
UNIVERSIDADE D
COIMBRA

Sara Joana Fino dos Santos Rodrigues de Carvalho

# Stable and Convergent Numerical Methods for Nonlinear Parabolic Systems: Application to Sunspots

Tese no âmbito do Programa Interuniversitário de Doutoramento em Matemática, orientada pelo Professor Doutor José Augusto Mendes Ferreira e pela Doutora Maria Teresa de Abrunhosa Barata e apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

fevereiro de 2021

# Stable and Convergent Numerical Methods for Nonlinear Parabolic Systems: Application to Sunspots

**Sara Joana Fino dos Santos Rodrigues de Carvalho**

UNIVERSIDADE D COIMBRA

U.PORTO

UC|UP Joint PhD Program in Mathematics

Programa Interuniversitário de Doutoramento em Matemática

PhD Thesis | Tese de Doutoramento

August 2021

# Acknowledgements

This PhD thesis is the result of the effort and support of several people and institutions to whom I am extremely grateful.

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Dr José Augusto Ferreira for his guidance, encouragement, and for sharing his exceptional scientific knowledge throughout this long walk. It was great privilege and honour to work and learn with him. My deepest gratitude also to my supervisor Dr Teresa Barata, for her constant support, for all the advices, for everything she taught me, for being an example of dedication and persistence without never leaving behind extraordinary human qualities, and because she was actually the one responsible for the final push for me to embark on this PhD.

I am very grateful to Dr Ricardo Gafeira not only for his willing collaboration and advices offering a valuable interdisciplinary view, but also for his unconditional friendship. I would also like to acknowledge Dr Gonçalo Pena for his precious collaboration in this work.

I acknowledge the financial support for my research study of the Portuguese Foundation for Science and Technology through the PhD fellowship SFRH/BD/107894/2015, under POCH funds, co-financed by the European Social Fund and Portuguese National Funds from MCTES. Also, a special thanks to the Centre for Mathematics of the University of Coimbra, and Centre for Earth and Space Research of the University of Coimbra for financially supporting my participation in scientific events.

I also want to express my gratitude to the Department of Mathematics of the University of Coimbra, and the Geophysical and Astronomical Observatory of the University of Coimbra especially to all the people I came across and who contributed, directly or indirectly, to my growth.

To the MaxPlanck Institute for Solar System Research for the opportunity to visit its headquarters and to learn over there a bit more about the Sun, especially Dr Jörn Warneke for the availability for scientific discussions.

I acknowledge also Dr Pedro Pina from Instituto Superior Técnico for his support in scientific questions.

I would like to thank my PhD mates for their companionship and for providing a friendly working atmosphere. Special thanks to Daniela Jordão for the support she gave me throughout this journey, both academically and personally, and to Carla Jesus for all the moments of good mood, thanks for your friendship girls. I cannot forget my colleagues of the Astronomical Observatory of the University of Coimbra, namely Dr Nuno Peixinho, Dr Ana Maria Lourenço, Claudino Romeiro, Adriana Garcia, and Dr João Fernandes for their friendship, encouragement, and all the good times we spent together.

To Zé Paulo, Miguel and Filipe, and to my lovely nephews Manel and João, I would like to thank for being part of my life and making it even complete. I cannot find words to sufficiently express how

thankful I am to my mother and my sisters for their endless love, encouragement, assistance, support, and for being always there, in every good and bad time of my life. To my father for being a constant inspiration. We will always be five. Finally, I would like to thank Sérgio for all the love, patience and assistance throughout this adventure, and to thank my beloved son Francisco for every smile, for every hug, for the happiness, strength and motivation that he gives me every day. This thesis is dedicated to my family.




# Abstract

The computation of the magnetic field in the umbra- the darkest part of a sunspot, was the main motivation of this work. After several simplifications of the physical scenario to concretize this ambitious goal, it was realized that one should start by constructing stable and convergent numerical methods for nonlinear parabolic equations.

In this thesis, numerical methods for initial boundary value problems (IBVPs) of nonlinear parabolic equations with Dirichlet or Neumann boundary conditions are proposed, and their stability and convergence analysis are established. These methods, defined in nonuniform partitions, can be seen simultaneously as finite difference methods (FDMs) and piecewise linear finite element methods (FEMs). Error estimates showing that the semi-discretization errors are second-order convergent with respect to a discrete version of the usual $H^1$-norm are established. These error estimates show that the proposed methods lead to a second-order approximation for the solution and for its gradient.

In the scope of the FDM, these error estimates are supraconvergent. This means that although they present spatial truncation error with first-order with respect to the norm $\|\cdot\|_\infty$, the corresponding error is second-order convergent with respect to a discrete version of the $H^1$-norm. On the other hand, in the finite element community, these error estimates can be seen as superconvergent estimates. In fact, piecewise linear FEMs lead, for linear elliptic equations, to first-order approximations with respect to the usual $H^1$-norm, and second-order approximations with respect to the usual $L^2$-norm. Although this fact, the second-order convergence is concluded with respect to a discrete version of the $H^1$-norm.

It should be pointed out that for differential problems with Dirichlet boundary conditions in two-dimensional domains or Neumann boundary conditions in the one-dimensional case, the error estimates are constructed assuming solution in $C^4$. For differential problems with Dirichlet boundary conditions in the one-dimensional case, lower smoothness assumptions are imposed.

The application of the developed methods to simulate strong and vertical magnetic fields is also an objective of the present work. Given this scenario, the numerical simulation is considered only the vertical component of the magnetic fields and one horizontal component. Dirichlet boundary conditions are assumed in a rectangular domain and are defined using numerical data, as well as the initial condition. The velocity field is also assumed to be known and obtained from numerical data too. As the quality of the magnitude of the magnetic field deteriorates with time due to the convective-dominated regime, a stabilization improvement is considered. Due to limitations on computational time, data handling, and availability of sunspot simulation, the numerical experiment is performed on a Network region where, on a shorter spatial scale, the condition is very similar to the one presented on the umbra of the sunspots.

Another goal of the present thesis is the automatic detection and geometric definition of sunspots, including the limits of umbra and penumbra, in solar images. An image processing algorithm based

on mathematical morphology is proposed, and its performance to detect and segment sunspots is analyzed. For this purpose, the Geophysical and Astronomical Observatory of the University of Coimbra database was used. In the near future, those results will be used to define the computational domain for the sunspots magnetic field evolution.

**Keywords:** Sunspots, magnetic field, nonlinear parabolic equations, finite difference methods, piecewise linear finite element methods, stability, convergence.

# Resumo

A evolução do campo magnético na umbra- a parte mais escura de uma mancha solar, foi a motivação central para este trabalho. Depois de várias simplificações no cenário físico, iniciou-se o estudo de métodos numéricos para equações de derivadas parciais parabólicas não lineares.

Nesta tese são propostos métodos numéricos para problemas não lineares parabólicos com condições inicial e de fronteira do tipo Dirichlet ou Neumann e é estabelecida a sua análise numérica no que diz respeito à estabilidade e convergência. Os métodos propostos, definidos em partições não uniformes, podem ser vistos simultaneamente como métodos de diferenças finitas e métodos de elementos finitos segmentados lineares. São ainda construídas estimativas de segunda ordem para o erro associado à discretização espacial, relativamente a uma versão discreta da norma $H^1$. Estas estimativas mostram que os métodos propostos permitem construir aproximações de segunda ordem para a solução bem como para o seu gradiente.

No âmbito dos métodos de diferenças finitas, as estimativas de erro estabelecidas são estimativas supraconvergentes, isto é, o erro de truncatura associado à discretização espacial é de primeira ordem relativamente à norma $\|.\|_\infty$ e o correspondente erro global é de segunda ordem relativamente a uma norma que pode ser vista como uma versão discreta da norma usual de $H^1$. Na comunidade dos métodos de elementos finitos, estes resultados podem ser vistos como resultados de superconvergência. De facto, embora baseado no método de elementos finitos segmentado linear que apresenta, para equações elípticas lineares, ordem 2 relativamente à norma usual de $L^2$ e ordem 1 relativamente à norma de $H^1$, conclui-se ordem 2 relativamente a uma norma que pode ser vista como uma discretização da norma usual de $H^1$.

É de salientar que, quando o problema diferencial é complementado com condições de Dirichlet em domínios bidimensionais ou condições de Neumann num intervalo, as estimativas de erro são construídas assumindo que as soluções analíticas estão em $C^4$. Por outro lado, para problemas com condições de Dirichlet no caso unidimensional são impostas condições de regularidade mais fracas.

A aplicação dos métodos desenvolvidos na simulação do campo magnético na umbra é também um dos objetivos da presente trabalho. Neste contexto são consideradas na simulação numérica a componente vertical e uma componente horizontal. No que diz respeito às condições de fronteira, são assumidas condições de Dirichlet definidas a partir de dados numéricos, assim como a condição inicial. Supõe-se que o campo de velocidades é também conhecido. Atendendo a que o problema em questão é dominado pela convecção, observa-se que a qualidade do campo magnético se deteriora no tempo. Com o objetivo de contornar esta patologia numérica, é implementado um método de estabilização.

Outro objetivo da presente tese é a detecção automática e definição geométrica das manchas solares, incluindo os limites da umbra e da penumbra, em imagens do sol. Um algoritmo de processamento de

imagens baseado em morfologia matemática é proposto e seu desempenho na detecção e segmentação de manchas solares é analisado. Para o efeito, utiliza-se a base de dados do Observatório Geofísico e Astronómico da Universidade de Coimbra. Os resultados obtidos serão utilizados, num trabalho futuro, para definir o domínio computacional para o estudo da evolução do campo magnético de manchas solares.

**Palavras-Chave:** Manchas solares, campo magnético, equações parabólicas não lineares, métodos de diferenças finitas, métodos de elementos finitos segmentado linear, estabilidade, convergência.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Introduction

The evolution of technology and scientific knowledge presents us every day with challenging problems in the most diverse areas. In what concerns complex problems of physical nature, numerical models with good mathematical properties are computationally simulated, under different scenarios. These models allow mimic physical phenomena and study their evolution under certain assumptions. The numerical analysis thus plays a fundamental role in the most diverse areas, such as medicine, industry, or even in solar physics. Another current problem arising from the evolution of technology is the processing and analysis of image data, in a fast and efficient way. The application of image processing techniques and their robustness contributes to the development of full automatic tools, able to detect features in images. The automatic methods reveal many advantages over manual ones, namely processing speed, precision and objectivity.

The initial aim of this thesis was the simulation of the evolution of solar magnetic field (intensity and orientation) and the consequent emergence of sunspots in the solar surface. Note that this simulation involves several complex physical processes to characterize the system completely. Radiative transfer, turbulence, processes related to atomic and plasma physics, among others, should be taken into consideration. Actually, it is a huge challenge, of enormous complexity and computational demand, which numerous teams have been working for more than 15 years (see, for example, [12] and [78]). To develop robust and effective methods to solve this problem, it is necessary, on one hand, simulate the evolution of solar magnetic field as well as the plasma. On the other hand, it is also required to establish the scenarios that lead to the emergence of sunspots. These scenarios can be obtained simulating the emergency of magnetic field from deeper layers (very complex process), or/and using sunspots as a starting point for seeding magnetic field configuration. Therefore having well-defined sunspots in solar images has an important role in the definition of scenarios. The candidate had the opportunity to visit Max Planck Institute for Solar System Research, a known renown institute in solar physics, to gain some insights on this thematic.

Given this initial aim's boldness, two new objectives emerged to gain sensitivity and experience with these themes. Firstly, a numerical study of a simplistic model of equations applicable to regions characterized by a strong and vertical magnetic field is done. These conditions are valid and originate not just sunspots, but also several other solar structures like, for example, network, inter-network regions, sharing similarities in their behaviour although observed at different scales [81]. It should be pointed out that although complex physical processes are significant elements in a realistic simulation,

they can be, in many particular cases, neglected. Secondly, the development of an automatic tool based on image processing techniques that precisely define sunspots limits, and also separate it into umbra and penumbra. This second aim is related with the fact that the magnetic field evolution and its interplay with the plasma are crucial to the evolution of previously mentioned solar structures. Therefore imprints of those structures in the solar atmosphere can be used to complement and validate the magnetic field study.

## 1.1 The sun

All matter is made up of particles with specific properties according to the physical state in which belongs. Among the classic physical states we have solid, liquid, and gas. Frequently, plasma is considered the fourth state of matter, and is the most common state in the universe. Daily, plasmas can be found inside plasma TV's, fluorescent lamps, or just looking at the sun. Basically plasma is a hot ionized gas with positive and negative charged particles, approximately in the same number, that are in constant activity of ionization and recombination. Because it is ionized, this state allows the matter to be a good conductor of electric current, responding to the interaction with electromagnetic forces.

Essentially, the sun is a huge ball made of plasma where magnetic and electric fields are everywhere. The sun, as Figure 1.1. illustrates, can be divided into three inner layers, which are the core, radiative zone, and convective zone, and three outer layers, the photosphere, chromosphere, and corona.

The core is where nuclear fusion occurs. During this physical process, a large amount of energy is released that then is propagated to more external layers. Since the inner part of the sun cannot be observed, everything that is known about its interior results from the combination of theoretical models and observations of its outer layers. Despite of being the energy source of the sun, the core does not play an important role in the interaction between plasma and magnetic field that is seen in the outer layers. The layer above the core is the radiative zone where energy is transported solely through radiation. It is also characterized by a rotation profile as a rigid body contrasting with the next layer. In convective zone, last inner layer of the sun, as its name suggests, the energy is mostly transported through convection. Here, the sun has a differential rotation where the equator rotating about 30% faster than the polar regions [75]. At the top of the convective zone there is a very thin layer called photosphere. It is from here that the radiation in the visible spectrum is able to escape and reach Earth, therefore it is the deepest layer of the sun that can be observed. The photosphere is called the solar surface, although it is not a solid surface, as it happens on Earth. The chromosphere is the next layer. Even though the plasma density in this layer is lower than in the photosphere, there is still energy transport by radiation and advection. Corona is the sun's outermost layer, and is characterized by such a low density of the plasma that the particles very rarely interact to each other. The set of the three outer layers of the sun make up the solar atmosphere.

It is known that the interaction of the plasma with the solar magnetic field generates several phenomena that can be seen in the different layers of the sun. The study of such phenomena (sunspots, filaments, plages, flares, etc.) is crucial to understand the physical processes present in the sun. This knowledge is of great importance since changes in the radiation emitted by the sun, in the solar magnetic field, energy explosions, etc., could have several impacts on the Earth. Such events can

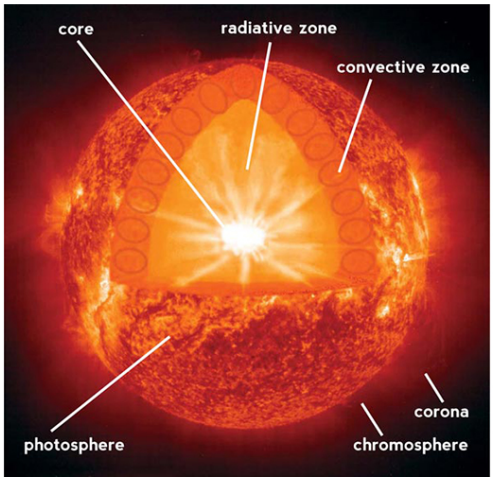



Fig. 1.1 Layers of the sun ©NASA

perturb precision navigation systems, interrupt satellite functions, and be dangerous for astronauts health. When they are more severe may even result in perturbations in the electric power system and cause the loss of satellites through damaged electronics [67]. Two examples of this kind of events are following pointed out. The first one occurred on 13th March 1989 and caused a blackout in Quebec, Canada [17]. The other one occurred on 4th November 2015, the flights disappear from the radar screens of the Swedish air traffic control towers which led to the closure of the country's airspace for more than an hour [58]. Changes in the Earth's environment that can be traced from physical processes occurring in the space, namely in the sun, are called Space Weather. It can be seen as a natural hazard where scientific knowledge should be used to understand how it affects the wider society, as well as to identify actions to mitigate the risks [36]. Therefore, the influence of the sun on the Earth makes the study of this star highly relevant today.

Despite having several phenomena resulting from solar activity, the study done in this thesis focuses only on sunspots since they are the first visible manifestation of the solar magnetic field in the solar atmosphere.

### 1.1.1 Sunspots

Sunspots were the first solar phenomena to be observed, by naked eye, and recorded, as drawings, dating from around 364 BC, by Chinese observers [45]. However, the first instrumental observations that allowed the study of the sun, from the physical point of view and in a systematic way, were carried out by Galilei, Scheiner and others, around 1611 [74]. In 1908, Hale discovered that sunspots have a magnetic nature [35]. Actually, sunspots are manifestations of strong magnetic field intensities visible in the photosphere. These structures are temporary, with a lifetime from hours to months, vary in diameter from thousand kilometers to several dozens of thousand kilometers [74].

Sunspots are composed by a dark central area, the umbra, which spans from 3900 to 4800 K in temperature ($K$ denotes Kelvin), and a less dark area that surrounds the umbra, the penumbra, where temperature varies from 5400 to 5500 $K$ (see Figure 1.2). In the umbra the magnetic field, which is nearly perpendicular to the surface, reaches values between approximately 1800 $G$ and 3700 $G$ ($G$

denotes Gauss). In penumbra the magnetic field is more horizontal, and its strength goes around 700 $G$ to 1000 $G$ (for more details please see [74]). Not always the magnetic field configuration leads to the formation of a fully formed sunspot. In these case only a umbra like structure, called pore, is created. Over the years, different models have been proposed to explain the mechanism for sunspot formation

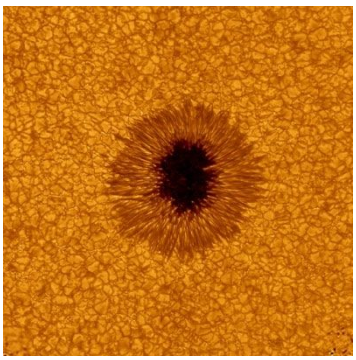


Fig. 1.2 Image taken on July, 2010 by the New Solar Telescope at the Big Bear Solar Observatory reveals up-close details of a sunspot like never before (Image: ©BBSO).

in the solar photosphere. The theory purposed by Parker [59] is widely accepted. This theory is based on the rise of magnetic flux tubes. According to it, magnetic field lines are approximated to magnetic flux tubes. With the increase of the magnetic field inside the tube, its magnetic pressure raises. To keep the total pressure inside the tube (the sum of gas and magnetic pressures) equals to its surroundings, the gas pressure drops leading to a decrease in density. The difference between the densities inside and around the tube makes the tube rises, and erupts to the surface due to a force that is exerted by the surrounding fluid, the so-called buoyancy (the same force that allows us to float on the sea). With rise of large flux tubes of great magnetic intensity results sunspots which are the intersection of them with the solar surface. A scheme representing this theory is shown in Figure 1.3. Sunspots appear darker and less warm than the surroundings because the strong magnetic fields that originated them suppress the convection mechanism. Consequently, the plasma coming from the deepest layers of the sun does not emerge and does not supply thermal energy to the upper layers, which end up cooling through radiation. The magnetic buoyancy instability, that in astrophysics is usually called Parker instability, is one of the most used instabilities created in the models to simulate the appearance of sunspots.

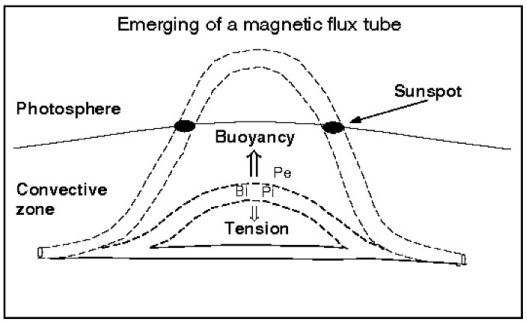



Fig. 1.3 A scheme representing the emergence of a magnetic flux tube (image from [80]).

According to rising flux-tubes model, sunspots are the first observable manifestation of the solar magnetic field. Therefore, the study of these structures is of great importance both to understand solar dynamics and to monitor and forecast solar activity allowing the obtaining of results for the Space Weather studies.

### 1.1.2 Network and internetwork magnetic fields

Inside the solar convective zone, there are convective cells spanning from 100 km (solar granulation) up to 30000 km (super granulation). These convective motions can concentrate a strong magnetic field in the edged of their cells. In the case of solar granulation, the typical magnetic field that those cells can concentrate reaches several hundreds of *G*, presenting both polarities. The regions where this type of magnetic field is present is called inter-network. Concerning the supergranulation, mainly due to their long lifetime, the magnetic field goes up the few *kG* concentrated in nearly vertical flux tube, on both polarities, with a size that often exceeds 500 km [8]. The regions where this type of magnetic field is present is called network. An example of the configuration of the z-component of magnetic field in a network region is shown in Figure 1.4.

Although the magnetic field configuration in this type of regions is very similar to the one presented in the umbra of sunspots, these structures are difficult to observe due to their smaller size and their appearance similar to the background.

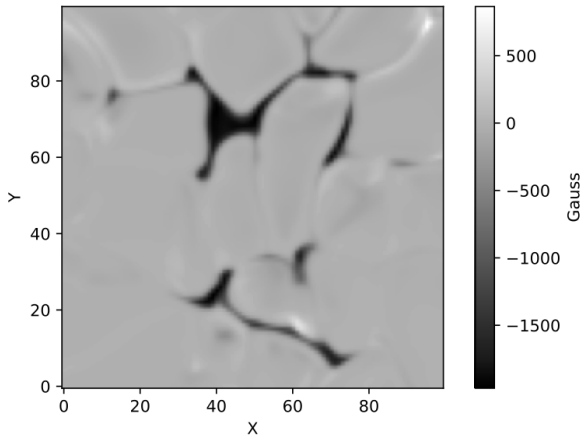



Fig. 1.4 Example of the z-component of the magnetic field in a network region.

### 1.1.3 Equations

Magnetohydrodynamics (MHD) is the branch of Physics that studies the dynamics of fluids that are electrical conductors in the presence of magnetic field. The nobel prize in Physics in 1970, Hannes Alfvén, was who started the study of MHD. Afterwards, a set of coupled equations were stated, the so-called MHD equations. This set of governing equations is often used to simulate magnetic fields, and therefore are also used in many studies of solar phenomena. Two of the equations that can be

used to describe the system are the induction equation

$$\frac{\partial B}{\partial t} = \nabla \times (u \times B) + \eta \Delta B, \tag{1.1}$$

and the Navier-Stokes equation,

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u = -\nabla(p_g/\rho) + \nu \Delta u + (J \times B)/\rho. \tag{1.2}$$

The induction equation describes the evolution of the magnetic field $B$ over time in a medium with magnetic diffusivity $\eta$, and a velocity field $u$. The Navier-Stokes equation describes the evolution of the velocity field $u$ in a medium with dynamic viscosity $\nu$, density $\rho$, being $p_g$ the gas pressure, and $J$ the electric current density. Please note that the following equality holds

$$J \times B = \frac{1}{\mu_0} B \cdot \nabla B - \nabla(B^2/2\mu_0),$$

where $\mu_0$ is the vacuum permeability, $B^2$ is given by the square of the euclidean norm of $B$, and $B^2/2\mu_0$ is called the magnetic pressure ($p_m$). For more details see [20].

Let assume $\nabla \cdot u = 0$, i.e. the particular case in which the plasma is non-relativistic and incompressible, that implies a velocity of the medium much less than the speed of sound. With the establishment of previous condition, can be considered without loss of generality that the density of the fluid follows a homogeneous distribution, and normalize it, getting $\rho = 1$ [9]. Let also assume that plasma motions are dominated by the magnetic field which implies that the gas pressure can be neglected [32, 75].

Therefore, taking into account the cross product rule

$$\nabla \times (X \times Y) = (\nabla \cdot Y + Y \cdot \nabla)X - (\nabla \cdot X + X \cdot \nabla)Y,$$

where $X$ and $Y$ are vector fields of the same dimension, the solenoidal constraint on $B$ (i. e., $\nabla \cdot B = 0$) [20], and the assumptions considered before, the equations (1.1) and (1.2) can be rewritten, respectively, as follows,

$$\frac{\partial B}{\partial t} + (u \cdot \nabla)B = (B \cdot \nabla)u + \eta \Delta B, \tag{1.3}$$

and

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u = -\nabla(B^2/2\mu_0) + \frac{1}{\mu_0}(B \cdot \nabla)B + \nu \Delta u. \tag{1.4}$$

The assumptions previously made are considered valid in the solar atmosphere. For example, the coupled equations (1.3) and (1.4) can be used to describe the temporal evolution of $B$ and $u$, in the umbra and pores [1].

The system (1.3)-(1.4) is a particular case of the system of second-order nonlinear parabolic equations

$$\frac{\partial U}{\partial t} + F(U, \nabla U) = \nabla \cdot (A(U)\nabla U) + G \quad \text{in } \Omega \times (0, T], \tag{1.5}$$

where $U = (B,u)$, $A$ is a diagonal matrix with entries $A_{ii} \geq A_0 > 0, i = 1,\ldots,6$, $\nabla U = (\nabla B, \nabla u)$, $F = (f_i, i = 1,\ldots,6)$ and $G = (g_i, i = 1,\ldots,6)$.

## 1.2 Outline

The organization of the present thesis is described in what follows.

Chapter 2, *Stability and convergence analysis for IBVP with Dirichlet boundary conditions* is devoted to the study of numerical methods for nonlinear parabolic IBVPs, defined in a real interval, and complemented with boundary conditions of Dirichlet type and an initial condition. The corresponding elliptic boundary value problems (BVPs) are also considered. For both classes of differential problems, methods that can be simultaneously seen as FDMs and piecewise linear FEMs are proposed, stability and convergence results are established. In what concerns convergence, error estimates are deduced considering discrete $H^1$-norms, and smooth and nonsmooth solutions. This last assumption requires different techniques. For smooth solution, estimates for the spatial truncation error have a central role in the final result while for nonsmooth solution the analysis is based on the use of the Bramble-Hilbert lemma [11]. In both cases, second-order error estimates are deduced for the numerical solution and for its gradient. Consequently, the results can be seen simultaneously as supraconvergent results, in the scope of FDMs, and superconvergent results, in the scope of FEMs.

Chapter 3, *An application to sunspots*, is motivated by the application of the methods studied in Chapter 2, for one-dimensional IBVPs, to the numerical simulation of the vertical solar magnetic field observed in some solar regions, namely in the umbra and also in the network. As the physical phenomenon is defined in a two-dimensional domain, this chapter starts with the natural extension of the results presented in Chapter 2 for the corresponding numerical methods defined in nonuniform partitions of a two-dimensional domain. It should be pointed out that the error estimates are obtained considering smooth solutions. The last part of this chapter is focused in the application of the studied numerical tools to the numerical simulation of the vertical magnetic field in umbra. To simplify is assumed that the velocity field is known, in all space and time domain. The initial condition is also known, as well as the magnetic field on the boundary of space and time domain. As the behaviour of the magnetic and velocity fields in the network are similar to the ones in umbra, the velocity field and the the magnetic fields used as input data were taken from the numerical simulation of a network region obtained with the Bifrost code [14, 34]. During the numerical experiences, it was realized that the numerical solution does not reproduce the reference solution due to the convection-dominated regime observed in the solar magnetic field. A stabilization term defined by an artificial diffusion was used to correct the observed numerical results' pathologic behaviour, improving its quality.

To obtain, in Chapter 3, the numerical experiments for the solar magnetic field, the magnetic field knowledge on the boundary of space and time domain is assumed, i.e., the solar magnetic field dynamics is complemented with Dirichlet boundary conditions. Indeed this assumption is not realistic, and conditions of Neumann or Robin type should be more adequate. Chapter 4, *Convergence analysis for IBVP with Neumann boundary conditions* intends to contribute to the study of numerical methods for nonlinear parabolic IBVPs with Neumann boundary conditions. It is clear that even for problems with smooth solutions, the construction of the discretization of the boundary conditions that leads to second-order error estimates is not a trivial question. Here, defining a convenient functional scenario,

are proposed numerical methods for a nonlinear parabolic IBVP and for its correspondent nonlinear elliptic equation that lead to second-order convergent numerical solutions with respect to a discrete $H^1$-norm. While in the previous chapters, only the semi-discrete approximations are studied, this chapter also provides the theoretical convergence support for fully discrete in time and space methods.

Chapter 5, *Detection of sunspots in spectroheliograms*, is focused on image processing techniques to detect and segment sunspots. For this purpose, an automatic algorithm based on mathematical morphology is developed for the data acquired in the Geophysical and Astronomical Observatory of the University of Coimbra. Its performance to detect and segment sunspots in solar images is analyzed.

This thesis finishes with Chapter 6, *Conclusions*, where main conclusions are drawn and open problems are raised, which will be object of future research.

# Chapter 2

# Stability and Convergence Analysis for IBVP with Dirichlet boundary conditions

## 2.1 Introduction

The main goal of the present chapter is to propose numerical discretizations of the following system of second-order nonlinear parabolic equations

$$\frac{\partial U}{\partial t} + F(U, \nabla U) = \nabla \cdot (A(U)\nabla U) + G \quad \text{in } \Omega \times (0,T], \tag{2.1}$$

that depends only on $x$ and $t$, and where, to simplify, $\Omega = (0,1)$, $U = (u_1, u_2)$, $A$ is a diagonal matrix with entries $\alpha(u_1) \geq \alpha_0 > 0$ and $\beta(u_2) \geq \beta_0 > 0$, $\nabla U = (\nabla u_1, \nabla u_2)$, $F = (f_1, f_2)$ and $G = (g_1, g_2)$. This system is complemented with homogeneous Dirichlet boundary conditions

$$U = 0 \quad \text{on } \partial\Omega \times (0,T], \tag{2.2}$$

with $\partial\Omega = \{0,1\}$, and initial conditions

$$U(x,0) = U_0(x), \quad x \in \Omega, \tag{2.3}$$

having $U_0 = (u_{1,0}, u_{2,0})$. In what follows the following notation is used: if $U : \overline{\Omega} \times [0,T] \mapsto \mathbb{R}^2$ then by $U(t)$ is represented the function $U(t) : \overline{\Omega} \mapsto \mathbb{R}^2, U(t)(x) = U(x,t)$.

The objective is to introduce numerical methods that can be seen simultaneously as fully discrete piecewise linear FEMs, defined with convenient quadrature rules, and as FDMs defined on nonuniform grids. We will be mainly focused on semi-discretization methods, that is in spatial discretizations, that converts the IBVP (2.1), (2.2), (2.3) in an initial value problem that can be solved using an efficient and accurate numerical methods for this kind of problems.

The proposed methods are based on piecewise linear FEMs, and it is well known that they lead to a first-order approximation with respect to the usual $H^1$-norm, and to a second-order approximation with respect to the usual $L^2$-norm. On the other hand, the methods that will be studied can be seen as FDMs defined on non-uniform grids characterized by a truncation error with first-order with respect to the norm $\|\cdot\|_\infty$. Following the lines of the research presented in [41, 71], it will be shown that

the proposed numerical methods lead to a second-order approximation with respect to a discrete version of the usual $H^1$-norm. Once it is obtained a higher-order approximation than the expected, the piecewise linear FEM is said superconvergent and, as FDM is said supraconvergent. Thus, the designation of super-supraconvergent for the methods studied here is used to explicitly state these facts. Recall that there are several procedures to define second-order approximations for the gradient of the solution using the piecewise linear finite element approximations. Without being exhaustive it is mention the classical papers [42, 44].

