is assumed that the vector field f of (2.5) is Lipschitz continuous, i.e., there exists $k \in \mathbb{R}_{>0}$ such that for every $x, \bar{x} \in \mathbb{R}^n$ the following is satisfied

$$\|f(x) - f(\bar{x})\| \leq k\|x - \bar{x}\|. \tag{2.25}$$

This assumption guarantees the existence a unicity of solutions for the continuous system and also guarantees that the discretization problem for the methods studied in those references is well posed.Furthermore, assuming the existence of a Lipschitz constant facilitates the demonstration of key properties of discretization methods, such as convergence and consistency, which will be addressed later in this section.es of numerical discretization methods. However, the Lipschitz condition is too restrictive for the systems studied in this thesis; instead, we will assume the following with respect to (2.5):

Assumption 1. *The vector field f is continuous on $\mathbb{R}^n \setminus \{0\}$ and for every initial condition $x_0 \in \mathbb{R}^n$, the solution of* (2.5) *exists and it is unique in forward time.*

Assumption 1 directly implies that the requirements for well-posedness are satisfied for the discretization problem of (2.5), using any method like (2.24). In addition to being well posed, any effective discretization method must satisfy other additional properties to guarantee precision and accuracy in the discrete-time approximation of solutions, these properties will be discussed in what follows.

Numerical convergence

It is expected that solutions of the discrete-time system obtained by a discretization method will not be too sensitive to small errors in computational calculations. The calculation error that accumulates at each discretization step should not be large enough to move the approximated discrete-time solution too far from the continuous-time solution.

A discretization method is said to be *numerically convergent* if the discrete-time solution $(x_k)_{k\in\mathbb{N}}$ of (2.23) converges to the continuous-time solution $x : [t_0, t_f] \longrightarrow \mathbb{R}^n$ of (2.5) as the discretization step h tends to zero. However, although the numerical convergence concept is intuitively easy to understand, proving this property is almost impossible in most cases, because the continuous solution is unknown. Numerical convergence is stated in terms of the global error in most of the well-known references as follows.

Definition 7 (Global error)**.** *The* ***global error*** e_k*, up to the instant $k \in K_h$ of the discretization scheme* (2.24) *is defined by*

$$e_k = |x(t_k) - x_k|. \tag{2.26}$$

The global error considers the cumulative effect of the errors committed in the first k discretization steps.

Definition 8 (Numerical convergence)**.** *Discretization method* (2.24) *is said to be numerically convergent if and only if for every initial value problem given by* (2.5) *and* (2.6)*, the condition*

$$\lim_{h\to 0} \max_{k\in K_h} |e_k| = 0$$

is satisfied.

Numerical consistency

Another property that is usually expected from discretization methods is the difference system (2.23) being a sufficiently accurate representation of system (2.5), which would be satisfied if when one substitutes each x_k for $x(t_k)$, with $k = 1, \ldots, k+1$, in (2.23), then the equations (2.23) continue to hold. The above is difficult to prove because the continuous-time solution is usually unknown; thus, consistency is defined then in terms of the local truncation error.

Definition 9. *The local truncation error of discretization scheme* (2.24)*, denoted by* T_{k+1}*, is defined by*

$$T_{k+1} = x(t_{k+1}) - x_{k+1}, \tag{2.27}$$

assuming $x(t_k) = x_k$*.*

The local truncation error computes the error that occurs at one step of discretization, when integrating from t_k to t_{k+1}. This error does not take into account the global error accumulated in the previous discretization steps, so in order to calculate it, one assumes $x(t_k) = x_k$.

Definition 10 (Consistency of order σ)**.** *The discretization method* (2.24) *is said to be consistent of order* $\sigma \in \mathbb{N}$ *if and only if there exists* $H \in \mathbb{R}_{>0}$ *such that for every IVP given by* (2.5) *and an initial condition* (2.6)*, the local truncation error satisfies*

$$\lim_{h \to 0} \frac{|T_{k+1}|}{h^\sigma} = 0, \tag{2.28}$$

for every $k \in \mathbb{N}$*.*

2.3.1 Single step methods

Single-step methods only require one previous sample of the solution in order to determine the next value. Some of the best known and most commonly used single-step methods for solving the discretization problem of dynamical systems, and many other computational problems, are listed below (see e.g. (Iserles, 2009), (Eugene & Keller, 1994), (Lambert, 1991)).

- **Euler's method.** Is the simplest discretization approach, but despite its simplicity, it remains significant in practice, to the extent that many other methods that will be discussed in this work are merely more complex generalizations of Euler's method. This method approximates (2.5) by

  $$x_{k+1} = x_k + hf(t_k, x_k). \tag{2.29}$$

  In words, x_{k+1} is approximated using the information provided by the slope of the tangent line at x_k, given by $\dot{x}(t_k)$. This procedure is illustrated in Figure 2.1. This

method is explicit, numerically convergent, and consistent of order $\sigma = 1$. There exists an implicit version of (2.29), which is called the *Backward Euler's method* and approximates (2.5) by

$$x_{k+1} = x_k + hf(t_{k+1}, x_{k+1}), \tag{2.30}$$

where in order to compute $f(t_{k+1}, x_{k+1})$, an implicit equation must to be solved to obtain x_{k+1}, for example using the Newton-Raphson method. Backward Euler's method is numerically convergent and consistent of order $\sigma = 1$.

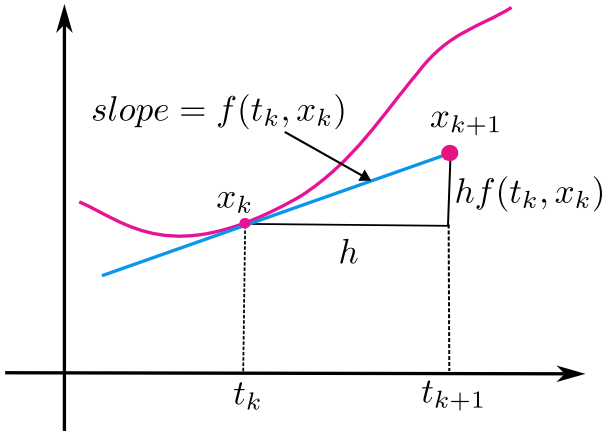

Figure 2.1: Graphic representation of one discretization step for the explicit Euler's method.

- **Midpoint method.** It is similar to the explicit Euler's method, but the slope of the tangent line to the midpoint between x_{k+1} and x_k is computed instead of the slope of the tangent line at x_k. This method is given by

$$x_{k+1} = x_k + hf\left(t_k + \frac{h}{2}, x_k + \frac{h}{2} f(t_k, y_k)\right), \tag{2.31}$$

where $x_k + \frac{h}{2} f(t_k, y_k)$ is the approximation of the midpoint between x_{k+1} and x_k, i.e., the approxiamtion of $x(t_k + \frac{h}{2})$. This method, illustrated in Figure 2.2, is explicit, numerically convergent, and consistent of order $\sigma = 2$.

- **Heun's method.** It is an explicit method, based on Euler's method, in which an intermediate value, obtained with the trapezoidal rule, is used to compute the tangent line to it, similar to the midpoint method. This method uses the averaging of the slopes of the tangent lines at $x(t_k)$ and $x(t_{k+1})$ and is given by

$$\begin{aligned} \tilde{x}_{k+1} &= x_k + hf(t_k, x_k) \\ x_{k+1} &= x_k + \frac{h}{2}\left(f(t_k, x_k) + f(t_{k+1}, \tilde{x}_{k+1})\right). \end{aligned} \tag{2.32}$$

It is numerically convergent and consistent of order $\sigma = 2$.

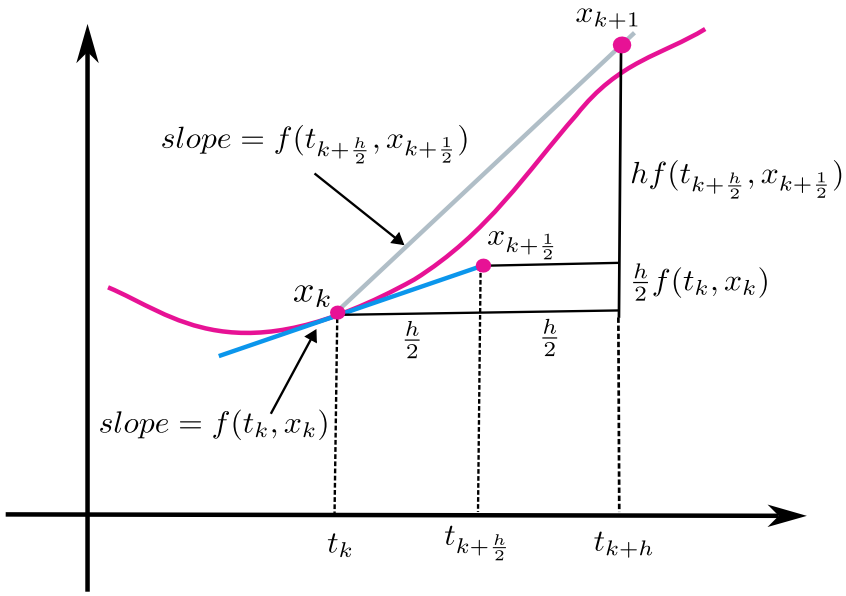

Figure 2.2: Graphic representation of one discretization step for the midpoint method.

- **Runge-Kutta family.** Is a family of nonlinear single-step methods that achieve higher consistency by preserving the possibly nonlinear dependence of f on (t_k, x_k, h). Any s-stage Runge-Kutta method can be written in the form

$$\begin{aligned} x_{k+1} &= x_k + \sum_{j=1}^{s} b_j P_j \\ P_j &= f\left(t_k + c_j h, x_k + h \sum_{i=1}^{s} a_{ij} P_i\right) \end{aligned}, \tag{2.33}$$

where $c_i = \sum_{j=1}^{s} a_{ij}$, and $a_{ij} \in \mathbb{R}$. Constants a_{ij} define whether the scheme is explicit or implicit. If the matrix given by the a_{ij}'s is lower triangular, with all the elements on its principal diagonal equal to zero, that is, with $a_{ij} = 0$ for $j = i, \ldots, s$, then the scheme is explicit and is implicit otherwise. One well-known member of this family, named the *classic Runge–Kutta method* or the *RK4 method*, is the four stage ($s = 4$) method given by

$$x_{k+1} = x_k + \tfrac{h}{6}\left(K_1 + 2K_2 + 2K_3 + K_4\right). \tag{2.34}$$

with

$$\begin{aligned} K_1 &= f(t_k, x_k), \\ K_2 &= f\left(t_k + \tfrac{h}{2}, x_k + \tfrac{h}{2}K_1\right), \\ K_3 &= f\left(t_k + \tfrac{h}{2}, x_k + \tfrac{h}{2}K_2\right), \\ K_4 &= f(t_k + h, x_k + hK_3). \end{aligned}$$

The RK4 method is explicit, numerically convergent, and consistent of order $\sigma = 4$.

2.3.2 Linear Multistep methods

Linear s-step methods

Multistep methods aim to improve efficiency by retaining and utilizing information from earlier steps instead of discarding it, as single-step methods do. As a result, these methods incorporate multiple previous points and their derivative values. Specifically, linear multistep methods use a linear combination of these past points and the derivatives at these points to approximate x_{k+1}. Any linear *s-step method* can be written in the form

$$\sum_{j=0}^{s} \alpha_j y_{k+j} = h \sum_{j=0}^{s} \beta_j f\left(t_{k+j}, y_{k+j}\right), \tag{2.36}$$

where coefficients α_i, β_i, for $i = 0, \ldots, s$, are real constants. To avoid degenerate cases, one shall assume that $\alpha_s \neq 0$ and that α_0 and β_0 are not both equal to zero. Method (2.36) is explicit when $\beta_s = 0$ and implicit otherwise. These are methods that depend linearly on the vector field f.

This thesis does not delve further into the theory for linear multistep methods as the schemes developed here do not take these types into account.

2.3.3 Example of discretization procedure

Let us consider the system

$$\begin{aligned} \dot{x} &= y - x^3 \\ \dot{y} &= -x^5 \end{aligned}. \tag{2.37}$$

By applying Heun's method to discretize (2.37) one obtains

$$\begin{aligned} \tilde{x}_{k+1} &= x_k + hf_1(x_k, y_k) \\ &= x_k + h(y_k - x_k^3) \\ \tilde{y}_{k+1} &= y_k + hf_2(x_k, y_k) \\ &= y_k - hx_k^5 \end{aligned} \tag{2.38}$$

$$
\begin{aligned}
x_{k+1} &= x_k + \frac{h}{2}\left(f_1(x_k, y_k) + f_1(\tilde{x}_{k+1}, \tilde{y}_{k+1})\right) \\
&= x_k + \frac{h}{2}\left(y_k - x_k^3 + y_k - hx_k^5 - \left(x_k + h(y_k - x_k^3)\right)^3\right) \\
y_{k+1} &= y_k + \frac{h}{2}\left(f_2(x_k, y_k) + f_2(\tilde{x}_{k+1}, \tilde{y}_{k+1})\right) \\
&= y_k + \frac{h}{2}\left(-x_k^5 - \left(x_k + h(y_k - x_k^3)\right)^5\right)
\end{aligned}
\tag{2.39}
$$

From (2.39) one obtains a discrete-time approximation of the continuous solution $x : [t_0, t_f] \longrightarrow \mathbb{R}^n$, with initial conditions (x_0, y_0), given by

$$
\begin{aligned}
x_{k+1} &= x_0 + \frac{h}{2}\sum_{i=0}^{k}\left(2y_i - x_i^3 - hx_i^5 - \left(x_i + h(y_i - x_i^3)\right)^3\right) \\
y_{k+1} &= y_0 + \frac{h}{2}\sum_{i=0}^{k}\left(-x_i^5 - \left(x_i + h(y_i - x_i^3)\right)^5\right).
\end{aligned}
\tag{2.40}
$$

For purposes relevant to this thesis, it is worth noting that, in sequences (2.39) and (2.40), x_{k+1} approximates the integrals

$$
\int_{t_k}^{t_{k+1}} \dot{x}(\tau)d\tau, \quad \text{and} \quad \int_{t_0}^{t_f} \dot{x}(\tau)d\tau, \tag{2.41}
$$

respectively. Anagously, in sequences (2.39) and (2.40), y_{k+1} approximates the integrals $\int_{t_k}^{t_{k+1}} \dot{y}(\tau)d\tau$ and $\int_{t_0}^{t_{k+1}} \dot{y}(\tau)d\tau$, respectively.

3

Discretization methods for r-homogeneous systems

This chapter summarizes several discretization schemes, found in the literature, that preserve some of the properties inherent to weighted homogeneous systems. These include, for example, the asymptotic stability of the equilibrium point, the manner in which solutions converge to the origin, the scalability of the solutions, and so on. Here, we will assume that system (2.5) is r-homogeneous of degree $\mu \in \mathbb{R}$, for some $r \in \mathbb{R}^n_{>0}$, and that the origin is an asymptotically stable equilibrium point for this system.

3.1 Improved Euler's scheme with a state-dependent discretization step

In (Efimov et al., 2017), the application of explicit and implicit Euler methods to approximate solutions of asymptotically stable homogeneous systems is studied. This research focuses on determining the conditions under which the discrete-time solutions of stable dynamical systems, obtained with Euler discretization schemes, converge to the origin, and analyzes the accuracy of both methods. It is highlighted that, while the explicit method may have global convergence limitations for certain degrees of homogeneity, the implicit Euler method offers more robust and global convergence under specific conditions, despite its higher computational complexity. The study also delves into how homogeneity properties impact the scalability of solutions and approximation errors.

In the following we will refer to the discrete-time solutions obtained by the explicit and implicit Euler methods simply as discrete-Euler solutions. With respect to both schemes (explicit and implicit), applied in homogeneous systems, the following was

proved by Efimov and co-authors:

- Homogeneity simplifies the analysis of some properties of the discrete-Euler solutions, because of the scalability between the approximations calculated for different initial conditions and discretization steps. In a strict sense, discrete-Euler solutions are invariant under positive weighted dilations, except for a rescaling on the discretization step. If $(x_k)_{k\in\mathbb{N}}$ is a discrete-Euler solution of (2.5), for an initial condition x_0 and a discretization step $h \in \mathbb{R}_{>0}$, then for every $\lambda \in \mathbb{R}_{>0}$, the sequence $(y_k)_{k\in\mathbb{N}}$, with $y_k = \Delta_\lambda^r(x_k)$, is a discrete-Euler solution of (2.5), for initial condition $y_0 = \Delta_\lambda^r(x_0)$ and discretization step $\bar{h} = \lambda^{-\mu}$, which implies that $(y_k)_{k\in\mathbb{N}}$ can be written in the form

  $$y_{k+1} = y_k + \lambda^{-\mu} h f(y_k) \tag{3.1}$$

  for the explicit-Euler method, and

  $$y_{k+1} = y_k + \lambda^{-\mu} h f(y_{k+1}) \tag{3.2}$$

  for the implicit-Euler method.

- When $\mu = 0$, the properties of every discrete-Euler solution $(x_k)_{k\in\mathbb{N}}$ depend on the size of the discretization step h. If one of the following is satisfied

  $$\sup_{i\geq 0}\|x_i\| < +\infty \tag{3.3}$$

  $$\lim_{i\to\infty} x_i = 0 \tag{3.4}$$

  then any other discrete solutions starting at any $y_0 \in \mathbb{R}^n$, will be bounded or converge to the origin too, for the same discretization step h. In other words, local existence of convergent discrete solutions of (2.5) for one discretization step implies global existence of convergent discrete solutions of (2.5) for the same discretization step. Thus, the convergence of discrete solutions may depend on the size of the chosen discretization step, so convergence to zero with one discretization step does not necessarily imply convergence with a different discretization step value.

- When $\mu \neq 0$, on a sphere the convergence to zero or the bounding (3.3) of the discrete-Euler solutions for a given discretization step implies the same property for a discretization step suitably selected as $\bar{h} = h_0 \left(\frac{\|y_0\|_r}{\rho_0}\right)^{-\mu}$, for any initial condition. Thus, local properties of discrete-Euler solutions for one discretization step, imply global properties for properly scaled discretization step.

3.1.1 Explicit Euler scheme

It is proved that when the homogeneity degree is equal to zero, there always exists a discretization step h for which the explicit-Euler scheme preserves the stability of the

system. This implies that the computed discrete system is homogeneous in the sense of (Tuna & Teel, 2004). When $\mu < 0$, discrete-Euler solutions converge globally to a ball centered at the origin, for sufficiently small discretization steps. Moreover, the radius of this neighborhood tend to zero as the discretization step approaches zero. When $\mu > 0$, discrete-Euler solutions converge locally to zero, for sufficiently small discretization steps. Moreover the local convergence becomes global in the limit when h tends to zero.

The explicit Euler method has some drawbacks when it is applied to discretize dynamical systems whose degree of homogeneity μ is non-zero. In the paper, it is stated that for any discretization step h, the explicit Euler scheme becomes unstable for initial conditions sufficiently small when $\mu < 0$ and for initial conditions sufficiently large when $\mu > 0$. Here, the instability of a discretization scheme refers to the fact that the behavior of the discrete-time solutions does not guarantee convergence to the origin or may even diverge, depending on the sign of μ and the initial conditions.

When the homogeneity degree is positive, the explicit Euler method can generate solutions that diverge to infinity for sufficiently large initial conditions. This means that, instead of approaching the origin (expected behavior of an asymptotically stable system), the numerical approximations diverge without limit. Although for sufficiently small steps, the approximations can converge locally to zero, covering the entire state space as the step tends to zero, the instability for large initial conditions is a significant drawback to the global approximation.

On the other hand, when the homogeneity degree is negative, the solutions approximated by the explicit Euler method do not converge asymptotically to zero. Instead, they converge to a neighborhood of the origin, and this neighborhood contracts as the discretization step tends to zero. The origin is unstable for sufficiently small initial conditions (near the origin), which implies that, although the continuous system converges to the origin in finite time, the discrete solutions will not do so accurately in the vicinity of the equilibrium point. Although convergence to a neighborhood of the origin may be acceptable in many applications, especially if this neighborhood is reduced with the discretization step, the lack of exact convergence to zero is a limitation.

3.1.2 Implicit Euler scheme

In, citeEfimov2017, it was proved that discrete-time solutions approximated with the implicit-Euler method exist for any initial condition and any discretization step under an additional mild condition: If the function f associated with (2.5) is continuously differentiable outside the origin and there exists $h_0 > 0$ such that

$$\det\left(I_n - h_0 \frac{\partial f}{\partial x}(x)\right) \neq 0, \quad \forall x \in \mathbb{R} \setminus \{0\}.$$

Moreover, the discrete-time solutions converge to zero, independently of the initial condition, whenever the level sets of the Lyapunov function associated with (2.5) are convex.

In summary, the explicit Euler method exhibits a significant limitation when applied to homogeneous systems with a nonzero degree of homogeneity, as it does not yield accurate or reliable global approximations of their trajectories. The behavior of the discrete-Euler solutions can be erratic or undesirable (divergence, convergence to a neighborhood instead of the origin) depending on the degree of homogeneity and the region of the state space. This is in contrast to the implicit Euler method, which is shown to be more robust and guarantees convergence of the approximate solutions to zero under smoother conditions, although at the cost of higher computational complexity.