To gain some insights into the design and convergence analysis of super-supraconvergent discretizations of second-order nonlinear equations, we start by considering the nonlinear elliptic BVP

$$-(A(u)u')' + f(u,u') = g \quad \text{in } \Omega, \tag{2.4}$$

with homogeneous Dirichlet boundary conditions

$$u = 0 \quad \text{on } \partial\Omega. \tag{2.5}$$

It is desired to extend the results obtained for the previous elliptic problem, to the following system of nonlinear elliptic equations

$$-\nabla \cdot (A(U)\nabla U) + F(U,\nabla U) = G \quad \text{in } \Omega, \tag{2.6}$$

where $U$ depends only on the spatial variable, with homogeneous Dirichlet boundary conditions

$$U = 0 \quad \text{on } \partial\Omega. \tag{2.7}$$

For the fully discrete piecewise linear finite element approximations for the solution of the previous nonlinear problems, it is proved second-order convergence for the gradient without any post-processing, a popular procedure followed in different contexts, that leads to an improvement of the accuracy of the gradient approximations (see for instance [44, 79]).

In what concerns linear elliptic BVPs, in [6, 23], numerical methods based on piecewise linear FEMs, that are equivalent to FDMs, were proposed. In these papers, the authors presented a new approach to analyse the convergence properties of a fully discrete in space piecewise linear FEM. In this new approach, the Bramble-Hilbert lemma [11] is the main tool in the convergence analysis. It allows reducing the smoothness assumptions on the solutions of the differential problems usually required when Taylor expansion is used: the usual smoothness assumption $u \in C^4(\overline{\Omega})$ is replaced by the weaker assumption $u \in H^3(\Omega) \cap H_0^1(\Omega)$. This new methodology has been largely used as can be seen, for instance, in [5, 7, 24–27, 41, 71]. It should be pointed out that it is intended to use the same approach here to establish the convergence results for the nonlinear problems introduced above for lower smooth solutions. However we start by establishing the same convergence order considering the particular structure of the truncation error and assuming that $u \in C^4(\overline{\Omega})$. Finally, it is observed that the convergence results are established analysing carefully the error equation. As we are dealing with nonlinear differential problems, the convergence estimates are not obtained using stability arguments.

In the FDMs, for linear time dependent problems, stability and consistency are equivalent to convergence [43]. When we go to nonlinear differential problems, stability is a local property [65]. In

fact, the stability of numerical methods applied to nonlinear problems is a local problem property in the sense that is defined for approximations in a ball centered in a numerical solution. The stability analysis usually requires some smoothness on such numerical solution, and the neighborhood can be defined with a fixed radius or the radius can be step-size dependent. Although the stability analysis for nonlinear methods can be seen as a classical problem where apparently no novelty can be introduced, it requires a careful treatment and deep understanding of the concept. In fact, as it will be seen in what follows, standard arguments can be used to show that, for instance, for the discretization of nonlinear elliptic equations, the bound for the difference between two solutions depends on the discrete $W^{1,\infty}$-norm of the fixed solution. To conclude stability, it is necessary to assume the boudness of such quantity. Such result can be proved provided that the discrete $W^{1,\infty}$-norm for the error is at least of first-order. This fact requires the study of the error. It is assumed that the diffusion coefficient depends on the dependent variable and this fact requires a careful discretization of this term to obtain a stable and convergent method. Although the methods introduced here, also seen as FDMs, present first-order truncation error, it will be proved that they are second-order accurate with respect to a discrete $H^1$-norm assuming lower smoothness assumptions on the solutions. These results allow obtaining the upper bounds that are crucial in the stability analysis. The stability will be established in a ball centered in the numerical solution with radius step-size dependent ([47], [65]). As for elliptic equations, for nonlinear parabolic problems it will be provided in this chapter the error estimates that allow the establishment of the upper bounds for the discrete solutions needed in the stability analysis.

The present chapter is organized as follows. In Section 2.2, the elliptic equation (2.4) is considered with homogeneous Dirichlet boundary conditions (2.5). In Section 2.2.1 some notations and important results are introduced. Section 2.2.2 is devoted to the stability analysis, considering a first approach where is realized that an additional assumption needs to be imposed in the numerical solution. In Section 2.2.3 are presented convergence results for smooth solutions. Firstly is concluded first-order of convergence for solutions in $C^3(\overline{\Omega})$ taking into account the expression of the truncation error, and using Taylor expansion. An improvement of the order of convergence is after achieved, for solutions in $C^4(\overline{\Omega})$, manipulating the expression of the truncation error which allows obtaining second-order estimates for the error. The stability is established in Section 2.2.4, based on the error estimates previously obtained, which is not a common approach in the literature. Section 2.2.5 is dedicated for the convergence analysis for nonsmooth solutions. Using Bramble-Hilbert Lemma as the main tool second-order of convergence is obtained for solutions in $H^3(\Omega) \cap H_0^1(\Omega)$. Numerical results illustrating the theoretical support developed in the previous section are in Section 2.2.6. The extension of the principal results to the BVP (2.6)-(2.7) is the objective of Section 2.3, where numerical simulations illustrating the order of convergence obtained theoretically for nonsmooth solutions are also presented. Section 2.4 is devoted to the main goal of the present chapter, the construction of a numerical method for the IBVP (2.1)-(2.3), and its convergence (Section 2.4.1) and stability (Section 2.4.2) analysis, following the same approaches carried out for the stationary problems. Numerical simulations illustrating the convergence result can also be found in this section. Finally, in Section 2.5 some conclusions are presented.

Remark that the analytical study concerning the existence and uniqueness of the solution of the BVPs or IBVPs studied in the scope of this thesis will not be considered.

## 2.2 A nonlinear elliptic equation

### 2.2.1 Notations and definitions

We start by introducing a non-uniform grid $\overline{\Omega}_h$, in $\overline{\Omega}$, of size $N$, induced by a vector $h=(h_1,...,h_N)$, $h_i>0, \forall i\in\{1,...,N\}$, with $\sum_{i=1}^{N} h_i = 1$. Let $\Lambda$ be a sequence of such vectors $h$, and $h_{\max} = \max_{i=1,...,N} h_i \to 0$. Let $\overline{\Omega}_h = \{x_i : i=0,...,N\}$ be the nonuniform grid in $\overline{\Omega}$, defined by $x_i = x_{i-1}+h_i$, $i=1,\ldots,N$, with $x_0=0$ and $x_N=1$. Let $\Omega_h = \Omega\cap\overline{\Omega}_h$ be the set of interior nodes, and the set of boundary points of the grid is denoted by $\partial\Omega_h = \partial\Omega\cap\overline{\Omega}_h$. Let $W_h$ and $W_{h,0}$ be, respectively, the space of grid functions defined in $\overline{\Omega}_h$, and the space of grid functions defined in $\overline{\Omega}_h$ and null on $\partial\Omega_h$.

In $W_{h,0}$ is introduced the following $L^2$-discrete inner product

$$(u_h,w_h)_h = \sum_{i=1}^{N-1} h_{i+1/2}u_h(x_i)w_h(x_i),\ u_h,w_h\in W_{h,0},$$

considering $h_{i+1/2}=\frac{h_i+h_{i+1}}{2}$. The norm induced by this inner product is denoted by $\|\cdot\|_h$. In this space, it is also considered the norm

$$\|v_h\|_{h,\infty} = \max_{i=1,...,N-1}|v_h(x_i)|,\ \forall v_h\in W_{h,0}.$$

The following notations are used in what follows

$$(u_h,w_h)_+ = \sum_{i=1}^{N} h_i u_h(x_i)w_h(x_i),\ u_h,w_h\in W_h,$$

and $\|u_h\|_+ = \sqrt{(u_h,u_h)_+}, u_h\in W_h$.

Note that $\|\cdot\|_h$ can be seen as a discrete version of $L^2$-norm, and also

$$\|v_h\|_{1,h} = \left(\|v_h\|_h^2 + \|D_{-x}v_h\|_+^2\right)^{1/2},\ v_h\in W_{h,0},$$

can be seen as a discrete version of the usual $H^1$-norm.

Let $D_{-x}$ and $D_x^*$ be the first-order finite difference operators defined by

$$D_{-x}u_h(x_i) = \frac{u_h(x_i)-u_h(x_{i-1})}{h_i},\ i=1,\ldots,N.$$

$$D_x^*u_h(x_i) = \frac{u_h(x_{i+1})-u_h(x_i)}{h_{i+1/2}},\ i=0,\ldots,N-1.$$

Recall some useful results regarding functions in $W_{h,0}$.

**Proposition 2.1.** *For $v_h\in W_{h,0}$ it holds*

$$\|v_h\|_h \le \|D_{-x}v_h\|_+ \tag{2.8}$$

*and*

$$\|v_h\|_{h,\infty} \leq \|D_{-x}v_h\|_+.$$

Inequality (2.8) is usually called Poincaré-Friedrichs inequality, and it can be seen as a discrete version of its continuous version.

Consider now the following finite difference operator

$$D_x^*(A(M_h u_h)D_{-x}u_h)(x_i) = \frac{A(M_h u_h(x_{i+1}))D_{-x}u_h(x_{i+1}) - A(M_h u_h(x_i))D_{-x}u_h(x_i)}{h_{i+1/2}}, \; i = 1, ..., N-1, \tag{2.9}$$

where $M_h$ is the following average operator

$$M_h u_h(x_i) = \frac{u_h(x_i) + u_h(x_{i-1})}{2}.$$

When $A$ is constant, the finite difference operator (2.9) is reduced to the second-order centered finite difference operator, $\Delta_h$, as the discrete version of the second derivative in space, defined by

$$\Delta_h u_h(x_i) = \frac{h_i u_h(x_{i+1}) - (h_i + h_{i+1})u_h(x_i) + h_{i+1}u_h(x_{i-1})}{h_i h_{i+1} h_{i+1/2}},$$

for $i = 1, \ldots, N-1$.

For the finite difference operator (2.9), the following result holds, which is an analogue of the integration by parts formula known in calculus.

**Proposition 2.2.** *For $u_h, v_h \in W_{h,0}$,*

$$(D_x^*(A(M_h u_h)D_{-x}u_h), v_h)_h = -(A(M_h u_h)D_{-x}u_h, D_{-x}v_h)_+.$$

It is also introduced the discrete operator $\nabla_h$ defined as

$$\nabla_h u_h(x_i) = \frac{h_i}{h_i + h_{i+1}} D_{-x}u_h(x_{i+1}) + \frac{h_{i+1}}{h_i + h_{i+1}} D_{-x}u_h(x_i), \; i = 1, ..., N-1.$$

The source term $g$ is discretized by

$$g_h(x_i) = \frac{1}{h_{i+1/2}} \int_{x_{i-1/2}}^{x_{i+1/2}} g(x)dx, \tag{2.10}$$

where $x_{i-1/2} = x_i - \frac{h_i}{2}$, and $x_{i+1/2} = x_i + \frac{h_{i+1}}{2}$.

Consider, for the elliptic equation (2.4), the weak formulation: find $u \in H_0^1(\Omega)$ such that

$$(A(u)u', v') + (f(u, u'), v) = (g, v), \; \forall v \in H_0^1(\Omega). \tag{2.11}$$

The piecewise linear finite element approximation $P_h u_h$ with $u_h \in W_{h,0}$, where $P_h u_h$ denotes the piecewise linear interpolator of $u_h$, is solution of the following problem: find $u_h \in W_{h,0}$ such that

$$(A(P_h u_h)(P_h u_h)', (P_h v_h)') + (f(P_h u_h, (P_h u_h)'), P_h v_h) = (g, P_h v_h), \; \forall v_h \in W_{h,0}.$$

The fully discrete piecewise linear approximation is then obtained taking into account the following approximations

$$(A(P_h u_h)(P_h u_h)', (P_h v_h)') \simeq (A(M_h u_h) D_{-x} u_h, D_{-x} v_h)_+$$

and

$$(f(P_h u_h, (P_h u_h)'), P_h v_h) \simeq (f(u_h, \nabla_h u_h), v_h)_h, \quad (g, P_h v_h) \simeq (g_h, v_h)_h.$$

Then the fully discrete finite element approximation is obtained solving the following problem: find $u_h \in W_{h,0}$ such that

$$(A(M_h u_h) D_{-x} u_h, D_{-x} v_h)_+ + (f(u_h, \nabla_h u_h), v_h)_h = (g_h, v_h)_h, \ \forall v_h \in W_{h,0}. \tag{2.12}$$

Taking into account Proposition 2.2, it is observed that the last problem is equivalent to the following one: find $u_h \in W_{h,0}$ such that

$$-D_x^*(A(M_h u_h) D_{-x} u_h) + f(u_h, \nabla_h u_h) = g_h \quad \text{in } \Omega_h, \tag{2.13}$$

complemented with the boundary condition

$$u_h = 0 \quad \text{on } \partial\Omega_h. \tag{2.14}$$

### 2.2.2 Stability analysis

To study the stability of the FDM (2.13)-(2.14) or equivalently of the fully discrete FEM (2.12) it is considered a solution $u_h \in W_{h,0}$, and its perturbation $v_h \in W_{h,0}$ induced by some different source data $\tilde{g}_h$, that satisfies

$$(A(M_h v_h) D_{-x} v_h, D_{-x} q_h)_+ + (f(v_h, \nabla_h v_h), q_h)_h = (\tilde{g}_h, q_h)_h, \ \forall q_h \in W_{h,0}. \tag{2.15}$$

Suppose that $A \geq A_0 > 0$, where $A_0$ is a constant, and $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a Lipschitz function with Lipschitz constant $C_L$. In what follows, suitable regularity conditions are assumed. In order to obtain stability it is necessary to impose the following smoothness assumption on $\Lambda$

$$\exists C_R > 0 : \frac{h_{max}}{h_{min}} \leq C_R, \forall h \in \Lambda, \tag{2.16}$$

where $h_{min} = \min\{h_i, i = 1, \ldots, N\}$.

Let $\omega_h = u_h - v_h \in W_{h,0}$. Following [65], the objective is to establish conditions which guarantee that if $\|g_h - \tilde{g}_h\|_h \to 0$ as $h_{max} \to 0$ then that $\|\omega_h\|_h \to 0$ as $h_{max} \to 0$.

Since $f$ is a Lipschitz function, using Proposition 2.1 for $\omega_h$, we have successively

$$\begin{aligned}(A(M_hv_h)D_{-x}\omega_h, D_{-x}\omega_h)_+ &= -((A(M_hu_h)-A(M_hv_h))D_{-x}u_h, D_{-x}\omega_h)_+ \\ &+(f(v_h,\nabla_h v_h)-f(u_h,\nabla_h u_h),\omega_h)_h + (g_h-\tilde{g}_h,\omega_h)_h \\ &\leq \|A\|_{C_B^1(\mathbb{R})}\|D_{-x}u_h\|_{h,\infty}\|\omega_h\|_h\|D_{-x}\omega_h\|_+ \\ &+C_L(1+\sqrt{2C_R})\|D_{-x}\omega_h\|_+^2 + \|g_h-\tilde{g}_h\|_h\|\omega_h\|_h \\ &\leq \left(\|A\|_{C_B^1(\mathbb{R})}\|D_{-x}u_h\|_{h,\infty}+C_L(1+\sqrt{2C_R})+\varepsilon^2\right)\|D_{-x}\omega_h\|_+^2 \\ &+\frac{1}{4\varepsilon^2}\|g_h-\tilde{g}_h\|_h^2,\end{aligned}$$

where $C_B^m(\mathbb{R})$ denotes the space of real differentiable functions with derivative of order $m$ bounded in $\mathbb{R}$, where is considered the norm $\|v\|_{C_B^m(\mathbb{R})} = \max\limits_{i=0,\ldots,m}\|v^{(i)}\|_\infty$. The last inequality leads to

$$\left(A_0-\left(\|A\|_{C_B^1(\mathbb{R})}\|D_{-x}u_h\|_{h,\infty}+C_L(1+\sqrt{2C_R})+\varepsilon^2\right)\right)\|D_{-x}\omega_h\|_+^2 \leq \frac{1}{4\varepsilon^2}\|g_h-\tilde{g}_h\|_h^2. \tag{2.17}$$

To conclude a sufficient condition for stability, it is necessary to ensure that there exists a positive constant $\varepsilon^2$, $h$ independent, such that

$$A_0-\left(\|A\|_{C_B^1(\mathbb{R})}\|D_{-x}u_h\|_{h,\infty}+C_L(1+\sqrt{2C_R})+\varepsilon^2\right)>0. \tag{2.18}$$

To guarantee that the last assumption holds it is necessary to impose the uniform boundness of $\|D_{-x}u_h\|_{h,\infty}$ for $h\in\Lambda$. It is observed that this property can follow from an error estimate as it will be stated in Proposition 2.6.

### 2.2.3 Convergence analysis: smooth solutions

#### 1. First approach

To simplify the presentation in what follows, we rewrite the BVP (2.4),(2.5) and (2.13), (2.14) by

$$\begin{cases} F(u)=g \text{ in } \Omega, \\ u=0 \text{ on } \partial\Omega, \end{cases} \tag{2.19}$$

and

$$\begin{cases} F_h(u_h)=g_h \text{ in } \Omega_h, \\ u_h=0 \text{ on } \partial\Omega_h. \end{cases} \tag{2.20}$$

Let $R_h : C(\overline{\Omega})\to W_h$ be the restriction operator defined by $R_hu(x)=u(x)$, $x\in\overline{\Omega}_h$, where $C(\overline{\Omega})$ denotes the space of continuous functions. Let $E_u$ be the discretization error, $E_u=u_h-R_hu$. By $T_h$ is represented the truncation error induced by the discretization $F_h$, and by $T_{h,g}=g_h-R_hg$ is represented the error induced by the replacement of $g$ by $g_h$. An error estimate can be obtained easily considering that

$$\begin{aligned}F_h(R_hu(x)) &= F(u(x))+T_h(x) \\ &= g(x)+T_h(x),\ x\in\Omega_h.\end{aligned}$$

Then

$$\begin{cases} F_h(u_h(x)) - F_h(R_h u(x)) = T_{h,g}(x) - T_h(x),\ x \in \Omega_h, \\ E_u = 0 \text{ on } \partial\Omega_h. \end{cases} \tag{2.21}$$

As we have

$$\begin{aligned} F_h(u_h(x)) - F_h(R_h u(x)) &= \int_0^1 \frac{d}{d\theta} F_h(R_h u(x) + \theta E_u(x)) d\theta \\ &= \int_0^1 JF_h(R_h u(x) + \theta E_u(x)) d\theta E_u(x),\ x \in \Omega_h. \end{aligned}$$

to obtain an error estimate for $E_u$ it is necessary to analyse the following system

$$\begin{cases} \int_0^1 JF_h(R_h u + \theta E_u) d\theta E_u = T_{h,g} - T_h \text{ in } \Omega_h \\ E_u = 0 \text{ on } \Omega_h. \end{cases} \tag{2.22}$$

To get the desired estimate, it is needed to guarantee that $\left[\int_0^1 JF_h(R_h u + \theta E_u) d\theta\right]^{-1}$ exists and it is bounded. In this case it is obtained that

$$\|E_u\|_h \leq C_b \left(\|T_h\|_h + \|T_{h,g}\|_h\right), \tag{2.23}$$

where

$$\left\| \left[\int_0^1 JF_h(R_h u + \theta E_u) d\theta\right]^{-1} \right\|_h = \sup_{0 \neq w_h \in W_{h,0}} \frac{\left\| \left[\int_0^1 JF_h(R_h u + \theta E_u) d\theta\right]^{-1} w_h \right\|_h}{\|w_h\|_h} \leq C_b, \tag{2.24}$$

Remark that if

$$\left\| \int_0^1 JF_h(R_h u + \theta E_u) d\theta w_h \right\|_h \geq \frac{1}{C_b} \|w_h\|_h,\ \forall w_h \in W_{h,0}, \tag{2.25}$$

then (2.24) holds. Moreover, if

$$(JF_h(R_h u + \theta E_u) v_h, v_h)_h \geq \frac{1}{C_b} \|v_h\|_h^2,\ \forall v_h \in W_{h,0}, \tag{2.26}$$

then it is also concluded (2.24). This means that (2.25) and (2.26) are sufficient conditions for (2.24) and consequently for the error estimate (2.23).

In what follows it is established an error estimate of the type (2.23) taking into account the particular structure of the error equation (2.21).

Given the equality

$$(F_h(u_h) - F_h(R_h u), E_u)_h = (T_{h,g} - T_h, E_u)_h,$$

for the particular definition of $F_h$ of the problem under analysis, and considering Proposition 2.2, the following expression is deduced,

$$\begin{aligned}(A(M_hu_h)D_{-x}u_h, D_{-x}E_u)_+ - (A(M_hR_hu)D_{-x}R_hu, D_{-x}E_u)_+ &= (f(R_hu, \nabla_hR_hu) - f(u_h, \nabla_hu_h), E_u)_h \\ &\quad + (T_{h,g} - T_h, E_u)_h,\end{aligned}$$

and it leads to

$$\begin{aligned}(A(M_hu_h)D_{-x}E_u, D_{-x}E_u)_+ &= (f(R_hu, \nabla_hR_hu) - f(u_h, \nabla_hu_h), E_u)_h \\ &\quad + ((A(M_hR_hu) - A(M_hu_h))D_{-x}R_hu, D_{-x}E_u)_+ \\ &\quad + (T_{h,g} - T_h, E_u)_h.\end{aligned} \tag{2.27}$$

We also have that

$$\begin{aligned}((A(M_hR_hu) - A(M_hu_h))D_{-x}R_hu, D_{-x}E_u)_+ &\leq \|D_{-x}R_hu\|_{h,\infty}\left(|A(M_hR_hu) - A(M_hu_h)|, |D_{-x}E_u|\right)_+ \\ &\leq \|D_{-x}R_hu\|_{h,\infty}\|A\|_{C^1_B(\mathbb{R})}\left(|M_hE_u|, |D_{-x}E_u|\right)_+ \\ &\leq \|D_{-x}R_hu\|_{h,\infty}\|A\|_{C^1_B(\mathbb{R})}\|E_u\|_h\|D_{-x}E_u\|_+ \\ &\leq \|u'\|_\infty\|A\|_{C^1_B(\mathbb{R})}\|D_{-x}E_u\|_+^2.\end{aligned} \tag{2.28}$$

Therefore, from (2.27) and (2.28), and given that $f$ is a Lipschitz function, it is obtained

$$\begin{aligned}(A(M_hu_h)D_{-x}E_u, D_{-x}E_u)_+ &\leq \|u'\|_\infty\|A\|_{C^1_B(\mathbb{R})}\|D_{-x}E_u\|_+^2 \\ &\quad + C_L\sqrt{2C_R}\|D_{-x}E_u\|_+^2 + C_L\|D_{-x}E_u\|_+^2 + (T_{h,g} - T_h, E_u)_h.\end{aligned}$$

Moreover, since $A_0$ is a lower bound of $A$, the following relation holds

$$\left(A_0 - \left(\|u'\|_\infty\|A\|_{C^1_B(\mathbb{R})} + C_L(\sqrt{2C_R} + 1)\right)\right)\|D_{-x}E_u\|_+^2 \leq (T_{h,g} - T_h, E_u)_h. \tag{2.29}$$

By the definition of the truncation error $T_h$, we have $T_h = \sum_{i=1}^{3} T_h^{(i)}$ with $T_h^{(1)}$ being induced by the discretization $-D_x^*(A(M_hR_hu)D_{-x}R_hu)$, and so it is given by

$$T_h^{(1)}(x_i) = (h_{i+1} - h_i)R(x_i),$$

where $R(x_i) = -\frac{1}{3}A(u(x_i))u^{(3)}(x_i) - u'(x_i)\left(u^{(2)}(x_i)A'(u(x_i)) + \frac{1}{4}A^{(2)}(u(x_i))u'(x_i)^2\right)$, with $T_h^{(2)}$ being induced by the discretization $f(u(x_i), \nabla_hu(x_i))$, and therefore

$$|T_h^{(2)}(x_i)| \leq \frac{1}{6}C_Lh_{max}^2\|u\|_{C^3(\overline{\Omega})},$$

and with $T_h^{(3)}$ being the remaining part of $T_h$, and so it is obtained

$$|T_h^{(3)}(x_i)| \leq Ch_{max}^2\|A\|_{C^3_B(\mathbb{R})}\|u\|_{C^4(\overline{\Omega})}.$$

Remark that assuming $u \in C^4(\overline{\Omega})$ than for $T_{h,g}$ is obtained the following representation $T_{h,g} = T_{h,g}^{(1)} + T_{h,g}^{(2)}$ where

$$T_{h,g}^{(1)}(x_i) = \frac{1}{4}(h_{i+1} - h_i)g'(x_i),$$

and

$$|T_{h,g}^{(2)}(x_i)| \leq \frac{1}{2}h_{max}^2 \|g''\|_\infty.$$

Note that, although the previous truncation errors representations were constructed assuming $u \in C^4(\overline{\Omega})$, to obtain only first-order representation it is enough to assume that $u \in C^3(\overline{\Omega})$.

From (2.29) and given the truncations errors representations, it is obtained

$$\left(A_0 - \left(\|u'\|_\infty \|A\|_{C_B^1(\mathbb{R})} + C_L(\sqrt{2C_R}+1)\right)\right)\|D_{-x}E_u\|_+^2 \leq Ch_{max}^2 + \varepsilon^2\|E_u\|_h^2$$

and finally, from the discrete Poincaré-Friedrichs, it is concluded the first-order estimate

$$\left(A_0 - \varepsilon^2 - \left(\|u'\|_\infty \|A\|_{C_B^1(\mathbb{R})} + C_L(\sqrt{2C_R}+1)\right)\right)\|D_{-x}E_u\|_+^2 \leq Ch_{max}^2, \tag{2.30}$$

provided that there exists a positive constant $\varepsilon^2$ such that

$$\left(A_0 - \varepsilon^2 - \left(\|u'\|_\infty \|A\|_{C_B^1(\mathbb{R})} + C_L(\sqrt{2C_R}+1)\right)\right) > 0. \tag{2.31}$$

The convergence result can be summarized in the following theorem.

**Proposition 2.3.** *Let us suppose that $u \in C^3(\overline{\Omega})$, there exists a positive constant $\varepsilon^2$ such that (2.31) holds, and $u_h \in W_{h,0}$ is defined by (2.12) or by (2.13)-(2.14), where the sequence of grids $\overline{\Omega}_h, h \in \Lambda$, satisfies (2.16). Then, for the error $E_u = u_h - R_h u$, there exists a positive constant $C_u$ such that*

$$\|E_u\|_{1,h} \leq C_u h_{max}. \tag{2.32}$$

Proposition 2.3 states first convergence order with respect to the norm $\|\cdot\|_{1,h}$ which can be seen as the discrete version of the usual $H^1$- norm. Although this result was established assuming $u \in C^3(\overline{\Omega})$, it can be shown, considering the truncation error representations obtained before, and for $u \in C^4(\overline{\Omega})$ that if the grid $\overline{\Omega}_h$ is uniform then $\|E_u\|_{1,h} \leq Ch^2$. In what follows it is shown that for nonuniform grids we still have $\|E_u\|_{1,h} \leq Ch_{max}^2$.

Recall the inequality (2.17) that leads to the stability of our discretization. To conclude the sufficient condition for stability, it is needed to guarantee that there exists a positive constant $\varepsilon \neq 0$ such that (2.18) holds. To guarantee such existence it is necessary to prove that $\|D_{-x}u_h\|_{h,\infty}$ for $h \in \Lambda$, is bounded. As

$$\|D_{-x}u_h\|_{h,\infty} \leq \|D_{-x}E_u\|_{h,\infty} + \|D_{-x}R_h u\|_{h,\infty}$$

and

$$\begin{aligned} |D_{-x}E_u(x_i)| &\leq \frac{1}{h_{min}}\sum_{j=1}^{N} h_j |D_{-x}E_u(x_j)| \\ &\leq \frac{1}{h_{min}}\|D_{-x}E_u\|_+, \end{aligned}$$

then from (2.32) we get

$$\|D_{-x}E_u\|_{h,\infty} \leq \frac{1}{h_{min}} C_u h_{max}. \tag{2.33}$$

and consequently, from (2.16), it is obtained

$$\|D_{-x}u_h\|_{h,\infty} \leq C_u C_R + \|u'\|_\infty, \tag{2.34}$$

that guarantees the boundness of $\|D_{-x}u_h\|_{h,\infty}$ for $h \in \Lambda$, that leads to the stability of the finite difference scheme (2.13)-(2.14), or equivalently, of the fully discrete piecewise linear FEM (2.12). It is observed that $C_u C_R$ can be very large depending on the smoothness of the solution $u$. We would like to obtain an upper bound such that $\|D_{-x}u_h\|_{h,\infty}$ for $h \in \Lambda$, is asymptotically bounded by $\|u'\|_\infty$. This will be done in the next section where is proved that $\|E_u\|_{1,h} \leq C_u h_{max}^2$.

### 2. Second approach

An improvement in Proposition 2.3 can be obtained if the term $(T_{h,g} - T_h, E_u)_h$, in (2.29), is be treated carefully. The following result will have an important role in the estimation of such term.