The conclusions obtained in (Efimov et al., 2017) served as the foundation in (Efimov et al., 2019) to propose discretization schemes for weighted homogeneous systems, which improve the explicit and the implicit Euler methods by introducing time rescaling that allows recovering finite-time or fixed-time convergence of the solutions. In (Efimov et al., 2019), it is explored how Euler methods (explicit and implicit), fundamental tools for approximating solutions of differential equations, can be adapted to preserve the asymptotic stability of homogeneous dynamical systems. The main innovation is the use of a discretization step that depends on the state of the system, which guarantees that the discrete-time solutions approach zero, as the continuous-time solutions do. It is shown that this modification not only preserves the convergence of these systems, even in finite or fixed time, but also maintains their robustness to external perturbations. Its authors prove that this technique offers superior accuracy in solution approximation, as validated through simulations. The two schemes proposed in this reference are of consistency order one.

Let $\Phi^f_{x_0} : \mathbb{R} \longrightarrow \mathbb{R}^n$ denotes an integral curve of f, for initial conditions $x_0 \in \mathbb{R}^n$. Let us define for every $x_0 \in \mathbb{R}^n$ the map $\psi^f_{x_0} : \mathbb{R} \longrightarrow \mathbb{R}$ as

$$\psi^f_{x_0}(t) = \int_0^t \|\Phi^f_{x_0}(s)\|_r^\mu ds, \tag{3.5}$$

which is well defined and is invertible when the trajectory stays outside of the origin, i.e., for every $t \in [0, T)$, where T is the time of convergence to the origin and can be infinity. The construction of the schemes proposed in (Efimov et al., 2019) is based on the fact that, when $\mu \neq 0$, a simple change of coordinates (rescaling) of the time for system (2.5) results in an r-homogeneous system of degree $\mu = 0$, which is topologically equivalent to (2.5). As mentioned in (Efimov et al., 2017), for systems with a homogeneity degree of zero, there always exists a discretization step for which the explicit-Euler scheme preserves stability. The main idea of the schemes proposed in (Efimov et al., 2019) is to approximate discrete-time solutions of a homogeneous system by means of applying

Euler methods, with a constant step size, to another system, homogeneous of degree zero and topologically equivalent to the original system, where stability is more manageable.

The topologically equivalent system is given as follows. For simplicity, we will define $x(t) = \Phi^f_{x_0}(t)$ and $y(\tau) = \Phi^f_{x_0}(\psi^{-1}(\tau))$, for $\tau \in [0, \psi_{x_0}(T))$, with the understanding that the above only makes sense if an initial condition is given. Then

$$\begin{aligned} \frac{dy}{dt}(\tau) &= \frac{d(x \circ \psi^{-1}_{x_0})}{dt}(\tau) \\ &= \frac{dx}{ds}(\psi^{-1}_{x_0}(\tau)) \frac{1}{\|x(\psi^{-1}_{x_0}(\tau))\|^\mu_r} \\ &= \frac{1}{\|y(\tau)\|^\mu_r} f(y(\tau)). \end{aligned} \tag{3.6}$$

Thus, asymptotical stability of the origin of (2.5) implies the same for (3.6), and since (3.6) is r-homogeneous of degree zero, for a properly chosen discretization step, discrete-Euler solutions of (3.6) converge asymptotically to the origin for every initial condition.

For the initial value problem $((2.5), x_0)$, with $x_0 \in \mathbb{R}^n$, and for a discretization step $h \in \mathbb{R}_{>0}$, the proposed schemes are given, for every $k \in \mathbb{N}$, by

- **State-dependent explicit Euler scheme:**

$$\begin{aligned} x_{k+1} &= x_k + h \frac{1}{\|x_k\|^\mu_r} f(x_k) \\ t_{k+1} &= t_k + h \frac{1}{\|x_k\|^\mu_r} \end{aligned} \tag{3.7}$$

- **State-dependent implicit Euler scheme:**

$$\begin{aligned} x_{k+1} &= x_k + h \frac{1}{\|x_{k+1}\|^\mu_r} f(x_{k+1}) \\ t_{k+1} &= t_k + h \frac{1}{\|x_{k+1}\|^\mu_r}, \end{aligned} \tag{3.8}$$

where $\|\cdot\|_r$ denotes the r-homogeneous norm defined as

$$\|x\|_r = \left(\sum_{i=1}^n |x_i|^{\frac{\rho}{r_i}} \right)^{\frac{1}{\rho}}, \tag{3.9}$$

for any $x \in \mathbb{R}^n$ and any $\rho > \max\{r_1, \ldots, r_n\}$. If $x_k = 0$ or $x_{k+1} = 0$ the discretization scheme stops at the origin. The rescaling of the discretization step is the key to recovering finite-time or fixed-time convergence rates by discrete-time solutions.

In the paper it is shown that there exists a discretization step $H \in \mathbb{R}_{>0}$ such that for any discretization step $h \in (0, H]$, the discrete-time solutions $(x_k)_{k\in\mathbb{N}}$, generated by the modified Euler methods, possess the following properties, for every initial condition $x_0 \in \mathbb{R}^n$:

- **Global boundedness.** There exists $\gamma \in \mathbb{R}_{>0}$ such that $\sup_{k\in\mathbb{N}}\|x_k\| \leq \gamma\|x_0\|_r$.
- **Convergence to the origin.** $\lim_{k\to\infty} x_k = 0$
- **Convergence types preserving.** If $\mu = 0$ the discrete solutions have an exponential rate. If $\mu < 0$ discrete-time solutions converge in finite time, with the settling time given by $T = \lim_{k\to+\infty} t_k$. If $\mu > 0$ discrete-time solutions converge to a ball centered at the origin in finite time.
- **Rescaling of solutions.** If $(x_k)_{k\in\mathbb{N}}$ is a discrete solution of (2.5), obtained trough either (3.7) or (3.8), at the discrete time instants $t_0, t_1, \ldots$, with discretization step $h \in \mathbb{R}_{>0}$, then for any $\lambda \in \mathbb{R}_{<0}$, the sequence $(y_k)_{k\in\mathbb{N}}$, defined by $y_k = \Delta^r_\lambda(x_k)$, is also a discrete solution of (2.5), constructed with (3.7) or (3.8), with initial state $y_0 = \Delta^r_\lambda(x_0)$, discretization step h, and discrete instant times $\hat{t}_k = \lambda^{-\mu}(t_k)$.

H is called the maximal admissible discretization step and is given by

$$H = \min\left\{1, \frac{a}{\mu}, \frac{2c_2}{a}\right\}, \tag{3.10}$$

where

$$a = \inf_{y\in S_r(1)} \frac{\partial V}{\partial y}(y)f(y), \qquad c_2 = \sup_{y\in S_r(1)} V(y) \tag{3.11}$$

$$v = \sup_{y\in S_r(1)} \sup_{\theta\in(0,1)} f^\top(y)\, \frac{\partial^2 V}{\partial \xi^2}(\xi)\bigg|_{\xi=y \;+\; 0\; f(y)} f(y), \tag{3.12}$$

where $S_r(1)$ denotes the ball of radius one given by $S_r(1) = \{x \in \mathbb{R}^n : \|x\|_r = 1\}$, with $\|\cdot\|_r$ the r-homogeneous norm defined by $\|x\|_r = \left(\sum_{i=1}^n |x_i|^{\frac{\rho}{r_i}}\right)^{\frac{1}{\rho}}$, for every $\rho \geq \max\{r_1, \ldots, r_n\}$.

While the existence of x_{k+1} is straightforward in the explicit case. The implicit method (3.8) requires an additional condition: there exists $h_0 > 0$ such that for every $h \in (0, h_0)$ equation (3.8) admits a solution x_{k+1}, for every $x_k \in \mathbb{R}^n$. This condition can be ensured if f is continuous and homogeneous, provided that the step size h is chosen sufficiently small.

Authors of (Efimov et al., 2019) studied the behavior of (3.7) and (3.8) on perturbed systems of the form

$$\dot{x}(t) = F\left(x(t), d(t)\right), \tag{3.13}$$

with $t \geq 0$, $F : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^n$ and $F(x,0) = f(x)$, with f given as in (2.5) and where $d : \mathbb{R}_{>0} \longrightarrow \mathbb{R}^m$ is a measurable and essentially bounded function. They demonstrated that both proposed schemes preserve the input-to-state stability of (3.13).

They proved that both of the schemes they proposed preserve the input-to-state stability of (3.13), with the same asymptotic gain, in the discrete-time solutions approximated.

In this reference, the numerical properties of the proposed discretization schemes are analyzed by examining the absolute and the relative errors. To define the absolute and relative errors, the authors of (Efimov et al., 2019) adopt a notational simplification, representing these errors as a function on $\mathbb{R}^n$, in terms of two functions Υ and $\hat{\Upsilon}$, where Υ corresponds to the exact solution at a given time instant and $\hat{\Upsilon}$ to its discrete approximation, that is $\Upsilon(x) = x(h\|x\|_r)$ and $\hat{\Upsilon}(x) = x + \frac{h}{\|x\|_r^\mu} f(x)$. The absolute error is defined for every discretization step h and every $x \in \mathbb{R}^n$ by

$$\Xi_r^h(x) = \|\Upsilon(x) - \hat{\Upsilon}(x)\|_r$$

and the relative error is given by

$$\xi_r^h(x) = \frac{\Xi_r^h(x)}{\|\Upsilon(x)\|_r}.$$

Efimov and colleagues showed that, for schemes (3.7) and (3.8), the absolute error function is homogeneous of degree one, whereas the relative error function is homogeneous of degree zero. They also remark that if for any initial condition x in the unitary homogeneous ball, the error $\xi_r^h(x)$ remains sufficiently small for a reasonable choice of h, i.e., if the one-step error of the standard Euler discretization is small on the sphere, then the explicit (3.7) and implicit (3.8) Euler schemes yield a globally uniformly bounded relative error. This result, combined with the assumption that the vector field f associated with the continuous-time system is continuously differentiable, implies that the absolute error decays faster than Kh^2, where K is a positive constant, thereby providing a justification for the numerical accuracy of (3.7). However, they do not provide a formal proof of the numerical convergence nor the consistency order of the schemes.

The authors of the aforementioned reference observed that if $\|x_k\|_r$ is sufficiently large and the homogeneity degree μ is negative, the time $t_{k+1} - t_k = \|x_k\|_r^{-\mu}$ may become excessively small. This can affect the accuracy of the approximation, since a more uniform sampling would be preferable in such cases. They recommend that, in this case, it is preferable to employ the conventional Euler methods (without scaling the time discretization step), which can yield a reasonable approximation efficiency for the solutions under a sufficiently regular sampling.

In summary, the schemes proposed by (Efimov et al., 2019) exploit the property of

the explicit and implicit Euler methods to preserve the convergence type of solutions for homogeneous systems of degree zero. Combined with the topological equivalence between any homogeneous dynamical system of nonzero degree and a degree-zero homogeneous system, this allows the construction of explicit and implicit discretization schemes that preserve the convergence type of the solutions (exponential, finite-time, or nearly fixed-time). In the case of the implicit improved-Euler method the existence of a discrete-time solution is not as straightforward as in the explicit case. Some drawbacks of these discretization schemes include the following: they may require relatively large time steps, which reduces accuracy for systems with negative degrees of homogeneity; if the system is subject to a nonzero input $u(t)$, the trajectory may approach the origin but fail to remain there due to the perturbation; and the maximum admissible step size that ensures preservation of the convergence type is restricted. Moreover, the state-dependent rescaling of the time discretization step enables the discrete-time approximations of the solutions to recover finite- or fixed-time convergence rates, only after an infinite number of iterations. In contrast, within a finite number of steps, the trajectories converge merely to a neighborhood of the origin.

3.2 Consistent discrezitation for r-homogeneous systems

The authors of (Polyakov et al., 2019) address the challenge of discretizing ODEs systems, particularly those that do not satisfy classical regularity assumptions (such as the Lipschitz condition), while preserving their inherent stability properties. The work focuses on systems that exhibit finite-time stability or practical fixed-time stability (nearly-fixed-time stability), desirable features in control and estimation because they ensure that system trajectories reach equilibrium within a bounded time. The main contribution lies in the development of implicit and semi-implicit discretization algorithms that preserve these stability properties for generalized homogeneous systems. This advancement is essential for the digital implementation of control and estimation algorithms, since inconsistent discretizations may lead to loss of accuracy, chattering effects, or even instability.

This work introduces consistent discretization schemes that rely on transforming the original system into an equivalent formulation, which can then be discretized using implicit or semi-implicit methods designed to preserve stability properties. In the context of this article, consistency refers to the requirement that the discretized version of the system should retain the key stability properties that the equilibrium point possessed in the original continuous system.

Here, dilation operators are considered in a more general sense than the one introduced in Chapter 2. In this regard, a *dilation* on $\mathbb{R}^n$ is a one parameter subgroup d that satisfies $\lim_{s\to-\infty}\|d(s)x\| = 0$ and $\lim_{s\to\infty}\|d(s)x\| = +\infty$ uniformly on the unit sphere $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$. Any dilation generate a continuous group of invertible linear

maps $d(s) \in \mathbb{R}^{n\times n}$ and the matrix

$$G_d = \lim_{s\to 0} \frac{d(s) - I_n}{s} \tag{3.14}$$

is called the *generator* of this group. Any anti-Hurwitz matrix $G_d \in \mathbb{R}^{n\times n}$ defines a dilation $d(s) = e^{G_d s}$ in $\mathbb{R}^n$. The *canonical homogeneous norm* $\|\cdot\|_d : \mathbb{R}^n \longrightarrow \mathbb{R}_{>0}$ is defined by

$$\|x\|_d = e^{s_x}, \tag{3.15}$$

where $s_x \in \mathbb{R}$ satisfies $\|d(-s)x\| = 1$ .Although the canonical homogeneous norm has a precise formal definition, its practical computation is highly challenging, in contrast to the more tractable homogeneous norm presented in Chapter 2.

Authors of (Polyakov et al., 2019) developed an implicit discretization scheme for homogeneous systems having a possible discontinuity only at the origin, with negative homogeneity degree and a semi-implicit scheme for systems with positive homogeneity degree. The construction of both schemes is based on transforming the original system into an equivalent one that admits an implicit or semi-implicit discretization scheme, preserving the stability properties of the original system.

Two dynamical systems are said to be topologically equivalent if there exists a homeomorphism that maps solutions of one system into solutions of the other system and vice versa. In general, Euler's implicit scheme does not preserve neither the type of convergence of the solutions nor the topological equivalence between dynamical systems, so the discrete solutions of two topological equivalent homogeneous systems, obtained with the implicit Euler's method may not could be mapped onto each other by means of the homeomorphism Ω that characterizes the topological equivalence, in other words, the following diagram does not necessarily commutes

$$\begin{array}{ccc} \dot{x} = f(x) & \xrightarrow{y=\Omega(x)} & \dot{y} = \tilde{f}(y) \\ \Big\downarrow \textit{Implicit Euler's} & & \Big\downarrow \textit{Implicit Euler's} \\ x_{k+1} = x_k + hf(x_{k+1}) & \xrightarrow{y_{k+1}=\Omega(x_{k+1})} & y_{k+1} = y_k + h\tilde{f}(y_{k+1}) \end{array}$$

The methods proposed in (Polyakov et al., 2019) are based on the principle that a discretization method for (2.5) can retain its asymptotic stability and convergence properties if an equivalent system is considered, for which the implicit Euler's discretization preserves these characteristics. By doing so, one can derive the discretization of (2.5) by using the discretization of the equivalent system and the homeomorphism that demonstrates the equivalence between these two systems. The main idea of the proposed

schemes is illustrated in the following diagram:

$$\begin{array}{ccc} \dot{x} = f(x) & \xrightarrow{y=\Omega(x)} & \dot{y} = \tilde{f}(y) \\ \Big\downarrow \textit{Proposed Scheme} & & \Big\downarrow \textit{Implicit Euler} \\ x_{k+1} = x_k + hf(x_{k+1}) & \xleftarrow{x_{k+1}=\Omega^{-1}y_{k+1})} & y_{k+1} = y_k + h\tilde{f}(x_{k+1}) \end{array}$$

In (Polyakov, 2017) and (Polyakov, 2019) it was proved that any generalized homogeneous ODE is topologically equivalent (homeomorphic on $\mathbb{R}^n$ and diffeomorphic on $\mathbb{R}^n \setminus \{0\}$) to a standard homogeneous one, and any asymptotically stable homogeneous system is topologically equivalent to a quadratically stable system. Here, the consistency of a discretization scheme is understood in terms of the following definitions.

Definition 11 (Consistent finite time stable approximation.)**.** *A possibly set-valued mapping*

$$Q : \mathbb{R}_{>0} \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \tag{3.16}$$

is said to be a consistent discrete-time approximation of the globally uniformly finite-time stable system (2.5) *if it satisfies*

- ***Existence property.*** *For any $\tilde{x} \in \mathbb{R}^n$ and any $h \in \mathbb{R}_{>0}$, there exists $\tilde{x}_h \in \mathbb{R}^n$ s.t. $0 \in Q(h, \tilde{x}, \tilde{x}_h)$ and $\tilde{x}_h = 0$ is the unique solution to $0 \in Q(h, 0, \tilde{x}_h)$.*

- ***Finite-time convergence property.*** *For any $h \in \mathbb{R}_{>0}$, each sequence $(x_k)_{k\in\mathbb{N}}$ generated by the inclusion $0 \in Q(h, x_i, x_{i+1})$, for $i = 0, 1, 2, \ldots$, converges to zero in a finite number of steps, i.e., for any $x_0 \in \mathbb{R}^n \setminus \{0\}$ there exists $i* > 0$ such that $x_i = 0$ for $i > i^*$.*

- ***Approximation property.*** *For any $\epsilon \in \mathbb{R}_{>0}$ and any $R > \epsilon$, there exists $\omega \in \mathcal{K}$ such that any sequence $(x_k)_{k\in\mathbb{N}}$ generated by the inclusion $0 \in Q(h, x_i, x_{i+1})$ satisfies $\|\phi(h, x_i) - x_{i+1}\| \leq h\omega(h)$, provided that $\|x_{i+1}\|, \|x_i\| \in [\epsilon, R]$, where $\phi(\cdot, x_i)$ is a solution to* (2.5) *with initial condition $x(0) = x_i$.*

Definition 12 (Consistent practical fixed-time approximation.)**.** *A map*

$$q : \mathbb{R}_{>0} \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \tag{3.17}$$

is said to be a consistent discrete-time approximation of the globally practically fixed-time stable system (2.5) *if the existence and approximation properties of Definition 11 hold and*

- ***Practical fixed-time convergence property.*** *For any $r > 0$ there exists $N(r) > 0$ such that any sequence $(x_k)_{k\in\mathbb{N}}$, with $x_0 \neq 0$, generated by the equation $q(h, x_i, x_{i+1}) = 0$, for $i = 0, 1, 2, \ldots$, satisfies $\|x_i\| \leq r$ for $i \geq N(r)$, independently of x_0.*

The consistent implicit discretization for r-homogeneous systems like (2.5), with finite-time convergence, proposed in (Polyakov et al., 2019) is given by

$$\begin{aligned} Q(h, x_i, x_{i+1}) &= \tilde{Q}\left(h, \Omega(x_i), \Omega(x_{i+1})\right), \\ \Omega(x) &= \|x\|_d d\left(-ln\|x\|_d\right) x, \end{aligned} \tag{3.18}$$

where $h > 0$ and

$$\begin{aligned} \tilde{Q}(h, y_i, y_{i+1}) &= y_{i+1} - y_i - h\tilde{F}(y_{i+1}), \\ \tilde{F}(y) &= \cap_{\varepsilon>0} \mathrm{co}\tilde{f}\left(y + \varepsilon\mathcal{B} \setminus \{0\}\right), \end{aligned} \tag{3.19}$$

where co denotes the convex closure, $\mathcal{B}$ denotes the unit ball on $\mathbb{R}^n$ and

$$\tilde{f}(y) := \left(\frac{(I_n - G_d)yy^\top P}{y^\top G_d P y} + I_n\right) f\left(\frac{y}{\sqrt{y^\top P y}}\right), \qquad y \in \mathbb{R}^n \setminus \{0\}, \tag{3.20}$$

with G_d given as in (3.14) and P a symmetric matrix satisfying $PG_d + G_d^\top P > 0$ and $P = P^\top > 0$. The authors of (Polyakov et al., 2019) proved that (3.18) is a consistent discrete-time approximation of (2.5), whenever the origin is finite-time stable for this system and the condition

$$z^\top I_n^\top P I_n \left(\frac{(I_n - G_d)zz^\top P}{z^\top P G_d z} + I_n\right) f\left(\frac{z}{\|z\|}\right) < 0 \tag{3.21}$$

is satisfied. They also proved that the existence of a consistent implicit discretization for (2.5), with a finite-time stable equilibrium point, depends on the vector field f satisfying $f(-x) = -f(x)$.