**Proposition 2.4.** *If $q \in C^1(\overline{\Omega})$ and $q_h(x_i) = (h_{i+1} - h_i)q(x_i), i = 1, \ldots, N-1$, then*

$$\begin{aligned}(q_h, v_h)_h \quad &\leq \tfrac{1}{16\varepsilon^2} h_{max}^4 \left(2\|q\|_{H^1}^2 + \|q\|_{C(\overline{\Omega})}^2\right) \\ &+2\varepsilon^2 \|D_{-x}v_h\|_+^2,\end{aligned} \tag{2.35}$$

*for $v_h \in W_{h,0}$. In (2.35), being $\varepsilon \neq 0$ an arbitrary constant.*

*Proof.* We have successively the following

$$\begin{aligned}(q_h, v_v)_h \quad &= \frac{1}{2}\sum_{i=1}^{N-1} (h_{i+1}^2 - h_i^2) q(x_i) v_h(x_i) \\ &= \frac{1}{2}\sum_{i=1}^{N} h_i^2 \big(q(x_{i-1})v_h(x_{i-1}) - q(x_i)v_h(x_i)\big) \\ &= -\frac{1}{2}\sum_{i=1}^{N} h_i^2 q(x_{i-1})\big(v_h(x_i) - v_h(x_{i-1})\big) \\ &-\frac{1}{2}\sum_{i=1}^{N} h_i^2 \big(q(x_i) - q(x_{i-1})\big) v_h(x_i) \\ &= -\frac{1}{2}\sum_{i=1}^{N} h_i^3 q(x_{i-1}) D_{-x}v_h(x_i) \\ &-\frac{1}{2}\sum_{i=1}^{N} h_i^2 \int_{x_{i-1}}^{x_i} q'(x)dx\, v_h(x_i) \\ &:= Q_1 + Q_2,\end{aligned}$$

where

$$Q_1 = -\frac{1}{2}\sum_{i=1}^{N} h_i^3 q(x_{i-1}) D_{-x}v_h(x_i),$$

and

$$Q_2 = -\frac{1}{2}\sum_{i=1}^{N} h_i^2 \int_{x_{i-1}}^{x_i} q'(x)dx\, v_h(x_i).$$

For $Q_1$ it is obtained

$$\begin{aligned} Q_1 &\leq \frac{1}{2}\|q\|_{C(\overline{\Omega})} h_{max}^2 \|D_{-x}v_h\|_+ \\ &\leq \frac{1}{16\varepsilon^2}\|q\|_{C(\overline{\Omega})}^2 h_{max}^4 + \varepsilon^2 \|D_{-x}v_h\|_+^2, \end{aligned}$$

where $\varepsilon \neq 0$ is an arbitrary constant, and for $Q_2$

$$\begin{aligned} Q_2 &\leq \frac{1}{2}\sum_{i=1}^{N} h_i^2 \left(\int_{x_{i-1}}^{x_i} (q'(x))^2 dx\right)^{1/2} \sqrt{h_i}|v_h(x_i)| \\ &\leq \frac{1}{2}\Big(\sum_{i=1}^{N} h_i^4 \|q'\|_{L^2(x_{i-1},x_i)}^2\Big)^{1/2} \sqrt{2}\|v_h\|_h \\ &\leq \frac{1}{8\varepsilon^2} h_{max}^4 \|q\|_{H^1}^2 + \varepsilon^2 \|v_h\|_h^2. \end{aligned}$$

The previous estimates for $Q_i, i = 1,2$, allow obtaining the following estimate for $(q_h, v_v)_h$

$$\begin{aligned} (q_h, v_h)_h &\leq \tfrac{1}{16\varepsilon^2} h_{max}^4 \Big(2\|q\|_{H^1}^2 + \|q\|_{C(\overline{\Omega})}^2\Big) \\ &\quad + \varepsilon^2 \Big(\|D_{-x}v_h\|_+^2 + \|v_h\|_h^2\Big). \end{aligned}$$

Furthermore, considering the discrete Poincaré -Friedrichs inequality, we establish (2.35).

□

**Proposition 2.5.** *If $p_h$ is a grid function defined in $\Omega_h$ such that there exists a positive constant $C$ satisfying*

$$|p_h(x_i)| \leq C h_{max}^2, i = 1, \ldots, N-1,$$

*then*

$$(p_h, v_h)_h \leq \frac{1}{4\varepsilon^2} C^2 h_{max}^4 + \varepsilon^2 \|D_{-x}v_h\|_+^2, \tag{2.36}$$

*for $v_h \in W_{h,0}$. In (2.36), $\varepsilon \neq 0$ is an abitrary constant.*

**Theorem 2.1.** *Let us suppose that the solution $u$ of the BVP (2.4), (2.5) belongs to $C^4(\overline{\Omega})$, $g \in C^2(\overline{\Omega})$, $A \in C_B^3(\mathbb{R})$, and $f$ is a Lipschitz function with Lipschitz constant $C_L$. Let $u_h \in W_{h,0}$ be defined by (2.12) or by (2.13)-(2.14), where the sequence of grids $\overline{\Omega}_h, h \in \Lambda$, satisfies (2.16). Then, for the error $E_u = u_h - R_h u$, there exists a positive constant $C_u$ such that*

$$\|E_u\|_{1,h}^2 \leq C_u h_{max}^4 \Big(\|g\|_{C^2(\overline{\Omega})}^2 + \|u\|_{C^4(\overline{\Omega})}^2 \big(1 + \|u\|_{C^4(\overline{\Omega})}^4\big)\Big) \tag{2.37}$$

*provided that there exist $\varepsilon \neq 0$ such that*

$$A_0 - \Big(7\varepsilon^2 + \|u'\|_\infty \|A\|_{C_B^1(\mathbb{R})} + C_L(\sqrt{2C_R} + 1)\Big) > 0. \tag{2.38}$$

*Proof.* To obtain an upper bound for $\|E_u\|_{1,h}$, it can be seen from (2.29) that it is necessary to get an estimate for $(T_{h,g} - T_h, E_u)_h$, which is done in what follows taking into account Propositions 2.4 and

2.5, and the definitions of $T_h$ and $T_{h,g}$. For $(T_h^{(1)}, E_u)_h$ and $(T_{h,g}^{(1)}, E_u)_h$, Proposition 2.4 leads to

$$-(T_h^{(1)}, E_u)_h \leq Ch_{max}^4 \|A\|_{C_B^3(\mathbb{R})}^2 \|u\|_{C^4(\overline{\Omega})}^2 \left(1 + \|u\|_{C^4(\overline{\Omega})}^4\right) + 2\varepsilon^2 \|D_{-x}E_u\|_+^2, \tag{2.39}$$

and

$$(T_{h,g}^{(1)}, E_u)_h \leq Ch_{max}^4 \|g\|_{C^2(\overline{\Omega})}^2 + 2\varepsilon^2 \|D_{-x}E_u\|_+^2, \tag{2.40}$$

where $C$ is a positive constant depending on $\varepsilon^2$, with $\varepsilon \neq 0$ is an arbitrary constant.

Proposition 2.5 is now applied to $(T_h^{(2)}, E_u)_h$, $(T_h^{(3)}, E_u)_h$, and $(T_{h,g}^{(2)}, E_u)_h$ leading to

$$-(T_h^{(2)}, E_u)_h \leq Ch_{max}^4 \|u\|_{C^3(\overline{\Omega})}^2 + \varepsilon^2 \|D_{-x}E_u\|_+^2, \tag{2.41}$$

$$-(T_h^{(3)}, E_u)_h \leq Ch_{max}^4 \|A\|_{C_B^3(\mathbb{R})}^2 \|u\|_{C^4(\overline{\Omega})}^2 + \varepsilon^2 \|D_{-x}E_u\|_+^2, \tag{2.42}$$

$$(T_{h,g}^{(2)}, E_u)_h \leq Ch_{max}^4 \|g\|_{C^2(\overline{\Omega})}^2 + \varepsilon^2 \|D_{-x}E_u\|_+^2. \tag{2.43}$$

Inserting the previous estimates in (2.29) we get

$$\begin{aligned} &\left(A_0 - \left(7\varepsilon^2 + \|u'\|_\infty \|A\|_{C_B^1(\mathbb{R})} + C_L(\sqrt{2C_R} + 1)\right)\right) \|D_{-x}E_u\|_+^2 \\ &\quad \leq Ch_{max}^4 \left(\|g\|_{C^2(\overline{\Omega})}^2 + \|u\|_{C^4(\overline{\Omega})}^2 \left(1 + \|u\|_{C^4(\overline{\Omega})}^4\right)\right). \end{aligned} \tag{2.44}$$

Assuming that there exists a positive constant $\varepsilon^2$ such that (2.38) holds we conclude (2.37).

□

### 2.2.4 Stability analysis revisited

Theorem 2.1 states that the finite difference scheme (2.13)-(2.14) or equivalently, the fully discrete piecewise linear FEM (2.12), is second-order convergent with respect to the norm $\|\cdot\|_{1,h}$ which is a discrete version of the usual $H^1$-norm.

To obtain stability of the nonlinear finite difference scheme (2.13)-(2.14), or equivalently, of the fully discrete piecewise linear FEM (2.12), it is needed to establish the uniform boundness of $\|D_{-x}u_h\|_{h,\infty}, h \in \Lambda$. Remark that this was obtained in (2.34) taking into account the first error estimate established in Proposition 2.3. As it was mentioned before, to control the upper bound of $\|D_{-x}u_h\|_{h,\infty}, h \in \Lambda$, it is shown in what follows that this sequence is asymptotically bounded by $\|u'\|_\infty$.

**Corollary 2.1.** *If $u_h \in W_{h,0}$ is defined by the finite difference scheme* (2.13)-(2.14)*, or equivalently, by the fully discrete piecewise linear FEM* (2.12)*, then under the assumptions of the Theorem 2.1, it holds*

$$\|u_h\|_{h,\infty} \leq C_u h_{max}^2 + \|u\|_\infty, \ h \in \Lambda, \tag{2.45}$$

$$\|D_{-x}u_h\|_{h,\infty} \leq C_u C_R h_{max} + \|u'\|_\infty, \ h \in \Lambda. \tag{2.46}$$

*Proof.* Given that

$$u_h(x_i) = E_u(x_i) + R_h u(x_i),\ i = 1,\dots,N-1,$$

and consequently,

$$\begin{aligned}\|u_h\|_{h,\infty} &\leq \|E_u\|_{h,\infty} + \|u\|_\infty \\ &\leq \|D_{-x}E_u\|_+ + \|u\|_\infty \\ &\leq C_u h_{max}^2 + \|u\|_\infty, h \in \Lambda,\end{aligned}$$

which concludes the proof of (2.45).

On the other hand, since

$$D_{-x}u_h(x_i) = D_{-x}E_u(x_i) + D_{-x}R_h u(x_i),\ i = 1,\dots,N,$$

and

$$\begin{aligned}|D_{-x}E_u(x_i)| &\leq \frac{1}{h_{min}} \sum_{j=1}^{N} h_j |D_{-x}E_u(x_j)| \\ &\leq \frac{1}{h_{min}} \|D_{-x}E_u\|_+,\end{aligned}$$

from (2.37) it is obtained

$$\begin{aligned}|D_{-x}E_u(x_i)| &\leq C_u \frac{h_{max}^2}{h_{min}} \\ &\leq C_u C_R h_{max},\end{aligned}$$

which concludes the proof of (2.46).

□

From (2.17) we are able to guarantee the stability around $u_h \in W_{h,0}$ provided that there exists a positive constant $C_s$ such that

$$\|D_{-x}u_h\|_{h,\infty} \leq C_s, h \in \Lambda, \tag{2.47}$$

and

$$A_0 - \left(\|A\|_{C_B^1(\mathbb{R})} C_s + C_L(1+\sqrt{2C_R}) + \varepsilon^2\right) > 0, \tag{2.48}$$

and therefore, we have the stability inequality

$$\left(A_0 - \left(\|A\|_{C_B^1(\mathbb{R})} C_s + C_L(1+\sqrt{2C_R}) + \varepsilon^2\right)\right) \|D_{-x}\omega_h\|_+^2 \leq \frac{1}{4\varepsilon^2} \|g_h - \tilde{g}_h\|_h^2, \tag{2.49}$$

for $h \in \Lambda$ with $h_{max}$ small enough. The stability result can be summarized as follows.

**Proposition 2.6.** *Under the assumptions of Theorem 2.1, if $u_h \in W_{h,0}$ satisfies* (2.13)-(2.14)*, then there exists $\varepsilon \neq 0$ such that for all $v_h \in W_{h,0}$ satisfying* (2.13)-(2.14)*, with $g_h$ replaced by $\tilde{g}_h$, we have (2.49) provided that (2.48) holds, and where $C_s$ satisfies (2.47).*

Proposition 2.6 states the local stability of the FDM (2.13)-(2.14) or equivalently of the fully discrete FEM (2.12) in the numerical approximation $u_h \in W_{h,0}$ defined by (2.13)-(2.14) ([65]). Let denote $w_h, v_h$ solutions of (2.13)-(2.14), with $g_h$ replaced by $\tilde{g}_h, g_h^*$, respectively. In fact if $\tilde{g}_h, g_h^* \in B_{\hat{h}_{max}}(g_h)$ are such that $\|\tilde{g}_h - g_h^*\|_h \leq \hat{h}_{max}$, then $\|w_h - v_h\|_{1,h} \leq h_{max}$, where $\hat{h}_{max} = \Big(A_0 - \big(\|A\|_{C_B^1(\mathbb{R})} C_s + C_L(1 + \sqrt{2C_R}) + \varepsilon^2\big)\Big)^{1/2} \sqrt{2}\varepsilon h_{max}$.

To compare our results with some results presented in the literature, in what concerns stability for nonlinear problems, we realise that for instance in [47] $v_h, w_h$ are chosen on a ball centered in $R_h u \in W_{h,0}$ while here such ball is centered in the numerical solution $u_h$ as in [65]. In particular, for approximations on the ball centered in $R_h u \in W_{h,0}$ holds the same result.

### 2.2.5 Convergence analysis: nonsmooth solutions

Theorem 2.1 states that if $u \in C^4(\overline{\Omega})$, and $g \in C^2(\overline{\Omega})$, then $\|E_u\|_{1,h} \leq C h_{max}^2$. Using the approach introduced in [6, 23], it is intended to improve the previous result, obtaining an analogous estimate but requiring less regularity on the solution, assuming $u \in H^3(\Omega) \cap H_0^1(\Omega)$. The main tool in the convergence analysis is the Bramble-Hilbert Lemma [11]. The basic idea behind this convergence analysis for the nonlinear problem under study is the following. Starting from

$$(F_h(u_h), E_u)_h = (g_h, E_u)_h$$

that can be rewritten in the equivalent form

$$B_h(u_h)(u_h, E_u) = (g_h, E_u)_h, \tag{2.50}$$

where $B_h(u_h)(.,.) : W_{h,0} \times W_{h,0} \to \mathbb{R}$ is a bilinear form. Since

$$\begin{aligned}(g_h, E_u)_h &= \sum_{i=1}^{N-1} \int_{x_{i-1/2}}^{x_{i+1/2}} F(u(x))dx E_u(x_i) \\ &= B_h(R_h u)(R_h u, E_u) + \left( \sum_{i=1}^{N-1} \int_{x_{i-1/2}}^{x_{i+1/2}} F(u(x))dx E_u(x_i) - B_h(R_h u)(R_h u, E_u) \right),\end{aligned}$$

from (2.50) it is deduced the following equality

$$B_h(u_h)(u_h, E_u) - B_h(R_h u)(R_h u, E_u) = \sum_{i=1}^{N-1} \int_{x_{i-1/2}}^{x_{i+1/2}} F(u(x))dx E_u(x_i) - B_h(R_h u)(R_h u, E_u). \tag{2.51}$$

The desired estimate for $E_u$ is obtained proving that

$$B_h(u_h)(u_h, E_u) - B_h(R_h u)(R_h u, E_u) \geq C_e \|E_u\|_{1,h}^2, \tag{2.52}$$

where $C_e > 0$, and

$$\sum_{i=1}^{N-1} \int_{x_{i-1/2}}^{x_{i+1/2}} F(u(x))dx E_u(x_i) - B_h(R_h u)(R_h u, E_u) \leq C h_{max}^4 \|u\|_{H^3(\Omega)}^2 + \gamma^2 \|E_u\|_{1,h}^2, \tag{2.53}$$

where $\gamma \neq 0$ is an arbitrary constant and $C$ depends on $\gamma$. In fact, (2.51), (2.52), and (2.53) allow deducing

$$(C_e - \gamma^2)\|E_u\|_{1,h}^2 \leq Ch_{max}^4 \|u\|_{H^3(\Omega)}^2,$$

and the result follows immediately.

Following the previous lines it can be to proved the following result.

**Theorem 2.2.** *Let us suppose that the sequence of grids $\Lambda$ satisfies* (2.16)*, the solution $u$ of* (2.11) *belongs to $H^3(\Omega) \cap H_0^1(\Omega)$, $A \in C_B^1(\mathbb{R})$ with $A \geq A_0 > 0$, and $f : \mathbb{R}^2 \to \mathbb{R}$ is a Lipschitz function, with Lipschitz constant $C_L$, such that $f(u,u') \in H^2(\Omega)$. Let $u_h \in W_{h,0}$ be a solution of* (2.12) *or* (2.13)-(2.14)*, and let $E_u$ be the discretization error, $E_u = u_h - R_h u$. If*

$$C_L\left(1 + \sqrt{2C_R}\right) + \|A\|_{C_B^1(\mathbb{R})}\|u'\|_\infty < A_0$$

*then there exists a positive constant $C$, $h$-independent, such that*

$$\|D_{-x}E_u\|_+^2 \leq C\sum_{i=1}^{N} h_i^4 \left(\|u\|_{H^3(x_{i-1},x_{i+1})}^2 + \|f(u,u')\|_{H^2(x_{i-1},x_i)}^2\right), \tag{2.54}$$

*for $h \in \Lambda$, $h_{max}$ small enough.*

*Proof.* It can be shown that

$$\begin{aligned}(A(M_h u_h)D_{-x}u_h, D_{-x}E_u)_+ &= (g_h, E_u)_h - (f(u_h, \nabla_h u_h), E_u)_h \\ &= (f(R_h u, \nabla_h R_h u), E_u)_h + (A(M_h R_h u)D_{-x}R_h u, D_{-x}E_u)_+ \\ &+ \sum_{i=0}^{3} T_h^{(i)} - (f(u_h, \nabla_h u_h), E_u)_h,\end{aligned} \tag{2.55}$$

where

$$\begin{aligned}T_h^{(0)} &= ((A(\hat{R}_h u) - A(M_h R_h u))\hat{R}_h u', D_{-x}E_u)_+, \\ T_h^{(1)} &= (A(M_h R_h u)(\hat{R}_h u' - D_{-x}R_h u), D_{-x}E_u)_+, \\ T_h^{(2)} &= ((f(u,u'))_h - R_h f(u,u'), E_u)_h, \\ T_h^{(3)} &= (R_h f(u,u') - f(R_h u, \nabla_h R_h u), E_u)_h,\end{aligned}$$

where $\hat{R}_h : C(\overline{\Omega}) \to W_h$ is defined by $\hat{R}_h v(x_i) = v(x_{i-1/2})$, $i = 1,\dots,N$, $\hat{R}_h v(x_0) = v(x_0)$. To show (2.55) observe that we have successively

$$\begin{aligned}
(g_h, E_u)_h &= \sum_{i=1}^{N-1} \int_{x_{i-1/2}}^{x_{i+1/2}} (-(A(u)u')' + f(u,u'))\,dx E_u(x_i) \\
&= \sum_{i=1}^{N} h_i \left(A(M_h u(x_i))(u'(x_{i-1/2}) - D_{-x}u(x_i)\right) D_{-x}E_u(x_i) \\
&\quad + \sum_{i=1}^{N} h_i (A(u(x_{i-1/2})) - A(M_h u(x_i)))u'(x_{i-1/2}) D_{-x}E_u(x_i) \\
&\quad + \sum_{i=1}^{N} h_i A(M_h u(x_i)) D_{-x}u(x_i) D_{-x}E_u(x_i) \\
&\quad + \sum_{i=1}^{N-1} \left( \int_{x_{i-1/2}}^{x_{i+1/2}} f(u,u')dx - h_{i+1/2} f(u(x_i), u'(x_i)) \right) E_u(x_i) \\
&\quad + \sum_{i=1}^{N-1} h_{i+1/2} \left( f(u(x_i), u'(x_i)) - f(u(x_i), \nabla_h u(x_i)) \right) E_u(x_i) \\
&\quad + \sum_{i=1}^{N-1} h_{i+1/2} f(u(x_i), \nabla_h u(x_i)) E_u(x_i).
\end{aligned}$$

In what follows $T_h^{(i)}$, $i = 0,1,2,3$ are estimated separately.

Estimation of $T_h^{(0)}$

As $A$ has a bounded derivative, therefore is a Lipschitz function with constant L,

$$|A(\hat{R}_h u) - A(M_h R_h u)| \le L|\hat{R}_h u - M_h R_h u|.$$

Let consider

$$|\hat{R}_h u - M_h R_h u| = \left| \omega(1/2) - \frac{\omega(1)+\omega(0)}{2} \right| = |\lambda(\omega)|,\ \omega \in W^{2,1}(\Omega),$$

with $\omega(\xi) = u(x_{i-1} + \xi h_i)$, $\xi \in [0,1]$, $\lambda$ is the linear functional $\lambda : W^{2,1}(\Omega) \to \mathbb{R}$ such that

$$\lambda(g) = g(1/2) - \frac{g(1)+g(0)}{2},\ g \in W^{2,1}(\Omega).$$

Since $\lambda(g) = 0$ for $g = 1, \xi$, and $\lambda$ is bounded in $W^{2,1}(\Omega)$, by Bramble-Hilbert Lemma is obtained

$$|\lambda(g)| \le C \int_0^1 |g''(\xi)| d\xi,$$

which leads to

$$|\hat{R}_h u - M_h R_h u| \le C h_i \int_{x_{i-1}}^{x_i} |u''(s)| ds,$$

taking into account the smoothness assumption on $u$, and being $C$ a positive constant. Since $u \in H^3(\Omega)$, $u'$ is bounded, thus $\|\hat{R}_h u'\|_{h,\infty} \leq \|u'\|_\infty$. Therefore, for a positive constant $C_0$, it is obtained the following estimate for $T_h^{(0)}$

$$\begin{aligned} |T_h^{(0)}| &\leqslant C_0 \left( \sum_{i=1}^{N} h_i^4 \|u''\|^2_{L^2(x_{i-1},x_i)} \right)^{1/2} \|D_{-x}E_u\|_+ \\ &\leqslant C_0 \left( \sum_{i=1}^{N} h_i^4 \|u\|^2_{H^2(x_{i-1},x_i)} \right)^{1/2} \|D_{-x}E_u\|_+ \\ &\leqslant \frac{C_0}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|u\|^2_{H^2(x_{i-1},x_i)} + \varepsilon^2 \|D_{-x}E_u\|^2_+, \end{aligned}$$

where $\varepsilon$ is an arbitrary nonzero real constant.

Estimation of $T_h^{(1)}$

Note that $A(M_h R_h u)$ is bounded, which leads to

$$|T_h^{(1)}| \leq C \left( |\hat{R}_h u' - D_{-x} R_h u|, |D_{-x}E_u| \right)_+,$$

with $C$ a positive constant.

Let consider

$$h_i |\hat{R}_h u' - D_{-x} R_h u| = |\omega'(1/2) - \omega(1) + \omega(0)| = |\lambda(\omega)|, \ \omega \in W^{3,1}(\Omega),$$

with $\omega(\xi) = u(x_{i-1} + \xi h_i)$, $\xi \in [0,1]$, $\lambda$ is the linear functional $\lambda : W^{3,1}(\Omega) \to \mathbb{R}$ defined by

$$\lambda(g) = g'(1/2) - g(1) + g(0), \ g \in W^{3,1}(\Omega).$$

As $\lambda(g) = 0$, for $g = 1, \xi, \xi^2$, and $\lambda$ is bounded in $W^{3,1}(\Omega)$, by Bramble-Hilbert Lemma, it is concluded that

$$|\lambda(g)| \leq C \int_0^1 |g'''(\xi)| d\xi,$$

which implies, under the smoothness assumption on $u$, that

$$h_i |\hat{R}_h u' - D_{-x} R_h u| \leq C h_i^2 \int_{x_{i-1}}^{x_i} |u'''(s)| ds.$$

For a positive constant $C_1$, the estimate of $T_h^{(1)}$ is obtained as follows

$$\begin{aligned} |T_h^{(1)}| &\leqslant C_1 \sum_{i=1}^{N} h_i^2 \int_{x_{i-1}}^{x_i} |u^{(3)}(x)| \, dx \, |D_{-x}E_u(x_i)| \\ &\leqslant C_1 \left( \sum_{i=1}^{N} h_i^4 \|u^{(3)}\|^2_{L^2(x_{i-1},x_i)} \right)^{1/2} \|D_{-x}E_u\|_+ \\ &\leqslant \frac{C_1}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|u\|^2_{H^3(x_{i-1},x_i)} + \varepsilon^2 \|D_{-x}E_u\|^2_+, \end{aligned}$$

for $\varepsilon \neq 0$ an arbitrary constant.

Estimation of $T_h^{(2)}$

To obtain an estimate for $T_h^{(2)}$ it is followed the procedure developed in [6]. Therefore, considering $w(x) = f(u(x), u'(x))$, $T_h^{(2)}$ can be rewritten as follows

$$T_h^{(2)} = \sum_{i=1}^{N-1} \left( \int_{x_{i-1/2}}^{x_{i+1/2}} w(x)dx - h_{i+1/2} w(x_i) \right) E_u(x_i).$$

Note that for

$$\hat{T}_h^{(2)} = \sum_{i=0}^{N-1} \left( -\frac{h_{i+1}}{2}(w(x_i) + w(x_{i+1})) + \int_{x_i}^{x_{i+1}} w(x)dx \right) (E_u(x_i) + E_u(x_{i+1})),$$

and

$$\tilde{T}_h^{(2)} = \sum_{i=0}^{N-1} \left( \frac{h_{i+1}}{2}(w(x_i) - w(x_i+1)) - \int_{x_i}^{x_{i+1/2}} w(x)dx + \int_{x_{i+1/2}}^{x_{i+1}} w(x)dx \right) (E_u(x_{i+1}) - E_u(x_i)),$$

it holds the following equality

$$\left| T_h^{(2)} \right| = \left| \frac{\hat{T}_h^{(2)} + \tilde{T}_h^{(2)}}{2} \right|. \tag{2.56}$$

To get an estimate of $\hat{T}_h^{(2)}$, it is considered, for $\omega \in W^{2,1}(\Omega)$ that

$$\left| -\frac{w(x_i) + w(x_{i+1})}{2} + \frac{1}{h_{i+1}} \int_{x_i}^{x_{i+1}} w(x)dx \right| = \left| -\frac{\omega(0) + \omega(1)}{2} + \int_0^1 \omega(\mu)d\mu \right| = |\lambda(\omega)|,$$

with $\omega(\xi) = u(x_i + \xi h_{i+1})$, $\xi \in [0,1]$, $\lambda$ is the linear functional $\lambda : W^{2,1}(\Omega) \to \mathbb{R}$ defined by

$$\lambda(g) = \frac{g(0) + g(1)}{2} - \int_0^1 g(\mu)d\mu,\ g \in W^{2,1}(\Omega).$$

As $\lambda(g) = 0$, for $g = 1, \xi$, and $\lambda$ is bounded in $W^{2,1}(\Omega)$, by Bramble-Hilbert Lemma, it is concluded that there exists a positive constant $\hat{C}_2$ such that

$$|\lambda(g)| \leq \hat{C}_2 \int_0^1 |g''(\xi)|d\xi,$$

which implies, under the smoothness assumption on $w$, that

$$h_{i+1} \left| -\frac{w(x_i) + w(x_{i+1})}{2} + \frac{1}{h_{i+1}} \int_{x_i}^{x_{i+1}} w(x)dx \right| \leq \hat{C}_2 h_{i+1}^2 \int_{x_{i-1}}^{x_i} |w''(s)|ds.$$

Consequently, the following estimate is obtained

$$|\hat{T}_h^{(2)}| \leqslant \frac{\hat{C}_2}{4\varepsilon^2} \left( \sum_{i=1}^{N} h_{i+1}^4 \|w\|_{H^2(x_{i-1}, x_i)}^2 \right)^{1/2} + \varepsilon^2 \|D_{-x} E_u\|_+^2 \tag{2.57}$$

with $\varepsilon$ an arbitrary constant different from 0.

To get an estimate of $\tilde{T}_h^{(2)}$, it is considered, for $\omega \in W^{1,1}(\Omega)$ that

$$\left| \frac{(w(x_i)-w(x_{i+1})}{2} - \frac{1}{h_{i+1}} \int_{x_i}^{x_{i+1/2}} w(x)dx + \frac{1}{h_{i+1}} \int_{x_{i+1/2}}^{x_{i+1}} w(x)dx \right| = \left| \frac{\omega(0)-\omega(1)}{2} - \int_0^{1/2} \omega(\mu)d\mu + \int_{1/2}^{1} \omega(\mu)d\mu \right| = |\lambda(\omega)|,$$

with $\omega(\xi) = u(x_i + \xi h_{i+1})$, $\xi \in [0,1]$, $\lambda$ is the linear functional $\lambda : W^{1,1}(\Omega) \to \mathbb{R}$ defined by

$$\lambda(g) = \frac{g(1)-g(0)}{2} + \int_0^{1/2} g(\mu)d\mu - \int_{1/2}^{1} g(\mu)d\mu,\ g \in W^{1,1}(\Omega).$$

As $\lambda(g) = 0$ for $g = 1$, and $\lambda$ is bounded in $W^{1,1}(\Omega)$, by Bramble-Hilbert Lemma, it is concluded that there exist a positive constant $\tilde{C}_2$ such that,

$$|\lambda(g)| \leq \tilde{C}_2 \int_0^1 |g'(\xi)|d\xi,$$

which implies, under the smoothness assumption on $w$, that

$$h_{i+1} \left| (w(x_i) + w(x_{i+1})) - \frac{1}{h_{i+1}} \int_{x_i}^{x_{i+1}} w(x)dx \right| \leq \tilde{C}_2 h_{i+1}^2 \int_{x_{i-1}}^{x_i} |w'(s)|ds.$$

Consequently, the following estimate is obtained

$$|\tilde{T}_h^{(2)}| \leqslant \frac{\tilde{C}_2}{4\varepsilon^2} \left( \sum_{i=1}^{N} h_{i+1}^4 \|w\|^2_{H^1(x_{i-1},x_i)} \right)^{1/2} + \varepsilon^2 \|D_{-x}E_u\|_+^2, \tag{2.58}$$

with $\varepsilon$ an arbitrary constant different from 0.

Finally, from (2.56), (2.57), and (2.58), the following estimate is obtained for $T_h^{(2)}$

$$|T_h^{(2)}| \leqslant \frac{C_2}{4\varepsilon^2} \sum_{i=1}^{N} h_{i+1}^4 \|f(u,u')\|^2_{H^2(x_{i-1},x_i)} + \varepsilon^2 \|D_{-x}E_u\|_+^2,$$

providing that $f(u,u') \in H^2(\Omega)$.

Estimation of $T_h^{(3)}$

Taking into account that $f$ is a Lipschitz function with Lipschitz constant $C_L$, it follows that

$$|T_h^{(3)}| \leqslant C_L \sum_{i=1}^{N-1} h_{i+1/2} |u'(x_i) - \nabla_h u(x_i)||E_u(x_i)|.$$

As in [62], is considered now, for $\omega \in W^{3,1}(\Omega)$, the following representation

$$|u'(x_i) - \nabla_h u(x_i)| = \frac{1}{h_{i+1}+h_i} \left| \omega'(\rho) - \left[ \hat{\rho}\,(\omega(1)-\omega(\rho)) + \frac{1}{\hat{\rho}}\,(\omega(\rho)-\omega(0)) \right] \right| = \frac{1}{h_{i+1}+h_i} |\lambda(\omega)|,$$

where $\omega(\xi) = u(x_{i-1} + \xi(h_i + h_{i+1}))$, $\xi \in [0,1]$, $\rho = \frac{h_i}{h_i+h_{i+1}}$, $\hat{\rho} = \frac{h_i}{h_{i+1}}$, and $\lambda$ is the linear functional $\lambda : W^{3,1}(\Omega) \to \mathbb{R}$ defined by

$$\lambda(g) = g'(\rho) - \left[\hat{\rho}\,(g(1) - g(\rho)) + \frac{1}{\hat{\rho}}\,(g(\rho) - g(0))\right],\ g \in W^{3,1}(\Omega).$$

Since the functional $\lambda$ is bounded in $W^{3,1}(\Omega)$, and $g = 1, \xi, \xi^2$ implies $\lambda(g) = 0$, then, by Bramble-Hilbert lemma, there exists a positive constant $C$ such that

$$|\lambda(g)| \leq C \int_0^1 |g'''(\xi)| d\xi,$$

which leads to

$$|u'(x_i) - \nabla_h u(x_i)| \leq C(h_i + h_{i+1}) \int_{x_{i-1}}^{x_{i+1}} |u'''(s)| ds.$$

The previous inequality leads to the following estimate for $T_h^{(3)}$,

$$\begin{aligned} |T_h^{(3)}| &\leqslant C_3 \left( \sum_{i=1}^{N} h_i^4 \|u^{(3)}\|^2_{L^2(x_{i-1},x_i)} \right)^{1/2} \|D_{-x}E_u\|_+ \\ &\leqslant \frac{C_3}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|u\|^2_{H^3(x_{i-1},x_i)} + \varepsilon^2 \|D_{-x}E_u\|^2_+, \end{aligned}$$

with $C_3$ a positive constant, and $\varepsilon$ an arbitrary nonzero constant.

Taking into account representation (2.55) it follows successively

$$\begin{aligned} (A(M_h u_h) D_{-x}E_u, D_{-x}E_u)_+ - 4\varepsilon^2 \|D_{-x}E_u\|^2_+ &\leq (f(R_h u, \nabla_h R_h u) - f(u_h, \nabla_h u_h), E_u)_h \\ &\quad + ((A(M_h R_h u) - A(M_h u_h)) D_{-x} R_h u, D_{-x}E_u)_+ + T_h \\ &\leq C_L \left( \sqrt{2C_R} \|E_u\|_h \|D_{-x}E_u\|_+ + \|E_u\|^2_h \right) \\ &\quad + \|A\|_{C^1_B(\mathbb{R})} \|D_{-x}R_h u\|_{h,\infty} \|E_u\|_h \|D_{-x}E_u\|_+ + T_h \\ &\leq \left( C_L \left( 1 + \sqrt{2C_R} \right) + \|A\|_{C^1_B(\mathbb{R})} \|u'\|_\infty \right) \|D_{-x}E_u\|^2_+ + T_h, \end{aligned}$$

where the last inequality was established taking into account Proposition 2.1, and having the following estimate for $T_h$

$$T_h \leq \left( \sum_{j=0}^{3} \frac{C_j}{4\varepsilon^2} \right) \sum_{i=1}^{N} h_i^4 \left( \|u\|^2_{H^3(x_{i-1},x_i)} + \|f(u,u')\|^2_{H^2(x_{i-1},x_i)} \right).$$

Therefore, it is obtained that

$$\left( A_0 - 4\varepsilon^2 - \left( C_L \left( 1 + \sqrt{2C_R} \right) + \|A\|_{C^1_B(\mathbb{R})} \|u'\|_\infty \right) \right) \|D_{-x}E_u\|^2_+ \leq T_h. \tag{2.59}$$

If $C_L \left(1 + \sqrt{2C_R}\right) + \|A\|_{C^1_B(\mathbb{R})} \|u'\|_\infty < A_0$, then there exists $\varepsilon \neq 0$ such that

$$A_0 - 4\varepsilon^2 - \left( C_L \left( 1 + \sqrt{2C_R} \right) + \|A\|_{C^1_B(\mathbb{R})} \|u'\|_\infty \right) > 0,$$

which leads to the existence of a positive constant $C$ satisfying (2.54). □

From Proposition 2.1 and Theorem 2.2 it can be concluded for the norm $\|\cdot\|_{1,h}$ of the error the following estimate.

**Corollary 2.2.** *Under the assumptions of Theorem 2.2, there exists a positive constant $C_u$, h-independent, such that the following bound holds for the error $E_u = u_h - R_h u$,*

$$\|E_u\|_{1,h} \leq C_u h_{max}^2.$$

### 2.2.6 Numerical simulations

In what follows, are presented numerical examples to illustrate the main convergence result, Theorem 2.2. These numerical experiments also allow showing the sharpness of the smoothness assumptions imposed in these result. To simplify is considered $\Omega = (0,1)$, and $A(u) = 1$. It is also used random nonuniform grids for the spatial discretization.

As a first example it is considered

$$f(x_1,x_2) = \cos(x_1) + \sin(x_2),\ x_1,x_2 \in \mathbb{R},$$

and $g$ such that the BVP (2.4)-(2.5) has the solution

$$u(x) = (e^x - 1)(x-1),\ x \in \overline{\Omega}.$$

In this case $u \in H^3(\Omega)$, and $f$ satisfies the assumptions of the Theorem 2.2.

In what concerns the second example it is considered

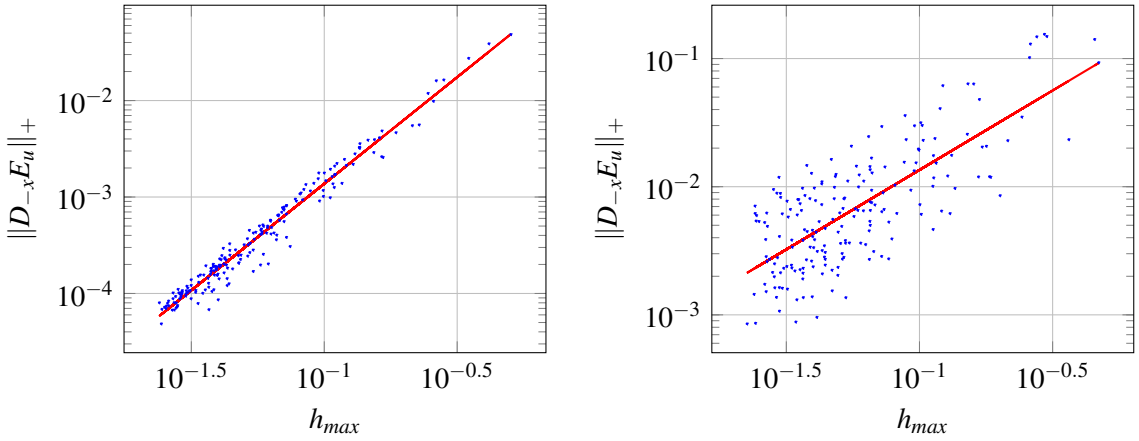
$$f(x_1,x_2) = \cos(x_1) + \sin(x_2),\ x_1,x_2 \in \mathbb{R},$$

and $g$ is such that

$$u(x) = |2x-1|^{1.6} - 1,\ x \in \overline{\Omega},$$

is solution of the BVP (2.4)-(2.5). In this case $u \in H^2(\Omega)$.

The discrete errors in these two situations are reported in Figure 2.1.




(a) $u \in H^3(\Omega)$ and the line has a slope of 2.21.

(b) $u \in H^2(\Omega)$ and the line has a slope of 1.24.

Fig. 2.1 Log-log plots of $\|D_{-x}E_u\|_+$ *versus* $h_{max}$ for the elliptic equation. The solid lines represent least-squares fittings.

The slope of the line in the left-hand side of Figure 2.1 is 2.21, which confirms that if it is assumed the smoothness assumption specified in Theorem 2.2, i. e. $u \in H^3(\Omega)$, then it is concluded a second-order convergence rate. The slope of the line in the plot in the right-hand side is 1.24, which shows that considering a weaker assumption of smoothness of the solution ($u \in H^2(\Omega)$), the convergence rate decreases.

## 2.3 A nonlinear system of elliptic equations

### 2.3.1 Convergence analysis

In this section Theorem 2.2 is extended for the solution of the following FDM

$$-D_x^*(A(M_hU)D_{-x}U_h) + F(U_h, \nabla_h U_h) = G_h \text{ in } \Omega_h, \tag{2.60}$$

where $U_h = (u_{1,h}, u_{2,h})$, with boundary conditions

$$U_h = 0 \quad \text{on } \partial\Omega_h, \tag{2.61}$$

that leads to an approximation for the solution of the differential system (2.6)-(2.7). In (2.60), for $V_h = (v_{1,h}, v_{2,h})$, $D_{-x}V_h = (D_{-x}v_{1,h}, D_{-x}v_{2,h})$, $D_x^*V_h = (D_x^*v_{1,h}, D_x^*v_{2,h})$, $\nabla_h V_h = (\nabla_h v_{1,h}, \nabla_h v_{2,h})$, $A(M_hV_h)$ is a diagonal matrix with diagonal entries $\alpha(M_hv_{1,h}), \beta(M_hv_{2,h})$, and $G_h = (g_{1,h}, g_{2,h})$ with $g_{\ell,h}$ defined by (2.10) with $g$ replaced by $g_\ell$, $\ell = 1, 2$.

Remark that the FEM (2.60)-(2.61) is equivalent to the following fully discrete piecewise linear FEM:

find $U_h \in [W_{h,0}]^2$ such that

$$(A(M_hU_h)D_{-x}U_h, D_{-x}Q_h)_+ + (F(U_h, \nabla_h U_h), Q_h)_h = (G_h, Q_h)_h, \ \forall Q_h \in [W_{h,0}]^2 \tag{2.62}$$

where the inner product $(\cdot,\cdot)_h$ in $[W_{h,0}]^2$ is defined in the usual way. In (2.62), if $Q_h=(q_{1,h},q_{2,h}), V_h=(v_{1,h},v_{2,h})$, then $(Q_h,V_h)_+=\sum_{i=1,2}(q_{i,h},v_{i,h})_+$.

The fully discrete variational problem is a fully discrete version of the variational problem:

find $U\in[H_0^1(\Omega)]^2$ such that

$$((AU)\nabla U,\nabla Q)+(F(U,\nabla U),Q)=(G,Q),\ \forall Q\in[H_0^1(\Omega)]^2. \tag{2.63}$$

For $V_h=(v_{1,h},v_{2,h})\in[W_{h,0}]^2$ the following notation is used $\|D_{-x}V_h\|_+^2=\sum_{i=1,2}\|D_{-x}v_{i,h}\|_+^2$, and $\|D_{-x}V_h\|_{h,\infty}=\max_{\ell=1,2}\max_{i=1,\ldots,N}|D_{-x}v_{\ell,h}(x_i)|$.

Furthermore, it is also used the notation: if $A$ is a 2-diagonal matrix with entries $\alpha$ and $\beta$ then

$$\|A\|_{C_B^1(\mathbb{R})}=\max\{\|\alpha\|_{C_B^1(\mathbb{R})},\|\beta\|_{C_B^1(\mathbb{R})}\}.$$

**Theorem 2.3.** *Let us suppose that the sequence of grids $\Lambda$ satisfies* (2.16)*, the solution $U=(u_1,u_2)$ of* (2.63) *belongs to $[H^3(\Omega)\cap H_0^1(\Omega)]^2$, $F(U,\nabla U)\in[H^2(\Omega)]^2$, $f_i, i=1,2$, are Lipschitz functions with Lipschitz constant $C_L$, the diagonal entries of $A$, $\alpha$ and $\beta$, have a lower positive bound $A_0$, and $\alpha,\beta\in C_B^1(\mathbb{R})$. Let $U_h=(u_{1,h},u_{2,h})\in[W_{h,0}]^2$ be solution of the FDM* (2.60)-(2.61) *or, equivalently, of the fully discrete piecewise FEM* (2.62)*, and let $E_U=U_h-R_hU$ be the discretization error. If*

$$A_0-\left(2\left(1+\sqrt{2C_R}\right)C_L+\|A\|_{C_B^1(\mathbb{R})}\|D_{-x}R_hU\|_{h,\infty}\right)>0. \tag{2.64}$$

*then there exists a positive constant $C$, $h$-independent, such that*

$$\|D_{-x}E_U\|_+^2\le C\sum_{i=1}^N h_i^4\left(\|U\|_{[H^3(x_{i-1},x_i)]^2}^2+\|F(U,\nabla U)\|_{[H^2(x_{i-1},x_i)]^2}^2\right), \tag{2.65}$$

*for $h\in\Lambda$ and $h_{max}$ small enough.*

*Proof.* The proof of this result follows the proof of Theorem 2.2.

Let $E_U=(E_1,E_2)$. It is observed that for $Q_1:=(f_1(R_hU,\nabla_hR_hU)-f_1(U_h,\nabla_hU_h),E_1)_h$ we have successively the following inequalities

$$\begin{aligned}
Q_1&\le C_L\left(\left(\|E_1\|_h+\sqrt{2C_R}\|D_{-x}E_1\|_++\|E_2\|_h+\sqrt{2C_R}\|D_{-x}E_2\|_+\right)\|E_1\|_h\right)\\
&\le C_L\left((1+\sqrt{2C_R})\|D_{-x}E_1\|_+^2+(1+\sqrt{2C_R})\|D_{-x}E_2\|_+\|D_{-x}E_1\|_+\right)\\
&\le C_L\left(\frac{3}{2}(1+\sqrt{2C_R})\|D_{-x}E_1\|_+^2+\frac{1}{2}(1+\sqrt{2C_R})\|D_{-x}E_2\|_+^2\right).
\end{aligned}$$

Analogously, for $Q_2:=(f_2(R_hU,\nabla_hR_hU)-f_2(U_h,\nabla_hU_h),E_2)_h$ it is easily get

$$Q_2\le C_L\left(\frac{1}{2}(1+\sqrt{2C_R})\|D_{-x}E_1\|_+^2+\frac{3}{2}(1+\sqrt{2C_R})\|D_{-x}E_2\|_+^2\right).$$

Then for $Q_1+Q_2$ it is deduced

$$Q_1+Q_2\le 2C_L\left(1+\sqrt{2C_R}\right)\|D_{-x}E_U\|_+^2.$$

Consequently, following the proof of Theorem 2.2 it is obtained

$$\left(A_0 - 4\varepsilon^2 - \left(2(1+\sqrt{2C_R})C_L + \|A\|_{C_B^1(\mathbb{R})}\|D_{-x}R_hU\|_{h,\infty}\right)\right)\|D_{-x}E_U\|_+^2$$
$$\leq C\sum_{i=1}^{N} h_i^4\left(\|U\|^2_{[H^3(x_{i-1},x_i)]^2} + \|F(U,\nabla U)\|^2_{[H^2(x_{i-1},x_i)]^2}\right),$$

where $\varepsilon \neq 0$.

If (2.64) holds, then there exists a positive constant $C$ such that (2.65) holds.

□

**Corollary 2.3.** *Under the assumptions of Theorem 2.3, there exists a positive constant C, h-independent, such that*

$$\|E_U\|_{1,h} \leq C\,h_{max}^2,$$

*for $h \in \Lambda$ with $h_{max}$ small enough.*

Under the assumptions of Theorem 2.3, and following the proof of Corollary 2.1 the next result is easily proved.

**Corollary 2.4.** *Let $U_h \in [W_{h,0}]^2$ be solution of the FDM* (2.60)-(2.61)*, or equivalently, of the fully discrete piecewise FEM* (2.62)*. Then, under the conditions of Theorem 2.3, there exists a positive constant $C_s$, $h-$independent, such that*

$$\|U_h\|_{h,\infty} \leq C_s$$

$$\|D_{-x}U_h\|_{h,\infty} \leq C_s$$

*for $h \in \Lambda$ with $h_{max}$ small enough.*

The stability analysis for the FDM (2.60)-(2.61), or equivalently, of the fully discrete piecewise FEM (2.62) can be carried out following the steps of the analysis presented before for the case of the nonlinear elliptic equation.

### 2.3.2 Numerical simulations

In the numerical experiments concerning the scheme (2.60)-(2.61) it is considered, as before, $\Omega = (0,1)^2$, $A$ is the identity matrix, and it is also used random nonuniform grids for the spatial discretization. The nonlinear system is numerically solved using the fixed-point method, with iteration function $R(U) = B^{-1}(G - F(U))$, where $B$ is the tridiagonal matrix by blocks related to the discretization $-D_x^*(A(M_hU_h)D_{-x}U_h)$, $F$ is the vector representing the nonlinear term, and G is the vector representing the source term.

In the first example, it is considered $F = (f_1, f_2)$ defined by

$$f_1(x_1,x_2,x_3,x_4) = \cos(x_1) + \sin(x_2) + \cos(x_3) + \sin(x_4),\ x_1,x_2,x_3,x_4 \in \mathbb{R},$$

$$f_2(x_1,x_2,x_3,x_4) = x_1 + \sin(x_2) + x_3 + \sin(x_4),\ x_1,x_2,x_3,x_4 \in \mathbb{R}.$$

The function $G$ is such that $U = (u_1, u_2)$ with

$$u_1(x) = (e^x - 1)(x-1),\ x \in \overline{\Omega}, \tag{2.66}$$

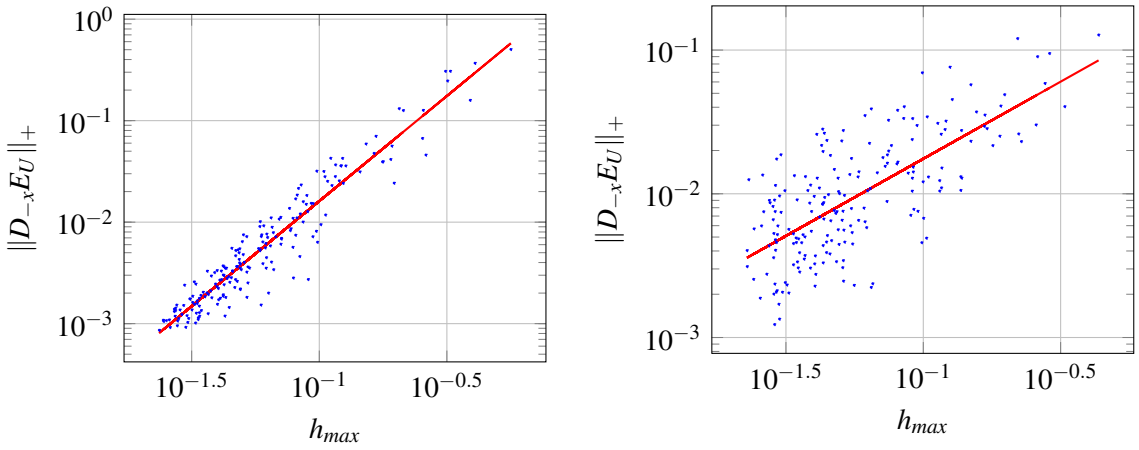
$$u_2(x) = |2x-1|^4 - 1,\ x \in \overline{\Omega},$$

is solution of the BVP (2.6)-(2.7). It is observed that $U \in [H^3(\Omega)]^2$, and $F$ satisfies the assumptions of Theorem 2.3.

In the second example $F$ is considered as before, $u_1$ is given by (2.66) and $u_2$ defined by

$$u_2(x) = |2x-1|^{1.52} - 1,\ x \in \overline{\Omega}.$$

In this case $U \in [H^2(\Omega)]^2$.




(a) $U \in [H^3(\Omega)]^2$ and the line has a slope of 2.07.

(b) $U \in [H^2(\Omega)]^2$ and the line has a slope of 1.07.

Fig. 2.2 Log-log plots of $\|D_{-x}E_U\|_+$ *versus* $h_{max}$ for the system of elliptic equations. The solid lines represent least-squares fittings.

Figure 2.2 illustrates the behaviour of the numerical method (2.60)-(2.61), in the last two scenarios: the results in the plot in the left-hand side were obtained with $U \in [H^3(\Omega)]^2$, and the estimated convergence rate is 2.07, while in the plot in the right-hand side we take $U \in [H^2(\Omega)]^2$, and the estimated convergence rate is 1.07. These results illustrate the sharpness of Theorem 2.3 on smoothness assumptions on the solutions. In fact, the second-order convergence rate is lost when the solution is in $[H^2(\Omega)]^2$.

## 2.4 A nonlinear system of parabolic equations

### 2.4.1 Convergence analysis

To compute an approximation for the IBVP (2.1)-(2.3), it is proposed the following semi-discrete scheme

$$\frac{dU_h}{dt}(t) + F(U_h(t), \nabla_h U_h(t)) = D_x^*(A(M_h U_h(t))D_{-x}U_h(t)) + G_h(t), t \in (0,T] \tag{2.67}$$

with boundary and initial conditions

$$U_h = 0 \text{ on } \partial\Omega_h \times (0,T], \tag{2.68}$$

$$U_h(0) = R_h U_0 \text{ in } \Omega_h. \tag{2.69}$$

Remark that the semi-discrete scheme is equivalent to the following fully discrete piecewise linear FEM

find $U_h(t) \in [W_{h,0}]^2$ such that for all $t \in (0,T]$ and $Q_h \in [W_{h,0}]^2$

$$\left(\frac{dU_h}{dt}(t), Q_h\right)_h + (A(M_h U_h(t))D_{-x}U_h(t), D_{-x}Q_h)_+ + (F(U_h(t), \nabla_h U_h(t)), Q_h)_h = (G_h(t), Q_h)_h, \tag{2.70}$$

with

$$(U_h(0), Q_h)_h = (R_h U_0, Q_h)_h, \ \forall Q_h \in [W_{h,0}]^2. \tag{2.71}$$

This fully discrete method is obtained from the piecewise linear FEM

find $U_h(t) \in [W_{h,0}]^2$ such that for all $t \in (0,T]$ and $Q_h \in [W_{h,0}]^2$

$$\left(\frac{dP_h U_h}{dt}(t), P_h Q_h\right) + (A(P_h U_H(t))\nabla P_h U_h(t), \nabla P_h Q_h) + (F(P_h U_h(t), \nabla P_h U_h), P_h Q_h) = (G(t), P_h Q_h),$$

with

$$(P_h U_h(0), P_h Q_h) = (U_0, P_h Q_h), \forall Q_h \in [W_{h,0}]^2$$

considering suitable quadrature rules.

**Theorem 2.4.** *Let us suppose that the sequence of grids $\Lambda$ satisfies* (2.16)*, and $U(t) \in [H^3(\Omega) \cap H_0^1(\Omega)]^2$, $U \in [C^1([0,T], C(\overline{\Omega}))]^2$, $F(U(t), \nabla U(t)) \in [H^2(\Omega)]^2$, $F$ is a Lipschitz function with Lipschitz constant $C_L$, the diagonal entries of $A$, $\alpha$ and $\beta$, have a lower positive bound $A_0$, and $\alpha, \beta \in C_B^1(\mathbb{R})$. Let $U_h(t) \in [W_{h,0}]^2$ be a solution of the initial value problem* (2.67)-(2.69) *or, equivalently, solution of the fully discrete piecewise linear FEM* (2.70)- (2.71)*, and suppose that $U_h \in [C^1([0,T], W_{h,0})]^2$. Let $E_U = U_h - R_h U$ be the discretization error. Then, for any $\varepsilon \neq 0$ there exists a positive constant $C$, $h$ and $t$ independent such that*

$$\begin{aligned} \|E_U(t)\|_h^2 + D\int_0^t e^{\int_s^t S(U(\mu))d\mu} \|D_{-x}E_U(s)\|_+^2 ds \ &\leq e^{\int_0^t S(U(\mu))d\mu} \|E_U(0)\|_h^2 \\ &+ \int_0^t e^{\int_s^t S(U(\mu))d\mu} T_h(s)ds, \ t \in [0,T], \end{aligned} \tag{2.72}$$

*where*

$$T_h(t) = C\sum_{i=1}^N h_i^4 \left( \left\| \frac{\partial U}{\partial t}(t) \right\|^2_{[H^2(x_{i-1},x_i)]^2} + \|U(t)\|^2_{[H^3(x_{i-1},x_i)]^2} + \|F(U(t), \nabla U(t))\|^2_{[H^2(x_{i-1},x_i)]^2} \right),$$

*the coefficients D and S being given by*

$$D = 2(A_0 - 6\varepsilon^2), \tag{2.73}$$

*and*

$$S(U(t)) \quad = \frac{1}{2\varepsilon^2}\|A\|^2_{C^1_B(\mathbb{R})}\|D_{-x}R_hU(t)\|^2_{h,\infty} + C_L\left(4 + 2\frac{C_LC_R}{\varepsilon^2}\right), \tag{2.74}$$

*respectively.*

*Proof.* Let $U(t) = (u_1(t), u_2(t))$ and $E_U(t) = (E_1(t), E_2(t))$. It can be shown that

$$\begin{aligned}
(g_{1,h}(t), E_1(t))_h = &\left(R_h\frac{\partial u_1}{\partial t}(t), E_1(t)\right)_h + \left(\left(\frac{\partial u_1}{\partial t}(t)\right)_h - R_h\frac{\partial u_1}{\partial t}(t), E_1(t)\right)_h \\
&+ (\alpha(M_hu_{1,h}(t))D_{-x}R_hu_1(t), D_{-x}E_1(t))_+ \\
&+ \left(\left(\alpha(\hat{R}_hu_1(t)) - \alpha(M_hR_hu_1(t))\right)\hat{R}_h\frac{\partial u_1}{\partial x}(t), D_{-x}E_1(t)\right)_+ \\
&+ \left(\alpha(M_hR_hu_1(t))\left(\hat{R}_h\frac{\partial u_1}{\partial x}(t) - D_{-x}R_hu_1(t)\right), D_{-x}E_1(t)\right)_+ \\
&+ \left(\Big(\alpha(M_hR_hu_1(t)) - \alpha(M_hu_{1,h}(t))\Big)D_{-x}R_hu_1(t), D_{-x}E_1(t)\right)_+ \\
&+ (f_1(R_hU(t), \nabla_hR_hU(t)), E_1(t))_h \\
&+ (f_{1,h}(t) - f_1(t), E_1(t))_h \\
&+ (f_1(t) - (f_1(R_hU(t), \nabla_hR_hU(t)), E_1(t))_h,
\end{aligned}$$

where, to simplify, the following notation is used

$$f_1(t) = f_1(R_hU(t), R_h\nabla U(t)),$$

and $f_{1,h}(t)$ is defined by (2.10) with $g$ replaced by $f_1(t)$.

Taking this into account, from the first equation of (2.70) with $q_{1,h} = E_1(t)$, it is easily obtained

$$\begin{aligned}
\left(\frac{dE_1}{dt}(t), E_1(t)\right)_h \quad &+(\alpha(M_hu_{1,h}(t))D_{-x}E_1(t), D_{-x}E_1(t))_+ \\
&= (f_1(R_hU(t), \nabla_hR_hU(t)) - f_1(U_h(t), \nabla_hU_h(t)), E_1(t))_h \\
&\quad + ((\alpha(M_hR_hu_1(t)) - \alpha(M_hu_{1,h}(t)))D_{-x}R_hu_1(t), D_{-x}E_1(t))_+ \\
&\quad + \sum_{i=0}^{4} T_h^{(i)},
\end{aligned} \tag{2.75}$$

where

$$
\begin{aligned}
T_h^{(0)} &= \left( \left( \alpha(\hat{R}_h u_1(t)) - \alpha(M_h R_h u_1(t)) \right) \hat{R}_h \frac{\partial u_1}{\partial x}(t), D_{-x} E_1(t) \right)_+, \\
T_h^{(1)} &= \left( \alpha(M_h R_h u_1(t)) \left( \hat{R}_h \frac{\partial u_1}{\partial x}(t) - D_{-x} R_h u_1(t) \right), D_{-x} E_1(t) \right)_+, \\
T_h^{(2)} &= (f_{1,h}(t) - f_1(t), E_1(t))_h, \\
T_h^{(3)} &= (f_1(t) - f_1(R_h U(t), \nabla_h R_h U(t)), E_1(t))_h, \\
T_h^{(4)} &= \left( R_h \frac{\partial u_1}{\partial t}(t) - \left( \frac{\partial u_1}{\partial t} \right)_h, E_1(t) \right)_h.
\end{aligned}
$$

It can be shown that there exist positive constants $C_i, i = 0, \ldots, 4$, $h$ and $t$ independent, such that

$$
\begin{aligned}
|T_h^{(0)}| &\le C_0 \left( \sum_{i=1}^{N} h_i^4 \, \|u_1(t)\|^2_{H^2(x_{i-1}, x_i)} \right)^{1/2} \|D_{-x} E_1(t)\|_+ \\
&\le \frac{C_0}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \, \|u_1(t)\|^2_{H^2(x_{i-1}, x_i)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+ \\
&= \hat{T}_h^{(0)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+
\end{aligned}
$$

$$
\begin{aligned}
|T_h^{(1)}| &\le C_1 \left( \sum_{i=1}^{N} h_i^4 \|u_1(t)\|^2_{H^3(x_{i-1}, x_i)} \right)^{1/2} \|D_{-x} E_1(t)\|_+ \\
&\le \frac{C_1}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|u_1(t)\|^2_{H^3(x_{i-1}, x_i)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+ \\
&= \hat{T}_h^{(1)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+
\end{aligned}
$$

$$
\begin{aligned}
|T_h^{(2)}| &\le C_2 \left( \sum_{i=1}^{N} h_i^4 \|f_1(U(t), \nabla U(t))\|^2_{H^2(x_{i-1}, x_i)} \right)^{1/2} \|D_{-x} E_1(t)\|_+ \\
&\le \frac{C_2}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|f_1(U(t), \nabla U(t))\|^2_{H^2(x_{i-1}, x_i)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+ \\
&= \hat{T}_h^{(2)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+
\end{aligned}
$$

$$
\begin{aligned}
|T_h^{(3)}| &\le C_3 \left( \sum_{i=1}^{N} h_i^4 \|U(t)\|^2_{[H^3(x_{i-1}, x_i)]^2} \right)^{1/2} \|D_{-x} E_1(t)\|_+ \\
&\le \frac{C_3}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \|U(t)\|^2_{[H^3(x_{i-1}, x_i)]^2} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+ \\
&= \hat{T}_h^{(3)} + \varepsilon^2 \|D_{-x} E_1(t)\|^2_+
\end{aligned}
$$

$$
\begin{aligned}
|T_h^{(4)}| &\leq C_4 \left( \sum_{i=1}^{N} h_i^4 \left\| \frac{\partial u_1}{\partial t}(t) \right\|^2_{H^2(x_{i-1},x_i)} \right)^{1/2} \|D_{-x}E_1(t)\|_+ \\
&\leq \frac{C_4}{4\varepsilon^2} \sum_{i=1}^{N} h_i^4 \left\| \frac{\partial u_1}{\partial t}(t) \right\|^2_{H^2(x_{i-1},x_i)} + \varepsilon^2 \|D_{-x}E_1(t)\|_+^2 \\
&= \hat{T}_h^{(4)} + \varepsilon^2 \|D_{-x}E_1(t)\|_+^2
\end{aligned}
$$

Note that the estimates of $T_h^{(i)}$, for $i = 0, ..., 3$, are obtained exactly in the same way as, respectively, $T_h^{(i)}$, for $i = 0, ..., 3$, defined for the case of the elliptic equation in Section 2.2.5. Considering $w(x) = \dfrac{\partial u_1}{\partial t}(t)$, $T_h^{(4)}$ can be rewritten as follows

$$
T_h^{(4)} = \sum_{i=1}^{N-1} \left( h_{i+1/2} w(x_i) - \int_{x_{i-1/2}}^{x_{i+1/2}} w(x)dx \right) E_1(x_i),
$$

and therefore, the estimate of $T_h^{(4)}$ is obtained following the same procedure as the one used to obtain the estimate of $T_h^{(2)}$.