The origin of is said to be *practically fixed-time stable* if it is an equilibrium point for (2.5) and for any $r > 0$ there exists $\overline{T}(r) > 0$ such that for any initial condition x_0 the solution $x(\cdot)$ of (2.5) satisfies $\|x(t)\| \leq r$, for $t \geq \overline{T}(r)$, i.e, if solutions converge to a neighborhood of the origin in finite time. The consistent semi-implicit discretization for r-homogeneous systems like (2.5), with nearly fixed-time convergence, proposed in (Polyakov et al., 2019) is given by

$$q(h, x_i, x_{i+1}) = \bar{q}\left(h, \Phi(x_i), \Phi(x_{i+1})\right), \tag{3.22}$$

where $h > 0$ and

$$\Phi(x) = \|x\|_d d(-\ln\|x\|_d)x, \quad x \in \mathbb{R}^n. \tag{3.23}$$

$$\bar{q}(h, y_i, y_{i+1}) = y_{i+1} - y_i - h\|y_i\|\|y_{i+1}\|\tilde{f}(y_{i+1}), \tag{3.24}$$

with

$$\tilde{f}(z) = \left(\frac{(I_n - G_d) z z^\top P}{z^\top G_d P z} + I_n \right) f\left(\frac{z}{\|z\|} \right), \qquad z \in \mathbb{R}^n \setminus \{0\}, \tag{3.25}$$

with G_d given as in (3.14) and P a symmetric matrix satisfying $PG_d + G_d^\top P > 0$ and $P = P^\top > 0$. In (Polyakov et al., 2019) it was proved that if f is uniformly continuous on the unitary sphere S, the origin is a practically fixed-time equilibrium point for (2.5), and $S \subset W_\alpha(\mathbb{R}^n)$, where $W_\alpha(z) = z - \alpha \|z\| \tilde{f}(z)$, then (3.22) is a consistent discrete-time approximation of (2.5). They also explain how discretizations (3.18) and (3.22) may be applied to controllers and observers. Moreover, they prove that (3.18) and (3.22) preserve the stability and convergence type for r-homogeneous systems with positive and negative homogeneity degree.

In summary, although the proposed methods provide a robust and theoretically sound framework for the consistent discretization of finite- and fixed-time stable systems, their practical implementation may be more demanding due to the complexity of the required transformations, the need to solve implicit equations, and the specific conditions that the systems must satisfy.

3.3 Projection methods preserving Lyapunov functions

In (Grimm & Quispel, 2005), the problem of constructing projection-based numerical integrators (discretization schemes) for dynamical systems, preserving a known Lyapunov function, is addressed. It is mentioned that before this work, geometric integrators capable of preserving Lyapunov functions were at most of second consistency order. The main contribution of the article is the introduction of numerical integrators of arbitrary consistency order, such that when generating a discrete system, they keep a discrete Lyapunov function, which ensures that the schemes reflect the correct dynamic behavior of the system. The main idea behind this scheme is the projection: instead of just calculating the next point in the trajectory and then checking if the Lyapunov function was preserved, the method forces the preservation of the Lyapunov function through a projection step. The key contribution of this approach lies in the integrate-then-project strategy, as it decouples the task of numerical integration (aiming for high-order accuracy) from the task of preserving the system's geometric structure—in this case, the Lyapunov function. Initial value problems considered for the methodology are in the form

$$\dot{y}(t) = f(y(t)), \quad y(t_0) = y_0 \in \mathbb{R}^n, \tag{3.26}$$

for which the origin is assumed to be a stable or asymptotically stable equilibrium. Stability ensures the existence of a Lyapunov function $V : \mathbb{R}^n \longrightarrow \mathbb{R}_{>0}$ satisfying $\dot{V}(y) \leq 0$ for every $y \neq 0$. In this reference, the discretization of (3.26) by a numerical method

is abbreviated in the form

$$y_{k+1} = \Psi_h(y_k), \tag{3.27}$$

where the subindex h indicates the size of the discretization step. The methodology consists of three main steps: initial approximation with a standard Runge-Kutta method, approximation of the Lyapunov function in the next step, and projection of the initial approximation. The authors assert that their schemes preserve the Lyapunov function and the consistency order of the discretization method applied to discretize (3.26) and $\dot{V}$. The proposed methodology is developed solely for the first discretization instance; the authors state that its extension to the following steps is straightforward. Two types of discretization schemes are proposed: projection-based schemes that preserve the decreasing nature of the Lyapunov function, i.e. $V(x_k) \leq V(x_{k+1})$, and symmetric projection schemes that preserve the property $\Phi_t \circ \Phi_{-t}(y_0) = y_0$ of the phase flow of (3.26), i.e, that satisfies $\Psi_h \circ \Psi_{-h}(y_0) = y_0$, for every $y_0 \in \mathbb{R}^n$. A brief overview of the proposed methodology is presented below.

3.3.1 Projection schemes

- **Step 1. Discretization with a standard scheme.** System (3.26) is approximated using any single-step RK discretization method of consistency order p of s-stages. Thus, the approximation of y at the instant time t_{k+1} with this scheme is denoted by $\tilde{y}_{k+1}$ and given by

$$\begin{aligned} \tilde{y}_1 &= \tilde{y}_0 + h \sum_{i=1}^{s} b_i f(g_i), \\ g_i &= y_0 + h \sum_{j=1}^{s} a_{ij} f(g_j), \quad i = 1, 2, \ldots, s. \end{aligned}$$

  The coefficients b_i must satisfy $b_i > 0$, for every $i = 1, \ldots, s$, to guarantee that the Lyapunov function is preserved.

- **Step 2. Approximation of V in the next step.** $V(y(t_{k+1}))$ is approximated also through the same RK method that was used to approximate $\tilde{y}_{k+1}$. Thus

$$V_1 = V(y_0) + h \sum_{i=1}^{s} b_i W(g_i) \tag{3.28}$$

  Where $\dot{V} = -W$.

- **Step 3. Orthogonal projection of $\tilde{y}_{k+1}$ onto a level set.** Finally, y_{k+1} is computed by projecting $\tilde{y}_{k+1}$ on the level set $S = \{x \in \mathbb{R}^n : V(x) = V_1\}$, using the orthogonal projection.

3.3.2 Symmetric projection methods

Numerical methods that satisfies $\Psi_h \circ \Psi_{-h}(y_0) = y_0$, for every $y_0 \in \mathbb{R}^n$, are said to be symmetric or self-adjoint. A reversing symmetry of the phase flow Φ_t of (3.26) is an invertible map R that satisfies

$$R \circ \Phi_t = \Phi_t^{-1} \circ R.$$

The proposed symmetric projections are able to preserve linear reversing symmetries in the sense that they satisfy

$$R \circ \Psi_h = \Psi_h^{-1} \circ R,$$

for every invertible linear transformation R, when $V(R(y)) = rV(y)$ holds for $r \neq 0$. Combining the approach given by (Ascher & Reich, 1999) to enforce conservation of energy and in more general contexts (Hairer, 2000), the authors of (Grimm & Quispel, 2005) proposed the projection method based on a symmetric s-stage Runge-Kutta method given by

$$\tilde{y}_0 = y_0 + \frac{\partial V}{\partial y}^{\top}(y_0)\lambda \tag{3.29}$$

$$g_i = \tilde{y}_0 + h\sum_{j=1}^{s} a_{i,j} f(g_j) \tag{3.30}$$

$$\tilde{y}_1 = y_0 + h\sum_{j=1}^{s} b_{s+1-i} f(g_i) \tag{3.31}$$

$$y_1 = \tilde{y}_1 + \frac{\partial V}{\partial y}^{\top}(y_1)\lambda \tag{3.32}$$

$$V(y_1) = V(y_0) + h\sum_{i=1}^{s} b_i W(g_i). \tag{3.33}$$

The RK scheme used to compute y_1 and $V(y_1)$ must be symmetric, which implies its coefficients satisfies

$$a_{i,j} = b_{s+1-j} - a_{s+1-i,s+1-j}, \qquad b_{s+1-i} = b_i.$$

The authors do not go into detail on how λ is defined.

Several advantages of the proposed schemes, such as the ability to achieve an arbitrarily high order of consistency, compatibility with any one-step method -even beyond the Runge-Kutta family— and the preservation of a Lyapunov function, among others, are emphasized. Some numerical implementations are included to illustrate selected benefits of the method.

The discretization scheme developed by Grimm et al. has some limitations. For example, for the projection-based schemes that are proposed in this reference, the coefficients b_i must be non-negative to maintain the monotonic decrease of the Lyapunov function, thereby ruling out the application of several efficient high-order methods—such as the commonly used Dormand–Prince method. Additionally, determining the value of g_i may involve solving an implicit equation at every time step, which leads to a high computational burden. Moreover, the orthogonal projection is made on a different level set at each discretization step, which in turn adds to the overall computational cost of the scheme.

The authors of (M. P. Calvo et al., 2010) also tackled the problem of numerically integrating ordinary differential equations (ODEs) that have a known Lyapunov function. The methods they proposed are an extension and improvement of the previous work by (Grimm & Quispel, 2005). While Grimm and Quispel's methods were limited to Runge-Kutta formulas with nonnegative quadrature coefficients ($b_i \geq 0$), (M. P. Calvo et al., 2010) introduces a modification that allows the use of more general RK methods, even those with negative coefficients, by incorporating a dense output and an independent Gaussian quadrature formula for the approximation of the Lyapunov function.

The proposed schemes are based on the use of explicit Runge-Kutta methods and are implemented by means of a three-step algorithm:

- **Step 1. Discretization with a standard RK scheme.** An explicit RK scheme of s-stages, with consistency order p is used to approximate $\tilde{y}_{k+1}$ as

$$\begin{aligned} \tilde{y}_{k+1} &= y_k + h \sum_{i=1}^{s} b_i(1) g_i \\ g_i &= f\left(y_k + h \sum_{j=1}^{i-1} a_{ij} g_j \right), \end{aligned}$$

  where the g_i's and the b_i's are given by the dense output of order $\bar{p}$ (usually $p \geq p-1$), with $\bar{s} \geq s$stages:

$$\begin{aligned} y_{k+\theta} &= y_k + h \sum_{i=1}^{\bar{s}} b_i(\theta) g_i, & \theta \in [0,1] \\ g_i &= f\left(y_k + h \sum_{j=1}^{i-1} a_{ij} g_j \right), & i = 1, \ldots, \bar{s} \end{aligned}$$

  This step imposes no restrictions on the coefficients b_i of the underlying RK method, since in the following step a separate Gaussian quadrature formula, whose coefficients $\hat{b}_i$ are always positive, is used in order to preserve the decreasing nature of V.

- **Step 2. Approximation of** $V(x(t_{k+1}))$ This approximation is given by

$$V_{k+1} = V(y_k) + h\sum_{i=1}^{m} \hat{b}_i W(y_{k+\hat{c}_i}),$$

where $\hat{b}_i$ and $\hat{c}_i$ are the coefficients and the nodes of the Gaussian quadrature formula in $[0, 1]$, with m nodes and $y_{k+\hat{c}_i} \approx y(t_k + \hat{c}_i h)$ is computed with the dense output of the scheme. No details are provided on the values that m can take.

- **Step 3. Project** $\tilde{y}_{k+1}$**.** Finally, y_{k+1} is obtained by projecting $\tilde{y}_k + 1$ onto the level set $S_{k+1} = \{x \in \mathbb{R}^n : V(x) = V_{k+1}\}$, that is

$$x_{k+1} = P(\tilde{x}_{k+1}). \tag{3.34}$$

where P is used in (3.34) to denote a projection of $\mathbb{R}^n$ onto S_{k+1}. P must satisfy

$$P(\tilde{x}_{k+1}) = \tilde{x}_{k+1} + \lambda_k(h, x_k)\omega_k(h, x_k), \tag{3.35}$$

where $\omega_k : \mathbb{R}_{>0} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defines the direction of the projection and $\lambda_k : \mathbb{R}_{>0} \times \mathbb{R}^n \longrightarrow \mathbb{R}$ ensures that x_{k+1} belongs to S_{k+1}. Two examples of projections of the type (3.35) are the standard projection defined in (Hairer, Wanner, & Lubich, 2006), which computes $P(\tilde{x}_{k+1})$ by solving the constrained problem of minimizing $\|x_{k+1} - \tilde{x}_{k+1}\|$ subject to $g(x_{k+1}) = 0$, and the projection proposed in (M. Calvo et al., 2006) where, after choosing the direction ω_k, in order to compute λ_k, one has to solve a non-linear equation by using some iterative scheme, for example the Newton's method.

As in the work of Grimm, the discussion in this reference regarding the numerical properties of the schemes and the preservation of V is neither detailed nor rigorous. Nonetheless, examples of implementations that exhibit these properties are provided. A notable drawback of the schemes proposed by Calvo and colleagues lies in the projection step, which requires solving a nonlinear equation at every iteration. The need to compute dense output at every iteration further raises the computational cost of the scheme. Moreover, while these schemes preserve a discrete Lyapunov function associated with the system, they do not ensure that the nature of the convergence to the equilibrium point is maintained.

4

A Lyapunov-Function-Based family of discretization schemes for homogeneous systems

This chapter outlines the key contributions and findings of the thesis. It begins by revisiting method of (Sanchez et al., 2020), which is a a two stage discretization scheme, designed to preserve the stability of the origin, which incorporates the explicit Euler scheme at one stage as an auxiliary method. Subsequently, we introduce a set of assumptions related to system (2.5), and define a collection of auxiliary methods, including the explicit Euler method. Each one of the auxiliary methods gives rise to a specific scheme within our proposed family, which generalizes the method of (Sanchez et al., 2020). Finally, we prove that every member in the proposed family is numerically convergent, consistent of some order σ, depending on the consistency order of the auxiliary method, and maintains the stability characteristics of the equilibrium point, preserving thus the type of convergence of the solutions to the origin, and a Lyapunov function.

4.1 Order one LFB method with explicit Euler

Authors of (Sanchez et al., 2020) proposed a one step discretization method that approximates the solutions of r-homogeneous systems with an asymptotically stable equilibrium point, while preserving the stability of the equilibrium and thus the type of convergence of the solutions. It is a two stage scheme that makes use of a known homogeneous Lyapunov function and a one-parameter family of dilations associated with the homogeneous system being discretized, it is applicable to initial value problems of the form ((2.6), (2.5)) satisfying the following assumption.

Assumption 2. *The following holds for System* (2.5)*:*

- *There exists* $r \in \mathbb{R}^n$ *such that* (2.5) *is* r*-homogeneous of degree* $\mu \in \mathbb{R}$.

- *The origin is an asymptotically stable equilibrium point.*

Because Assumption 2 is satisfied, the existence of an r-homogeneous of degree $m \in \mathbb{R}_{>0}$ and of class C^p (with p being arbitrarily chosen) Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ for (2.5) is ensured by Theorem 5.8 in (Rosier & Bacciotti, 2005). For the implementation of the scheme described in this section, it is essential to have an explicitly known function V of class C^1, because the key idea of the discretization scheme involves obtaining a numerical solution of the projection of System (2.5) on an speciffic level set, and subsequently lifting the approximated values back to $\mathbb{R}^n$, using the information provided by V. This requirement might seem very restrictive for implementing the scheme in real dynamical control systems, however, we consider it justified, as it is a tailor-made method for autonomous, homogeneous, open-loop or closed-loop, systems, for which a formal analysis of certain qualitative properties—such as the stability of the origin or the degree of homogeneity of the system—is typically performed as part of a control design procedure. It is important to highlight that the aim of this method is to serve as a tool for studying quantitative properties of the solutions or of the control laws, ensuring that the properties due to the homogeneity of the system are not lost during the discretization process. Although multiple Lyapunov functions may exist for one particular system, the preference of one Lyapunov function over another does not alter the method's properties, as long as that V is r-homogeneous of positive degree and belongs to class C^1. Additionally, the design of Lyapunov functions for homogeneous systems is extensively addressed in the literature, with examples including works by (Rosier, 1992), (Nakamura et al., 2002), (Moreno & Osorio, 2012), (Polyakov, 2020) and (Sanchez & Moreno, 2019).

The discretization scheme for the IVP ((2.6), (2.5)), proposed in (Sanchez et al., 2020), is given by

$$x_{k+1} = \begin{cases} \Delta^r_{v_{k+1}^{\frac{1}{m}}}(z_{k+1}), & x_k \neq 0, \\ 0, & x_k = 0, \end{cases} \tag{4.1}$$

where:

$$z_{k+1} = \Delta^r_{V^{\frac{-1}{m}}(\tilde{z}_{k+1})}(\tilde{z}_{k+1}), \tag{4.2a}$$

$$\tilde{z}_{k+1} = z_k + h v_k^{\frac{\mu}{m}} \left(f(z_k) + \frac{1}{m} W(z_k) R z_k \right), \quad z_0 = \Delta^r_{V^{\frac{-1}{m}}(x_0)}; \tag{4.2b}$$

$$v_{k+1} = \begin{cases} v_k \exp\left(-W(z_k)h\right), & \mu = 0, \\ \frac{v_k}{\left(1+\frac{\mu}{m} v_k^{\frac{\mu}{m}} W(z_k)h,\right)^{\frac{m}{\mu}}}, & \mu > 0, \\ \begin{cases} \left(v_k^{\frac{-\mu}{m}} - \frac{-\mu h}{m} W(z_k)\right)^{\frac{-m}{\mu}}, & \frac{-\mu h}{m} W(z_k) < v_k^{\frac{-\mu}{m}} \\ 0, & \frac{-\mu h}{m} W(z_k) \geq v_k^{\frac{-\mu}{m}} \end{cases}, & \mu < 0. \end{cases} \tag{4.2c}$$

This subsection is dedicated to explain how (4.2b) and (4.2c) are defined, as well as to highlight some of the properties, advantages and disadvantages of the method. The main idea of the discretization procedure consists in projecting the dynamics of (2.5) onto the Lyapunov level surface S_1 (with S_1 as defined in (2.7)), in order to obtain a discretization of the projected dynamics, using the explicit Euler method, and the lifting it to a discretization of (2.5), using the information about the decreasing nature of the solutions provided by the Lyapunov function V. Let $x \in \mathbb{R}^n \setminus \{0\}$, given the change of variable $z = \Delta^r_{V^{\frac{-1}{m}}(x)}(x)$ (with $\Delta^r_{V^{\frac{-1}{m}}(x)}$ as defined in (2.12)), the derivative of z along the trajectories of (2.5) is given by

$$\dot{z}_i = \sum_{j=1}^{n} \frac{\partial z_i}{\partial x_j}(x) f_i(x), \tag{4.3}$$

where $\frac{\partial z_i}{\partial x_j}(x)$ is computed, for every $i = 1, \ldots, n$, as

$$\frac{\partial z_i}{\partial x_j}(x) = \frac{\partial \left(V^{\frac{-r_i}{m}}(x) x_i \right)}{\partial x_j} = V^{\frac{-r_i}{m}}(x) \delta^i_j - \frac{r_i}{m} V^{\frac{-r_i - m}{m}}(x) \frac{\partial V}{\partial x_j}(x) x_i$$

Recalling that f, V, and W are r-homogeneous of degrees μ, m and $m + \mu$ respectively, one obtains

$$
\begin{aligned}
\dot{z}_i &= \sum_{j=1}^{n} \left[V^{\frac{-r_i}{m}}(x)\delta_j^i - \frac{r_i}{m} V^{\frac{-r_i-m}{m}}(x) \frac{\partial V}{\partial x_j}(x) x_i \right] f_j(x) \\
&= V^{\frac{-r_i}{m}}(x) f_i(x) + \frac{r_i}{m} V^{\frac{-r_i-m}{m}}(x) x_i \sum_{j=1}^{n} -\frac{\partial V}{\partial x_j}(x) f_j(x) \\
&= V^{\frac{-r_i}{m}}(x) f_i(x) + \frac{r_i}{m} V^{\frac{-r_i-m}{m}}(x) x_i W(x) \\
&= \left[\Delta^r_{V^{\frac{-1}{m}}(x)}(f(x)) \right]_i + \frac{r_i}{m} \left[\Delta^r_{V^{\frac{-1}{m}}(x)}(x) \right]_i V^{-1}(x) W(x) \\
&= \frac{f_i(\Delta^r_{V^{\frac{-1}{m}}}(x))}{(V^{\frac{-1}{m}}(x))^{\mu}} + \frac{r_i}{m} z_i V^{-1}(x) \frac{W\left(\Delta^r_{V^{\frac{-1}{m}}(x)}(x) \right)}{V^{\frac{-m-\mu}{m}}(x)} \\
&= V^{\frac{\mu}{m}}(x) \left[f_i(z) + \frac{1}{m} W(z) r_i z_i \right],
\end{aligned}
$$

where δ_j^i is the Kronecker delta. It follows that the dynamics of (2.5), projected onto S_1, is given by

$$
\dot{z} = v^{\frac{\mu}{m}} \left[f(z) + \frac{1}{m} W(z) R z \right], \tag{4.4}
$$

under the use of the state variable $v = V(x)$, where R is the diagonal matrix whose principal diagonal is given by the components of the vector of weights r. Since V is positive definite and r-homogeneous of positive degree m, raising it to the power of $\frac{1}{m}$ yields an r-homogeneous norm. Hence, given a representative element $x \in \mathbb{R}^n \backslash \{0\}$, the application of a dilation with "homogeneous" scale factor equal to the reciprocal of its r-homogeneous norm effectively projects x onto the level set S_1.