Taking into account that $f$ is a Lipschitz function (with Lipschitz constant $C_L$), it is also obtained the following

$$
\begin{aligned}
&(f_1(R_hU(t), \nabla_h R_h U(t)) - f_1(U_h(t), \nabla_h U_h(t)), E_1(t))_h \\
&\quad \leq C_L \left( \|E_1(t)\|_h + \|E_2(t)\|_h + \sqrt{2C_R} \left( \|D_{-x}E_1(t)\|_+ + \|D_{-x}E_2(t)\|_+ \right) \right) \|E_1(t)\|_h \\
&\quad \leq C_L \left( \frac{3}{2} + \frac{C_L C_R}{\varepsilon^2} \right) \|E_1(t)\|_h^2 + \frac{C_L}{2} \|E_2(t)\|_h^2 + \varepsilon^2 \left( \|D_{-x}E_1(t)\|_+^2 + \|D_{-x}E_2(t)\|_+^2 \right),
\end{aligned}
$$

and, as $\alpha \in C_B^1(\mathbb{R})$, it follows that

$$
\begin{aligned}
&((\alpha(M_h R_h u_1(t)) - \alpha(M_h u_{1,h}(t)))D_{-x}R_h u_1(t), D_{-x}E_1(t))_+ \\
&\quad \leq \|A\|_{C_B^1(\mathbb{R})} \|D_{-x}R_h u_1(t)\|_{h,\infty} \|E_1(t)\|_h \|D_{-x}E_1(t)\|_+ \\
&\quad \leq \frac{1}{4\varepsilon^2} \|A\|^2_{C_B^1(\mathbb{R})} \|D_{-x}R_h u_1(t)\|^2_{h,\infty} \|E_1(t)\|_h^2 + \varepsilon^2 \|D_{-x}E_1(t)\|_+^2,
\end{aligned}
$$

where $\varepsilon \neq 0$ is an arbitrary positive constant.

Considering the previous estimates in (2.75) it is deduced

$$
\begin{aligned}
&\frac{d}{dt}\|E_1(t)\|_h^2 + 2\left(A_0 - 7\varepsilon^2\right) \|D_{-x}E_1(t)\|_+^2 \\
&\quad \leq \left( \frac{1}{2\varepsilon^2} \|A\|^2_{C_B^1(\mathbb{R})} \|D_{-x}R_h u_1(t)\|^2_{h,\infty} + C_L \left( 3 + 2\frac{C_L C_R}{\varepsilon^2} \right) \right) \|E_1(t)\|_h^2 \\
&\quad + C_L \|E_2(t)\|_h^2 + 2\varepsilon^2 \|D_{-x}E_2(t)\|_+^2 + T_{1,h}(t),
\end{aligned}
\tag{2.76}
$$

where

$$T_{1,h}(t) = 2\sum_{i=0}^{4}\hat{T}_h^{(i)} = \frac{\tilde{C}_1}{2\varepsilon^2}\left(\sum_{i=1}^{N} h_i^4\left(\left\|\frac{\partial u_1}{\partial t}(t)\right\|^2_{H^2(x_{i-1},x_i)} + \|U(t)\|^2_{[H^3(x_{i-1},x_i)]^2} + \|f_1(U(t),\nabla U(t))\|^2_{H^2(x_{i-1},x_i)}\right)\right),$$

for some positive constant $\tilde{C}_1$, $h$ and $t$ independent.

Analogously, from the second equation of (2.70) with $q_{2,h} = E_2(t)$ we can establish for $\varepsilon \neq 0$,

$$\begin{aligned}\frac{d}{dt}\|E_2(t)\|_h^2 &+ 2\left(A_0 - 7\varepsilon^2\right)\|D_{-x}E_2(t)\|_+^2 \\ &\leq \left(\frac{1}{2\varepsilon^2}\|A\|^2_{C_B^1(\mathbb{R})}\|D_{-x}R_h u_2(t)\|^2_{h,\infty} + C_L\left(3 + 2\frac{C_L C_R}{\varepsilon^2}\right)\right)\|E_2(t)\|_h^2 \\ &+ C_L\|E_1(t)\|_h^2 + 2\varepsilon^2\|D_{-x}E_1(t)\|_+^2 + T_{2,h}(t),\end{aligned} \tag{2.77}$$

with

$$T_{2,h}(t) \leq \frac{\tilde{C}_2}{2\varepsilon^2}\left(\sum_{i=1}^{N} h_i^4\left(\left\|\frac{\partial u_2}{\partial t}(t)\right\|^2_{H^2(x_{i-1},x_i)} + \|U(t)\|^2_{[H^3(x_{i-1},x_i)]^2} + \|f_2(U(t),\nabla U(t))\|^2_{H^2(x_{i-1},x_i)}\right)\right),$$

for some positive constant $\tilde{C}_2$, $h$ and $t$ independent.

Combining (2.76) and (2.77) we get the following differential inequality

$$\begin{aligned}\frac{d}{dt}\|E_U(t)\|_h^2 &+ 2\left(A_0 - 6\varepsilon^2\right)\|D_{-x}E_U(t)\|_+^2 \\ &\leq \left(\frac{1}{2\varepsilon^2}\|A\|^2_{C_B^1(\mathbb{R})}\|D_{-x}R_h U(t)\|^2_{h,\infty} + C_L\left(4 + 2\frac{C_L C_R}{\varepsilon^2}\right)\right)\|E_U(t)\|_h^2 \\ &+ T_{1,h}(t) + T_{2,h}(t)\end{aligned}$$

that leads to (2.72).

□

**Corollary 2.5.** *Under the assumptions of Theorem 2.4, if $\varepsilon$ is fixed such that*

$$A_0 - 6\varepsilon^2 > 0, \tag{2.78}$$

*then there exists a positive constant $C$, $h$ and $t$ independent, such that holds the following*

$$\|E_U(t)\|_h^2 + \int_0^t \|D_{-x}E_U(s)\|_+^2 ds \leq C h_{max}^4,$$

*for $h \in \Lambda$ and $h_{max}$ small enough.*

### 2.4.2 Stability analysis

The stability of the initial value problem (2.67)-(2.69) in $(U_h(t))_{h\in\Lambda}$, where $U_h(t) \in [W_{h,0}]^2$ is a solution of (2.67)-(2.69), is now discussed in the following sense: for all $\sigma_s > 0$ there exists a positive constant $\sigma_i > 0$ such that, for all $V_h(t) \in [W_{h,0}]^2$ that satisfies (2.67) and $V_h(0) \in [W_{h,0}]^2$, if $\|V_h(0) - U_h(0)\|_h \leq \sigma_i$, then we have $\|V_h(t) - U_h(t)\|_h \leq \sigma_s$, for $t \in [0,T]$, $h \in \Lambda$ with $h_{max}$ small enough.

We start by considering $\omega_h(t) = U_h(t) - V_h(t)$. We have

$$\begin{aligned}\left(\frac{d\omega_h}{dt}(t), \omega_h(t)\right)_h + \quad & (A(M_h V_h(t))D_{-x}\omega_h(t), D_{-x}\omega_h(t))_+ \\ & = -((A(M_h U_h(t)) - A(M_h V_h(t)))D_{-x}U_h(t), D_{-x}\omega_h(t))_+ \\ & +(F(V_h(t), \nabla_h V_h(t)) - F(U_h(t), \nabla_h U_h(t)), \omega_h(t))_h.\end{aligned}$$

Proceeding as in the proof of Theorem 2.4, it can be shown that

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\omega_h(t)\|_h^2 + \quad & A_0\|D_{-x}\omega_h(t)\|_+^2 \leq \|A\|_{C_B^1(\mathbb{R})}\|D_{-x}U_h(t)\|_{h,\infty}\|\omega_h(t)\|_h\|D_{-x}\omega_h(t)\|_+ \\ & +C_L\left(1 + \frac{C_L C_R}{2\varepsilon^2}\right)\|\omega_h(t)\|_h^2 + \varepsilon^2\|D_{-x}\omega_h(t)\|_+^2,\end{aligned}$$

that leads to

$$\begin{aligned}\frac{d}{dt}\|\omega_h(t)\|_h^2 + \quad & 2(A_0 - 2\varepsilon^2)\|D_{-x}\omega_h(t)\|_+^2 \\ & \leq \left(\frac{1}{2\varepsilon^2}\|A\|_{C_B^1(\mathbb{R})}^2\|D_{-x}U_h(t)\|_{h,\infty}^2 + C_L\left(2 + \frac{C_L C_R}{\varepsilon^2}\right)\right)\|\omega_h(t)\|_h^2,\end{aligned} \tag{2.79}$$

for $t \in (0,T]$, $h \in \Lambda$ and $h_{max}$ small enough.

From the inequality (2.79) we obtain

$$\|\omega_h(t)\|_h^2 + D\int_0^t e^{\int_s^t S(U_h(\mu))d\mu}\|D_{-x}\omega_h(s)\|_+^2 ds \leq e^{\int_0^t S(U_h(\mu))d\mu}\|\omega_h(0)\|_h^2, \tag{2.80}$$

for $t \in [0,T]$, $h \in \Lambda$ and $h_{max}$ small enough. In (2.80), $D = 2(A_0 - 2\varepsilon^2)$, and

$$S(U_h(t)) = \frac{1}{2\varepsilon^2}\|A\|_{C_B^1(\mathbb{R})}^2\|D_{-x}U_h(t)\|_{h,\infty}^2 + C_L\left(2 + \frac{C_L C_R}{\varepsilon^2}\right). \tag{2.81}$$

The stability of the initial value problem (2.67)-(2.69) in $(U_h(t))_{h\in\Lambda}$ can be easily concluded if there exists a positive constant $C_s$, $h$ and $t$ independent, such that

$$\int_0^t \|D_{-x}U_h(\mu)\|_{h,\infty}^2 d\mu \leq C_s,\ t \in [0,T], \tag{2.82}$$

for $h \in \Lambda$ with $h_{max}$ small enough.

Theorem 2.4 is now used to guarantee the existence of such positive constant $C_s$. As we have

$$\begin{aligned}\int_0^t |D_{-x}u_{\ell,h}(x_i,s)|^2 ds &\leq 2\int_0^t |D_{-x}E_\ell(x_i,s)|^2 ds + 2\int_0^t |D_{-x}u_\ell(x_i,s)|^2 ds \\ &\leq \frac{2}{h_{min}}\int_0^t \|D_{-x}E_\ell(s)\|_+^2 ds + 2\int_0^t \left\|\frac{\partial u_\ell}{\partial x}(s)\right\|_\infty^2 ds,\end{aligned}$$

for $\ell = 1,2$, from Corollary 2.5 it is successively obtained

$$\begin{aligned}\int_0^t \|D_{-x}U_h(s)\|_{h,\infty}^2 ds &\leq \frac{2}{h_{min}} C h_{max}^4 + 2\int_0^t \left\|\frac{\partial U}{\partial x}(s)\right\|_\infty^2 ds \\ &\leq C_s,\end{aligned}$$

because $u_\ell(t) \in H^3(\Omega)$ and $H^3(\Omega)$ is imbedded in $C^1(\overline{\Omega})$, $\ell = 1,2$.

Then, there exists a positive constant $\hat{C}_s$ such that $S(U_h(t)) \leq 2\hat{C}_s, t \in [0,T]$. Consequently,if we fix $\sigma_s > 0$, then for $\sigma_i < e^{-\hat{C}_s T}\sigma_s$, we have

$$\forall V_h(0) \in [W_{h,0}]^2 : \|V_h(0) - U_h(0)\|_h < \sigma_i \Longrightarrow \|V_h(t) - U_h(t)\|_h < \sigma_s, t \in [0,T].$$

The stability result is now summarized in what follows.

**Corollary 2.6.** *Under the assumptions of Theorem 2.4, the solution $U_h(t) \in [W_{h,0}]^2$ of (2.67)-(2.69), is a stable solution in $[0,T]$.*

### 2.4.3 Numerical simulations

Regarding the implementation of the numerical method (2.67)-(2.69), to simplify, it is considered $\Omega = (0,1)^2$, and it is used random nonuniform grids for the spatial discretization. Furthermore, let introduce, in the time domain $[0,T]$, the uniform grid $\{t_n = n\Delta t,\ n = 0,\ldots,\mathscr{M}\}$ with $t_{\mathscr{M}} = T$, and $\Delta t = t_{n+1} - t_n$ denoting the step size.

A first-order implicit-explicit Euler method for the time discretization is applied. The linear part is implicitly discretized while the nonlinear part is explicitly discretized. It is denoted by $u_{1,h}(t_n)$ and $u_{2,h}(t_n)$ the numerical approximations for $u_1(t_n)$ and $u_2(t_n)$, respectively. The fully discrete numerical scheme reads as

$$\frac{U_h^{n+1} - U_h^n}{\Delta t} - A\Delta_h U_h^{n+1} + F(U_h^n, \nabla_h U_h^n) = G_h^{n+1},\ n = 0,1,\ldots,\mathscr{M}-1,$$

with initial conditions

$$U_h^0 = R_h U_0 \text{ in } \Omega_h,$$

and boundary conditions

$$U_h^n = 0 \text{ on } \partial\Omega_h,\ n = 0,...,\mathscr{M}.$$

In order to estimate the convergence rate numerically, it is defined the error

$$\|E_U\|_\square = \sqrt{\|E_1\|_\square^2 + \|E_2\|_\square^2},$$

where

$$\|E_1\|_{\square}^2 = \max_{n=1,\ldots,\mathscr{M}} \left\{ \|E_1^n\|_h^2 + \Delta t \sum_{i=1}^{n} \|D_{-x}E_1^i\|_+^2 \right\},$$

and $\|E_2\|_{\square}$ is defined analogously.

Now is considered the diagonal entries of $A$, $\alpha = \beta = 0.1$, and is defined $F$ as

$$f_1(x_1,x_2,x_3,x_4) = \cos(x_1) + \sin(x_2) + \cos(x_3) + \sin(x_4),\ x_1,x_2,x_3,x_4 \in \mathbb{R}$$

and

$$f_2(x_1,x_2,x_3,x_4) = x_1 + \sin(x_2) + x_3 + \sin(x_4),\ x_1,x_2,x_3,x_4 \in \mathbb{R}.$$

Finally, the time step is defined $\Delta t = 10^{-6}$, which is small enough so the first-order error from the time discretization does not pollute the convergence rate with reference to the space variable.

In the first example it is considered $U = (u_1,u_2)$ defined by

$$u_1(x,t) = -e^{-t}(e^x - 1)(x-1),$$

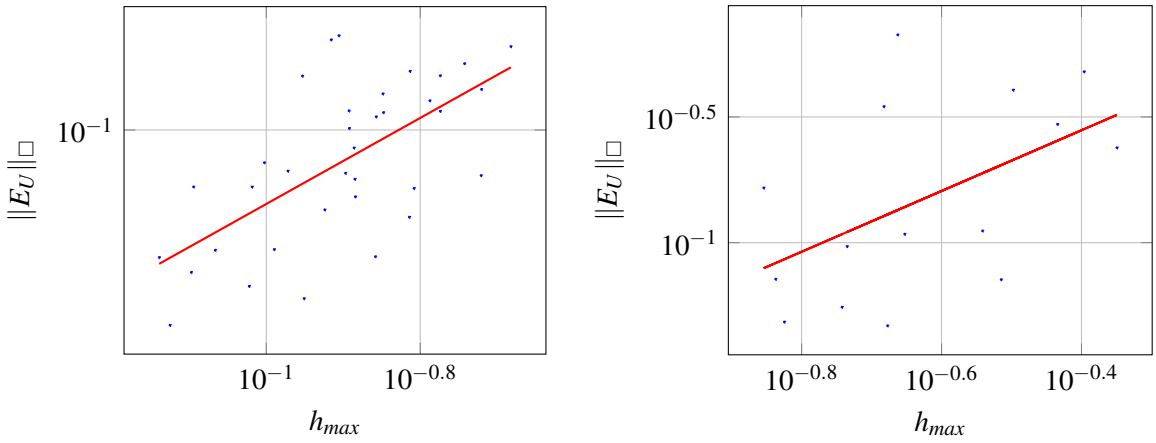
and

$$u_2(x,t) = e^{-t}((2x-1)^4 - 1),$$

for $x \in \overline{\Omega}$, $t \in [0,1]$ and determine $G$ such that $U$ is solution of the IBVP (2.1)-(2.3). In this case $U(t) \in [H^3(\Omega)]^2$ for $t \in [0,1]$, and $F$ satisfies the assumptions of Theorem 2.4.

In the second example, $U$ is such that $u_1$ is the same as before, and $u_2$ is given by

$$u_2(x,t) = e^{-t}|2x-1|^{1.52} - 1.$$




(a) $U(t) \in [H^3(\Omega)]^2$ and the line has a slope of 1.87. (b) $U(t) \in [H^2(\Omega)]^2$ and the line has a slope of 1.21.

Fig. 2.3 Log-log plots of $\|E_U\|_{\square}$ *versus* $h_{max}$ for the system of parabolic equations. The solid lines represent the least-squares fitting.

Figure 2.3 illustrates the error estimate in Theorem 2.4. The results included in the left figure were obtained with $U(t) \in [H^3(\Omega)]^2$. The slope of the linear regression, in this case, is 1.87 but it reduces to 1.2 when we take $u_2 \in H^2(\Omega)$.

## 2.5 Conclusions

In this chapter was considered the nonlinear system of parabolic equations (2.1). This system is a generalization of the system of equations (1.3)-(1.4), which describes the evolution of the magnetic field in the umbra of sunspots. To simplify, were assumed in this generalization homogeneous Dirichlet boundary conditions (2.2), and the initial conditions (2.3). To gain some intuition, the chapter started considering the nonlinear elliptic problem (2.4)-(2.5), and its extension to the nonlinear elliptic system (2.6)-(2.7). For the three problems, piecewise linear FEM, that can be seen as FDM, were proposed. The goal was to establish conditions that allowed to prove the stability of the methods, and also conditions that allowed obtaining second-order approximations in space with respect to a discrete $H^1$-norm.

The main results of this chapter are, for the elliptic equation, Theorem 2.2 and Proposition 2.6, for the elliptic system, Theorem 2.3, and finally for the parabolic system, Theorem 2.4, and Corollary 2.6. In these results, under certain assumptions, for the numerical approximation of $u \in H^3(\Omega) \cap H_0^1(\Omega)$, $U \in [H^3(\Omega) \cap H_0^1(\Omega)]^2$, and $U(t) \in [H^3(\Omega) \cap H_0^1(\Omega)]^2$ respectively, the second-order of convergence and stability were proved .

In addition, for each numerical method, numerical experiments illustrating the convergence results for nonsmooth solutions were included.

# Chapter 3

# An application to sunspots

## 3.1 Introduction

This chapter aims to extend some of the results presented before to an IBVP defined in a two-dimensional domain. Besides, it intends to show our studies' application to a particular situation involving the time and space evolution of the solar magnetic field in one region characterized by a strong and vertical magnetic field like is present in sunspot umbra or Network regions. In fact, results presented in Chapter 2 are extended to a two-dimensional problem. Thus, the needed natural discrete functional scenario is introduced in Section 3.2, and, in Section 3.3, the main convergence result of this section is established - Theorem 3.1. It should be recalled that this result is proved assuming that the solution of the IBVP (3.1), (3.2), (3.3) is smooth. In Section 3.3, *Numerical simulations - application to sunspots*, some numerical experiments in the context of the solar magnetic field in the umbra are presented. In the umbra, it is assumed that the evolution of the magnetic field and velocity are well described by the coupled equations (1.3) and (1.4). As the methods proposed in the previous sections are meant to be simulated, it is assumed that the velocity is known in the computational domain, and the magnetic field is also known in the boundary of such domain, as well as the initial condition. As a matter of fact, such assumptions are not realistic, but they are prompted by the analytical support developed before in the sense that the ultimate aim is to apply the methods studied before. It is believed that the extension of the results presented here, considering the boundary conditions used in the next chapter, may lead to more realistic results. As far as the magnetic field equation is concerned, the $z$ component of the magnetic field is the strongest one in the darkest zone of the umbra, as reported in the literature ([10]). In what follows, the behaviour of the two magnetic field components is simulated: the $z$ and $y$ components. Surely, component $x$ could replace $y$ component. In what concerns the computational domain, a two-dimensional domain is considered in $x$ and $y$ directions. The numerical experiments presented in this chapter may well be acknowledged as preliminary work; as such, it should be point out that further numerical work on the coupling of the magnetic field (1.3) the velocity equations (1.4) should be considered in the future.

## 3.2 The numerical scheme and some results

Motivated by the application that we would like to consider and the corresponding simplification, which is the numerical illustration of the behaviour of the magnetic field in the umbra considering only two components, in this section is considered the system of equations (1.5) that has as a particular case the one defined by the equations (1.3) and (1.4), one of each with two unknowns, and that can be rewritten in the following equivalent form

$$\begin{cases} \dfrac{\partial U_i}{\partial t} + f_i(U, \nabla U) = \nabla . \Big( \alpha_{2i-1,2i}(U_i(t)) \nabla U_i(t) \Big) + g_i \ \text{ in } \Omega \times (0,T], \\ i = 1,2,3,4, \end{cases} \tag{3.1}$$

with $\Omega = (0,1)^2$, $\alpha_{2i-1,2i}(U_i(t))$, $i = 1,2,3,4$, are 2-diagonal matrix with diagonal entries $\alpha_{2i-1}(U_i(t))$ and $\alpha_{2i}(U_i(t))$. In (3.1), $\nabla U = (\nabla U_i, i = 1,2,3,4)$ and $\nabla U_i = \left( \dfrac{\partial U_i}{\partial x}, \dfrac{\partial U_i}{\partial y} \right)$. Observe that, in the context of the differential system (1.5), $(U_1, U_2)$ denotes the magnetic field and $(U_3, U_4)$ represents the velocity field.

It is considered that system (1.5) is complemented with

$$U(t) = 0 \text{ on } \partial\Omega \times (0,T], \tag{3.2}$$

and

$$U(0) = U_0 \text{ in } \Omega. \tag{3.3}$$

It is clear that for (1.3) and (1.4), the previous boundary conditions are not realistic. In fact, as the domain $\Omega$ is bounded, Dirichlet boundary conditions are not appropriate to describe the behaviour of the magnetic and the velocity fields on the boundary. In this case, Neumann boundary conditions or Robin boundary conditions seem to be more realistic. However, to make easier the numerical analysis, homogeneous Dirichlet boundary conditions were adopted. Remark that, as mentioned before, in the application to sunspots, it is assumed that the velocity field is known which reduces the system (3.1) to a system for the magnetic components with known magnetic field on the boundary.

Let $\Lambda$ be a sequence of vectors $H = (h,k)$, where $h = (h_1, \ldots, h_N)$ and $k = (k_1, \ldots, k_M)$ have positive entries such that $\sum_{i=1}^{N} h_i = \sum_{j=1}^{M} k_j = 1$. Let $H_{max} = \max\{h_{max}, k_{max}\}$, where $h_{max} = \max_{i=1,\ldots,N} h_i$, being $k_{max}$ be defined analogously. It is assumed that $H_{max} \to 0$. Let $\{x_i, i = 1, \ldots, N\}$ and $\{y_j, j = 1, \ldots, M\}$ be two partitions of $[0,1]$ induced by $h$ and $k$, respectively, and let $\overline{\Omega}_H$ be defined by $\overline{\Omega}_H = \{x_\ell, \ell = 0, \ldots, N\} \times \{y_j, j = 0, \ldots, M\}$. Let $\Omega_H = \Omega \cap \overline{\Omega}_H$, $\partial\Omega_H = \partial\Omega \cap \overline{\Omega}_H$. By $W_{H,0}$ is represented the space of grid functions defined in $\overline{\Omega}_H$ that are null on $\partial\Omega_H$, and the space of grid functions defined in $\overline{\Omega}_H$ is denoted by $W_H$.

Let $U_H(t) \in W_{H,0}$ be the solution of the following system of differential equations

$$\begin{cases} U'_{i,H}(t) + f_i(U_H(t), \nabla_{H,a} U_H(t)) = \nabla^*_H \Big( \alpha_{2i-1,2i,H}(U_{i,H}(t)) \nabla_{H,-} U_{i,H}(t) \Big) + R_H g_i \ \text{ in } \Omega_H \times (0,T], \\ i = 1,2,3,4. \end{cases} \tag{3.4}$$

In (3.4),

$$\nabla_{H,a}U_H(t) = (\nabla_{H,a}U_{i,H}(t),\ i = 1,2,3,4),$$

$$\nabla_{H,a}U_{i,H}(t) = (\nabla_h U_{i,H}(t), \nabla_k U_{i,H}(t)),\ i = 1,2,3,4,$$

with

$$\nabla_h U_{i,H}(x_\ell, y_j, t) = \frac{1}{h_\ell + h_{\ell+1}}\Big(h_\ell D_{-x}U_{i,H}(x_{\ell+1}, y_j, t) + h_{\ell+1}D_{-x}U_{i,H}(x_\ell, y_j, t)\Big),\ (x_\ell, y_j) \in \Omega_H,$$

being $\nabla_k$ defined analogously. Moreover,

$$\nabla_{H,-}U_H(t) = (\nabla_{H,-}U_i, i = 1,2,3,4),$$

$$\nabla_{H,-}U_{i,H}(t) = (D_{-x}U_{i,H}(t), D_{-y}U_{i,H}(t)),$$

$$\nabla^*_H(v_{1,H}, v_{2,H}) = D^*_x v_{1,H} + D^*_y v_{2,H},\ v_{i,H} \in W_H,\ i = 1,2,$$

where,

$$D_{-x}v_{1,H}(x_\ell, y_j, t) = \frac{v_{1,H}(x_\ell, y_j, t) - v_{1,H}(x_{\ell-1}, y_j, t)}{h_\ell},$$

$$D^*_x v_{1,H}(x_\ell, y_j, t) = \frac{v_{1,H}(x_{\ell+1}, y_j, t) - v_{1,H}(x_\ell, y_j, t)}{h_{\ell+1/2}},$$

and $D_{-y}$ and $D^*_y$ are defined analogously. Furthermore, $M_h$ represents the average operator

$$M_h v_H(x_\ell, y_j) = \frac{1}{2}\big(v_H(x_{\ell-1}, y_j) + v_H(x_\ell, y_j)\big),\ (x_\ell, y_j) \in \Omega_H,$$

being $M_k$ defined analogously. The 2-diagonal matrix $\alpha_{2i-1,2i,H}(U_{i,H}(t))$ has the diagonal entries $\alpha_{2i-1}(M_h U_{i,H}(t))$ and $\alpha_{2i}(M_k U_{i,H}(t))$.

The differential system (3.4) is complemented with the boundary condition

$$U_H = 0 \text{ on } \partial\Omega_H \times (0,T], \tag{3.5}$$

and initial condition

$$U_H(0) = R_H U_0 \text{ in } \Omega_H. \tag{3.6}$$

In (3.6), $R_H : [C(\overline{\Omega})]^4 \to [W_{H,0}]^4$ is the restriction operator $R_H U(x,y) = U(x,y), (x,y) \in \overline{\Omega}_H$. To simplify, the same notation is used to represent the restriction operator $R_H : C(\overline{\Omega}) \to W_{H,0}$.

The convergence properties of the semi-discrete approximation defined by (3.4), (3.5) and (3.6), can now be easily established using convenient norms.

Let introduce the inner product

$$(u_H, v_H)_H = \sum_{\ell=1}^{N-1}\sum_{j=1}^{M-1} h_{\ell+1/2}k_{j+1/2}u_H(x_\ell, y_j)v_H(x_\ell, y_j),\ u_H, v_H \in W_{H,0},$$

and the corresponding norm denoted by $\|\cdot\|_H$. As in the one-dimensional case, it is introduced the discrete $H^1$-norm

$$\|u_H\|_{1,H} = \sqrt{\|u_H\|_H^2 + \|\nabla_{H,-}u_H\|_{H,+}^2},$$

where

$$\|\nabla_{H,-}u_H\|_{H,+}^2 = \|D_{-x}u_H\|_{x,+}^2 + \|D_{-y}u_H\|_{y,+}^2,$$

with

$$\|D_{-x}u_H\|_{x,+}^2 = \sum_{\ell=1}^{N}\sum_{j=1}^{M-1} h_\ell k_{j+1/2}\big(D_{-x}u_H(x_\ell,y_j)\big)^2$$

being $\|D_{-y}u_H\|_{y,+}^2$ defined analogously. The following notations are also used

$$(u_H,v_H)_{x,+} = \sum_{\ell=1}^{N}\sum_{j=1}^{M-1} h_\ell k_{j+1/2}u_H(x_\ell,y_\ell)v_H(x_\ell,y_j),\ u_H,v_H\in W_H,$$

being $(u_H,v_H)_{y,+}$ defined analogously.

In this scenario the Poincaré-Friedrichs inequality also holds, and is given by

$$\|u_H\|_H^2 \le \frac{1}{2}\|\nabla_{H,-}u_H\|_{H,+}^2,\ \forall u_H\in W_{H,0}. \tag{3.7}$$

In fact, as in the one-dimensional version of the Poincaré-Friedrichs inequality presented before, we have

$$\|u_H\|_H^2 \le \|D_{-x}u_H\|_{x,+}^2,\ \|u_H\|_H^2 \le \|D_{-y}u_H\|_{y,+}^2,$$

and consequently it arises (3.7).

Let $A$ be a two-diagonal matrix with diagonal entries $a$ and $b$. Observe that holds the following equality

$$(\nabla_H^*(A(v_H)\nabla_{H,-}u_H,w_H)_H = -(a(M_hv_H)D_{-x}u_H,D_{-x}w_H)_{x,+} - (b(M_kv_H)D_{-y}u_H,D_{-y}w_H)_{y,+}, \tag{3.8}$$

for all $u_H,v_H,w_H\in W_{H,0}$.

### 3.2.1 Convergence analysis

The next result establishes that the error of the semi-discrete approximation $U_H(t)\in[W_{H,0}]^4$, defined by (3.4), (3.5), (3.6), is of second-order with respect to the norm $\|.\|_{1,H}$, provided that the solution $U(t)$ of the IVBP (3.1), (3.2) and (3.3) belongs to $[C^4(\overline{\Omega})]^4$. It is assumed that the sequence of grids $\overline{\Omega}_H, H\in\Lambda$, satisfies the following: there exists a positive constant $C_R$ such that

$$\frac{h_{max}}{h_{min}} \le C_R,\ \frac{k_{max}}{k_{min}} \le C_R. \tag{3.9}$$

For $U_H\in[W_{H,0}]^4$ the following notations are used

$$\|U_H\|_H^2 = \sum_{i=1}^{4}\|U_{H,i}\|_H^2,$$

$$\nabla_{H,-}U_H = (\nabla_{H,-}U_{H,i}, i = 1,\ldots,4),$$

$$\|\nabla_{H,-}U_H\|^2_{H,+} = \sum_{i=1}^{4} \|\nabla_{H,-}U_{H,i}\|^2_{H,+}.$$

Let $E_U(t) = U_H(t) - R_H U(t) \in [W_{H,0}]^4$ be the spatial discretization error where $U$ is solution of the IBVP (3.1), (3.2), (3.3), and $U_H(t) \in [W_{H,0}]^4$ is the the correspondent semi-discrete approximation defined by (3.4), (3.5), (3.6).

The error $E_U(t) = (E_i(t), i = 1,\ldots,4)$ is solution of the following problem

$$\begin{cases} (E_i'(t), E_i(t))_H & = -(\alpha_{2i-1,2i,H}(U_{i,H}(t))\nabla_{H,-}E_i(t), \nabla_{H,-}E_i(t))_{H,+} \\ & +((\alpha_{2i-1,2i,H}(R_H U_i(t)) - \alpha_{2i-1,2i,H}(U_{i,H}(t)))\nabla_{H,-}R_H U_i(t), \nabla_{H,-}E_i(t))_{H,+} \\ & +(f_i(R_H U(t), \nabla_{H,a}R_H U(t)) - f_i(U_H(t), \nabla_{H,a}U_H(t)), E_i(t))_H - (T_{i,H}(t), E_i(t))_H \\ & \text{in } (0,T], \\ & i = 1,2,3,4, \end{cases} \tag{3.10}$$

$$E_U(x,y,t) = 0, (x,y) \in \partial\Omega_H \times (0,T], \tag{3.11}$$

and

$$E_U(x,y,0) = 0, (x,y) \in \Omega_H. \tag{3.12}$$

In (3.10), $T_{i,H}(t)$ denotes the spatial truncation error that admits the following representation

$$T_{i,H}(t) = T_{i,x}(t) + T_{i,y}(t)$$

with

$$T_{i,x}(t) = T_{i,x}^{(1)}(t) + T_{i,x}^{(2)}(t)$$

and

$$T_{i,x}^{(1)}(x_\ell, y_j, t) = (h_{\ell+1} - h_\ell)R_{i,x}(x_\ell, y_j, t),$$

where

$$\begin{aligned} R_{i,x}(x,y,t) &= -\frac{\partial U_i}{\partial x}(x,y,t)\Big(\frac{\partial^2 U_i}{\partial x^2}(x,y,t)\alpha_{2i-1}'(U_i(x,y,t)) + \frac{1}{4}\alpha_{2i-1}''(U_i(x,y,t))\big(\frac{\partial U_i}{\partial x}(x,y,t)\big)^2\Big) \\ &\quad -\frac{1}{3}\alpha_{2i-1}(U_i(x,y,t))\frac{\partial^3 U_i}{\partial x^3}(x,y,t). \end{aligned}$$

Moreover,

$$|T_{i,x}^{(2)}(x_\ell, y_j, t)| \le C h_{max}^2 \Big( \|U_i(t)\|_{C^4(\overline{\Omega})} + \|U(t)\|_{[C^3(\overline{\Omega})]^4} \Big),$$

where $C$ is a $H$, $t$ and $U_i$ independent constant. The truncation error with respect to $y$, $T_{i,y}$, is defined analogously.

**Theorem 3.1.** *Let $U \in \big(C^1([0,T], C(\overline{\Omega}))\big)^4 \cap \big(C([0,T], C^4(\overline{\Omega}))\big)^4$ be solution of the IBVP (3.1), (3.2) and (3.3) and let $U_H \in \big(C^1([0,T], W_{H,0})\big)^4$ be the the correspondent semi-discrete approximation defined by (3.4), (3.5), (3.6). Let $E_U(t) = U_H(t) - R_H U(t) \in [W_{H,0}]^4, t \in [0,T]$, be the spatial discretization error. It is assumed that $f_i, i = 1,2,3,4$, are Lipschitz functions with Lipschitz constant $C_L$.*

*Let A be the diagonal block matrix, where each block i has as entries $\alpha_{2i-1} \geq A_0 > 0$ and $\alpha_{2i} \geq A_0 > 0$ in $\mathbb{R}$, for $i = 1,\ldots,4$, being $A_{jj} \in C_B^3(\mathbb{R}), j = 1,\ldots,8$. If the sequence of grids $\overline{\Omega}_H, H \in \Lambda$, satisfies (3.9), then there exists a positive constant $C_U$, H and t independent, such that*

$$\|E_U(t)\|_H^2 + \beta \int_0^t e^{\int_s^t \theta(U(\mu))d\mu} \|\nabla_{H,-} E_U(s)\|_{H,+}^2 ds \leq C_U H_{max}^4 \int_0^t e^{\int_s^t \theta(U(\mu))d\mu} \|U(s)\|_{[C^4(\overline{\Omega})]^4}^2 ds, \tag{3.13}$$

*for $t \in [0,T], H \in \Lambda$. In (3.13) $\varepsilon, \gamma \neq 0$ are arbitrary constants, and $\beta$ and $\theta$ are given by*

$$\beta = 2(A_0 - 6\varepsilon^2), \tag{3.14}$$

*and*

$$\theta(U(t)) = \frac{1}{2\varepsilon^2}\|U(t)\|_{[C^1(\overline{\Omega})]^4}^2 \|A\|_{C_B^1(\mathbb{R})}^2 + 2C_L^2\left(\frac{1}{\gamma^2} + \frac{C_R}{2\varepsilon^2}\right) + 10\gamma^2. \tag{3.15}$$

*Proof.* Notice that

$$\begin{aligned}\Big(\big(\alpha_{2i-1,2i,H}(R_H U_i(t)) - \quad & \alpha_{2i-1,2i,H}(U_{i,H}(t))\big)\nabla_{H,-} R_H U_i(t), \nabla_{H,-} E_i(t)\Big)_{H,+} \\ & \leq \frac{1}{4\varepsilon^2}\|U_i(t)\|_{C^1(\overline{\Omega})}^2 \|\alpha_{2i-1,2i}\|_{C_B^1(\mathbb{R})}^2 \|E_i(t)\|_H^2 \\ & \quad + \varepsilon^2 \|\nabla_{H,-} E_i(t)\|_{H,+}^2,\end{aligned} \tag{3.16}$$

where $\varepsilon \neq 0$ is an arbitrary constant. For $(f_i(R_H U(t), \nabla_{H,a} R_H U(t)) - f_i(U_H(t), \nabla_{H,a} U_H(t)), E_i(t))_H$ follows successively that

$$\begin{aligned}\big(f_i(R_H U(t), \nabla_{H,a} R_H U(t)) \quad & - f_i(U_H(t), \nabla_{H,a} U_H(t)), E_i(t)\big)_H \\ & \leq C_L\left(\sum_{i=1}^4 \|E_i(t)\|_H + \sqrt{2C_R}\sum_{i=1}^4 \|\nabla_{H,-} E_i(t)\|_{H,+}\right)\|E_i(t)\|_H \\ & \leq \frac{1}{4\gamma^2} C_L^2 \|E_U(t)\|_H^2 + \left(\frac{C_R C_L^2}{2\varepsilon^2} + \gamma^2\right)\|E_i(t)\|_H^2 \\ & \quad + \varepsilon^2 \|\nabla_{H,-} E_U(t)\|_{H,+}^2,\end{aligned} \tag{3.17}$$

where $\gamma \neq 0$ is an arbitrary constant. Taking into account the particular structure of $T_{i,x}(t)$, and following the proof of Proposition 2.4, it can be shown that

$$\begin{aligned}-(T_{i,x}^{(1)}(t), E_{i,H}(t))_H \quad & \leq h_{max}^4\Big(\frac{1}{4\varepsilon^2}\|R_{i,x}(t)\|_{C(\overline{\Omega})}^2 + \frac{1}{4\gamma^2}\|R_{i,x}(t)\|_{C^1(\overline{\Omega})}^2\Big) \\ & \quad + \varepsilon^2 \|D_{-x} E_i(t)\|_+^2 + \gamma^2 \|E_i(t)\|_H^2\end{aligned}$$

and also

$$-(T_{i,x}^{(2)}(t), E_{i,H}(t))_H \leq C\frac{1}{4\gamma^2} h_{max}^4 \|U_i(t)\|_{C^4(\overline{\Omega})}^2 + \gamma^2 \|E_i(t)\|_H^2.$$

As for that $(T_{i,y}(t),E_i(t))_H$ an analogous estimate could be obtained, for $(T_{i,H}(t),E_{i,H}(t))$ it follows that

$$\begin{aligned}-(T_{i,H}(t),E_{i,H}(t))_H &\leq H_{max}^4\Big(\frac{1}{4\varepsilon^2}\big(\|R_{i,x}(t)\|^2_{C(\overline{\Omega})}+\|R_{i,y}(t)\|^2_{C(\overline{\Omega})}\big)\\ &+\frac{1}{4\gamma^2}\big(\|R_{i,x}(t)\|^2_{C^1(\overline{\Omega})}+\|R_{i,y}(t)\|^2_{C^1(\overline{\Omega})}\big)\Big)\\ &+\varepsilon^2\|\nabla_{H,-}E_i(t)\|^2_{H,+}+4\gamma^2\|E_i(t)\|^2_H\\ &+C\tfrac{1}{2\gamma^2}H_{max}^4\|U_i(t)\|^2_{C^4(\overline{\Omega})}.\end{aligned} \tag{3.18}$$

Considering now the upper bounds (3.16), (3.17) and (3.18) in (3.10), we guarantee the existence of a positive constant $C_U$, $H$, $t$ and $U$ independent, such that

$$\tfrac{d}{dt}\|E_U(t)\|^2_H+2(A_0-6\varepsilon^2)\|\nabla_{H,-}E_U(t)\|^2_{H,+}\leq\theta(U(t))\|E_U(t)\|^2_H+C_UH_{max}^4\|U(t)\|^2_{[C^4(\overline{\Omega})]^4}, \tag{3.19}$$

where $\theta(U(t))$ is defined by (3.15). Attending that $E_U(0)=0$, the inequality (3.19) leads to (3.13).

□

**Corollary 3.1.** *Under the assumptions of Theorem 3.1, if $\varepsilon$ is fixed such that*

$$A_0-6\varepsilon^2>0$$

*then there exists a positive constant $C_U$, $H$, $t$ independent, such that*

$$\|E_U(t)\|^2_H+\int_0^t\|\nabla_{H,-}E_U(s)\|^2_{H,+}ds\leq C_UH_{max}^4\int_0^t\|U(s)\|^2_{[C^4(\overline{\Omega})]^4}ds,$$

*for $t\in[0,T],H\in\Lambda$.*

### 3.2.2 Application to sunspots- numerical experiments

This section is dedicated to the simulation of the numerical method proposed to solve the IBVP of parabolic equations (3.1)-(3.3), in the context of strong magnetic field concentrations like the conditions present for example in the umbra- the most central zone of the sunspots.

The governing equations to simulate magnetic fields under the conditions existing in the umbra, (1.3) and (1.4), can be described by the system (1.5) taking in account some assumptions. To simplify, the velocity field is assumed to be known and, consequently, only the magnetic field components needed to be computed. It is clear that this assumption is not realistic but simplifies the problem because the numerical computation of the numerical solution involving the Navier-Stokes equation (1.4) is a challenge requiring a significant time investment. In what concerns the magnetic field components, the vertical one, $B_z$, which, as mentioned in the introduction, is the most significant, and also the $y$ direction component, $B_y$, are considered. Another choice is, of course, possible.

The plasma velocity, used as input in the numerical experiments, was taken from the numerical simulation of a network region [14] obtained with the Bifrost code (a general and flexible 3D Radiation MHD code developed in Oslo, described in detail in [34]) (http://sdc.uio.no/search/simulations). The choice fell on the simulation of a network region and not a sunspot since there is a lack of availability of open-to-use sunspot simulations, and also because of computational cost and even data handling.

To define the magnetic components on the boundary $\partial\Omega$ the results from such simulation are also used.

While the original Bifrost simulation considered a fully three-dimensional problem, accounting for the velocity, $u$, and magnetic field, $B$, this simulation assumes that the velocity is known at certain time steps (every 10 seconds), and model only the $y$ and $z$ components of the magnetic field. To describe the time and space evolution of the magnetic field $B = (B_y, B_z) = (B_1, B_2)$ that depends on the velocity field $u = (u_y, u_z) = (u_1, u_2)$ the following system of partial differential equations is considered

$$\frac{\partial B}{\partial t} + (u \cdot \nabla)B - (B \cdot \nabla)u = \eta \Delta B, \text{ in } \Omega \times (0,T], \tag{3.20}$$

with

$$B(x,y,t) = B_{\partial\Omega}(x,y,t),\ (x,y) \in \partial\Omega, t \in (0,T], \tag{3.21}$$

and

$$B(x,y,0) = B_0(x,y),\ (x,y) \in \Omega. \tag{3.22}$$

where $B_{\partial\Omega}$ and $B_0$ are known functions.

The semi-discrete approximation $B_H$ for $B$ is then defined by the differential problem

$$\begin{cases} B'_{1,H}(x,y,t) & +u_{1,H}(x,y,t)\nabla_h B_{1,H}(x,y,t) + u_{2,H}(x,y,t)\nabla_k B_{1,H}(x,y,t) \\ & -B_{1,H}(x,y,t)\nabla_h u_{1,H}(x,y,t) - B_{2,H}(x,y,t)\nabla_k u_{1,H}(x,y,t) \\ & = \eta \Delta_H B_{1,H}(x,y,t) \\ B'_{2,H}(x,y,t) & +u_{1,H}(x,y,t)\nabla_h B_{2,H}(x,y,t) + u_{2,H}(x,y,t)\nabla_k B_{2,H}(x,y,t) \\ & -B_{1,H}(x,y,t)\nabla_h u_{2,H}(x,y,t) - B_{2,H}(x,y,t)\nabla_k u_{2,H}(x,y,t) \\ & = \eta \Delta_H B_{2,H}(x,y,t) \\ & \\ & (x,y) \in \Omega_H, t \in (0,T], \end{cases} \tag{3.23}$$

where $\Delta_H = D_{2,x} + D_{2,y}$, being $D_{2,x}$ and $D_{2,y}$ the standard second-order centered finite difference operator in $x$ and $y$ directions, respectively.

The system (3.23) is completed with

$$B_H(x,y,t) = B_{\partial\Omega}(x,y,t),\ (x,y) \in \partial\Omega_H, t \in (0,T], \tag{3.24}$$

and

$$B_H(x,y,0) = B_0(x,y),\ (x,y) \in \Omega_H. \tag{3.25}$$

The time integration of the differential problem (3.23), (3.24), (3.25) is a challenging problem in what concerns the stability in time due to the presence of the convective terms and reaction terms. In what follows is considered that the convective terms are explicitly discretized, while the diffusion and reaction terms are implicitly discretized.

To define a fully discrete scheme in space and time, it is introduced in $[0,T]$ the grid $\{t_n, n = 0, \ldots, \mathscr{M}\}$ where $t_0 = 0$, $t_{\mathscr{M}} = T$ and $t_{n+1} - t_n = \Delta t$. The differential problem (3.23), (3.24), (3.25) is

then replaced by the following implicit-explicit fully discrete method

$$\begin{cases} \dfrac{B_{1,H}^{n+1}(x,y)-B_{1,H}^{n}(x,y)}{\Delta t} & +u_{1,H}(x,y,t_{n+1})\nabla_h B_{1,H}^{n}(x,y)+u_{2,H}(x,y,t_{n+1})\nabla_k B_{1,H}^{n}(x,y) \\ & -B_{1,H}^{n+1}(x,y)\nabla_h u_{1,H}(x,y,t_{n+1})-B_{2,H}^{n+1}(x,y)\nabla_k u_{1,H}(x,y,t_{n+1}) \\ & =\eta\Delta_H B_{1,H}^{n+1}(x,y) \\ \dfrac{B_{2,H}^{n+1}(x,y)-B_{2,H}^{n}(x,y)}{\Delta t} & +u_{1,H}(x,y,t_{n+1})\nabla_h B_{2,H}^{n}(x,y)+u_{2,H}(x,y,t_{n+1})\nabla_k B_{2,H}^{n}(x,y) \\ & -B_{1,H}^{n+1}(x,y)\nabla_h u_{2,H}(x,y,t_{n+1})-B_{2,H}^{n+1}(x,y)\nabla_h u_{2,H}(x,y,t_{n+1}) \\ & =\eta\Delta_H B_{2,H}^{n+1}(x,y) \\ & (x,y)\in\Omega_H,\ n=0,\ldots,\mathscr{M}-1, \end{cases} \tag{3.26}$$

with

$$B_H^j(x,y)=B_{\partial\Omega}(x,y,t_j),\ (x,y)\in\partial\Omega_H,\ j=1,\ldots,\mathscr{M}, \tag{3.27}$$

and

$$B_H^0(x,y)=B_0(x,y),\ (x,y)\in\Omega_H. \tag{3.28}$$

Let $\Omega=(3360)\times(3800)$ $(km^2)$, $T=1560$ $(s)$, $\eta=0.2$ $(km^2/s)$ ([30]). The units of the magnetic field and of the velocity are $\mu T$ (microtesla) and $km/s$, respectively. The domain $\Omega$ is discretized with a uniform grid of rectangles with spacings $\Delta x=48$, and $\Delta y=19$. It was opted to use the same grid as the original Bifrost simulation to avoid spatial interpolation procedures. Regarding the temporal discretization, we take $\Delta t=0.1$. Since $B$ and $u$ are known for $t_i=10i, i=0,...,156$, it is considered linear interpolants of these quantities across time to allow a simulation with the chosen value of $\Delta t$.

Figure 3.1 illustrates the behaviour of the magnetic field for $t=100$ computed with (3.26), (3.27), (3.28). Figure 3.2 includes the plot of the magnetic field obtained with the Bifrost code. Note that $\|\cdot\|$ denotes the euclidean norm.

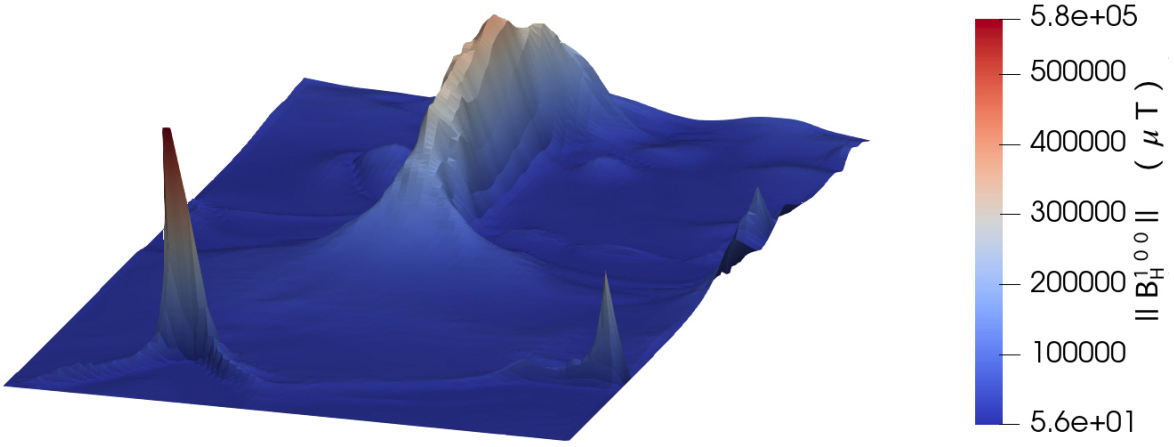



Fig. 3.1 $\|B_H^{100}\|$ obtained with (3.26), (3.27), (3.28).

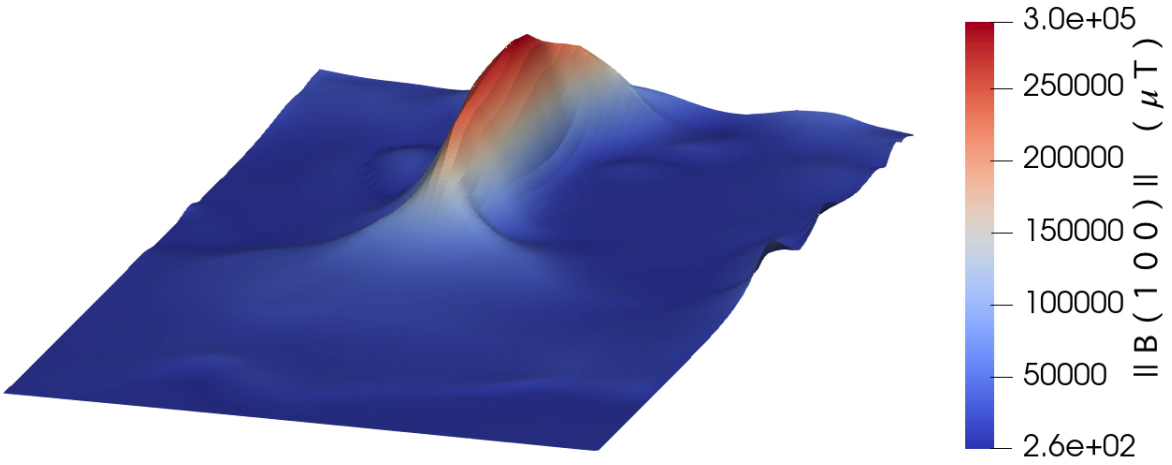



Fig. 3.2 $\|B(100)\|$ obtained with the Bifrost code.

Taking into account the results plotted in Figure 3.2 as reference solution, it is concluded that the convective dominance effect destroys the numerical solution plotted in Figure 3.1. To justify the last sentence the local Péclet number associated with the method given by

$$Pe = \frac{\max\{\Delta x, \Delta y\}\|u\|_{L^\infty(\Omega)}}{2\eta}, \tag{3.29}$$

is computed. As the magnitude of the velocity field is of order 20, then a quick calculation shows that $Pe$ is of order $10^3$ that confirms our remark.

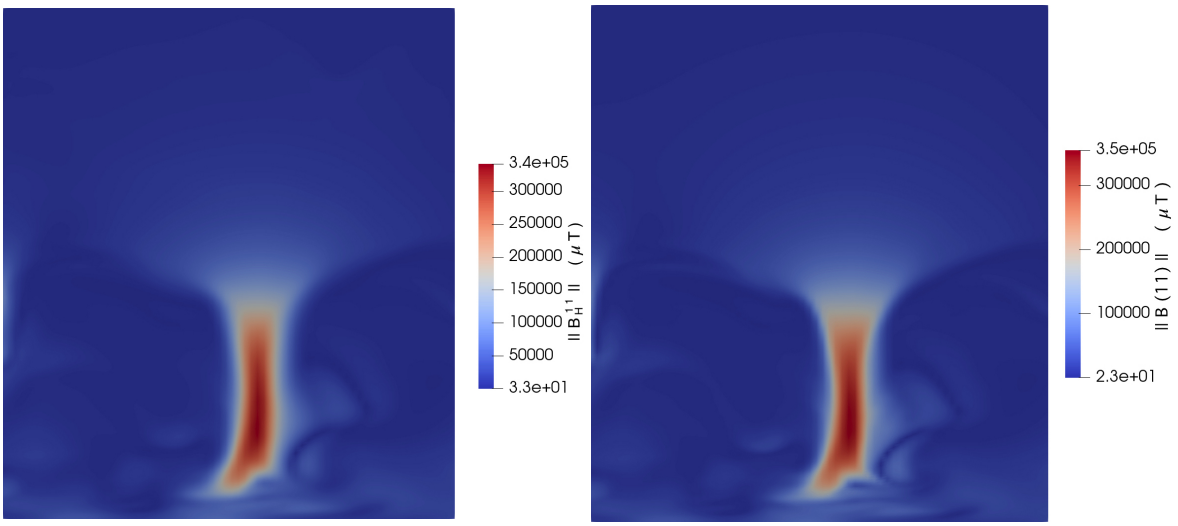
To circumvent this issue, a strategy of numerical diffusion by adding to the right hand side of (3.20) the term $\max\{\Delta x, \Delta y\}\|u\|_{L^\infty(\Omega)}\Delta B$ was adopted. This is a classical technique for stabilizing this type of problems. Although it allows stabilizing the convection-dominated regime, the technique overly diffuses the numerical solution, up to the point that it does not explode or oscillate, but bears little resemblance with the reference solution of the Bifrost simulation as time increase, which can be seen in Figures 3.3, 3.4 and 3.5.

From the numerical results presented in Figures 3.3, 3.4 and 3.5, it is clear that the study of methods well adapted to convection-dominated problems for solar magnetic field, even in the umbra scenario, will be considered in the near future (see for instance [13, 40, 56, 64] and the reference [78] in solar applications).

## 3.3 Conclusions

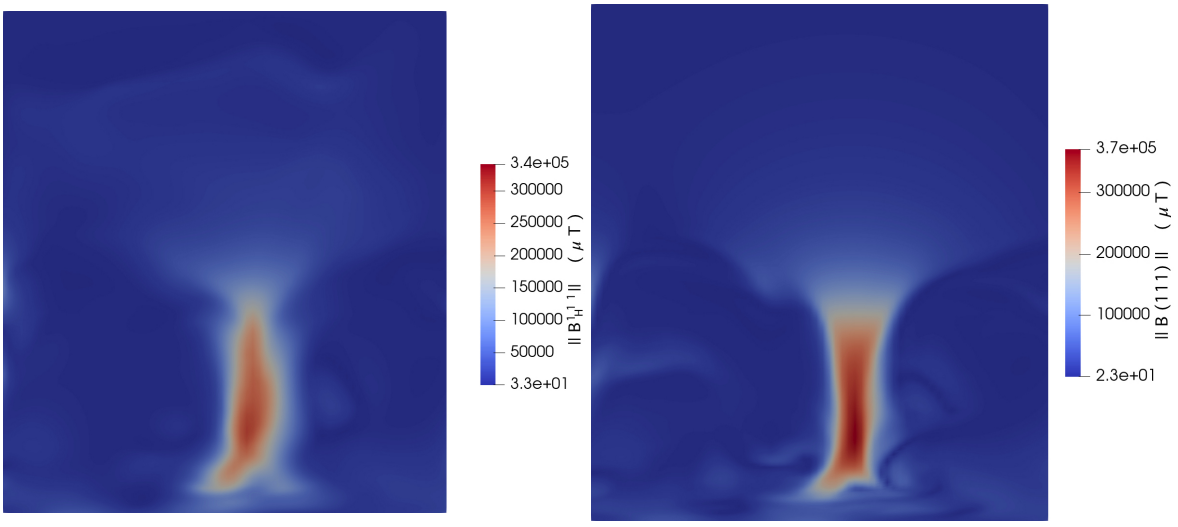
This chapter aims to extend some of the results presented in Chapter 2 to a system of parabolic equations defined in a two-dimensional domain. It also intends to apply the proposed methods to illustrate the behaviour of the solar magnetic field in the umbra.

In what concerns the first objective, in Theorem 3.1 is proved that the natural extension to the two-dimensional case of the semi-discrete approximation, studied before for system of parabolic



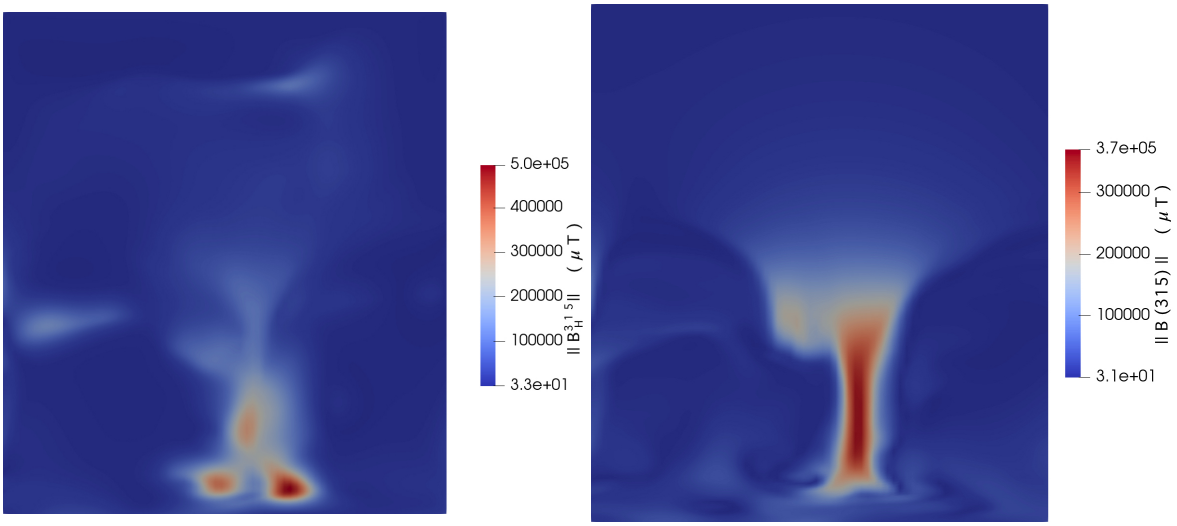

(a) $\|B_H^{11}\|$ obtained with (3.26), (3.27), (3.28). (b) $\|B(11)\|$ obtained with the Bifrost code.

Fig. 3.3 Magnetic field magnitude at $t = 11$.




(a) $\|B_H^{111}\|$ obtained with (3.26), (3.27), (3.28). (b) $\|B(111)\|$ obtained with the Bifrost code.

Fig. 3.4 Magnetic field magnitude at $t = 111$.




(a) $\|B_H^{315}\|$ obtained with (3.26), (3.27), (3.28).

(b) $\|B(315)\|$ obtained with the Bifrost code.

Fig. 3.5 Magnetic field magnitude at $t = 315$.

equations defined in an interval, also leads to a second-order approximation. This result is established assuming that the solution of the differential problem is smooth in the sense that, for each time, it belongs to $[C^4(\overline{\Omega})]^4$. Of course, that is expected to prove the same result when the solution is in $[H^3(\Omega) \cap H_0^1(\Omega)]^4$ using the Bramble-Hilbert approach.

The second objective was partially fulfilled in the sense that from the results presented in Figures 3.3, 3.4, and 3.5, it is observed that the numerical method with a very simple stabilization term is not able to illustrate the magnetic field for large times. In fact, as time increases, the numerical solution deteriorates. This fact can be avoided by investing in stabilization techniques to be coupled with the proposed method.

In the context of the application to the solar magnetic field, it is realized that additional investment is needed in the extension of the results presented here to problems with boundary conditions of Neumann or Robin type. The next chapter starts by giving one first answer to this question considering a parabolic equation with Neumann boundary conditions.