Because the function V remains constant on S_1, solutions of (4.4) starting at S_1 remain on S_1 as time evolves and never leave this set. Applying the explicit Euler's method to discretize (4.4), with discretization step $h \in \mathbb{R}_{>0}$, one obtains

$$
z_{k+1} = z_k + h v^{\frac{\mu}{m}} \left[f(z) + \frac{1}{m} W(z) R z \right]. \tag{4.5}
$$

Note that solutions of discrete system (4.5) may leave the set S_1 for some index k. To illustrate this fact, let us consider the system

$$
\begin{aligned}
\dot{x}_1 &= -2\lceil x_1 \rfloor^{\frac{3}{2}} + x_2 \\
\dot{x_2} &= -\lceil x_1 \rfloor^2
\end{aligned} \tag{4.6}
$$

which is r-homogeneous of degree $\mu = 1$ with $r = (2, 3)$. The function $V(x) = \frac{4}{5}|x_1|^{\frac{5}{2}} - x_1 x_2 + \frac{6}{5}|x_2|^{\frac{5}{3}}$ is r homogeneous of degree $m = 5$ with $r = (2, 3)$ an is a Lyapunov function

for (4.6). In this case $W(x) = \left(2\lceil x_1 \rfloor^{\frac{3}{2}} - x_2\right)^2 + 2\lceil x_1 \rfloor^2 \lceil x_2 \rfloor^{\frac{2}{3}} - |x_1|^3$. With discretization step $h = 0.1$ and initial conditions $x_0 = (1, 0)$, from (4.5) one obtains

$$\begin{aligned} z_0 &= \Delta^r_{V^{\frac{-1}{m}}(x_0)}(x_0) = \begin{pmatrix} 1.093 \\ 0 \end{pmatrix} \\ z_1 &= \begin{pmatrix} 1.038 \\ -0.114 \end{pmatrix} \end{aligned}$$

Then $V(z_0) = 1$ and $V(z_1) = 1.0307$, which implies that $v_1 \notin S_1$. This example illustrates that explicit Euler's method, might not faithfully reflect the fact that (4.4) is the projected dynamics of (2.5), which is fundamental in the method of (Sanchez et al., 2020) in order to be able to lift (4.5) to a discretization of (2.5). In this sense, the authors of (Sanchez et al., 2020) proposed the following discretization of (2.5)

$$z_{k+1} = \Delta^r_{V^{\frac{-1}{m}}(\tilde{z}_{k+1})}(\tilde{z}_{k+1}), \quad \tilde{z}_{k+1} = z_k + h v_k^{\frac{\mu}{m}} \left(f(z_k) + \frac{1}{m} W(z_k) R z_k \right), \tag{4.7}$$

for every $k \in \mathbb{N}$ and every $h \in \mathbb{R}_{>0}$, ensuring that the projection is preserved, where (4.7) is nothig but the projection on S_1 of the Euler's explicit discretization of (4.4). The following assumption is made to ensure that this discretization is well defined, since when $\tilde{z}_{k+1}$ is zero, it cannot be projected onto S_1.

Assumption 3. *For every $z \in S_1$ and every $\tau \in \mathbb{R}_{>0}$, the following is satisfied:*

$$z + \tau \left(f(z) + \frac{1}{m} W(z) R z \right) \neq 0.$$

In (Sanchez et al., 2020) are given some sufficient conditions to satisfy Assumption 3, for example the set $\{x \in \mathbb{R}^n : V(x) \leq 1\}$ being convex.

From the change of variable $z = \Delta^r_{V^{\frac{-1}{m}}(x)}(x)$ one obtains $x = \Delta^r_{V^{\frac{1}{m}}(x)} z$, for every $x \in \mathbb{R}^n \setminus \{0\}$. Because W is r-homogeneous of degree $m + \mu$, the derivative of V along the trajectories of (2.5) may be rewriten as

$$\dot{V}(x) = -W\left(\Delta^r_{V^{\frac{1}{m}}(x)}(z)\right) = -V^{\frac{m+\mu}{m}}(x) W(z), \tag{4.8}$$

Considering the previously defined variable $v = V(x)$ one obtains

$$\dot{v} = -v^{\frac{m+\mu}{m}} W(z). \tag{4.9}$$

Solutions of (4.9) may be easily computed in terms of the integral

$$\hat{W}_0(t) = \int_0^t W\left(z(\tau)\right) d\tau, \tag{4.10}$$

for instance, using separation of variables (and the change of variables formula for integrals), one obtains:

- If $\mu = 0$, then for every $v_0 = v(0) \in \mathbb{R}_{\geq 0}$ and every $t \in \mathbb{R}$:

$$\begin{aligned} \frac{dv}{d\tau} &= -v(\tau)W(z(\tau)) \\ v(t) &= v_0 \exp\left(-\hat{W}_0(t)\right) \end{aligned} \tag{4.11}$$

- If $\mu > 0$, then for every $v_0 = v(0) \in \mathbb{R}_{\geq 0}$ and every $t \in \mathbb{R}$:

$$\begin{aligned} \frac{dv}{d\tau} &= -v^{\frac{m+\mu}{m}}(\tau)W(z(\tau)) \\ v(t) &= \frac{v_0}{\left(1 + \frac{\mu}{m} v_0^{\frac{-\mu}{m}} \hat{W}_0(t)\right)^{\frac{m}{\mu}}} \end{aligned} \tag{4.12}$$

- If $\mu < 0$. Let T be the settling time for (2.5). Then the right-hand side of (4.4) is well defined on the interval $[0, T]$ and the solution of (4.9) is given, for every $v_0 = v(0) \in \mathbb{R}_{\geq 0}$ and every $t \in \mathbb{R}_{>0}$, by

$$\begin{aligned} \frac{dv}{d\tau} &= -v^{\frac{m+\mu}{m}}(\tau)W(z(\tau)) \\ v(t) &= \begin{cases} \left(v_0^{\frac{-\mu}{m}} + \frac{\mu}{m}\hat{W}_0(t)\right)^{\frac{-m}{\mu}}, & \hat{W}_0(t) > \frac{m}{\mu} v_0^{\frac{-\mu}{m}} \\ 0, & t > T \end{cases} \end{aligned} \tag{4.13}$$

Given a discretization step $h \in \mathbb{R}_{>0}$, let us define $t_k = t_0 + kh$, for $k \in \mathbb{N}$. By integrating (4.9) from t_k to t_{k+1} one obtains:

$$v(t_{k+1}) = v(t_k) \exp\left(-\widehat{W}(t_{k+1})\right), \qquad \mu = 0, \tag{4.14a}$$

$$v(t_{k+1}) = \frac{v(t_k)}{\left(1 + \frac{\mu}{m} v(t_k)^{\frac{-\mu}{m}} \widehat{W}(t_{k+1})\right)^{\frac{m}{\mu}}}, \qquad \mu > 0, \tag{4.14b}$$

$$v(t_{k+1}) = \begin{cases} \left(v(t_k)^{\frac{-\mu}{m}} - \frac{-\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}}, & v(t_k)^{\frac{-\mu}{m}} < \frac{\mu}{m}\widehat{W}(t_{k+1}) \\ 0, & v(t_k)^{\frac{-\mu}{m}} \geq \frac{\mu}{m}\widehat{W}(t_{k+1}) \end{cases}, \qquad \mu < 0, \tag{4.14c}$$

where

$$\widehat{W}(t_{k+1}) = \int_{t_k}^{t_{k+1}} W\left(z(\tau)\right) d\tau. \tag{4.15}$$

The integral in (4.15) may be numerically approximated using the explicit Euler's method as follows. Let us consider the dynamical system

$$\dot{\omega} = W(z). \tag{4.16}$$

Explicit Euler's approximation of (4.16), with discretization step $h \in \mathbb{R}_{>0}$, gives us

$$\omega_{k+1} = \omega_k + hW(z(t_k)). \tag{4.17}$$

Note that by integrating (4.17) one obtains

$$\int_{t_k}^{t_{k+1}} W(z(\tau))d\tau = W(z(t_{k+1})) - W(z(t_k)) \approx \omega_{k+1} - \omega_k \tag{4.18}$$

then the explicit Euler's approximation of (4.15), with discretization step h is given, for every $k \in \mathbb{N}$, by

$$\widehat{W}(t_k) \approx hW(z_k). \tag{4.19}$$

The discrete time approximation of v, proposed in (Sanchez et al., 2020), given by (4.2c), follows by substituting (4.19) in (4.14).

The following theorems, proved in (Sanchez et al., 2020), shows the existence of an homeomorphism $\Phi : \mathbb{R}^n \backslash \{0\} \longrightarrow \mathbb{R}_{>0} \times S_1$, that allows one to map solutions of (2.5) into solutions of (4.9) and (4.4), as well as to prove other interesting properties of the discretization scheme.

Theorem 3. *The map $\Phi : \mathbb{R}^n \backslash \{0\} \longrightarrow \mathbb{R}_{>0} \times S_1$, given for every $x \in \mathbb{R}^n \backslash \{0\}$ by*

$$\Phi(x) = \begin{pmatrix} V(x) \\ \Delta^r_{V^{\frac{-1}{m}}(x)}(x) \end{pmatrix}, \tag{4.20}$$

is a homeomorphism and its inverse is given, for every $(v, z) \in \mathbb{R}_{\geq 0} \times S_1$, by

$$\Phi^{-1}(v, z) = \Delta^r_{v^{\frac{1}{m}}}(z).$$

Theorem 4. *If Assumptions 2 and 3 are satisfied, then the origin of $\mathbb{R}^n$ is an asymptotically stable equilibrium point for (4.1) and V is a Lyapunov function for this discrete-time system. Moreover (4.1) preserves the type of convergence of the solutions, i.e., for every $k \in \mathbb{N}$ and every $x_0 = x(0) \in \mathbb{R}^n$, if one considers α as given in (2.20), the following conditions hold:*

1. *If $\mu = 0$ then*

$$V(x_k) \leq V(x_0) \exp(-\alpha h k), \tag{4.21}$$

and the origin is exponentially stable with respect to the homogeneous norm given by $V^{\frac{1}{m}}$.

2. *If* $\mu > 0$ *then*

$$V(x_k) \leq \frac{V(x_0)}{\left(1 + \frac{\mu}{m} V^{\frac{\mu}{m}}(x_0)\alpha h k\right)^{\frac{m}{\mu}}}, \tag{4.22}$$

and the origin is nearly-fixed time stable.

3. *If* $\mu < 0$ *then*

$$V(x_k) \leq \begin{cases} \left(V^{\frac{-\mu}{m}}(x_0) - \frac{-\mu h}{m}\alpha h k\right)^{\frac{m}{-\mu}}, & k < \frac{m}{-\mu\alpha h} V^{\frac{-\mu}{m}}(x_0) \\ 0, & k \geq \frac{m}{-\mu\alpha h} V^{\frac{-\mu}{m}}(x_0), \end{cases} \tag{4.23}$$

and the origin is finite-time stable.

Theorem 5. *If Assumptions 2 and 3 are satisfied, then the discretization scheme given by* (4.1) *is order 1 consistent.*

We exhibit below some r-homogeneous systems, presented in (Sanchez et al., 2020), that exemplify some of the advantages of (4.1) over other well-known discretization schemes. Let us consider the dynamical system

$$\begin{aligned} \dot{x}_1 &= -k_1 \lfloor x_1 \rceil^{\frac{3}{2}} + x_2 \\ \dot{x}_2 &= -k_2 \lfloor x_1 \rceil^2, \end{aligned} \tag{4.24}$$

which is r-homogeneous of degree $\mu = 1$ with $r = (2, 3)$. It can be proven that for every $k_1 \in \mathbb{R}_{>0}$, there exist $\beta \in \mathbb{R}_{>0}$ and $k_2 \in \mathbb{R}_{>0}$ such that the function $V : \mathbb{R}^2 \longrightarrow \mathbb{R}$, given for every $x \in \mathbb{R}^2$ by

$$V(x) = \frac{2}{5} k_1 |x_1|^{\frac{5}{2}} - x_1 x_2 + \frac{3}{5}\beta |x_2|^{\frac{5}{3}}, \tag{4.25}$$

is a Lyapunov function for (4.24). Thus the origin is nearly-fixed time stable. Computing W for (4.25) one obtains

$$W(x) = \left(k_1 \lfloor x_1 \rceil^{\frac{3}{2}} - x_2\right)^2 + k_2 \left(\alpha \lfloor x_1 \rceil^2 \lfloor x_2 \rceil^{\frac{2}{3}} - |x_1|^3\right). \tag{4.26}$$

Although Euler's method converges numerically, when used to approximate the solution of the initial value problem given by (4.24) and $x_0 \in \mathbb{R}^n$, this method produces unbounded trajectories for sufficiently large initial conditions and discretization step, as discussed for example in (Efimov et al., 2017) and (Levant, 2013). Let us consider the parameters $k_1 = 2$, $k_2 = 1$ and $\alpha = 2$. With this parameter values, W is positive definite, which implies that the discretization method described by (4.1) can be applied and the

discrete-time approximation of the solution of (4.24), using (4.1), is given by

$$\begin{aligned} v_{k+1} &= \frac{v_k}{\left(1+\frac{h}{5}v_k^{\frac{1}{5}}\right)^5 W(z_k)} \\ \tilde{z}_{k+1} &= z_k + h v_k^{\frac{1}{5}} g(z_k), \qquad g(z_k) = \begin{pmatrix} \frac{2}{5} z_k^1 W(z_k) - k_1 \lfloor z_k^1 \rceil^{\frac{3}{2}} + z_k^2 \\ \frac{3}{5} z_k^2 W(z_k) - k_2 \lfloor z_k^1 \rceil^2 \end{pmatrix} \\ z_{k+1} &= \operatorname{diag}\left(V^{\frac{-2}{5}}(\tilde{z}_{k+1}), V^{\frac{-3}{5}}(\tilde{z}_{k+1})\right) \tilde{z}_{k+1} \\ x_{k+1} &= \operatorname{diag}\left(v_{k+1}^{\frac{2}{5}}, v_{k+1}^{\frac{3}{5}}\right) z_{k+1} \end{aligned} \tag{4.27}$$

Figure 4.1 shows the output produced by an implementation of (4.27), with discretization step $h = 0.01$ and initial conditions $x_1(0) = 10^5$ and $x_2(0) = 0$. The graph illustrates that the method of (Sanchez et al., 2020) preserves asymptotic stability even with large initial conditions, while Euler's explicit method does not.

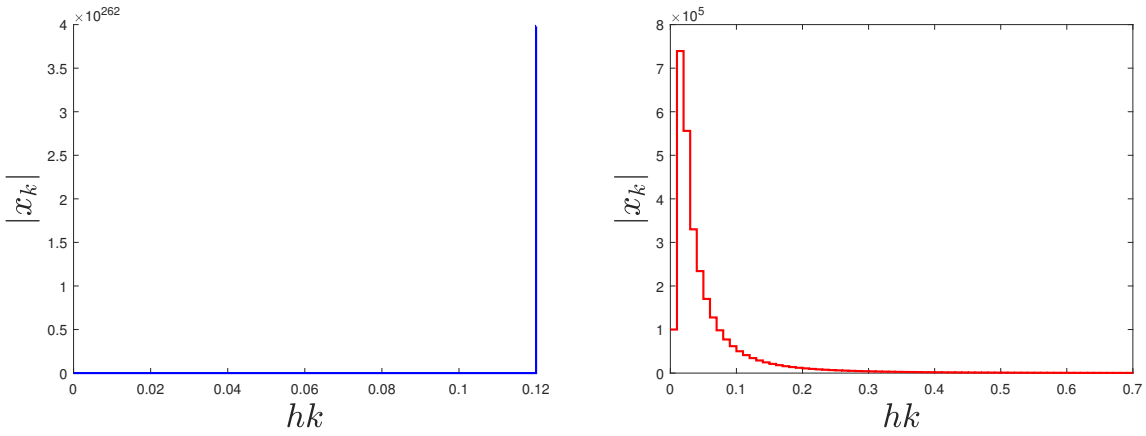

Figure 4.1: Comparison between approximate solutions of (4.24) with explicit Euler's method (left) and with the method of (Sanchez et al., 2020) (right).

Now let us consider the scalar system

$$\dot{x} = -\beta \operatorname{sign}(x), \tag{4.28}$$

with $\beta \in \mathbb{R}_{>0}$. The vector field given by $-\beta \operatorname{sign}(x)$ is discontinuous at the origin, which is asymptotically stable, with Lyapunov function given by

$$V(x) = x^2. \tag{4.29}$$

For $r = 1$, V is r-homogeneous of degree $m = 2$ and (4.28) is r-homogeneous of degree $\mu = -1$, which implies that its solutions converge to the origin in finite-time. Computing

W one obtains

$$W(x) = 2\beta|x|. \tag{4.30}$$

Thus the discrete-time approximation of the solution of (4.28), using (4.1), is given by

$$\begin{aligned} v_{k+1} &= \begin{cases} \left(\sqrt{v_k} - h\beta\right)^2, & v_k > h^2\beta^2 \\ 0, & v_k \leq h^2\beta^2 \end{cases} \\ z_{k+1} &= 1 \\ x_{k+1} &= \sqrt{v_{k+1}}\text{sign}(x_k). \end{aligned} \tag{4.31}$$

Figure 4.2 shows an implementation of the discrete approximation of solutions of (4.28), with both (4.31) and explicit Euler's method, for $\beta = 3$, $h = 0.1$ and $x_0 = 5$. It can be appreciated that Euler's explicit scheme produces numerical chattering.

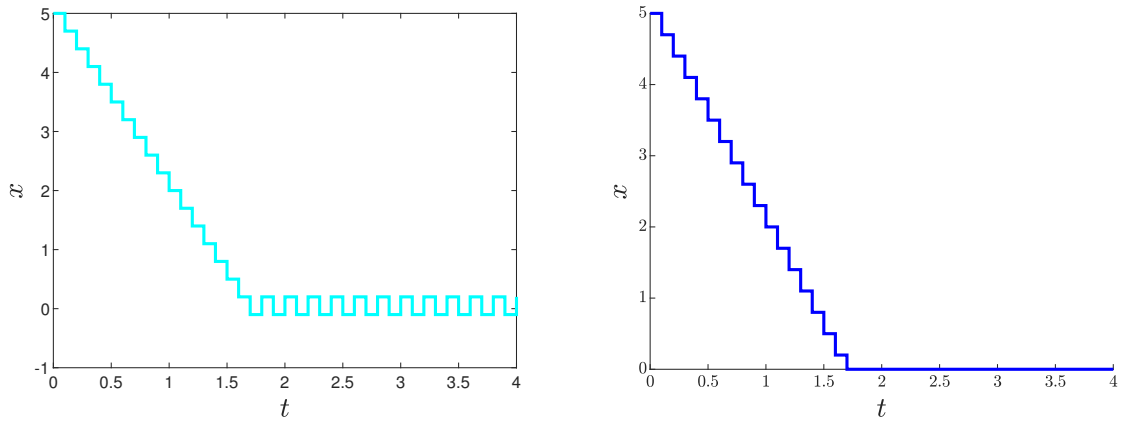

Figure 4.2: Comparison between approximate solutions of (4.61) with explicit Euler's method (left) and with the method of (Sanchez et al., 2020) (right).

4.2 LFB family of discretization schemes

It was demonstrated in (Sanchez et al., 2020) that the scheme (4.1) possesses a consistency order of one, mirroring the order of the explicit Euler method applied to (4.15) and (4.4). Consequently, it is logical to anticipate that employing higher-order consistent methods for the approximation of W and z will yield enhancements over the strategy detailed in (Sanchez et al., 2020). In this section, we introduce a refinement to the scheme from (Sanchez et al., 2020), generating a family of numerically convergent discretization schemes with varying consistency orders. These schemes maintain the stability of the origin, thereby preserving the convergence type of the solutions. We derive this family by generalizing the approximation of (4.15) and (4.4) to include any single-step discretization method that is consistent, numerically convergent, and compliant with specific regularity conditions, rather than limiting the approach to the explicit

Euler method. This modification significantly enhances the performance and computational efficiency found in (Sanchez et al., 2020). Analogous to the method in (Sanchez et al., 2020), our proposed family avoids numerical chattering in discrete-time solutions. Our method prevents the sustained oscillations characteristic of chattering by preserving both the Lyapunov function and the convergence type of the solutions. We reported the results presented in this section in (Silva et al., 2025).