# Chapter 4

# Convergence Analysis for IBVP with Neumann boundary conditions

## 4.1 Introduction

In the last chapters are studied discretizations of nonlinear elliptic and nonlinear parabolic systems considering Dirichlet boundary conditions, and it was proved that the corresponding numerical solutions are second-order convergent with respect to a discrete $H^1$-norm. The study of numerical methods for Neumann BVPs or Neumann IBVPs is not being developed even for problems with smooth solutions. Without being exhaustive is mentioned [28] where, for a second-order linear differential equation, involving the Laplace operator, defined in a square of $\mathbb{R}^2$, a FDM based on nonuniform grids was proposed, and its convergence analyzed following the approach taken in the last chapter for smooth solutions.

In what follows is considered the one-dimensional IBVP

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(A(u)\frac{\partial u}{\partial x}\right) - qu + g \ \text{ in } \Omega \times (0,T], \tag{4.1}$$

where $\Omega = (0,1)$, $q(x) \geq q_0 > 0$, $x \in \overline{\Omega}$, complemented with the homogeneous Neumann boundary condition

$$A(u(x,t))\frac{\partial u}{\partial x}(x,t) = 0, \ (x,t) \in \partial\Omega \times (0,T], \tag{4.2}$$

and with the initial condition

$$u(x,0) = u_0(x), \ x \in \Omega. \tag{4.3}$$

The simplified model (4.1)-(4.3) studied in what follows gives some insights in the generalization of the results presented in the two previous chapters to different boundary conditions. The analysis presented is well adapted for problems with smooth solutions. An extension of the Bramble-Hilbert analysis used before will be an object of analysis in the near future.

This chapter starts by considering in Section 4.2 the elliptic problem

$$-\big(A(u)u'\big)' + qu = g \ \text{ in } \Omega, \tag{4.4}$$

with the boundary condition

$$A(u)u' = 0 \ \text{ on } \partial\Omega, \tag{4.5}$$

and for this BVP, a FDM, which can be seen as fully discrete piecewise linear FEM, is proposed, and it leads to a second-order approximation with respect to a discrete $H^1$-norm. The main difficulty associated with this problem is the definition of the discretizations of (4.4) and (4.5) that are compatible in the sense that a discrete version of

$$(A(u)u', v') + (qu, v) = (g, v), \ \forall v \in H^1(\Omega)$$

holds. Then, in Section 4.3 is considered the IBVP (4.1), (4.2) and (4.3). For this problem is presented the stability and convergence analysis of a finite difference semi-discretization method, which can be seen as a semi-discrete method based on the piecewise linear finite element approximation. Moreover, a fully discrete scheme in space and time of the implicit-explicit type is studied, and its convergence is concluded. Conclusions are presented in Section 4.4.

## 4.2 An elliptic nonlinear BVP

### 4.2.1 Some notations and results

Let $\Lambda$ be a sequence of vectors $h = (h_1, \ldots, h_N)$ with positive entries such that $\sum_{i=1}^{N} h_i = 1$. It is also introduced the following grids

$$\overline{\Omega}_h = \{x_i : \quad i = 0, \ldots, N, \quad x_0 = 0, \quad x_N = 1, \quad x_i = x_{i-1} + h_i, \quad i = 1, \ldots, N\},$$

$$\overline{\Omega}_h^* = \{x_i : \quad i = -1, \ldots, N+1, \quad x_{-1} = -h_1, \quad x_{N+1} = 1 + h_N, \quad x_i \in \overline{\Omega}_h, \quad i = 0, \ldots, N\},$$

and $\partial\Omega_h = \partial\Omega \cap \overline{\Omega}_h$.

By $W_h$ is represented the space of grid functions defined in $\overline{\Omega}_h$ and the space of grid functions defined in $\overline{\Omega}_h^*$ is denoted by $W_h^*$. In $W_h$ is considered the following inner product

$$(v_h, w_h)_h = \frac{1}{2} h_1 v_h(x_0) w_h(x_0) + \sum_{i=1}^{N-1} h_{i+1/2} v_h(x_i) w_h(x_i) + \frac{1}{2} h_N v_h(x_N) w_h(x_N), v_h, w_h \in W_h,$$

and the correspondent norm is denoted by $\|\cdot\|_h$. As before, the following notation is used

$$\|D_{-x} v_h\|_+ = \left( \sum_{i=1}^{N} h_i v_h(x_i)^2 \right)^{1/2}, \ v_h \in W_h.$$

By $\|\cdot\|_{1,h}$ is denoted the norm in $W_h$

$$\|v_h\|_{1,h} = \left( \|v_h\|_h^2 + \|D_{-x} v_h\|_+^2 \right)^{1/2}, \ v_h \in W_h,$$

that can be seen as a discrete version of the usual $H^1$-norm.

The following result is an important tool in the convergence analysis.

**Proposition 4.1.** *For $v_h \in W_h$ we have*

$$v_h(x)^2 \leq 2\|v_h\|_{1,h}^2, \ x \in \partial\Omega_h. \tag{4.6}$$

*Proof.* Given

$$v_h(x_0) = -\sum_{j=1}^{i} h_j D_{-x} v_h(x_j) + v_h(x_i),$$

then

$$v_h(x_0)^2 \leq 2\left(\|D_{-x}v_h\|_+^2 + v_h(x_i)^2\right).$$

Consequently,

$$\sum_{i=1}^{N-1} h_{i+1/2} v_h(x_0)^2 \leq 2\left(\|D_{-x}v_h\|_+^2 + \sum_{i=1}^{N-1} h_{i+1/2} v_h(x_i)^2\right),$$

that leads to (4.6) for $x = x_0$. The proof of (4.6) for $x = x_N$ follows the same steps.

□

It is introduced now the operator that will be used to discretize the boundary conditions. By $D_\eta$ it is denoted the following operator

$$\begin{aligned} D_\eta v_h(x_0) &= -\frac{1}{2}\Big(A(M_h v_h(x_1))D_{-x}v_h(x_1) + A(M_h v_h(x_0))D_{-x}v_h(x_0)\Big), \\ D_\eta v_h(x_N) &= \frac{1}{2}\Big(A(M_h v_h(x_{N+1}))D_{-x}v_h(x_{N+1}) + A(M_h v_h(x_N))D_{-x}v_h(x_N)\Big), \end{aligned} \tag{4.7}$$

defined for $v_h \in W_h^*$.

The following proposition relates $-D_x^*(A(M_h u_h)D_{-x}u_h)$ with the boundary operator $D_\eta$. Its proof can be done by direct verification using summation by parts.

**Proposition 4.2.** *For $u_h \in W_h^*$ and $v_h \in W_h$ we have*

$$(-D_x^*(A(M_h u_h)D_{-x}u_h), v_h)_h = (A(M_h u_h)D_{-x}u_h, D_{-x}v_h)_+ - D_\eta u_h(x_0)v_h(x_0) - D_\eta u_h(x_N)v_h(x_N). \tag{4.8}$$

The weak solution $u \in H^1(\Omega)$ for the BVP (4.4) and (4.5) satisfies

$$(A(u)u', v') + (qu, v) = (g, v), \ \forall v \in H^1(\Omega). \tag{4.9}$$

The piecewise linear approximation $P_h u_h$ where, as before, $P_h$ denotes de piecewise linear interpolation operator and $u_h \in W_h$, is then obtained, solving the following variational problem

$$(A(P_h u_h)P_h u_h', P_h v_h') + (qP_h u_h, P_h v_h) = (g, P_h v_h), \ \forall v_h \in W_h. \tag{4.10}$$

To compute $u_h \in W_h$, it is considered the discrete problem

$$(A(M_h u_h)D_{-x}u_h, D_{-x}v_h)_+ + (R_h q u_h, v_h)_h = (R_h g, v_h)_h, \ \forall v_h \in W_h. \tag{4.11}$$

Assuming that $u_h \in W_h^*$, by Proposition 4.2, the last problem can be rewritten as

$$\begin{aligned}(-D_x^*\big(A(M_h u_h)D_{-x}u_h\big), v_h)_h + (R_h q u_h, v_h)_h &= (R_h g, v_h)_h \\ &\quad - D_\eta u_h(x_0)v_h(x_0) - D_\eta u_h(x_N)v_h(x_N),\end{aligned} \tag{4.12}$$

for all $v_h \in W_h$. If it is assumed that

$$-D_x^*\big(A(M_h u_h)D_{-x}u_h\big) + R_h q u_h = R_h g \ \text{ in } \overline{\Omega}_h \tag{4.13}$$

then

$$D_\eta u_h(x_0) = D_\eta u_h(x_N) = 0. \tag{4.14}$$

Moreover, if $u_h \in W_h^*$ is solution of (4.13), (4.14), then $u_h$ satisfies (4.11).

### 4.2.2 Convergence analysis for smooth solutions

The convergence analysis is presented in what follows assuming that the solution of the BVP (4.4) and (4.5) belongs to $C^4(\overline{\Omega}^*)$ where $\overline{\Omega}^* = \bigcup_{h\in\Lambda} [-x_1, x_{N+1}]$.

Let $T_h$ be the truncation error induced by the discretization (4.13) in $\overline{\Omega}_h$. Observe that

$$|T_h(x_0)| \le Ch_1^2 \|u\|_{C^4(\overline{\Omega}^*)}, \quad |T_h(x_N)| \le Ch_N^2 \|u\|_{C^4(\overline{\Omega}^*)},$$

and for $i = 1, \ldots, N-1$, $T_h(x_i) = T_h^{(1)}(x_i) + T_h^{(2)}(x_i)$, where $T_h^{(\ell)}$, $\ell = 1,2$ are induced by the discretization $-D_x^*\big(A(M_h R_h u)D_{-x}R_h u\big)$. Hence it follows that

$$T_h^{(1)}(x_i) = (h_{i+1} - h_i)R(x_i),$$

where, as before, $R(x_i) = -\frac{1}{3}A(u(x_i))u^{(3)}(x_i) - u'(x_i)\big(u^{(2)}(x_i)A'(u(x_i)) + \frac{1}{4}A^{(2)}(u(x_i))u'(x_i)^2\big)$, and

$$|T_h^{(2)}(x_i)| \le Ch_{max}^2 \|u\|_{C^4(\overline{\Omega})}.$$

If $T_{h,\partial\Omega}$ denotes the truncation error induced by the discretization of the boundary conditions defined by the boundary operator $D_\eta$, then

$$|T_{h,\partial\Omega}(x_0)| \le Ch_1^2 \|u\|_{C^3(\overline{\Omega}^*)}$$

and

$$|T_{h,\partial\Omega}(x_N)| \le Ch_N^2 \|u\|_{C^3(\overline{\Omega}^*)}.$$

Remark that

$$D_\eta u(x_0) = -A(u(x_0))u'(x_0) + T_{h,\partial\Omega}(x_0),$$

and as we are dealing with homogeneous Neumann conditions we get

$$D_\eta u(x_0) = T_{h,\partial\Omega}(x_0). \tag{4.15}$$

Analogously, we also have

$$D_\eta u(x_N) = T_{h,\partial\Omega}(x_N). \tag{4.16}$$

**Theorem 4.1.** *Let $u \in C^4(\overline{\Omega}^*)$ be the solution of the BVP (4.4) and (4.5) with $A \geq A_0 > 0$ in $\mathbb{R}$, and $q \geq q_0 > 0$ in $\overline{\Omega}$. Let $u_h \in W_h^*$ be the solution of the BVP (4.13), (4.14) and let $E_u = u_h - \tilde{R}_h u \in W_h^*$ be the error, where $\tilde{R}_h : C(\overline{\Omega}^*) \to W_h^*$ denotes the restriction operator. Then there exists a positive constant $C_u$, h independent, such that*

$$\|E_u\|_{1,h} \leq C_u h_{max}^2 \|u\|_{C^4(\overline{\Omega}^*)}, \tag{4.17}$$

*for $h \in \Lambda$, provided that there exists a $\varepsilon \neq 0$ such that*

$$A_0 - 6\varepsilon^2 > 0 \tag{4.18}$$

*and*

$$q_0 - 6\varepsilon^2 - \frac{1}{4\varepsilon^2}\|u'\|_\infty^2 \|A\|_{C_B^1(\mathbb{R})}^2 > 0. \tag{4.19}$$

*Proof.* For the numerical solution $u_h \in W_h^*$ we have

$$\begin{aligned}(A(M_h u_h)D_{-x}u_h, D_{-x}E_u)_+ + (R_h q u_h, E_u)_h &= (R_h g, E_u)_h \\ &\quad + \sum_{i=0,N} D_\eta u_h(x_i)E_u(x_i).\end{aligned} \tag{4.20}$$

For $u \in C^4(\overline{\Omega}^*)$ that satisfies (4.4) and (4.5) we also have

$$\begin{aligned}(A(M_h\tilde{R}_h u)D_{-x}\tilde{R}_h u, D_{-x}E_u)_+ + (R_h q R_h u, E_u)_h &= (R_h g, E_u)_h + (T_h, E_u)_h \\ &\quad + D_\eta u(x_0)E_u(x_0) + D_\eta u(x_N)E_u(x_N),\end{aligned}$$

and then, taking into account (4.15) and (4.16), we deduce

$$\begin{aligned}(A(M_h\tilde{R}_h u)D_{-x}\tilde{R}_h u, D_{-x}E_u)_+ + (R_h q R_h u, E_u)_h &= (R_h g, E_u)_h + (T_h, E_u)_h \\ &\quad + \sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i).\end{aligned} \tag{4.21}$$

From (4.20) and (4.21), it is concluded the following equation for $E_u$

$$\begin{aligned}(A(M_h u_h)D_{-x}u_h, D_{-x}E_u)_+ \quad &-(A(M_h\tilde{R}_h u)D_{-x}\tilde{R}_h u, D_{-x}E_u)_+ + (R_h q E_u, E_u)_h \\ &= -(T_h, E_u)_h - \sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i).\end{aligned} \tag{4.22}$$

As in Chapter 2, it can be shown that holds the following

$$\begin{aligned}(A(M_h u_h D_{-x}E_u, D_{-x}E_u)_+ + (R_h q E_u, E_u)_h &= ((A(M_h\tilde{R}_h u) - A(M_h u_h))D_{-x}\tilde{R}_h u, D_{-x}E_u)_+ \\ &\quad -(T_h, E_u)_h - \sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i) \\ &\leq \|D_{-x}\tilde{R}_h u\|_{h,\infty}\|A\|_{C_B^1(\mathbb{R})}\|E_u\|_h\|D_{-x}E_u\|_+ \\ &\quad -(T_h^{(1)}, E_u)_h - (T_h^{(2)}, E_u)_h \\ &\quad - \sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i).\end{aligned} \tag{4.23}$$

Remark that we are considering Neumann boundary conditions at $x = x_0, x_N$, and then $E_u(x_i), i = 0, N$, are not necessarily zero. This fact do not allow the use of Proposition 2.4 to get an estimate for $(T_h^{(1)}, E_u)_h$. Another question is related with the estimation of $\sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i)$ where $T_{h,\partial\Omega}(x_i)$, $i = 0, N$ are second-order terms.

Following the proof of Proposition 2.4 of Chapter 2, it can be shown that

$$\begin{aligned} -(T_h^{(1)}, E_u)_h &\leq h_{max}^2 \left( \tfrac{1}{2}\|R\|_{h,\infty}\|D_{-x}E_u\|_+ + \tfrac{\sqrt{2}}{2}\|R\|_{H^1}\|E_u\|_h \right) \\ &+ \sum_{i=0,N} \frac{h_i}{2}\left( |T_h^{(1)}(x_i)| + h_i|R(x_i)| \right)|E_u(x_i)| \\ &\leq h_{max}^4 \frac{1}{16\varepsilon^2}\left( \|R\|_{h,\infty}^2 + 2\|R\|_{H^1}^2 \right) \\ &+ \varepsilon^2\|D_{-x}E_u\|_+^2 + \varepsilon^2\|E_u\|_h^2 \\ &+ \sum_{i=0,N} \frac{1}{8\varepsilon^2}\left( h_i^2 T_h^{(1)}(x_i)^2 + h_i^4 R(x_i)^2 \right) + \sum_{i=0,N} \varepsilon^2 E_u(x_i)^2, \end{aligned} \tag{4.24}$$

where, to simplify the presentation, is considered in this chapter $h_0 = h_1$. Considering now Proposition 4.1, it is obtained the following estimate

$$\begin{aligned} -(T_h^{(1)}, E_u)_h &\leq h_{max}^4 \tfrac{1}{16\varepsilon^2}\left( \|R\|_{h,\infty}^2 + 2\|R\|_{H^1}^2 \right) \\ &+ \varepsilon^2\|D_{-x}E_u\|_+^2 + \varepsilon^2\|E_u\|_h^2 \\ &+ \sum_{i=0,N} \frac{1}{8\varepsilon^2}\left( h_i^2 T_h^{(1)}(x_i)^2 + h_i^4 R(x_i)^2 \right) + 2\varepsilon^2\|E_u\|_{1,h}^2. \end{aligned} \tag{4.25}$$

Using again Proposition 4.1 in $\sum_{i=0,N} T_{h,\partial\Omega}(x_i)E_u(x_i)$ we get

$$\sum_{i=0,N} -T_{h,\partial\Omega}(x_i)E_u(x_i) \leq \sum_{i=0,N} \frac{1}{4\varepsilon^2} T_{h,\partial\Omega}(x_i)^2 + 2\varepsilon^2\|E_u\|_{1,h}^2. \tag{4.26}$$

Furthermore, the following estimates are obtained

$$(A(M_h u_h)D_{-x}E_u, D_{-x}E_u)_+ \geq A_0\|D_{-x}E_u\|_+^2,$$

$$(R_h q E_u, E_u)_h \geq q_0\|E_u\|_h^2,$$

$$\begin{aligned} &\|D_{-x}\tilde{R}_h u\|_{h,\infty}\|A\|_{C_B^1(\mathbb{R})}\|E_u\|_h\|D_{-x}E_u\|_+ \\ &\quad \leq \frac{1}{4\varepsilon^2}\|u'\|_\infty^2\|A\|_{C_B^1(\mathbb{R})}^2\|E_u\|_h^2 + \varepsilon^2\|D_{-x}E_u\|_+^2, \end{aligned}$$

and

$$-(T_h^{(2)}, E_u)_h \leq \frac{1}{4\varepsilon^2}\|T_h^{(2)}\|_h^2 + \varepsilon^2\|E_u\|_h^2.$$

Inserting the last estimates, and the estimates (4.25) and (4.26) in (4.23), it is concluded that

$$\begin{aligned}\Big(A_0-6\varepsilon^2\Big)\|D_{-x}E_u\|_+^2 \quad &+\Big(q_0-6\varepsilon^2-\tfrac{1}{4\varepsilon^2}\|u'\|_\infty\|A\|_{C_B^1(\mathbb{R})}\Big)\|E_u\|_h^2\\ &\leq \frac{h_{max}^4}{16\varepsilon^2}\left(\|R\|_{h,\infty}^2+2\|R\|_{H^1}^2\right)\\ &+\sum_{i=0,N}\frac{1}{8\varepsilon^2}\left(h_i^2T_h^{(1)}(x_i)^2+h_i^4R(x_i)^2\right)\\ &+\sum_{i=0,N}\frac{1}{4\varepsilon^2}T_{h,\partial\Omega}(x_i)^2+\frac{1}{4\varepsilon^2}\|T_h^{(2)}\|_h^2.\end{aligned}\tag{4.27}$$

Fixing $\varepsilon\neq 0$ satisfying (4.18) and (4.19), and attending that $T_h^{(1)}(x_i)$, $T_{h,\partial\Omega}(x_i)$, $i=0,N$, are second order terms, conclude (4.17). □

## 4.3 A parabolic IBVP

### 4.3.1 A semi-discrete approximation

This section aims to extend the results of the previous section to the IBVP (4.1), (4.2), (4.3). Let $u_h(t)\in W_h^*$, $t\in[0,T]$, be the semi-discrete approximation for the previous IBVP defined by the spatial discretization studied in the last section that is, let $u_h(t)\in W_h^*$ defined by the following initial value problem (IVP)

$$\begin{cases}u_h'(t)=D_x^*(A(M_hu_h(t))D_{-x}u_h(t))-R_hq(t)u_h(t)+R_hg \text{ in } \overline{\Omega}_h\times(0,T],\\ D_\eta u_h(t)=0 \text{ on } \partial\Omega_h\times(0,T],\\ u_h(0)=R_hu_0 \text{ in } \overline{\Omega}_h.\end{cases}\tag{4.28}$$

To study the stability of the IVP (4.28) it is necessary to consider a solution $\tilde{u}_h(t)\in W_h^*$ of (4.28) with initial condition $\tilde{u}_h(0)\in W_h$. In the next result is established an estimate for $u_h(t)-\tilde{u}_h(t)$. Recall that $C^m([0,T],V)$, $m\in\mathbb{N}_0$, where $V$ is a normed vector space, denotes the space of functions $v:[0,T]\to V$ such that $v^{(j)}:[0,T]\to V$, $j=0,\ldots,m$, are continuous functions.

**Proposition 4.3.** *Let $u_h,\tilde{u}_h\in C([0,T],W_h^*)\cap C^1([0,T],W_h)$ be solutions of (4.28) with initial conditions $u_h(0),\tilde{u}_h(0)\in W_h$, respectively. Let $\omega_h(t)=u_h(t)-\tilde{u}_h(t)$. Then*

$$\begin{aligned}\|\omega_h(t)\|_h^2 \quad &+2(A_0-\varepsilon^2)\int_0^t e^{2q_0(s-t)+\int_s^t\frac{1}{2\varepsilon^2}\|A\|_{C_B^1(\mathbb{R})}^2\|D_{-x}u_h(\mu)\|_{h,\infty}^2d\mu}\|D_{-x}\omega_h(s)\|_+^2ds\\ &\leq e^{-2q_0t+\int_0^t\frac{1}{2\varepsilon^2}\|A\|_{C_B^1(\mathbb{R})}^2\|D_{-x}u_h(\mu)\|_{h,\infty}^2d\mu}\|\omega_h(0)\|_h \text{ in } [0,T],\end{aligned}\tag{4.29}$$

*where $\varepsilon\neq 0$ is an arbitrary constant.*

*Proof.* We start by remarking that $\omega_h(t)$ is solution of the following differential problem

$$\begin{cases}\omega_h'(t)=D_x^*(A(M_hu_h(t))D_{-x}u_h(t))-D_x^*(A(M_h\tilde{u}_h(t))D_{-x}\tilde{u}_h(t))-R_hq(t)\omega_h(t) \text{ in } \overline{\Omega}_h\times(0,T],\\ D_\eta u_h(t)=D_\eta\tilde{u}_h(t)=0 \text{ on } \partial\Omega_h\times(0,T],\\ \omega_h(0)=R_hu_0-\tilde{u}_h(0) \text{ in } \overline{\Omega}_h.\end{cases}\tag{4.30}$$

From (4.30), taking into account (4.8), it is obtained the following equality

$$\begin{aligned}(\omega_h'(t),\omega_h(t))_h &= -(A(M_hu_h(t))D_{-x}u_h(t)-A(M_h\tilde{u}_h(t))D_{-x}\tilde{u}_h(t),D_{-x}\omega_h(t))_+ \\ &\quad -(R_hq(t)\omega_h(t),\omega_h(t))_h, t\in(0,T].\end{aligned} \tag{4.31}$$

Moreover

$$\begin{aligned}-(A(M_hu_h(t))D_{-x}u_h(t) &\quad -A(M_h\tilde{u}_h(t))D_{-x}\tilde{u}_h(t),D_{-x}\omega_h(t))_+ \\ &\leq (-A_0+\varepsilon^2)\|D_{-x}\omega_h(t)\|_+^2 \\ &\quad +\frac{1}{4\varepsilon^2}\|D_{-x}u_h(t)\|_{h,\infty}^2\|A\|_{C_B^1(\mathbb{R})}^2\|\omega_h(t)\|_h^2,\end{aligned}$$

where $\varepsilon\neq 0$ is an arbitrary constant, and consequently, from (4.31), it is obtained

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\omega_h(t)\|_h^2 &\quad +(A_0-\varepsilon^2)\|D_{-x}\omega_h(t)\|_+^2 \\ &\leq \Big(-q_0+\frac{1}{4\varepsilon^2}\|D_{-x}u_h(t)\|_{h,\infty}^2\|A\|_{C_B^1(\mathbb{R})}^2\Big)\|\omega_h(t)\|_h^2 \text{ in } (0,T],\end{aligned}$$

that leads to (4.29). □

Inequality (4.29) guarantees that the IVP (4.28) has at most one solution in $C([0,T],W_h^*)\cap C^1([0,T],W_h)$. In fact, if $u_h,\tilde{u}_h$ are solutions in $C([0,T],W_h^*)\cap C^1([0,T],W_h)$, then, from (4.29), it is concluded that $u_h(t)=\tilde{u}_h(t)$ in $\overline{\Omega}_h$.

As for the Dirichlet boundary problem studied before, to obtain stability from (4.29), it is necessary to impose the uniform boundness of $\int_0^t\|D_{-x}u_h(\mu)\|_{h,\infty}d\mu, t\in[0,T], h\in\Lambda$. It is clear that such boundness cannot be concluded from (4.28). In fact from this last IVP it can be shown that

$$\begin{aligned}\|u_h(t)\|_h^2 &\quad +2A_0\int_0^t e^{2q_0(s-t)}\|D_{-x}u_h(s)\|_+^2ds \\ &\leq e^{-2q_0t}\left(\|R_hu_0\|_h^2+\frac{1}{q_0}\int_0^t e^{2q_0s}\|R_hg\|_h^2ds\right), t\in[0,T],\end{aligned} \tag{4.32}$$

and consequently $\int_0^t\|D_{-x}u_h(s)\|_+^2ds\leq C_b, t\in[0,T], h\in\Lambda$. This last upper bound allows concluding that there exists a positive constant $C_b$, $h$ and $t$ independent, such that

$$\int_0^t\|D_{-x}u_h(\mu)\|_{h,\infty}^2d\mu\leq\frac{1}{h_{min}}C_b, t\in[0,T], h\in\Lambda.$$

The boundness of $\int_0^t\|D_{-x}u_h(\mu)\|_{h,\infty}^2d\mu, t\in[0,T], h\in\Lambda$, will be established in what follows through the error analysis.

Let $E_u=u_h(t)-\tilde{R}_hu(t)\in W_h^*$ be the spatial discretization error where $u$ is solution of the IBVP (4.1), (4.2), (4.3) and $u_h(t)\in W_h^*$ is the semi-discrete approximation defined by (4.28). This error is solution of the IVP

$$\begin{cases}E_u'(t)=D_x^*(A(M_hu_h(t))D_{-x}u_h(t))-D_x^*(A(M_h\tilde{R}_hu(t))D_{-x}\tilde{R}_hu(t)) \\ \qquad -R_hq(t)E_u(t)-T_h(t) \text{ in } \overline{\Omega}_h\times(0,T], \\ D_\eta u_h(t)-D_\eta\tilde{R}_hu(t)=T_{h,\partial\Omega}(t) \text{ on } \partial\Omega\times(0,T], \\ E_u(0)=0 \text{ in } \overline{\Omega}_h,\end{cases} \tag{4.33}$$

where the expressions of $T_h(t)$ and $T_{h,\partial\Omega}(t)$ are obtained from the corresponding expressions $T_h$ and $T_{h,\delta\Omega}$ presented in the last section, replacing $u$ and their derivatives with respect to $x$ by $u(t)$ and their partial derivatives with respect to the spatial variable, respectively.

**Theorem 4.2.** *Let $E_u = u_h(t) - \tilde{R}_h u(t) \in W_h^*$ be the spatial discretization error where $u$ is solution of the IBVP (4.1), (4.2), (4.3) and $u_h(t) \in W_h^*$ is the semi-discrete approximation defined by (4.28). If $u \in C([0,T], C^4(\overline{\Omega}^*)) \cap C^1([0,T], C(\overline{\Omega}))$ then there exists a positive constant $C_u$, $h$ and $t$ independent, such that*

$$\begin{aligned} \|E_u(t)\|_h^2 \quad & +2(A_0 - 6\varepsilon^2)\int_0^t e^{\int_s^t 2\theta(u(\mu))d\mu}\|D_{-x}E_u(s)\|_+^2 ds \\ & \leq C_u h_{max}^4 \int_0^t e^{\int_s^t 2\theta(u(\mu))d\mu}\|u(s)\|_{C^4(\overline{\Omega}^*)}^2 ds, t \in (0,T], \end{aligned} \tag{4.34}$$

*with $\varepsilon \neq 0$ and*

$$\theta(u(t)) = q_0 - 6\varepsilon^2 - \frac{1}{4\varepsilon^2}\|u(t)\|_{C^1(\overline{\Omega})}^2 \|A\|_{C_B^1(\mathbb{R})}^2. \tag{4.35}$$

*Proof.* From (4.33) it is easily obtained the following relation

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|E_u(t)\|_h^2 \quad & \leq -(A(M_h u_h(t))D_{-x}u_h(t), D_{-x}E_u(t))_+ \\ & +(A(M_h\tilde{R}_h u(t))D_{-x}\tilde{R}_h u(t), D_{-x}E_u(t))_+ - q_0\|E_u(t)\|_h^2 \\ & -(T_h(t), E_u)_h - \sum_{i=0,N} T_{h,\partial\Omega}(x_i,t)E_u(x_i,t) \text{ in } (0,T] \end{aligned}$$

Following the proof of Theorem 4.1, it can be shown that there exists a positive constant $C_u$, $h$ and $t$ independent, such that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|E_u(t)\|_h^2 \quad & +(A_0 - 6\varepsilon^2)\|D_{-x}E_u(t)\|_+^2 \\ & \leq C_u h_{max}^4 \|u(t)\|_{C^4(\overline{\Omega}^*)}^2 - \theta(u(t))\|E_u(t)\|_h^2 \text{ in } (0,T], \end{aligned} \tag{4.36}$$

where $\theta$ is defined by (4.35). Inequality (4.36) leads to (4.34). □

**Corollary 4.1.** *Under the assumptions of Theorem 4.2, fixing $\varepsilon \neq 0$ such as*

$$A_0 - 6\varepsilon^2 > 0, \tag{4.37}$$

*then there exists a positive constant $C_u$, $h$ and $t$ independent, such that holds the following*

$$\|E_u(t)\|_h^2 + \int_0^t \|D_{-x}E_u(s)\|_+^2 ds \leq C_u h_{max}^4,$$

*for $h \in \Lambda$ and $h_{max}$ small enough.*

**Corollary 4.2.** *Under the assumptions of Theorem 4.2, the solution $u_h(t) \in W_h^*$ of (4.28) is a stable solution in $[0,T]$.*

*Proof.* The stability of the IBVP (4.28) in $(u_h(t))_{h\in\Lambda}$ can be concluded from (4.29) if there exists positive constant $C_s$, $h$ and $t$ independent, such that

$$\int_0^t \|D_{-x}u_h(s)\|_{h,\infty}^2 ds \leq C_s \; t \in [0,T],$$

for $h \in \Lambda$ with $h_{max}$ small enough. Theorem 4.2 is used to guarantee the existence of such constant. Given

$$\begin{aligned}\int_0^t \|D_{-x}u_h(s)\|_{h,\infty}^2 ds &\leq 2\int_0^t \|D_{-x}E_u(s)\|_{h,\infty}^2 ds + 2\int_0^t \left\|\frac{\partial u}{\partial t}(s)\right\|_\infty^2 ds \\ &\leq \frac{2}{h_{min}}\int_0^t \|D_{-x}E_u(s)\|_+^2 ds + 2\int_0^t \left\|\frac{\partial u}{\partial t}(s)\right\|_\infty^2 ds\end{aligned}$$

from Corollary 4.1 is consequently obtained that

$$\int_0^t \|D_{-x}u_h(s)\|_{h,\infty}^2 ds \leq 2C_u\frac{h_{max}^4}{h_{min}} + 2\int_0^t \left\|\frac{\partial u}{\partial t}(s)\right\|_\infty^2 ds.$$

□

### 4.3.2 A fully discrete scheme

This section is focused in the study of a fully implicit-explicit scheme to compute an approximation for the solution of the IBVP (4.1), (4.2), (4.3). Let $\{t_n, n = 0,\ldots,\mathscr{M}\}$ a uniform grid in $[0,T]$, with $t_0 = 0, t_{\mathscr{M}} = T$ and $t_{n+1} - t_n = \Delta t$. Let $u_h^n \in W_h^*$ be defined by

$$\begin{cases} \frac{u_h^{n+1}-u_h^n}{\Delta t} = D_x^*(A(M_h u_h^n)D_{-x}u_h^{n+1}) - R_h q(t_{n+1})u_h^{n+1} + R_h g(t_{n+1}) \text{ in } \overline{\Omega}_h, n = 0,\ldots,\mathscr{M}-1, \\ D_\eta u_h^j = 0 \text{ on } \partial\Omega_h, j = 1,\ldots,\mathscr{M}, \\ u_h^0 = R_h u_0 \text{ in } \overline{\Omega}_h, \end{cases} \tag{4.38}$$

where

$$D_\eta u_h^j(x_0) = -\frac{1}{2}\left(A(M_h u_h^{j-1}(x_1))D_{-x}u_h^j(x_1) + A(M_h u_h^{j-1}(x_0))D_{-x}u_h^j(x_0)\right),$$

$$D_\eta u_h^j(x_N) = \frac{1}{2}\left(A(M_h u_h^{j-1}(x_{N+1}))D_{-x}u_h^j(x_{N+1}) + A(M_h u_h^{j-1}(x_N))D_{-x}u_h^j(x_N)\right).$$

In what follows, $C^{n,m}(\overline{\Omega}^* \times [0,T])$ represents the space of functions defined in $\overline{\Omega}^* \times [0,T]$, with continuous partial derivatives with respect to $x$ and $t$ until order $n$ and $m$, respectively.