Consider the initial value problem given by ((2.5),(2.6)) under the proviso that (2.5) satisfies Assumption 2. The dynamics of (2.5), projected on the level set S_1, is given by (4.4), with the dynamics of v given by (4.9). Let $F(z) = v^{\frac{-\mu}{m}}\left[f(z) + \frac{1}{m}W(z)Rz\right]$. Let $\mathcal{M}$ denote an arbitrary explicit one-step discretization method in the form (2.24), which is assumed to be numerically convergent and consistent of order $\sigma \in \mathbb{N}$. Define $v_0 = V(x_0)$ and $z_0 = \Delta^r_{V^{\frac{-1}{m}}(x_0)}(x_0)$, for $x_0 \in \mathbb{R}^n$. For a discretization step $h \in \mathbb{R}_{>0}$, denote by

$$\overline{W}_{k+1} = \Psi_W(z_k, t_k, h), \tag{4.32a}$$

$$\overline{z}_{k+1} = \Psi_F(z_k, v_k, t_k, h) \tag{4.32b}$$

the discrete-time approximations of (4.15) and (4.4), respectively, obtained by means of $\mathcal{M}$. In (4.32a), $\overline{W}_{k+1}$ is computed by applying $\mathcal{M}$ to (4.16) and using (4.18), here $\overline{W}_0$ is assumed to be equal to zero. In (4.32b), instead of considering $\bar{z}_k$ as the previous value computed only with $\mathcal{M}$, one considers

$$z_k = \Delta^r_{V^{\frac{-1}{m}}(\bar{z}_k)}(\bar{z}_k), \tag{4.33}$$

i.e., the projection of $\bar{z}_k$ on S_1, to ensure that each element of the sequence $(\bar{z}_k)_{k\in K_h}$ do not stray too far from S_1; z_k is also defined in this way for (4.32a). Equation (4.33) is well defined only if $\mathcal{M}$ satisfies the following assumption.

Assumption 4. *For every $h \in \mathbb{R}_{>0}$ there exists $\varepsilon \in \mathbb{R}_{>0}$ such that*

$$\|\Psi_F(z, v, t, h)\| \leq \varepsilon \tag{4.34}$$

for every $z \in S_1$, $v \in \mathbb{R}_{>0}$, $t \in \mathbb{R}_{>0}$.

It can be demonstrated by induction that if Assumption 4 is satisfied, then for every $h \in \mathbb{R}_{>0}$ and every $k \in K_h$ one has $\overline{z}_k \neq 0$. Since we allow $\mathcal{M}$ to be a general explicit method, Ψ_F is not explicitly known in general, making it almost impossible to derive general sufficient or necessary conditions for $\mathcal{M}$ to satisfy Assumption 4. However, this becomes more manageable when a specific $\mathcal{M}$ is provided. For instance, consider the LFB explicit Euler's method in (Sanchez et al., 2020). For $z \in S_1$ (where $V(z) = 1$), $t \in \mathbb{R}_{>0}$, and $h \in \mathbb{R}_{>0}$ one obtains :

$$\Psi_F(z, v, t, h) = z + v^{\frac{\mu}{m}}h\left[f(z) + \frac{1}{m}W(z)Rz\right] = z + v^{\frac{\mu}{m}}hF(z). \tag{4.35}$$

By using the Euler's theorem (see e.g. Proposition (5.4) in (Rosier & Bacciotti, 2005)) one has

$$\begin{aligned}\frac{\partial V}{\partial z}(z)\cdot F(z) &= \frac{\partial V}{\partial z}(z)\cdot f(z)+\frac{1}{m}W(z)\frac{\partial V}{\partial z}(z)\cdot Rz\\ &= -W(z)+V(z)W(z)=0.\end{aligned}$$

Thus $F(z)$ is orthogonal to $\frac{\partial V}{\partial z}(z)$. A necessary condition to satisfy Assumption 4 is $\Psi(z,1,t,h)\neq 0$, which is not satisfied if z_0 and $F(z_0)$ are collinear. Therefore $\frac{\partial V}{\partial z}(z_0)\cdot z_0 \neq 0$ is a sufficient condition for Assumption 4 to hold. In (Sanchez et al., 2020), it was shown that the set $\{x\in\mathbb{R}^n : V(x)\leq 1\}$ being convex is another sufficient condition for Assumption 4 to hold when $\mathcal{M}$ is the explicit Euler's method. As the consistency order of $\mathcal{M}$ increases, finding conditions that satisfy Assumption 4 becomes very complex. Nevertheless, when implementing the scheme (4.37), one knows the explicit expression of the system (2.5), and selects the auxiliary scheme $\mathcal{M}$, and it is more feasible to verify that Assumption 4 holds for one specific $\mathcal{M}$ and one specific system, as we verified in examples 2 and 4 of Section 6. When both the system in (2.5) and the scheme $\mathcal{M}$ exhibit substantial complexity in their structures, we consider it more feasible and practical to numerically verify the condition at each implementation as follows: given one auxiliary scheme $\mathcal{M}$, one discretization step $h\in\mathbb{R}_{>0}$ and one final time $t_f\in\mathbb{R}_{>0}$, verify that $\overline{z}_k\neq 0$, for every $k\in K_h$. The above implies that Assumption 4 is satisfied for at least one auxiliary scheme and one value of discretization step, on a subset of S_1 and for $t\in[0,t_f]$. This numerical verification is done in examples 1 and 3 of section 6. The set of auxiliary methods employed to create the family of schemes, which represents the primary contribution of this work, is outlined by Assumption 4, together with the following assumption.

Assumption 5. *For every $h\in\mathbb{R}_{>0}$, there exists $\zeta\in\mathbb{R}_{>0}$ such that*

$$\zeta\leq\Psi_W(z,t,h), \tag{4.36}$$

for every $z\in S_1$ and every $t\in\mathbb{R}_{>0}$.

Although Assumption 5 might seem too strong at first sight, it is actually satisfied by a number of discretization methods in the present setting since the restriction W to S_1 is bounded below by α given in (2.20). Consider for example the Heun's method applied to approximate (4.15):

$$\Psi_W(z,t,h)=\frac{h}{2}\left(W(z_k)+W(z_{k+1})\right),$$

where $z_{k+1}=\Delta^r_{V\frac{-1}{m}(\overline{z}_{k+1})}$, with $\overline{z}_{k+1}=z_k+hF(z_k)$. Since $z_k,z_{k+1}\in S_1$ one is able to conclude $\Psi_W(z,t,h)\geq h\alpha$. Using a similar reasoning, it can be shown that all classical Runge–Kutta methods also satisfy this assumption.

By replacing $hW(z_k)$ and $\tilde{z}_{k+1}$ in (4.1) by (4.32a) and (4.32b), respectively, one

obtains a family of LFB discretizations of (2.5), written in the form

$$x_{k+1} = \begin{cases} \Delta^r_{v_{k+1}^{\frac{1}{m}}}(z_{k+1}) & \text{if} \quad x_k \neq 0, \\ 0 & \text{if } x_k = 0, \end{cases} \tag{4.37}$$

for every initial conditions $x_0 \in \mathbb{R}^n$, $z_0 = \Delta^r_{V^{\frac{-1}{m}}(x_0)}(x_0)$, and $v_0 = V(x_0)$, where

$$v_{k+1} = v_k \exp\left(-\overline{W}_{k+1}\right), \qquad \mu = 0, \tag{4.38a}$$

$$v_{k+1} = \frac{v_k}{\left(1 + \frac{\mu}{m} v_k^{\frac{\mu}{m}} \overline{W}_{k+1,}\right)^{\frac{m}{\mu}}}, \qquad \mu > 0, \tag{4.38b}$$

$$v_{k+1} = \begin{cases} \left(v_k^{\frac{-\mu}{m}} - \frac{-\mu}{m}\overline{W}_{k+1}\right)^{\frac{-m}{\mu}}, & \frac{-\mu}{m}\overline{W}_{k+1} < v_k^{\frac{-\mu}{m}}, \\ 0, & \frac{-\mu}{m}\overline{W}_{k+1} \geq v_k^{\frac{-\mu}{m}}, \end{cases} \qquad \mu < 0, \tag{4.38c}$$

and

$$z_{k+1} = \begin{cases} \Delta^r_{V^{\frac{-1}{m}}(\overline{z}_{k+1})}(\overline{z}_{k+1}) & \text{if} \quad v_{k+1} \neq 0, \\ z_k & \text{if} \quad v_{k+1} = 0, \end{cases} \tag{4.39}$$

for every $k \in \mathbb{N}$. Note that when the method $\mathcal{M}$ is the explicit Euler's method, (4.37) recovers the discretization proposed in (Sanchez et al., 2020). Although the structure of (4.37) remains similar to the one proposed in (Sanchez et al., 2020), its flexibility in the choice of the method $\mathcal{M}$ renders the analytical study of (4.37) and its properties a more elaborate endeavor. In (Sanchez et al., 2020) it was shown that the scheme they proposed is consistent of order 1, but the convergence of such a scheme was not proven. In the subsequent we will prove that the consistency order of (4.37) is equal to the order σ of the auxiliary method used to compute (4.38) and (4.39), and that the discretization scheme given by (4.37) is convergent, which implies that the method of (Sanchez et al., 2020) is also convergent. Moreover, we will prove that (4.37) preserves V and the stability of the origin. This results are reported in (Silva et al., 2025).

Theorem 6 (Stability preservation of the origin)**.** *Under Assumption 2, let $\mathcal{M}$ be a convergent discretization scheme, which is consistent of order $\sigma \in \mathbb{N}$ and satisfies Assumption 5. If $\mathcal{M}$ is used to compute* (4.32a) *and* (4.32b) *in* (4.37)*, then the following propositions hold :*

1. *The function V is a Lyapunov function for the discrete-time system given by* (4.37)*.*
2. *If $\mu = 0$ then the origin is exponentially stable with respect to the homogeneous norm $V^{\frac{1}{m}}$ and*

$$V(x_k) \leq V(x_0) \exp\left(-k\zeta_h\right), \tag{4.40}$$

for every $k \in \mathbb{N}$.

3. If $\mu > 0$ then the origin is nearly fixed-time stable and

$$V(x_k) \leq \frac{V(x_0)}{\left(1 + \frac{k\mu}{m} V(x_0)^{\frac{\mu}{m}} \zeta_h\right)^{\frac{m}{\mu}}}, \tag{4.41}$$

for every $k \in \mathbb{N}$.

4. If $\mu < 0$ then the origin is finite-time stable and

$$V(x_k) \leq \begin{cases} \left(V(x_0)^{\frac{-\mu}{m}} - \frac{-k\mu}{m}\zeta_h\right)^{\frac{m}{-\mu}}, & \frac{-k\mu}{m}\zeta_h < V(x_0)^{\frac{-\mu}{m}} \\ 0, & \frac{-k\mu}{m}\zeta_h \geq V(x_0)^{\frac{-\mu}{m}} \end{cases}, \tag{4.42}$$

for every $k \in \mathbb{N}$.

Proof. By construction z_k belongs to S_1 for every $k \in \mathbb{N}$. Moreover, V is r-homogeneous of degree m. It follows that for every $k \in \mathbb{N}$ one has

$$V(x_k) = V\left(\Delta^r_{v_n^{\frac{1}{m}}}(z_k)\right) = v_k V(z_k) = v_k. \tag{4.43}$$

Since v satisfies (4.38) and W is bounded below, it is a strictly decreasing function, which implies $v_{k+1} < v_k$ for every $k \in \mathbb{N}$, when $\mu \geq 0$ and for every $k \in [0, \tau]$, with $\tau = \lfloor \frac{\theta(v_0, z_0)}{h} \rfloor$, when $\mu < 0$. Therefore

$$V(x_{k+1}) - V(x_k) = v_{k+1} - v_k < 0.$$

From Lyapunov's stability theorem for discrete-time systems it follows that V is a strict Lyapunov function for (4.37). To prove the convergence types, consider the solutions of the discrete-time equations (4.38):

$$v_k = v_0 \exp\left(-\overline{W}^k_{sum}\right), \qquad \mu = 0, \tag{4.44a}$$

$$v_k = \frac{v_0}{\left(1 + \frac{\mu}{m} v_0^{\frac{\mu}{m}} \overline{W}^k_{sum}\right)}, \qquad \mu > 0, \tag{4.44b}$$

$$v_k = \begin{cases} \left(v_0^{\frac{-\mu}{m}} - \frac{-\mu}{m}\overline{W}^k_{sum}\right)^{\frac{m}{-\mu}}, & \frac{-\mu}{m}\overline{W}^k_{sum} < v_0^{\frac{-\mu}{m}} \\ 0, & \frac{-\mu}{m}\overline{W}^k_{sum} \geq v_0^{\frac{-\mu}{m}} \end{cases}, \qquad \mu < 0, \tag{4.44c}$$

where $\overline{W}^k_{sum} = \sum_{j=1}^k \overline{W}_j$. By Assumption 5, $(\overline{W}_k)_{k\in\mathbb{N}}$ is bounded below by ζ, hence, inequalities (4.40)–(4.42) follow from (4.44) and (4.43). Since V is r-homogeneous of degree m, positive definite and continuous in $\mathbb{R}\backslash\{0\}$, there exists positive constants C_1 and C_2, such that (see, e.g., (Nakamura et al., 2002) and the proof of Corollary 5.4 in

(Rosier & Bacciotti, 2005))

$$C_2 \leq V(z) \leq C_1, \qquad \forall z \in S_1. \tag{4.45}$$

Note that every $x \in \mathbb{R}^n$ can be written as $x = \Delta^r_{V^{\frac{1}{m}}(x)}\left(\Delta^r_{V^{\frac{-1}{m}}(x)}(x)\right)$, moreover $z = \Delta^r_{V^{\frac{-1}{m}}(x)}(x)$ belongs to S_1 . The function $V^{\frac{1}{m}}$ is an r-homogeneous norm in $\mathbb{R}^n$, which we denote by $V^{\frac{1}{m}}(x) = \|x\|_v$. From r-homogeneity of V one has

$$V(x) = \|x\|_v^m V(z). \tag{4.46}$$

It follows from (4.46) and (4.45) that

$$C_2\|x\|_v^m \leq V(x) \leq C_1\|x\|_v^m. \tag{4.47}$$

The type of convergence of $(x_k)_{k\in K_h}$ is readily deduced from (4.40) to (4.42), and (4.47). □

Given the existence of the homeomorphism (4.20), which maps solutions of (2.5) into solutions of (4.9) and (4.4), and vice versa, the convergence and consistency of (4.37) are easily established by proving the convergence and consistency of (4.38) and (4.39). In (Sanchez et al., 2020), consistency of order one for (4.2c) was shown using L'Hopital's rule, which allowed the authors to calculate the local truncation error. An attempt to use the same method to prove higher-order consistency of (4.38), say of order $\sigma > 1$, would require additional differentiability conditions on the system dynamics f, as derivatives of higher order would then occur in the computations. We utilize a different approach to prove that (4.37) is convergent and consistent of order σ.

Theorem 7 (**Consistency order**). *Under Assumption 2, let $\mathcal{M}$ be an auxiliary method, used to compute* (4.32a) *and* (4.32b) *in* (4.37)*, consistent of order $\sigma \in \mathbb{N}$. Let us suppose that $\mathcal{M}$ satisfy Assumptions 4 and 5 . Then the discretization scheme* (4.37) *is consistent of order σ.*

Proof. Let $(\overline{W}_k)_{k\in K_h}$ and $(v_k)_{k\in K_h}$ be the sequences given by (4.32a) and (4.38). Let $v_0 = V(x_0)$ and $z_0 = \Delta^r_{V^{\frac{-1}{m}}(x_0)}(x_0)$. The local truncation error T^v_{k+1}, for (4.38), is given by

$$T^v_{k+1} = v(t_{k+1}) - v_{k+1},$$

with $v(t_n) = v_n$, by definition. Let us consider $\mu = 0$, leading to the conclusion derived

from (4.38a) and (4.14a) that

$$\begin{aligned} T_{k+1}^{v} &= v(t_k)\exp(-\widehat{W}\,(t_{k+1})) - v_k \exp\left(-\overline{W}_{k+1}\right) \\ &= v_k \exp\left(-\widehat{W}\,(t_{k+1})\right)\left(1 - \frac{\exp(-\overline{W}_{k+1})}{\exp(-\widehat{W}\,(t_{k+1}))}\right) \\ &= v_k \exp(-\widehat{W}\,(t_{k+1}))\left(1 - \exp\left(\widehat{W}\,(t_{k+1}) - \overline{W}_{k+1}\right)\right). \end{aligned}$$

Recall that the local truncation error is computed at each step of discretization, without considering previous errors, thus $z(t_k) - z_k = 0$, which implies that $\widehat{W}\,(t_{k+1}) - \overline{W}_{k+1}$ tends to zero as h tends to zero. Now, by applying the Maclaurin series expansion for the exponential function at zero one obtains

$$\exp(W_R) = 1 + W_R\left(1 + \frac{W_R}{2!} + \frac{W_3^2}{3!} + \cdots\right),$$

with $W_R = \widehat{W}\,(t_{k+1}) - \overline{W}_{k+1}$, thus

$$\begin{aligned} \lim_{h\to 0}\frac{\left|T_{k+1}^{v}\right|}{h^{\sigma}} &= \lim_{h\to 0}\frac{\left|v_k \exp(-\widehat{W}\,(t_{k+1}))W_R\left(-1 - \frac{W_R}{2!} - \frac{W_R^2}{3!} - \cdots\right)\right|}{h^{\sigma}} \\ = \left(\lim_{h\to 0}\frac{|W_R|}{h^{\sigma}}\right)&\left(\lim_{h\to 0}\left|v_k \exp\left(-\widehat{W}(t_{k+1})\right)\left(-1 - \tfrac{W_R}{2!} - \tfrac{W_R^2}{3!} - \cdots\right)\right|\right). \end{aligned}$$

Since the scheme $\mathcal{M}$ is asumed to be consistent of order σ, it follows that $\lim_{h\to 0}\frac{|W_R|}{h^{\sigma}} = 0$. Moreover, because $\widehat{W}(t_{k+1})$ and W_R tend to zero as h tends to zero, the term $v_k \exp\left(-\widehat{W}(t_{k+1})\right)\left(-1 - \frac{W_R}{2!} - \frac{W_R^2}{3!} - \cdots\right)$ tends to v_k as h tends to zero . It follows that $\lim_{h\to 0}\frac{\left|T_{k+1}^{v}\right|}{h^{\sigma}} = 0$, i.e., the discretization method given by (4.38a) is consistent of order σ. Now let us assume $\mu \neq 0$. The local truncation error T_{k+1}^{v} for the schemes given

by (4.38b) or (4.38c), is

$$\begin{aligned}
T_{k+1}^{v} &= \left(v^{\frac{-\mu}{m}}(t_k) - \frac{-\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} - \left(v_k^{\frac{\mu}{m}} - \frac{-\mu}{m}\overline{W}_{k+1}\right)^{\frac{-m}{\mu}} \\
&= \left(v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} \left(1 - \left(\frac{v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\overline{W}_{k+1}}{v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})}\right)^{\frac{-m}{\mu}}\right) \\
&= \left(v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} \left(1 - \left(\frac{v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\overline{W}_{k+1} - v_k^{\frac{-\mu}{m}} - \frac{\mu}{m}\widehat{W}(t_{k+1})}{v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})} + 1\right)^{\frac{-m}{\mu}}\right) \\
&= \left(v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} \left(1 - \left(\frac{\frac{\mu}{m}(-W_R)}{v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})} + 1\right)^{\frac{-m}{\mu}}\right) \\
&= \left(v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} \left(1 - \left(\frac{1}{\frac{m}{\mu}v_k^{\frac{-\mu}{m}} + \widehat{W}(t_{k+1})}(-W_R) + 1\right)^{\frac{-m}{\mu}}\right) \\
&= \left(v_k^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}} \left(1 - (-CW_R + 1)^{\frac{-m}{\mu}}\right),
\end{aligned}$$

where

$$C = \frac{1}{\frac{m}{\mu}v_k^{\frac{-\mu}{m}} + \widehat{W}(t_{k+1})}.$$

By a procedure similar to the one carried out for $\mu = 0$, one approximates $\left(C\left(\overline{W}_{k+1} - \widehat{W}(t_{k+1})\right) + 1\right)^{\frac{-m}{\mu}}$ by the generalized binomial Maclaurin series:

$$\left(C\left(\overline{W}_{k+1} - \widehat{W}(t_{k+1})\right) + 1\right)^{\frac{-m}{\mu}} = \sum_{j=0}^{\infty} \binom{\frac{-m}{\mu}}{j} \left(C\left(\overline{W}_{k+1} - \widehat{W}(t_{k+1})\right)\right)^j$$

where

$$\binom{\frac{-m}{\mu}}{j} = \frac{\frac{-m}{\mu}\left(\frac{-m}{\mu} - 1\right)\left(\frac{-m}{\mu} - 2\right)\cdots\left(\frac{-m}{\mu} - j + 1\right)}{j\ !},$$

therefore

$$\lim_{h\to 0}\frac{\left|T^v_{k+1}\right|}{h^\sigma}=\lim_{h\to 0}\frac{\left|\left(v_k^{\frac{-\mu}{m}}+\frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}}\left(1-\sum_{j=0}^{\infty}\binom{\frac{-m}{\mu}}{j}\left(-CW_R^j\right)\right)\right|}{h^\sigma}$$
$$=\lim_{h\to 0}\frac{\left|\left(v_k^{\frac{-\mu}{m}}+\frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}}\left(CW_R\sum_{j=2}^{\infty}\binom{\frac{-m}{\mu}}{j}\left(-CW_R^j\right)\right)\right|}{h^\sigma}$$
$$=\left(\lim_{h\to 0}\frac{|W_R|}{h^\sigma}\right)\left(\lim_{h\to 0}\left|-C\left(v_k^{\frac{-\mu}{m}}+\frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}}\sum_{j=2}^{\infty}\binom{\frac{-m}{\mu}}{j}\left(-CW_R\right)^j\right|\right).$$

Since W_R tends to zero as h tends to zero, the term

$$\left|-C\left(v_k^{\frac{-\mu}{m}}+\frac{\mu}{m}\widehat{W}(t_{k+1})\right)^{\frac{-m}{\mu}}\sum_{j=2}^{\infty}\binom{\frac{-m}{\mu}}{j}\left(-CW_R\right)^j\right|$$

also tends to zero as h tends to zero. From the consistency of order σ of the method $\mathcal{M}$, it follows that $\lim_{h\to 0}\frac{|W_R|}{h^\sigma}=0$, hence, $\lim_{h\to 0}\frac{\left|T^v_{k+1}\right|}{h^\sigma}=0$. Therefore, (4.38b) and (4.38c) are also consistent of order σ.

On the other hand, the i-th component of the local truncation error T^z_{k+1} of (4.39) is defined by

$$\begin{aligned}\left(T^z_{k+1}\right)_i &= z^i(t_{k+1})-V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\overline{z}^i_{k+1}\\ &= z^i(t_{k+1})-\overline{z}^i_{k+1}+\left(1-V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\right)\overline{z}^i_{k+1}.\end{aligned}\tag{4.48}$$

Because $z(t_{k+1})$ lies in S_1, it is known that $V(z(t_{k+1}))=1$. Then $\left(1-V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\right)$ can be rewritten as

$$\begin{aligned}\left(1-V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\right) &= \frac{1}{V^{\frac{r_i}{m}}(\overline{z}_{k+1})}\left(V^{\frac{r_i}{m}}(\overline{z}_{k+1})-1\right)\\ &= \frac{1}{V^{\frac{r_i}{m}}(\overline{z}_{k+1})}\left(V^{\frac{r_i}{m}}(\overline{z}_{k+1})-V^{\frac{r_i}{m}}(z(t_{k+1}))\right)\end{aligned}\tag{4.49}$$

$\mathcal{M}$ being convergent implies that for every $h\in\mathbb{R}$, there exists $\epsilon\in\mathbb{R}_{>0}$ such that for every $k\in\mathbb{N}$ one has $|z(t_{k+1})-\overline{z}_{k+1}|\leq\epsilon$, . By triangle inequality, it follows that

$$\begin{aligned}|\overline{z}_{k+1}| &= |\overline{z}_{k+1}+z(t_{k+1})-z(t_{k+1})|\\ &\leq |\overline{z}_{k+1}-z(t_{k+1})|+|z(t_{k+1})|\leq\epsilon+|z(t_{k+1})|\leq\epsilon+\sup_{z\in S_1}|z|.\end{aligned}$$

Note that $\sup_{z\in S_1}|z|$ exists because S_1 is compact. Let $\epsilon+\sup_{z\in S_1}=\overline{\varepsilon}$. Since $\mathcal{M}$ satisfies

Assumption 4, it follows that

$$\varepsilon \leq |\overline{z}_{k+1}| \leq \overline{\varepsilon}. \tag{4.50}$$

Since V is a homogeneous norm on $\mathbb{R}^n$, one is able to conclude from (4.50) that there exists constants $d_1, d_2 \in \mathbb{R}_{>0}$ such that $d_1 \leq V(\overline{z}_{k+1}) \leq d_2$. Because the set $[d_1, d_2]$ is compact, the function $g : [d_1, d_2] \longrightarrow \mathbb{R}_{>0}$, that maps $V(\overline{z}_{k+1})$ into $V^{\frac{r_i}{m}}(\overline{z}_{k+1})$ is Lipschitz continuous, and since V is Lipschitz continuous on S_1, it follows that there exist $L_1, L_2 \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} \left|V^{\frac{r_i}{m}}(\overline{z}_{k+1}) - V^{\frac{r_i}{m}}(z(t_{k+1}))\right| &\leq L_1 \left|V(\overline{z}_{k+1}) - V(z(t_{k+1}))\right| \\ &\leq L_1 L_2 \left|\overline{z}_{k+1} - z(t_{k+1})\right|. \end{aligned} \tag{4.51}$$

From (4.49) and (4.51) one obtains

$$\left(1 - V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\right) \leq c_i \left|\overline{z}_{k+1} - z(t_{k+1})\right|,$$

with $c_i = L_1 L_2 d_1^{\frac{-r_i}{m}}$. Moreover, since $\left|z_i(t_{n+1}) - \overline{z}^i_{n+1}\right| \leq |z(t_{n+1}) - \overline{z}_{n+1}|$, it follows by triangle inequality that

$$\begin{aligned} \left|(T^z_{k+1})_i\right| &\leq \left|z_i(t_{k+1}) - \overline{z}^i_{k+1}\right| + \left|\left(1 - V^{\frac{-r_i}{m}}(\overline{z}_{k+1})\right)\overline{z}^i_{k+1}\right| \\ &\leq \left|z_i(t_{k+1}) - \overline{z}^i_{k+1}\right| + c_i \left|z(t_{k+1}) - \overline{z}_{k+1}\right| \left|\overline{z}^i_{k+1}\right| \\ &\leq |z(t_{k+1}) - \overline{z}_{k+1}| + c_i \left|z(t_{k+1}) - \overline{z}_{k+1}\right| |\overline{z}_{k+1}| \\ &= (1 + c_i |\overline{z}_{k+1}|) \left|z(t_{k+1}) - \overline{z}_{k+1}\right|, \\ &\leq (1 + c_i \overline{\varepsilon}) \left|z(t_{k+1}) - \overline{z}_{k+1}\right|. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{h\to 0} \frac{\left|(T^z_{k+1})_i\right|}{h^\sigma} &\leq \lim_{h\to 0} \frac{(1 + c_i\overline{\varepsilon}) \left|z(t_{k+1}) - \overline{z}_{k+1}\right|}{h^\sigma} \\ &= (1 + c_i\gamma_i) \left(\lim_{h\to 0} \frac{|z(t_{k+1}) - \overline{z}_{k+1}|}{h^\sigma}\right) \end{aligned}$$

From the consistency of order σ of $\mathcal{M}$ we have that $\lim_{h\to 0} \frac{|z(t_{k+1}) - \overline{z}_{k+1}|}{h^\sigma} = 0$. Thus, $\lim_{h\to 0} \frac{\left|(T^z_{k+1})_i\right|}{h^\sigma} = 0$, for $i = 1, \ldots, n$. Therefore, the discretization method given by (4.39) is consistent of order σ.

Finally, the local truncation error of (4.37) is given by

$$T^x_{k+1} = x(t_{k+1}) - x_{k+1}. \tag{4.52}$$

Since (4.20) is an homeomorphim, then $x_{k+1} = \Phi^{-1}(v_{k+1}, z_{k+1}) = \Delta^r_{v^{\frac{1}{m}}_{k+1}}(z_{k+1})$ and

$x(t_{k+1}) = \Phi^{-1}\left(v(t_{k+1}), z(t_{k+1})\right) = \Delta^r_{v^{\frac{1}{m}}(t_{k+1})}(z(t_{k+1}))$, with $z(t_{k+1}) = \Delta^r_{V^{\frac{-1}{m}}(x(t_{k+1}))}x(t_{k+1})$. Thus

$$\left(T^x_{k+1}\right)_i = v^{\frac{r_i}{m}}(t_{k+1})\, z_i(t_{k+1}) - v^{\frac{r_i}{m}}_{k+1} z^i_{k+1}. \tag{4.53}$$

It follows that

$$\left(\lim_{h\to 0}\frac{|T^x_{k+1}|}{h^\sigma}\right)_i = \lim_{h\to 0}\frac{|v^{\frac{r_i}{m}}(t_{k+1})\, z_i(t_{k+1}) - v^{\frac{r_i}{m}}_{k+1} z^i_{k+1}|}{h^\sigma}. \tag{4.54}$$

But

$$v_R z_S + z_R v_S = 2\left(v^{\frac{r_i}{m}}(t_{k+1})\, z_i(t_{k+1}) - v^{\frac{r_i}{m}}_{k+1} z^i_{k+1}\right),$$

with $v_R = v^{\frac{r_i}{m}}(t_{k+1}) - v^{\frac{r_i}{m}}_{k+1}$, $v_S = v^{\frac{r_i}{m}}(t_{k+1}) + v^{\frac{r_i}{m}}_{k+1}$, $z_S = z_i(t_{k+1}) + z^i_{k+1}$ and $z_R = z_i(t_{k+1}) - z^i_{k+1}$. It follows that

$$|v^{\frac{r_i}{m}}(t_{k+1})\, z_i(t_{k+1}) - v^{\frac{r_i}{m}}_{k+1} z^i_{k+1}| \;\leq\; |v_R z_S| + |z_R v_S|. \tag{4.55}$$

From (4.55), and the σ consistency order of (4.38) and (4.39), it follows that $\lim_{h\to 0}\frac{|T^x_{k+1}|}{h^\sigma} = 0$. □

Theorem 8 (**Numerical convergence**)**.** *Under the assumptions of Theorem 7, the discretization* (4.37) *is numerically convergent.*

Proof. By a similar reasoning as the one carried out to compute the local truncation error T^v_{k+1} in Theorem 7, one obtains from (4.14) and (4.44) that the global error of (4.38) up to $k \in K_h$ (with K_h defined in Chapter 2) is given by

$$e^v_k = v_0 \exp\left(-\widehat{W}_0(t_k)\right)\left(\sum_{j=0}^{\infty}\frac{\left(\widehat{W}_0(t_k) - \overline{W}^k_{sum}\right)^j}{j!}\right), \qquad \mu = 0, \tag{4.56a}$$

$$e^v_k = \left(v_0^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}_0(t_k)\right)^{\frac{-m}{\mu}}\left(\frac{1}{\frac{m}{\mu}v_0^{\frac{-\mu}{m}} + \widehat{W}_0(t_k)}\sum_{j=0}^{\infty}\binom{\frac{-m}{\mu}}{j}\left(\overline{W}^k_{sum} - \widehat{W}_0(t_k)\right)^j\right), \quad \mu \neq 0. \tag{4.56b}$$

Since $\mathcal{M}$ is convergent then, by definition of global error, for every $h \in \mathbb{R}_{>0}$, there exists $\rho = \max_{k\in K_h}\left|\widehat{W}_0(t_k) - \overline{W}^k_{sum}\right|$, such that $\lim_{h\to 0}\rho = 0$. Thus, for every $k \in K_h$ one has

$$\left|\widehat{W}_0(t_k) - \overline{W}^k_{sum}\right| \leq \rho. \tag{4.57}$$

Since W restricted to S_1 is strictly positive, it follows that $0 \leq \widehat{W}_0(t_k)$. Moreover,

when $\mu \neq 0$ one has $\left(v_0^{\frac{-\mu}{m}} + \frac{\mu}{m}\widehat{W}_0(t_k)\right)^{\frac{-m}{\mu}} \leq v_0$. It follows from (4.56) and (4.57) that

$$\max_{k\in K_h} |e_k^v| \leq \left| v_0 \sum_{j=0}^{\infty} \frac{\rho^j}{j!} \right|, \qquad \mu = 0, \tag{4.58a}$$

$$\max_{k\in K_h} |e_k^v| \leq \left| \frac{\mu}{m} v_0^{\frac{m+\mu}{m}} \sum_{j=0}^{\infty} \binom{\frac{-m}{\mu}}{j} \rho^j \right|, \qquad \mu < 0 \text{ or } \mu > 0. \tag{4.58b}$$

One can deduce from (4.58) that $\lim_{h\to 0} (\max_{k\in K_h} |e_k^v|) = 0$, for every $\mu \in \mathbb{R}$. Thus, (4.38) is convergent.

On the other hand, since $\mathcal{M}$ is assumed to be convergent, from the definition of global error it follows that for every h there exists $\rho_{\overline{z}} = \max_{k\in K_h} |z(t_k) - \overline{z}_k|$, which satisfies $\lim_{h\to 0}(\rho_{\overline{z}}) = 0$. By a similar reasoning to that used to compute the local truncation error T_{k+1}^z in Theorem 7, the i-th component of the global error of (4.39) up to $k \in K_h$ is given by

$$\begin{aligned} |(e_k^z)_i| &= \left| z_i(t_k) - \overline{z}_k^i + \left(\frac{1}{V^{\frac{r_i}{m}}(\overline{z}_k)} \left(V^{\frac{r_i}{m}}(\overline{z}_k) - V\left(z(t_k)\right) \right) \right) \overline{z}_k^i \right| \\ &\leq (1 + c_i\gamma_i) \, |z(t_k) - \overline{z}_k| \end{aligned}$$

Thus $(\max_{k\in K_h} |e_k^z|)_i \leq (1 + c_i\gamma_i)\, \rho_{\overline{z}}$, and it follows that $\lim_{h\to 0} \max_{k\in K_h} |e_k^z| = 0$, i.e., (4.39) is convergent. As in Theorem 7, convergence of (4.37) follows from consistency of (4.38) and (4.39). □

4.2.1 Implementations

Implementation 1

Consider the following system on $\mathbb{R}^2$

$$\begin{aligned} \dot{x}_1 &= x_2 - x_1^3, \\ \dot{x}_2 &= -x_1^5, \end{aligned} \tag{4.59}$$

which is r-homogeneous of degree $\mu = 2$ for $r = [1, 3]^\intercal$. The function $V : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$V(x) = \tfrac{1}{6}x_1^6 - \tfrac{1}{2}x_1 \lceil x_2 \rfloor^{\frac{5}{3}} + \tfrac{1}{2}x_2^2 \tag{4.60}$$

is r- homogeneous of degree $m = 6$. By Young's inequality it can be proved that $\frac{x_1^6}{12} + \frac{x_2^2}{12} \leq V(x)$, for every $x \in \mathbb{R}^n$, then V is positive-definite and that its derivative along the trajectories of (4.59) is $\dot{V}(x) = -W(x)$, with

$$W(x) = x_1^8 + \tfrac{1}{2}|x_2|^{\frac{8}{3}} - \tfrac{5}{6}x_1^6|x_2|^{\frac{2}{3}} - \tfrac{1}{2}x_1^3|x_2|^{\frac{5}{3}}.$$

Using the homogeneity of W, along with Young's inequality, one concludes that $0.0351x_1^8 + 0.0731x_2^2 \leq W(x)$, for every $x \in \mathbb{R}^n$, which implies that W is positive-definite and hence V is a strict Lyapunov function for (4.59), so the origin is an asymptotically stable equilibrium point. In this example, we apply the discretization method (4.37) to numerically approximate the solutions of (4.59), using three different schemes to approximate (4.32a) and (4.32b): Explicit Euler, Heun and RK4. As mentioned in Chapter 2, the three methods are numerically convergent, though they differ in their consistency orders.

The three methods may be viewed as particular cases of the Runge–Kutta family of discretization techniques and they all satisfy Assumption 5. The parameter values used for the implementations are as follows: The total duration of the experiment is $t_f = 2$, and the initial condition at $t = 0$ is $x_0 = [-5, 0]^\intercal$. The main objective of this implementation is to show how the *order of consistency* of $\mathcal{M}$ influences *the global error* of the approximate solutions. Before executing the implementations, it was numerically verified that $\|z_k\| \neq 0$ for every $k \in K_h$ (Figure 4.3), across all selected values of h (i.e., h =0.001, 0.01, 0.05 and 0.1). This confirms that Assumption 4 is satisfied, at least for the chosen parameters and the given initial condition. Figure 4.4 displays the approximations for each scenario, while Figure 4.5 shows that as the discretization step increases, the discrepancy in the global error of the approximate solutions becomes more pronounced. This behavior is expected due to the order of consistency of $\mathcal{M}$ and, by extension, the LFB method. The maximum values of the global errors obtained for discretization steps $h = 0.001, 0.01, 0.05, 0.1$, using each of the three methods are provided in Table 4.1. To calculate the global errors, we used an approximation obtained with Matlab's ODE45 method as a reference, with a maximum discretization step of $h = 0.0001$ and absolute tolerance equal to 10^{-10}.

Table 4.1: Maximum global errors.

h	$e_{\max}^{LE}$	$e_{\max}^{LH}$	$e_{\max}^{LRK4}$
0.001	0.682566	0.308292	0.010251
0.01	7.062057	2.946040	0.994293
0.05	21.355208	11.859379	5.994323
0.1	25.481345	24.830830	12.208390

Maximum global error of (4.37), using explicit Euler's method.
Maximum global error of (4.37), using Heun's method.
Maximum global error of (4.37), using order 4 Runge–Kutta method.

Implementation 2

The following system

$$\dot{x} = -x^3, \tag{4.61}$$

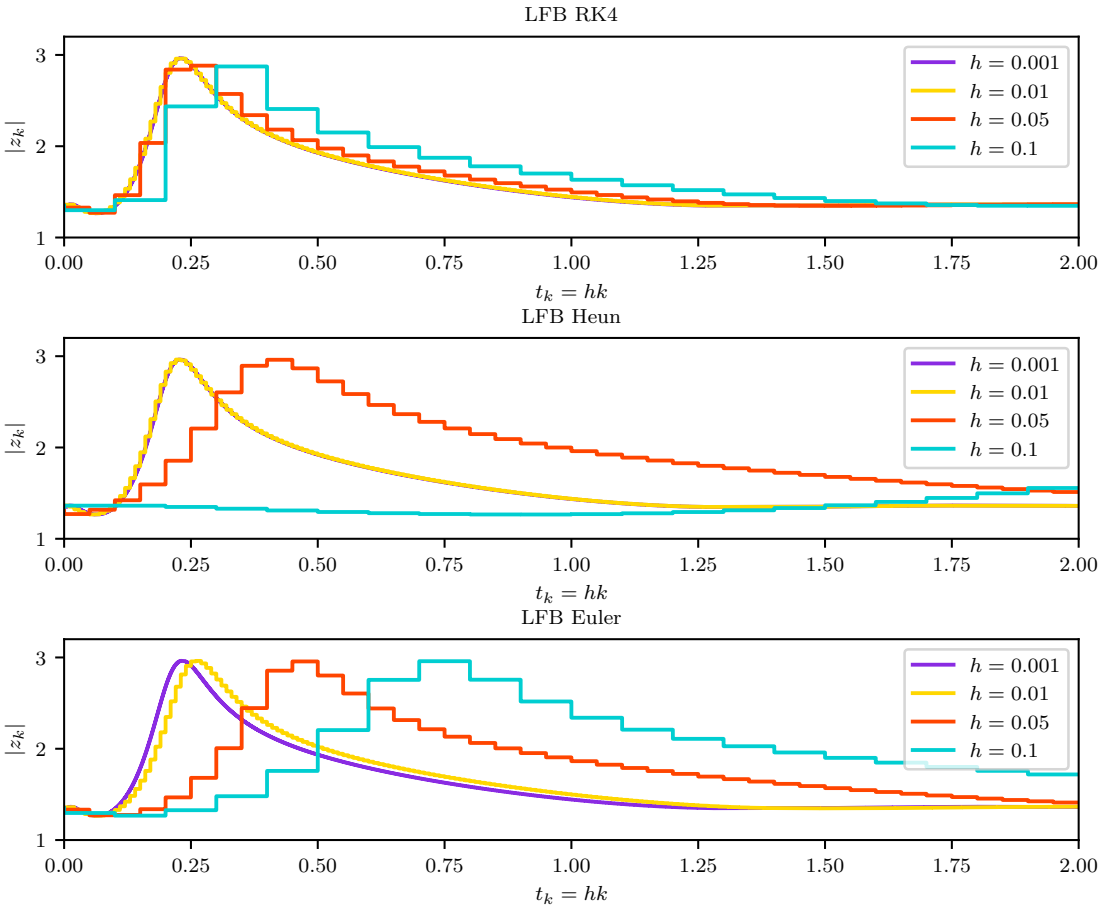

Figure 4.3: Implementation 1. Norm of z_k for different values of h, using three methods to compute (4.39). Numerical verification of Assumption 4.

is r-homogeneous of degree $\mu = 1$, for $r = \frac{1}{2}$. The function given by $V(x) = x^2$ is an strict Lyapunov function for this system, moreover is r-homogeneous of degree $m = 1$. From (4.4) one obtains

$$\dot{z} = v\left[z^5 - z^3\right], \tag{4.62}$$

and it is easy to prove that the explicit Euler's, the RK4 and the midpoint methods satisfy Assumption 4 for this system., beacuse $S_1 = \{1, -1\}$. Figure 4.6 displays the discrete-time solutions of (4.61), approximated using the LFB midpoint method and the explicit Euler's method. The discretization step size was set on $h = 0.005$. For this experiment, several initial conditions were defined as $x_0 = 5^k$, with k ranking from 0 to 2.5 on intervals of 0.25. For $x_0 \geq 5^2$, the explicit Euler method produces discrete-time solutions that diverge from the origin. In contrast, the solutions obtained via the LFB midpoint method converge in fixed-time to the ball $B_{0.298} = \{x \in \mathbb{R} : x \leq 0.298\}$, for each of the initial conditions set. To exhibit the nearly fixed time convergence, another

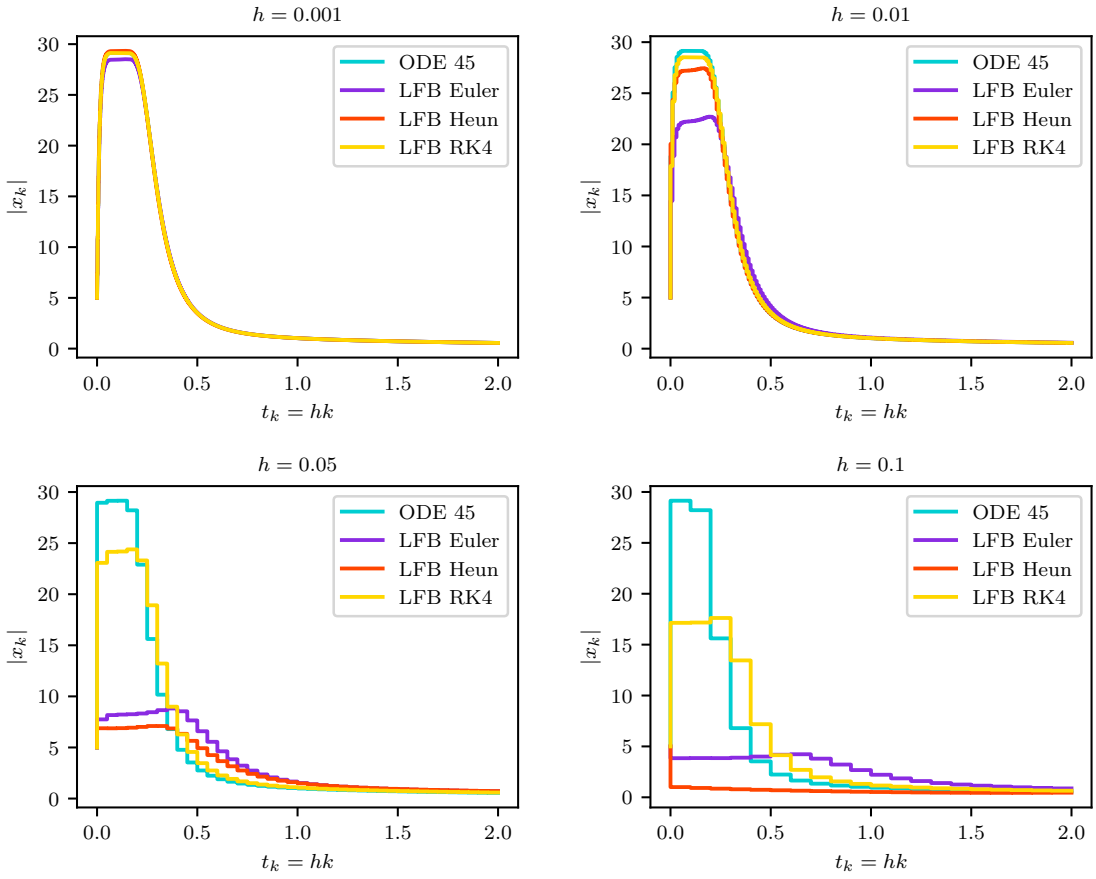

Figure 4.4: Implementation 1. Norm of the states of the LFB discretization of (4.59) for different sizes of h. The Matlbab ODE45 approximation is taken as reference.

implementation was carried out, in which the discretization step was set at $h = 0.01$, and several initial conditions were defined: $x_0 = 5, 50, 500, 5000$. The continuous-time solution of (4.61) is given by

$$x(t) = \frac{x_0}{\sqrt{1 + 2tx_0^2}}.$$

In Figure 4.7 it can be appreciated that the discrete-time solutions, obtained with the LFB Midpoint method for the initial conditions given, converge to the ball $S_2 = \{x \in \mathbb{R} : |x| \leq 0.2\}$ in $t \leq 0.2$, i.e., in fixed-time.

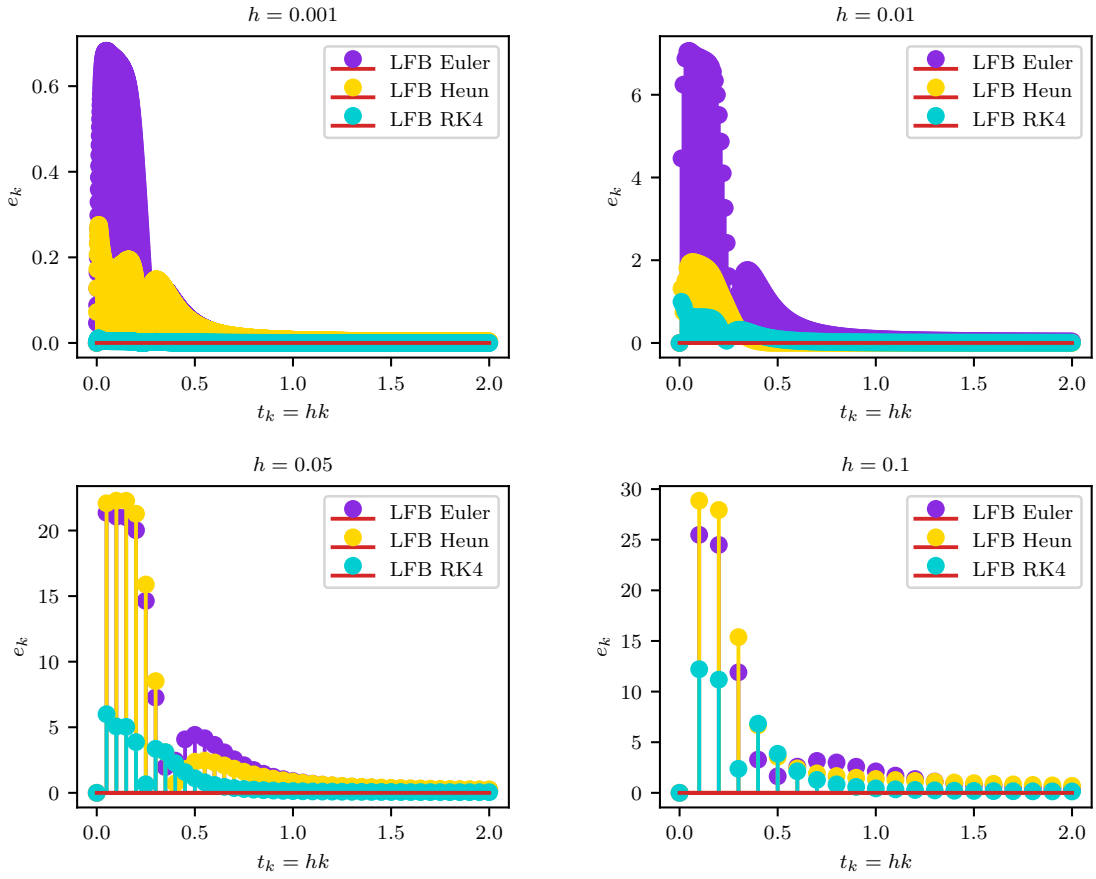

Figure 4.5: Implementation 1. Global errors up to k for the LFB discretization of (4.59) for different values of h.

Implementation 3

Let us now turn to another system on $\mathbb{R}^2$ given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\kappa \left\lceil \lceil x_1 \rfloor^{\frac{9}{4}} + 27 \lceil x_2 \rfloor^3 \right\rfloor^{\frac{2}{9}}. \end{aligned} \tag{4.63}$$

With $\kappa \geq 1$, the origin of $\mathbb{R}^2$ is an asymptotically stable equilibrium point for (4.63), which is readily established as follows. Taking $r = [4, 3]^\intercal$ one shows by inspection that (4.63) is r-homogeneous of degree $\mu = -1$, which implies that the solutions converge to the origin in finite time. The differentiable function

$$V(x) = |x_1|^3 + \lceil x_1 \rfloor^{\frac{9}{4}} x_2 + \tfrac{27}{4} |x_2|^4$$

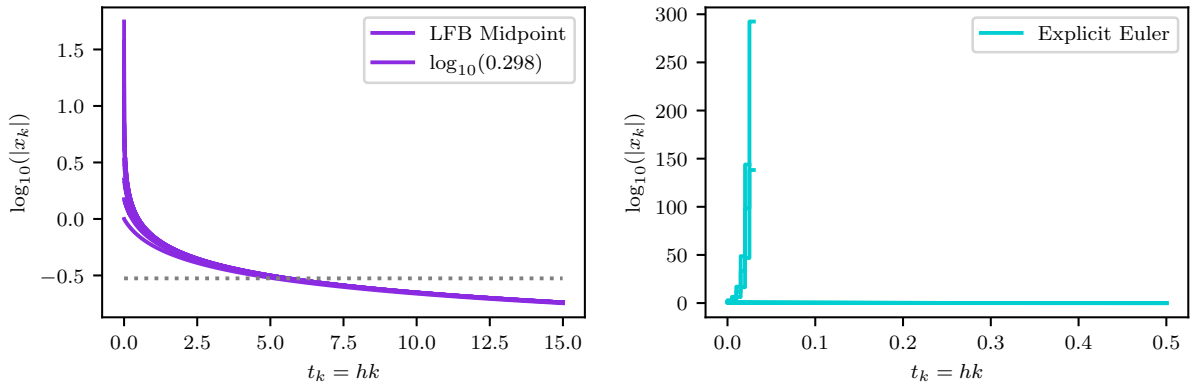

Figure 4.6: Implementation 2. $\log_{10}$ of the norm of the states of the discretization of (4.61) with LFB Midpint method (left) and Explicit Euler's method (right).

is r-homogeneous of degree $m = 12$. By applying Young's inequality to the second term of V we obtain

$$V(x) \geq (1 - \tfrac{3}{4}) |x_1|^3 + (\tfrac{27}{4} - \tfrac{1}{4}) |x_2|^4 .$$

Hence, V is positive-definite, and its derivative along the trajectories of (4.61) is given by $\dot{V}(x) = -W(x)$, with

$$W(x) = \kappa \left| \lceil x_1 \rfloor^{\frac{9}{4}} + 27 \lceil x_2 \rfloor^3 \right|^{\frac{11}{9}} - \left(3 \lceil x_1 \rfloor^2 + \tfrac{9}{4} |x_1|^{\frac{5}{4}} x_2 \right) x_2,$$

which is positive-definite for a sufficiently large value of $\kappa > 0$ (see for example Cruz-Zavala and Moreno, 2017, Lem. 12). Thus, V serves as a strict Lyapunov function for (4.63). We employ the LFB fourth–order Runge–Kutta method and the LFB Heun scheme, both described in (4.37), to numerically approximate the solutions of (4.63). The parameters used in these simulations are the following. The total simulation time is $t_f = 280$, and the initial condition at $t = 0$ is $x_0 = [-12^3, 14^4]^\intercal$. The discretization step is set to $h = 0.1$, with $\kappa = 1$. Prior to running the simulations, we numerically checked that $\|z_k\| \neq 0$ for these parameters (Figure 4.8), confirming that Assumption 4 is fulfilled, at least for this choice of values. The same verification was performed using a larger step size $h = 1$, for which Assumption 4 is not satisfied. Figure 4.9 displays the norm of the computed trajectories, showing that, in both cases, the approximated solutions reach zero exactly in finite time. The sequence $(x_k)_{k \in K_h}$ converges to zero at $t_k = 243.2$ when using the LFB Heun scheme, and at $t_k = 253$ when using the LFB RK4 method.

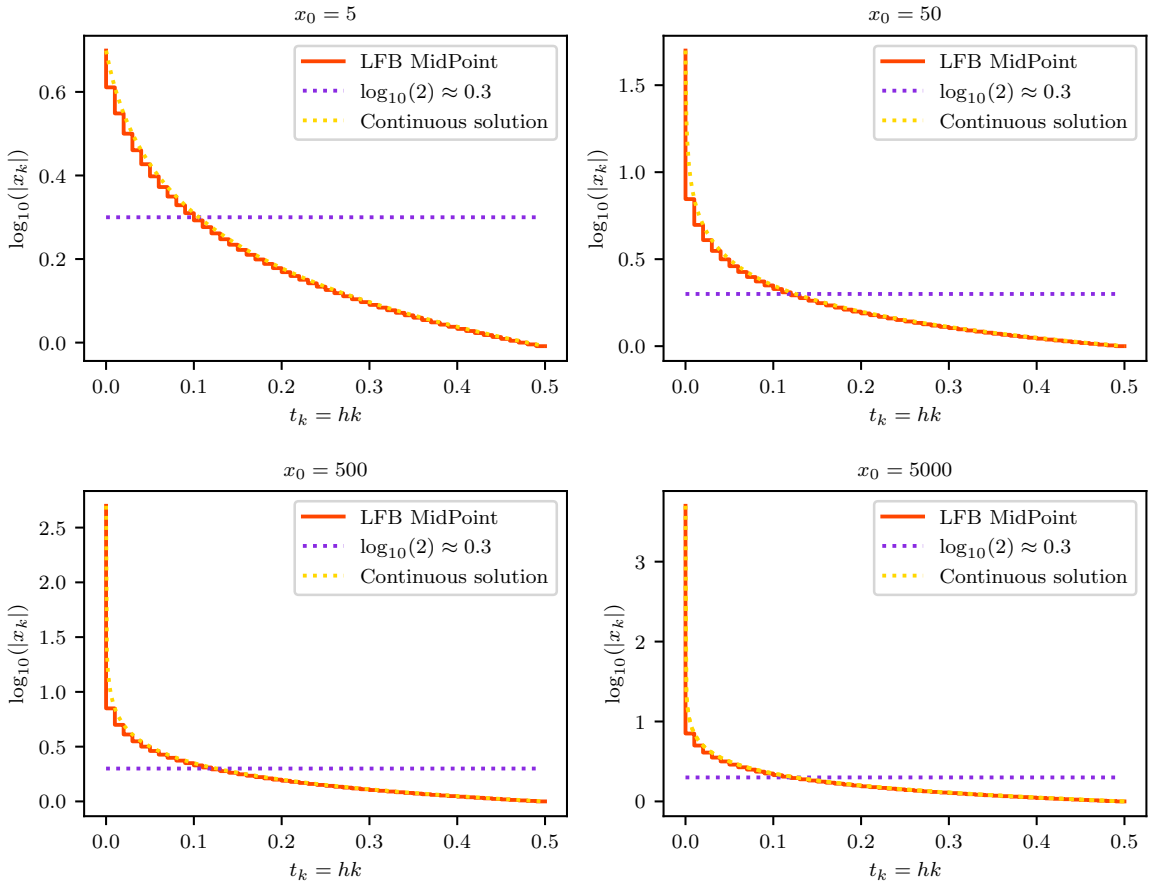

Figure 4.7: Implementation 2. $\log_{10}$ of the norm of the states of the LFB discretization of (4.61) with MidPoint method for different initial conditions.

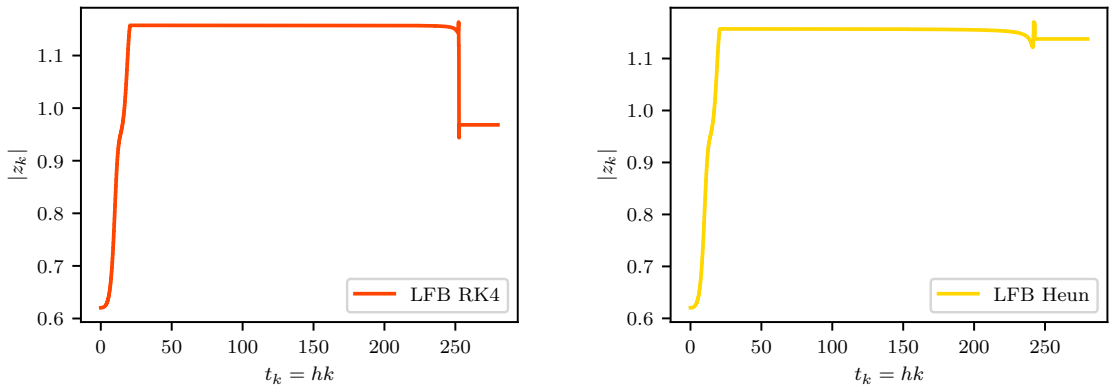

Figure 4.8: Implementation 3. Norm of z_k using two methods to compute (4.39). Numerical verification of Assumption 4.

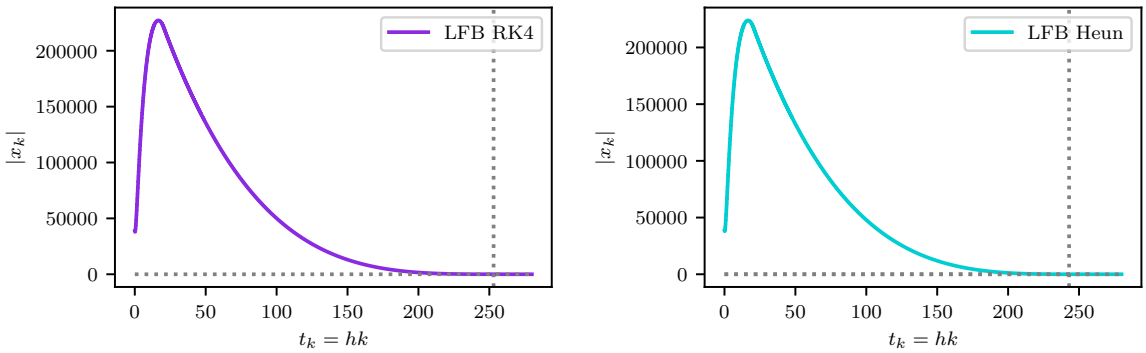

Figure 4.9: Implementation 3. Norm of the states of the discretization of (4.63), approximating (4.9) and (4.4) with two LFB methods.

Implementation 4

Now let us consider the scalar system

$$\dot{x} = -\beta \text{sign}(x), \tag{4.64}$$

with $\beta \in \mathbb{R}_{>0}$. Note that the vector field given by $-\beta \text{sign}(x)$ is discontinuous at the origin and that the origin is an asymptotically stable equilibrium point for (4.64) . A Lyapunov function for this system is given by

$$V(x) = x^2. \tag{4.65}$$

For $r = 1$, V is r-homogeneous of degree $m = 2$ and (4.64) is r-homogeneous of degree $\mu = -1$, which implies that its solutions converge to the origin in finite-time. Computing

W one obtains

$$W(x) = 2\beta|x|. \tag{4.66}$$

From (4.4) one has for this system

$$\dot{z} = v^{\frac{-1}{2}} \left[3|z|z - 3\text{sign}(z)\right], \tag{4.67}$$

Since $S_1 = 1, -1$, it can be readily shown that the explicit Euler, RK4, and Heun methods all satisfy Assumption 4 for this system. Figure 4.10 displays the discrete-time approximations of (4.64) obtained using the explicit Euler method, along with two schemes from the family proposed in this work: LFB Heun and LFB RK4. The parameters used were $\beta = 3$, discretization step $h = 0.1$, simulation horizon $t_f = 3$, and initial condition $x_0 = 4$. In Figure 4.10, it is clear that the explicit Euler method produces discrete-time trajectories that exhibit chattering, while the two LFB-based schemes do not. This difference arises because the explicit Euler method is not constructed to preserve stability of the origin, unlike the LFB schemes. Consequently, the discrete-time trajectories generated by the LFB methods cannot jump to level sets farther from the origin; instead, they evolve through progressively nested level sets that contract toward the origin.

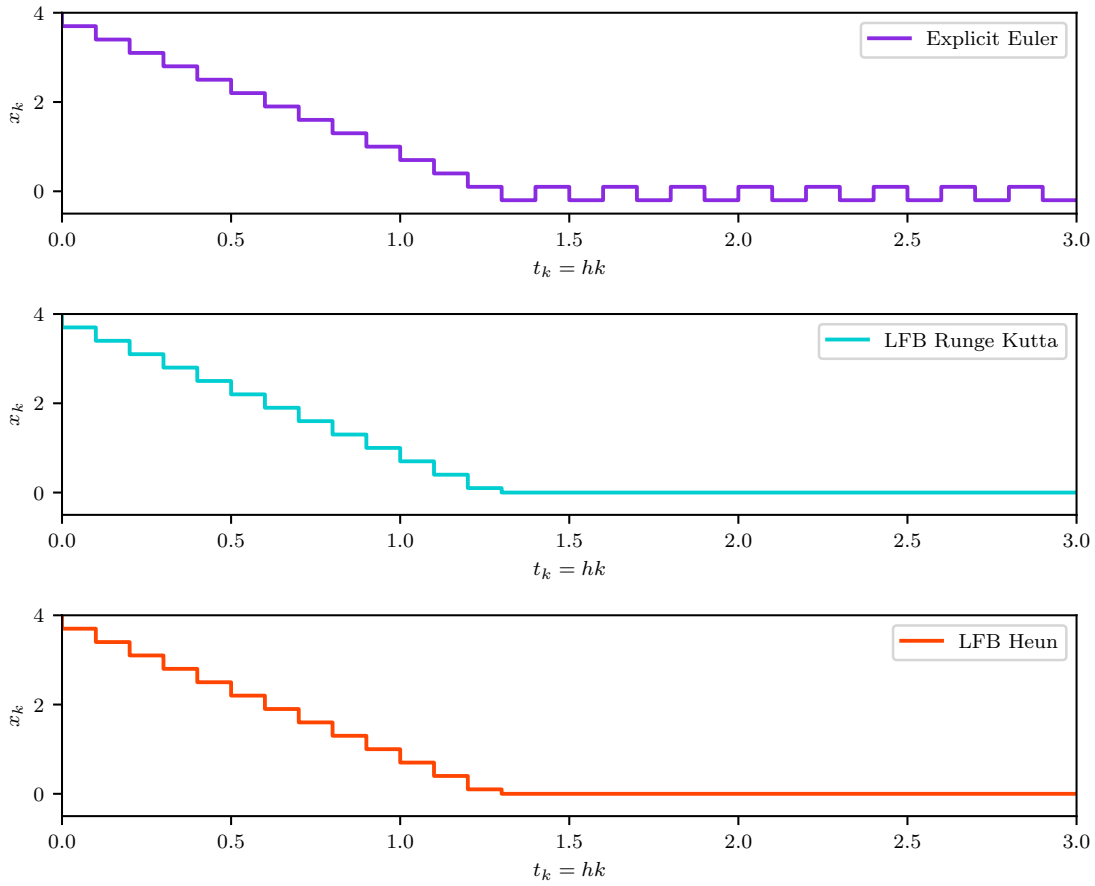

Figure 4.10: Implementation 4. Discretization of (4.64) for explicit Euler's, LFB Heun and LFB RK4.

5

Conclusions and Future Work

The central problem addressed in this thesis is the discretization of homogeneous systems while preserving the asymptotic stability of the origin, a challenge on which conventional schemes typically fall short. General-purpose methods, such as Runge–Kutta schemes, as well as the explicit Euler method, do not guarantee stability; they may produce discrete solutions that diverge or require overly restrictive step-size conditions that compromise computational efficiency. Moreover, they often fail to retain key qualitative properties of the continuous-time dynamics, including the nature of convergence—whether exponential, finite-time, or practically fixed-time.

On the other hand, methods that can guarantee stability, such as implicit Euler under some convexity assumptions, introduce a substantial computational burden. Because they require solving a system of equations at every iteration, their implementation becomes impractical for high-speed simulations or real-time control applications, where efficiency is essential.

Even advanced schemes tailored specifically to this problem, such as the one proposed in (Efimov et al., 2019), do not fully resolve this dichotomy. Although they preserve stability and the convergence type through time scaling, their performance still hinges critically on selecting a suitable step size. Another example is the squeme proposen in (Polyakov et al., 2019), which preserve the convergence type of the solutions, but their practical use is limited. Implementing them requires finding a homogeneous rule with specific properties, an extremely difficult task in general, and their implicit or semi-implicit nature leads to higher computational cost compared to explicit methods. The family proposed in (M. P. Calvo et al., 2010), which is a family of modified Runge–Kutta methods, preserves a given Lyapunov function, however, their numerical properties were

not established, and they do not preserve the convergence type of the solutions.

This highlights the need for for a discretization shceme that is both computationally efficient and robustly stable. This dissertation builds upon the scheme presented in (Sanchez et al., 2020), notable for maintaining key properties, including stability and the specific convergence type, while remaining computationally inexpensive. However, this approach was fundamentally limited by its low (first-order) consistency. Consequently, achieving high accuracy demanded extremely small step sizes, highlighting a clear need for the higher-order methods developed in this thesis.

This thesis introduces an extension to the discretization scheme proposed in (Sanchez et al., 2020), which allows, for the approximation of (4.15) and (4.4), using any one-step, convergent discretization method of order $\sigma \in \mathbb{N}$, rather than being restricted to the explicit Euler method considered in the original scheme. This leads to a family of numerically convergent discretization schemes with different orders of consistency, all of which preserve the convergence characteristics of solutions for asymptotically stable weighted homogeneous systems. Each method in this family is explicit, providing a computational advantage over implicit and semi-implicit methods found in existing literature for this class of systems, as it eliminates the need to solve implicit equations at each discretization step. The proposed schemes involve the explicit knowledge of a Lyapunov function, as well as a projection onto a specific function's level set, which are not required in other, more general and well-known schemes. However, due to these requirements, the schemes we propose do not exhibit inconsistencies that have been found in other schemes, such as Euler's methods, for example chattering, divergence of solutions, instead of consistency. These inconsistencies can be found, for example, when the system's vector field is discontinuous or non-smooth at the origin. Furthermore, the projection is consistently performed onto the same level set, thus eliminating the need to compute a new level set at each discretization step.

It was proven that the order of consistency of each of the schemes in the family we propose equals the order of consistency of the auxiliary method used to approximate $\overline{z}_{k+1}$ and v_{k+1}. Moreover, it was proved that every scheme is numerically convergent. Consistency order property, besides being essential for any discretization method, translates into a reduction of the number of discretization steps needed to achieve a desired level of accuracy for the discrete-time solutions: the higher the order, the fewer discretization steps required, thus reducing the computational cost. Chattering in the discrete-time solutions is observed as a sustained oscillation of the solution; the family of schemes we propose is specifically designed to prevent this situation, as it preserves the Lyapunov function and the type of convergence. Therefore, the discrete-time solutions always move from one level set S_c to inner ones, characterized by smaller values of c.

The family of schemes proposed in this work may be extended to incorporate some linear multi-step discretization methods, taking into account the necessary considera-

tions, without affecting the preservation property of the stability of the origin. These methods are a useful class of schemes for discretizing dynamic systems, especially when the differential equations defining the system are stiff. Stiff equations are those for which standard methods become unstable unless very small discretization steps are used. An interesting direction for future work would be to expand the range of dynamic systems for which the family of methods proposed in this thesis maintains its properties by applying them to non-autonomous systems. This would increase the methods' applicability, especially in control theory, such as in the discretization of observers. Another open problem is the design of a discretization scheme, or even a family of discretization schemes with different consistency orders, satisfying assumptions 4 and 5 for any discretization step, thus eliminating the dependence on a set of auxiliary methods, which may cause the properties of the schemes we propose to be satisfied only for discretization steps within certain value intervals.

Contributions

- Silva Cristina, Sanchez Tonametl, Lizárraga David, Zavala Arturo. (2025). "An Extension of a Lyapunov-Function-Based Discretization Method for Asymptotically Stable Homogeneous Systems". International Journal of Robust and Nonlinear Control. 10.1002/rnc.70290.
- Silva Cristina, Sanchez Tonametl, Lizarraga David, Zavala Arturo. (2024). "Study of a Lyapunov-based discretization method using the midpoint approach". Memorias del Congreso Nacional de Control Automático 2024, pp. 7-12.
- Esperanza Cuellar, Ana C. Silva, David A. Lizárraga, "Conditions to Flatten the Curvature Tensor Associated with a Symmetric Affine Connection", Journal of Mathematics, vol. 2022, Article ID 1528246, 11 pages, 2022.

Bibliography

Acary, V., Brogliato, B., & Orlov, Y. V. (2012). Chattering-free digital sliding-mode control with state observer and disturbance rejection. *IEEE Transactions on Automatic Control*, *57*(5), 1087–1101. https://doi.org/10.1109/TAC.2011.2174676

Acary, V., & Brogliato, B. (2008). *Numerical methods for nonsmooth dynamical systems: Applications in mechanics and electronics*. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75392-6

Acary, V., & Brogliato, B. (2010). Implicit euler numerical scheme and chattering-free implementation of sliding mode systems. *Systems & Control Letters*, *59*(5), 284–293. https://doi.org/10.1016/j.sysconle.2010.03.002

Andrieu, V., Praly, L., & Astolfi, A. (2008). Homogeneous approximation, recursive observer design, and output feedback. *SIAM Journal on Control and Optimization*, *47*(4), 1814–1850. https://doi.org/10.1137/060675861

Ascher, U. M., & Petzold, L. R. (1998). *Computer methods for ordinary differential equations and differential-algebraic equations*. Society for Industrial; Applied Mathematics. https://doi.org/10.1137/1.9781611971392

Ascher, U. M., & Reich, S. (1999). On some difficulties in integrating highly oscillatory hamiltonian systems. In P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, & R. D. Skeel (Eds.), *Computational molecular dynamics: Challenges, methods, ideas* (pp. 281–296). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58360-5_15

Bernuau, E., Polyakov, A., Efimov, D., & Perruquetti, W. (2013). Verification of ISS, iISS and IOSS properties applying weighted homogeneity. *Systems & Control Letters*, *62*(12), 1159–1167. https://doi.org/10.1002/rnc.5308

Calvo, M., Hernández-Abreu, D., Montijano, J. I., & Rández, L. (2006). On the preservation of invariants by explicit runge–kutta methods. *SIAM Journal on Scientific Computing*, *28*(3), 868–885. https://doi.org/10.1137/04061979X

Calvo, M. P., Laburta, M. P., & Montijano, J. I. (2010). Projection methods preserving Lyapunov functions. *BIT Numerical Mathematics*, *50*(2), 223–241. https://doi.org/10.1007/s10543-010-0259-3

Cruz-Zavala, E., & Moreno, J. A. (2017). Homogeneous high order sliding mode design: A Lyapunov approach. *Automatica*, *80*, 232–238. https://doi.org/10.1016/j.automatica.2017.02.039

Drakunov, S. V., & Utkin, V. I. (1989). On Discrete-Time Sliding Modes. *IFAC Proceedings Volumes*, *22*(3), 273–278. https://doi.org/10.1016/S1474-6670(17)53647-2

Efimov, D., Ushirobira, R., Moreno, J. A., & Perruquetti, W. (2018). Homogeneous Lyapunov functions: From converse design to numerical implementation. *SIAM Journal on Control and Optimization*, *56*(5), 3454–3477. https://doi.org/10.1137/17M113753X

Efimov, D., Polyakov, A., & Aleksandrov, A. (2019). Discretization of homogeneous systems using Euler method with a state-dependent step. *Automatica*, *109*, 108546. https://doi.org/10.1016/j.automatica.2019.108546

Efimov, D., Polyakov, A., Levant, A., & Perruquetti, W. (2017). Realization and discretization of asymptotically stable homogeneous systems. *IEEE Transactions on Automatic Control*, *62*(11), 5962–5969. https://doi.org/10.1109/TAC.2017.2699284

Epperson, J. (2002). *An introduction to numerical methods and analysis*. Wiley.

Eugene, I., & Keller, H. B. (1994). *Analysis of numerical methods*. Dover.

Grimm, V., & Quispel, G. R. W. (2005). Geometric integration methods that preserve lyapunov functions. *BIT Numerical Mathematics*, *45*(4), 709–723. https://doi.org/10.1007/s10543-005-0034-z

Grüne, L. (2000). Homogeneous state feedback stabilization of homogeneous systems. *SIAM Journal on Control and Optimization*, *38*(4), 1288–1308. https://doi.org/10.1137/S0363012998349303

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, *13*, 49–52.

Hahn, W. (1967). *Stability of motion*. Springer-Verlag. https://doi.org/10.1007/978-3-642-50085-5

Hairer, E. (2000). Symmetric projection methods for differential equations on manifolds. *BIT*, *40*, 726–734. https://doi.org/10.1023/A:1022344502818

Hairer, E., & Lubich, C. (2014). Energy-diminishing integration of gradient systems. *IMA Journal of Numerical Analysis*, *34*(2), 452–461. https://doi.org/10.1093/imanum/drt018

Hairer, E., Lubich, C., & Wanner, G. (2006). *Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations* (2nd ed., Vol. 31). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/3-540-30666-8

Hairer, E., Nørsett, S. P., & Wanner, G. (1993). *Solving ordinary differential equations i: Nonstiff problems* (2nd ed.). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78862-1

Hairer, E., Wanner, G., & Lubich, C. (2006). *Geometric numerical integration, structure-preserving algorithms for ordinary differential equations*. Springer Berlin, Heidelberg. https://doi.org/10.1007/3-540-30666-8

Hermes, H. (1986). Nilpotent approximations of control systems and distributions. *SIAM Journal on Control and Optimization*, *24*(4), 731–736. https://doi.org/10.1137/0324045

Hermes, H. (1991). Homogeneous coordinates and continuous asymptotically stabilizing feedback controls. In S. N. Elaydi (Ed.), *Differential equations: Stability and control* (pp. 249–260, Vol. 127). Marcel Dekker. http://eudml.org/doc/90504

Hong, Y., Huang, J., & Xu, Y. (1999). On an output feedback finite-time stabilisation problem. *Proceedings of the 38th IEEE Conference on Decision and Control (CDC)*, *2*, 1302–1307. https://doi.org/10.1109/CDC.1999.830117

Huber, O., Acary, V., & Brogliato, B. (2016). Lyapunov stability and performance analysis of the implicit discrete sliding mode control. *IEEE Transactions on Automatic Control*, *61*, 3016–3030. https://doi.org/10.1109/TAC.2015.2506991

Iserles, A. (2009). *A first course in the numerical analysis of differential equations*. Cambridge University Press. https://doi.org/10.1017/CBO9780511995569

Kawski, M. (1990). Homogeneous stabilizing feedback laws. *Control-Theory and Advanced Technology*, *6*(4), 497–516. https://doi.org/https://doi.org/10.1016/0167-6911(94)00035-T

Kawski, M. (1995). Geometric homogeneity and stabilization. *IFAC Proceedings Volumes*, *28*(14), 147–152. https://doi.org/https://doi.org/10.1016/b978-0-08-042371-5.50030-7

Koch, S., & Reichhartinger, M. (2018). Discrete-time equivalent homogeneous differentiators. *2018 15th International Workshop on Variable Structure Systems (VSS)*, 354–359. https://doi.org/10.1109/VSS.2018.8460284

Kumar, R., & Efimov, D. (2023). Finite/nearly fixed-time stability of nonlinear impulsive systems with destabilizing impulses and its application to neural networks. *Communications in Nonlinear Science and Numerical Simulation*, *125*, 107357. https://inria.hal.science/hal-04125430v1

Lambert, J. D. (1991). *Numerical methods for ordinary differential systems: The initial value problem*. John Wiley & Sons.

Levant, A. (2005). Homogeneity approach to high-order sliding mode design. *Automatica*, *41*(5), 823–830.

Levant, A. (2010). Chattering analysis. *IEEE Trans. Autom. Control.*, *55*(6), 1380–1389. https://doi.org/https://doi.org/10.1016/j.automatica.2004.11.029

Levant, A. (2013). On fixed and finite time stability in sliding mode control. *52nd IEEE Conference on Decision and Control*, 4260–4265. https://doi.org/10.1109/CDC.2013.6760544

Lyapunov, A. (1992). *The general problem of the stability of motion* (Vol. 55). Taylor & Francis. https://doi.org/https://doi.org/10.1080/00207179208934253

Marsden, J. E. (1992). *Lectures on mechanics* (Vol. 174). Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9780511624001

McLachlan, R. I., Quispel, G. R. W., & Robidoux, N. (1999). Geometric integration using discrete gradients. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *357*(1754), 1021–1045. https://doi.org/10.1098/rsta.1999.0363

Mendoza-Avila, J., Efimov, D., Fridman, L., & Moreno, J. A. (2023). On stability of homogeneous systems in presence of parasitic dynamics. *IEEE Transactions on Automatic Control*, *68*(11), 6852–6859. https://doi.org/10.1109/TAC.2023.3240958

Moreno, J. A., & Osorio, M. (2012). Strict Lyapunov functions for the Super–Twisting algorithm. *IEEE Transactions on Automatic Control*, *57*(4), 1035–1040. https://doi.org/10.1109/TAC.2012.2186179

Moreno, J. A. (2022). Arbitrary-order fixed-time differentiators. *IEEE Transactions on Automatic Control*, *67*(3), 1543–1549. https://doi.org/10.1109/TAC.2021.3071027

Nakamura, H., Yamashita, Y., & Nishitani, H. (2002). Smooth Lyapunov functions for homogeneous differential inclusions. *Proceedings of the 41st SICE Annual Conference*, 1974–1979.

Nekhoroshikh, A. N., Efimov, D., Polyakov, A., Perruquetti, W., & Furtat, I. B. (2020). On finite-time stabilization of a class of nonlinear time-delay systems: Implicit Lyapunov-Razumikhin approach. *CDC 2020 - 59th IEEE Conference on Decision and Control*. https://doi.org/10.1109/CDC42340.2020.9304421

Orlov, Y. (2004). Finite time stability and robust control synthesis of uncertain switched systems. *SIAM Journal on Control and Optimization*, *43*(4), 1253–1271. https://doi.org/10.1137/S0363012903425593

Polyakov, A. (2017). Quadratic-like stability of nonlinear homogeneous systems. *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 577–582. https://doi.org/https://doi.org/10.1109/CDC.2017.8263723

Polyakov, A. (2019). Sliding Mode Control Design Using Canonical Homogeneous Norm. *International Journal of Robust and Nonlinear Control*, *29*(3), 682–701. https://doi.org/https://doi.org/10.1002/rnc.4058

Polyakov, A. (2020). *Generalized homogeneity in systems and control*. Springer. https://doi.org/https://doi.org/10.1007/978-3-030-38449-4

Polyakov, A., Efimov, D., & Brogliato, B. (2019). Consistent discretization of finite-time and fixed-time stable systems. *SIAM Journal on Control and Optimization*, *57*(1), 78–103. https://doi.org/https://doi.org/10.1137/18M1197345

Polyakov, A., Efimov, D., & Perruquetti, W. (2016). Robust stabilization of MIMO systems in finite/fixed time. *International Journal of Robust and Nonlinear Control*, *26*(1), 69–90. https://doi.org/10.1002/rnc.3297

Polyakov, A., Efimov, D., & Ping, X. (2022). Consistent discretization of a homogeneous finite-time control for a double integrator. *2022 IEEE 61st Conference on Decision*

and Control (CDC), 3821–3826. https://doi.org/https://doi.org/10.1109/CDC51059.2022.9992689

Qian, C., & Lin, W. (2006). Recursive observer design, homogeneous approximation, and nonsmooth output feedback stabilization of nonlinear systems. *IEEE Trans. Automat. Contr.*, *51*(9), 1457–1471. https://doi.org/https://doi.org/10.1109/TAC.2006.880955

Rosier, L. (1992). Homogeneous Lyapunov function for homogeneous continuous vector field. *Systems and Control Letters*, *19*(6), 467–473. https://doi.org/https://doi.org/10.1016/0167-6911(92)90078-7

Rosier, L., & Bacciotti, A. (2005). *Liapunov functions and stability in control theory. Second edition*. Springer Verlag. https://doi.org/https://doi.org/10.1007/b139028

Sanchez, T., & Moreno, J. A. (2019). Design of Lyapunov functions for a class of homogeneous systems: Generalized forms approach. *Int. Journal of Robust and Nonlinear Control*, *29*(3), 661–681. https://doi.org/https://doi.org/10.1002/rnc.4274

Sanchez, T., Polyakov, A., & Efimov, D. (2020). Lyapunov-based consistent discretisation of stable homogeneous systems. *International Journal of Robust and Nonlinear Control*, *31*(9). https://doi.org/https://doi.org/10.1002/rnc.5308

Sanchez, T., Zavala-Río, A., & Zamora-Gómez, G. I. (2023). On the transient behavior and gain design for a class of second order homogeneous systems. *Automatica*, *149*, 110830. https://doi.org/https://doi.org/10.1016/j.automatica.2022.110830

Sanz-Serna, J. M. (1992). Symplectic integrators for hamiltonian problems: An overview. *Acta Numerica*, *1*, 243–286. https://doi.org/10.1017/S0962492900002282

Sepulchre, R., & Aeyels, D. (1996). Homogeneous Lyapunov functions and necessary conditions for stabilization. *Mathematics of Control, Signals and Systems*, *9*(1), 34–58. https://doi.org/https://doi.org/10.1007/BF01211517

Silva, C., Sanchez, T., Lizárraga, D., & Zavala, A. (2025). An extension of a lyapunov-function-based discretization method for asymptotically stable homogeneous systems. *International Journal of Robust and Nonlinear Control*. https://doi.org/https://doi.org/10.1002/rnc.70290

Stuart, A. M., & Humphries, A. R. (1996). *Dynamical systems and numerical analysis* (Vol. 2). Cambridge University Press.

Süli, E., & Mayers, D. F. (2003). *An introduction to numerical analysis*. Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9780511801181

Tuna, S., & Teel, A. (2004). Discrete-time homogeneous lyapunov functions for homogeneous difference inclusions. *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, *2*, 1606–1610 Vol.2. https://doi.org/10.1109/CDC.2004.1430274

Utkin, V. I. (1994). Sliding mode control in discrete-time and difference systems. In A. S. I. Zinober (Ed.), *Variable structure and lyapunov control* (pp. 87–107). Springer. https://doi.org/10.1007/BFb0033680

Zubov, V. I. (1964). *Methods of A. M. Lyapunov and their applications*. Groningen: P. Noordho: Limited.