Then is established an upper bound for the global error $E_u^n = u_h^n - \tilde{R}_h u(t_n) \in W_h^*$ that is solution of the following finite difference problem

$$\begin{cases} \frac{E_u^{n+1}-E_u^n}{\Delta t} = D_x^*(A(M_h u_h^n))D_{-x}u_h^{n+1} - R_h q(t_{n+1})E_u^{n+1} \\ \qquad -D_x^*(A(M_h\tilde{R}_h u(t_n))D_{-x}\tilde{R}_h u(t_{n+1})) - \tilde{T}_h(t_{n+1}) \text{ in } \overline{\Omega}_h, \\ \qquad n = 0,\ldots,\mathscr{M}-1, \\ D_\eta u_h^j - D_\eta\tilde{R}_h u(t_j) = T_{h,\partial\Omega_h}^{j-1,j} \text{ on } \partial\Omega_h,\ j = 1,\ldots,\mathscr{M}, \\ E_u^0 = 0 \text{ in } \overline{\Omega}_h, \end{cases} \tag{4.39}$$

where $T_{h,\partial\Omega_h}^{j-1,j}(x_i)$ denotes the truncation error induced by the discretization of the boundary conditions that admits the representation

$$T_{h,\partial\Omega_h}^{j-1,j}(x_i) = T_{h,\partial\Omega_h}(x_i,t_j) + T_{\partial\Omega_h,\Delta t}^{j-1,j}(x_i),$$

where $T^{j-1,j}_{\partial\Omega_h,\Delta t}$ is the truncation error induced by the replacement of $u(t_j)$ by $u(t_{j-1})$ in the coefficient function $A$. This error satisfies

$$|T^{j-1,j}_{\partial\Omega_h,\Delta t}(x_i)| \leq \Delta t \|A\|_{C^1_B(\mathbb{R})} \|u\|^2_{C^1(\overline{\Omega}^*\times[0,T])},\ i=0,N.$$

In (4.39), $\tilde{T}_h(t_{n+1})$ is given by

$$\tilde{T}_h(t_{n+1}) = T_h(t_{n+1}) + T^{n,n+1}_{h,\Delta t},$$

being $T_h(t_{n+1})$ the spatial truncation error at $t_{n+1}$ introduced before, and $T^{n,n+1}_{h,\Delta t}$ denotes the error induced by the discretization of the temporal derivative, and by the replacement of $A(M_h\tilde{R}_h u(t_{n+1}))$ by $A(M_h\tilde{R}_h u(t_n))$. This error satisfies

$$\begin{aligned} |T^{n,n+1}_{h,\Delta t}(x_i)| \leq\ & \frac{1}{2}\Delta t \|A\|_{C^1_B(\mathbb{R})} \|u\|_{C^1(\overline{\Omega}\times[0,T])} \Big(2\|u\|_{C^2(\overline{\Omega}\times[0,T])} + \|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\Big) \\ & + \frac{1}{2}\Delta t \|u\|_{C^{0,2}(\overline{\Omega}\times[0,T])}. \end{aligned} \tag{4.40}$$

**Theorem 4.3.** *Let $u \in C^{4,0}(\overline{\Omega}^*\times[0,T]) \cap C^2(\overline{\Omega}\times[0,T]) \cap C^1(\overline{\Omega}^*\times[0,T])$ be solution of the IBVP (4.1), (4.2), (4.3) and let $u^n_h \in W^*_h, n=0,\ldots,\mathscr{M}$, be defined by (4.38) and let $E^n_u = u^n_h - \tilde{R}_h u(t_n) \in W^*_h$ be the global error. It is assumed that $A \in C^3_B(\mathbb{R}), A \geq A_0 > 0$ in $\mathbb{R}$, and $q \geq q_0 > 0$ in $\overline{\Omega}$. Then there exists a positive constant $C_u$, $h$ and $n$ independent, such that*

$$\|E^n_u\|^2_h + \sigma\Delta t \sum_{j=1}^{n} e^{\theta(n-j)\Delta t} \|D_{-x}E^j_u\|^2_+ \leq \frac{1}{\theta}\Gamma\Big(e^{\theta n\Delta t} - 1\Big),\ n=1,\ldots,\mathscr{M}, \tag{4.41}$$

*where*

$$\begin{aligned} \Gamma \leq\ & C_u\Big(h^4_{max}\|u\|^2_{C^{4,0}(\overline{\Omega}^*\times[0,T])} + \Delta t^2\Big(\|u\|^4_{C^1(\overline{\Omega}^*\times[0,T])} + \|u\|^2_{C^2(\overline{\Omega}\times[0,T])} \\ & + \|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\Big(\|u\|^2_{C^2(\overline{\Omega}\times[0,T])} + \|u\|^4_{C^1(\overline{\Omega}\times[0,T])}\Big)\Big)\Big) \end{aligned} \tag{4.42}$$

$$\sigma = \frac{2(A_0 - 6\varepsilon^2)}{1 - 14\varepsilon^2\Delta t},$$

*and*

$$\theta = \frac{\frac{1}{2\varepsilon^2}\|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\|A\|^2_{C^1_B(\mathbb{R})} + 14\varepsilon^2}{1 - 14\varepsilon^2\Delta t}$$

*with $\varepsilon \neq 0$ satisfying*

$$1 - 14\Delta t\varepsilon^2 > 0 \text{ and } A_0 - 6\varepsilon^2 > 0. \tag{4.43}$$

*Proof.* From (4.39) it can be shown that holds the following

$$\left\{\begin{array}{ll} \|E^{n+1}_u\|^2_h & \leq (E^n_u, E^{n+1}_u)_h - \Delta t A_0 \|D_{-x}E^{n+1}_u\|^2_+ + \\ & + \Delta t((A(M_h\tilde{R}_h u(t_n)) - A(M_h u^n_h))D_{-x}\tilde{R}_h u(t_{n+1}), D_{-x}E^{n+1}_u)_+ \\ & - \Delta t q_0 \|E^{n+1}_u\|^2_h - \Delta t(T_h(t_{n+1}) + T^{n,n+1}_{h,\Delta t}, E^{n+1}_u)_h \\ & + \Delta t \sum_{i=0,N}\Big(T_{h,\partial\Omega_h}(x_i,t_{n+1}) + T^{n,n+1}_{\partial\Omega_h,\Delta t}(x_i)\Big)E^{n+1}_u(x_i) \\ & n = 0,\ldots,\mathscr{M}-1, \\ & E^0_u = 0 \text{ in } \overline{\Omega}_h, \end{array}\right. \tag{4.44}$$

Observe that

$$((A(M_h\tilde{R}_h u(t_n)) - A(M_h u_h^n))D_{-x}\tilde{R}_h u(t_{n+1}), D_{-x}E_u^{n+1})_+ \leq \frac{1}{2\varepsilon^2}\|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\|A\|^2_{C_B^1(\mathbb{R})}\|E_u^n\|_h^2 + \varepsilon^2\|D_{-x}E_u^{n+1}\|_+^2,$$

where $\varepsilon \neq 0$ is an arbitrary constant, and

$$\begin{aligned}
-(T_h(t_{n+1}), E_u^{n+1})_h &\leq C_u h_{max}^4 \|u(t_{n+1})\|^2_{C^4(\overline{\Omega}^*)} + \varepsilon^2\|D_{-x}E_u^{n+1}\|_+^2 + 2\varepsilon^2\|E_u^{n+1}\|_h^2 \\
&+ \sum_{i=0,N}\frac{h_i^2}{2}|T_h(x_i,t_{n+1})||E_u^{n+1}(x_i)| \\
&\leq C_u h_{max}^4 \|u(t_{n+1})\|^2_{C^4(\overline{\Omega}^*)} + \varepsilon^2\|D_{-x}E_u^{n+1}\|_+^2 + 2\varepsilon^2\|E_u^{n+1}\|_h^2 \\
&+ \sum_{i=0,N}\frac{h_i^2}{8\varepsilon^2}T_h(x_i,t_{n+1})^2 + 2\varepsilon^2\|E_u^{n+1}\|_{1,h}^2 \\
&\leq C_u h_{max}^4 \|u(t_{n+1})\|^2_{C^4(\overline{\Omega}^*)} + 3\varepsilon^2\|D_{-x}E_u^{n+1}\|_+^2 + 4\varepsilon^2\|E_u^{n+1}\|_h^2.
\end{aligned}$$

Moreover,

$$\begin{aligned}
-(T_{h,\Delta t}^{n,n+1}, E_u^{n+1})_h &\leq \frac{1}{16\varepsilon^2}\Big[\Big(\|A\|_{C_B^1(\mathbb{R})}\|u\|_{C^1(\overline{\Omega}\times[0,T])}\Big(2\|u\|_{C^2(\overline{\Omega}\times[0,T])} + \|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\Big)\Big)^2 \\
&+ \|u\|^2_{C^{0,2}(\overline{\Omega}\times[0,T])}\Big]\Delta t^2 + \varepsilon^2\|E_u^{n+1}\|_h^2,
\end{aligned}$$

and

$$\begin{aligned}
\sum_{i=0,N}\Big(T_{h,\partial\Omega_h}(x_i,t_{n+1}) + T_{\partial\Omega_h,\Delta t}^{n,n+1}(x_i)\Big)E_u^{n+1}(x_i) &\leq \sum_{i=0,N}\frac{1}{2\varepsilon^2}\Big(T_{h,\partial\Omega_h}(x_i,t_{n+1})^2 + T_{\partial\Omega_h,\Delta t}^{n,n+1}(x_i)^2\Big) \\
&+ 2\varepsilon^2\|E_u^{n+1}\|_{1,h}^2.
\end{aligned}$$

Taking in (4.44) the last estimates, it is obtained the following

$$\begin{aligned}
\big(1 - 14\Delta t\varepsilon^2\big)\|E_u^{n+1}\|_h^2 + 2\Delta t(A_0 - 6\varepsilon^2)\|D_{-x}E_u^{n+1}\|_+^2 \leq \big(1 + \Delta t\frac{1}{2\varepsilon^2}\|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\|A\|^2_{C_B^1(\mathbb{R})}\big)\|E_u^n\|_h^2 \\
+ \Delta t\Gamma,
\end{aligned} \tag{4.45}$$

where $\Gamma$ is defined by (4.42).

The sequence defined recursively by (4.45) can be seen, for $\varepsilon$ satisfying (4.43), as a particular case of the following one

$$a_{n+1} \leq -\alpha b_{n+1} + \beta a_n + \hat{\gamma},\ n = 0,\ldots,\mathscr{M}-1,\ a_0 = 0,$$

with $a_n \geq 0, b_n \geq 0$, and $\alpha,\beta,\hat{\gamma} > 0$. It can be shown that $(a_n)$ satisfies

$$a_n + \alpha\sum_{j=1}^{n}\beta^{n-j}b_j \leq \hat{\gamma}\sum_{j=0}^{n-1}\beta^j,\ n = 1,\ldots,\mathscr{M}. \tag{4.46}$$

Then, from (4.45), taking into account (4.46), we get

$$\begin{aligned}\|E_u^n\|_h^2 \quad &+\frac{\Delta t 2(A_0-6\varepsilon^2)}{1-14\varepsilon^2\Delta t}\sum_{j=1}^{n}\left(\frac{1+\Delta t\frac{1}{2\varepsilon^2}\|u\|^2_{C^1(\overline{\Omega}\times[0,T])}\|A\|^2_{C_B^1(\mathbb{R})}}{1-14\varepsilon^2\Delta t}\right)^{n-j}\|D_{-x}E_u^j\|_+^2\\ &\leq \frac{\Delta t}{1-14\varepsilon^2\Delta t}\Gamma\sum_{j=0}^{n-1}\left(\frac{1+\Delta t\frac{1}{2\varepsilon^2}\|u\|^2_{C^1(\overline{\Omega}\times 0,T])}\|A\|^2_{C_B^1(\mathbb{R})}}{1-14\varepsilon^2\Delta t}\right)^{j},\end{aligned}$$

for $j=1,\ldots,\mathscr{M}$, that leads to (4.41)

□

Theorem 4.3 establishes the following estimate for the error

$$\|E_u^n\|_h^2+\sigma\Delta t\sum_{j=1}^{n}e^{\theta(n-j)\Delta t}\|D_{-x}E_u^j\|_+^2\leq C_u\left(h_{max}^4+\Delta t^2\right),\ n=1,\ldots,\mathscr{M},$$

giving that the approximation (4.38) is of second-order in space, and of first-order in time.

## 4.4 Conclusions

The study of stability and convergence properties of a numerical method for the IBVP (4.1), (4.2), (4.3) was the main goal of this chapter. As a start point, it was considered the BVP (4.4) complemented with homogeneous Neumann boundary conditions (4.5). For this problem, a FDM, which can be seen as fully discrete piecewise linear FEM, was proposed, and in Theorem 4.1 was established its second-order of convergence, with respect to a discrete $H^1$-norm, assuming that the solution of the differential problem $u\in C^4(\overline{\Omega}^*)$. Note that the proof of these results followed the same techniques used in Chapter 2 with the respective adaptations.

In what concerns the study of the IBVP (4.1), (4.2), (4.3), a semi-discrete approximation defined by the FDM (4.28), which can be seen as a fully discrete piecewise linear FEM, was proposed, and its convergence and stability analysis were studied for solutions in $C([0,T],C^4(\overline{\Omega}^*))\cap C^1([0,T],C(\overline{\Omega}))$. In Theorem 4.2, was proved the second-order of convergence of the method, and then, as a consequence of the error analysis, the stability was concluded in Corollary 4.2. A fully discrete approximation for this IBVP was also presented and its convergence was studied, having been proved its second-order of convergence in space, and first-order of convergence in time. This result was stated in Theorem 4.3.

# Chapter 5

# Detection of sunspots in spectroheliograms

## 5.1 Introduction

The purpose of this chapter is the automatic detection and geometric definition of sunspots, including the limits of umbra and penumbra, in solar images. It is intended that with the experience in image processing techniques suitable for this type of solar phenomena, one can contribute, in a future work, to the study of magnetic field evolution in sunspots.

Concerning the image processing techniques to detect sunspots, the more oftenly used are threshold techniques [19, 39], edge detection [55, 84], region growing [85], mathematical morphology transforms [16, 21, 38, 83], neural networks [18], fuzzy sets [29, 30], and classification schemes [57, 63]. Hybrid methods, that include different approaches, have also been developed and can be found in [22, 50]. Another example of the integration of different methods could be found in [82], which combines morphological operators and region growing techniques. In [85], a comparation between automatic approaches and manual analysis is done, which proves the efficiency of the automatic techniques.

Once the Geophysical and Astronomical Observatory of the University of Coimbra has huge collection of solar images that was never been processed, this historical database was chosen to apply a sunspot detection and segmentation method developed within the scope of this work. However, the detection and segmentation of sunspots are not the primary objective of the present thesis, and as such, it has not invested in a new and innovative methodology. Since there is proven evidence of the success of the application of mathematical morphology (MM) in various types of images, such as medical images, remote sensing, as well as solar images, and also because of its versatility it was opted to apply this theory/technique of image analysis.

The organization of this chapter is described in what follows. In Section 5.2 are presented the basic concepts of the MM, which is the image processing technique chosen to be applied to detect and segment sunspots. In Section 5.3 is described the Geophysical and Astronomical Observatory of the University of Coimbra solar database once a subset of it was used in this work to construct and validate a tool to detect and segment sunspots. The developed algorithm is presented step-by-step in Section 5.4. Section 5.5 is reserved for the analysis and discussion of results, being presented

the metrics used to evaluate the performance of the algorithm as also as examples resulting from the application of the algorithm in some particular images. This chapter finishes with Section 5.6 of conclusions.

## 5.2 Mathematical morphology: basic concepts

The origin of MM dates back to the 1960s, at the École des Mines de Paris, when George Matheron and Jean Serra intended to study the geometry of porous medium [52]. Based on works of the mathematicians H. Minkowski and H. Hadwiger, they took the first steps in the creation of a new image analysis theory: the MM. Theoretical concepts and foundations have been developed over the years, and scientific works on the subject have been published [53, 68]. However, only in 1982, with the publication of the book "Image analysis and mathematical morphology" [69], MM starts to broaden its audience. MM was initially used only in binary images but, over time, it has been evolved and consolidated concepts. Several groups around the world have been contributing to its development, making MM a recognized and powerful framework for image analysis and processing. Currently, it is applied to all types of images, greyscale images, colour images, and even 3D images, defined in static or dynamic spaces. What makes MM powerful is not only its applicability to various types of images but also its versatility, giving the possibility to link basic operators and build more robust ones, and the possibility of extracting measurements (like areas, sinuosity, among others). MM has successful applications in several scientific fields [72] that require manipulation and analysis of images, such as medicine [54, 66, 76], geology [3, 51, 61], and solar physics [4, 33, 38], among others.

The essence of MM consists in comparing features to be analyzed with some known and simple shape, called the structuring element (SE). The SE must be chosen in a way that resembles the objects sought in the image and also depends on the purpose of the study that is intended to be carried out. If we are looking for, for example, solar filaments, the choice of the SE would fall on a line, since the shape of this structure is more elongated. Since it is intended to detect sunspots, the natural choice for SE is a disk. MM was initially developed to be applied only to binary images. In this sense, the foundations of MM are based on set theory, and the morphological operations were developed from order relations such as inclusion or intersection. The application of MM to greyscale images appeared later, making it necessary to adapt the morphological operators. If in binary images the range of pixels values is restricted to $\{0,1\}$, the same is not true for greyscale images which is extended to a larger subset of $\mathbb{N}_0$. More concretely, a 2D greyscale image $f$ is a function

$$f : D_f \subset \mathbb{Z}^2 \longrightarrow \{0,1,...,t_{max}\},$$

where $D_f$ the definition domain of $f$, and $t_{max}$ is the maximum value of the data type which depends on the image file format.

Since the set of images used in this work is composed of greyscale images, morphological operators in the scope of functions are then presented, having as reference [2, 60, 73].

MM consists of a larger number of operators, however only the morphological transformations used in the developed algorithm to detect sunspots are presented.

In what follows the translation of an image $f$ by a vector $b$, is denoted by $f_b$, and is giving by

$$f_b(x) = f(x-b).$$

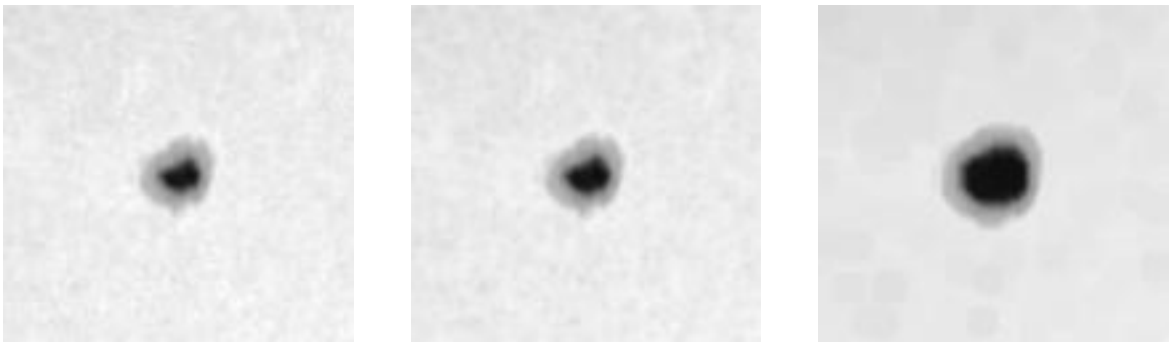
### 5.2.1 Erosion and dilation

The erosion and dilation are the most basic operations of MM. They are seen as the two letters of the morphological alphabet, once all the other operations can be expressed as a combinations of these two [73].

To greyscale images, the erosion of an image $f$ by a structuring element $B$ of size $\lambda$, denoted by $\varepsilon_B(f)$, is defined as the minimum of translations of $f$ by the vectors $-b$ of $B$:

$$\varepsilon_B(f) = \min_{b\in B} f_{-b}.$$

This operation is illustrated in Figure 5.1. As can be seen, the erosion causes an enlargement of the darker areas of the image, highlighting them. In the case of Figure 5.1, as the size of the SE increases, the umbra also increases. The dilation of an image ($\delta_B(f)$) of an image $f$ by a structuring



(a) Original image of a sunspot. (b) Erosion by a disk of size 1. (c) Erosion by a disk of size 6.

Fig. 5.1 Examples of the application of the erosion operator on a sunspot image.

element $B$ of size $\lambda$, is the maximum of translations of $f$ by the vectors $-b$ of $B$:

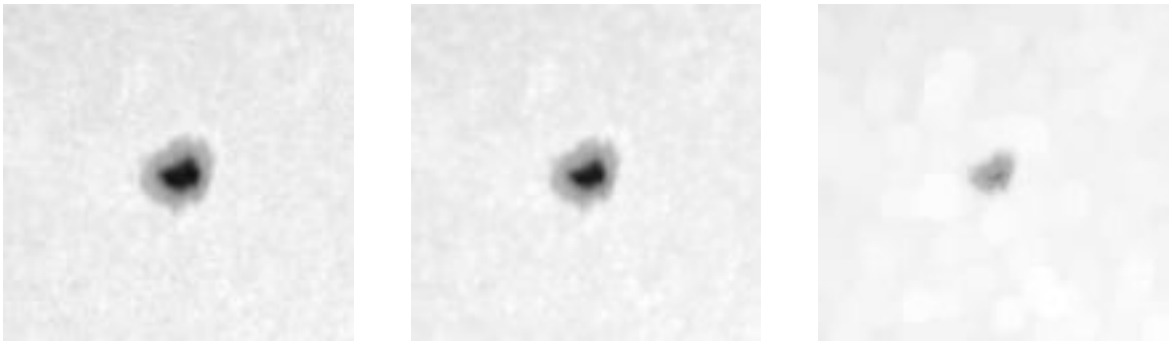
$$\delta_B(f) = \max_{b\in B} f_{-b}.$$

Unlike the erosion, the dilation leads to narrowing or even elimination of the darkest areas. In Figure 5.2 can be seen that, with the increase of the SE, a disk, darker areas (the entire sunspot) tend to disappear, and the lighter areas (the background) are highlighted.

The final results for both operators, erosion and dilation, are poor in details when compared with the initial image.

### 5.2.2 Morphological gradient

The objective of the morphological gradient is to enhance and extract the contours of the homogeneous regions of gray levels in an image.



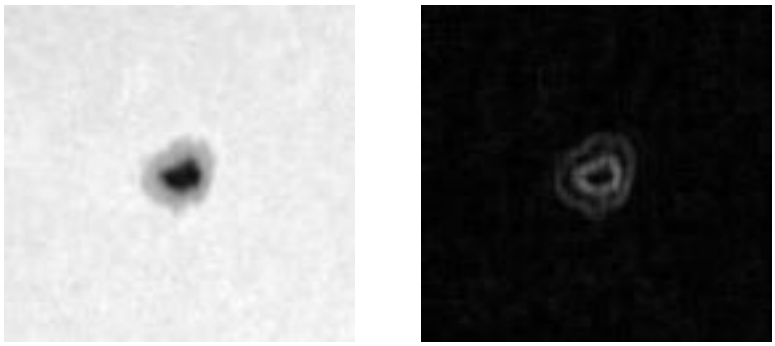
(a) Original image of a sunspot. (b) Dilation by a disk of size 1. (c) Dilation by a disk of size 6.

Fig. 5.2 Examples of the application of the dilation operator on a sunspot image.

Morphological gradient, also known as Beucher gradient, is defined as the arithmetic difference between the dilation and the erosion by the SE B of size $\lambda$,

$$\rho_B = \delta_B - \varepsilon_B.$$

The objective of the morphological gradient is to enhance and extract the contours of the homogeneous regions of grey levels in an image. The result of the aplication of the morphological gradient by a disk of size 1, to the image in Figure 5.3a, can be seen in Figure 5.3b, where the contours of umbra and penumbra were highlighted.



(a) Original image of a sunspot. (b) Gradient by a disk of size 1.

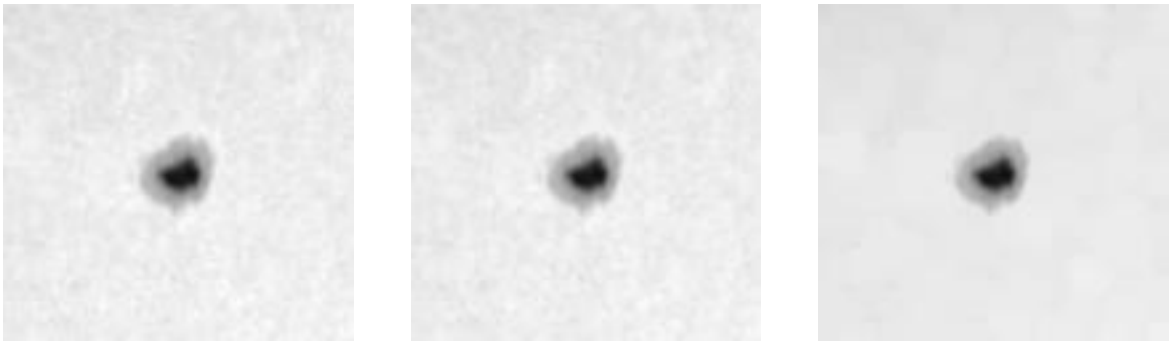
Fig. 5.3 Example of the application of the morphological gradient operator on a sunspot image.

### 5.2.3 Opening and closing

The erosion and dilation can be combined to perform two important transforms: opening and closing. The opening appears as a way to recover many structures lost in erosion. However, the structures that have been completely destroyed can never be recovered at all. So, the opening, denoted by $\gamma$, consists of the erosion of an image $f$ by the same structuring element $B$ of size $\lambda$ followed by a dilation with the reflection[1] of B,

$$\gamma_B(f) = \delta_{\check{\mathrm{B}}}[\varepsilon_B(f)].$$

[1]The reflection of set B is given by $\check{\mathrm{B}} = \{-b | b \in B\}$, with respect to the origin.



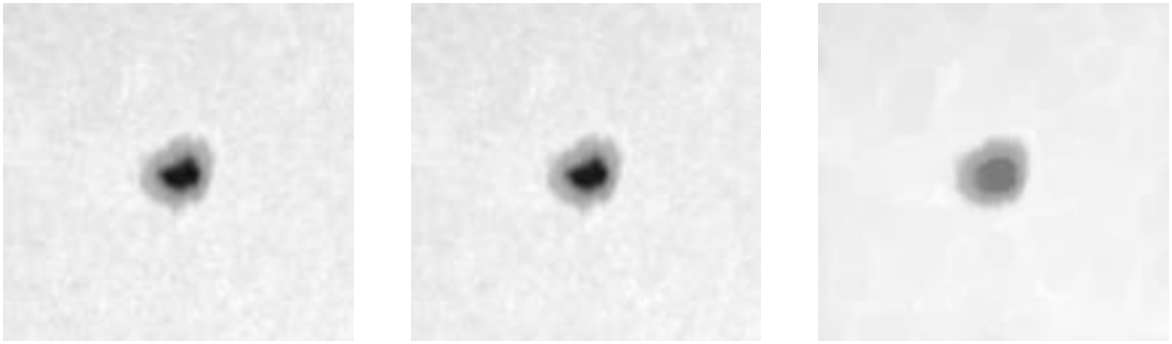
(a) Original image of a sunspot. (b) Opening by a disk of size 1. (c) Opening by a disk of size 6.

Fig. 5.4 Examples of the application of the opening operator on a sunspot image.

The application of the opening results in the cut of peaks (lighter zones of the image) and removes small object protuberances. What happens is that darker areas (valleys) widen due to the reduction of the peaks, thus losing details of the image. The Figure 5.4 illustrates the application of openings. As can be seen in the referred figure, when the size of the SE increases, the darker areas (sunspots) are enhanced, and the lighter structures (sun's granulation) tend to disappear.

The closing transform, $\phi$, is the dual operation of the opening, and consists in applying a dilation by a SE $B$ to an image $f$ followed by an erosion by $\check{\mathrm{B}}$,

$$\phi_B(f) = \varepsilon_{\check{\mathrm{B}}}[\delta_B(f)]$$



(a) Original image of a sunspot. (b) Closing by a disk of size 1. (c) Closing by a disk of size 6.

Fig. 5.5 Examples of the application of the closing operator on a sunspot image.

Figure 5.5 shows that the application of a closing operation wide the lighter zones and the darker ones narrowed, and tend to disappear as the size of the SE increases.

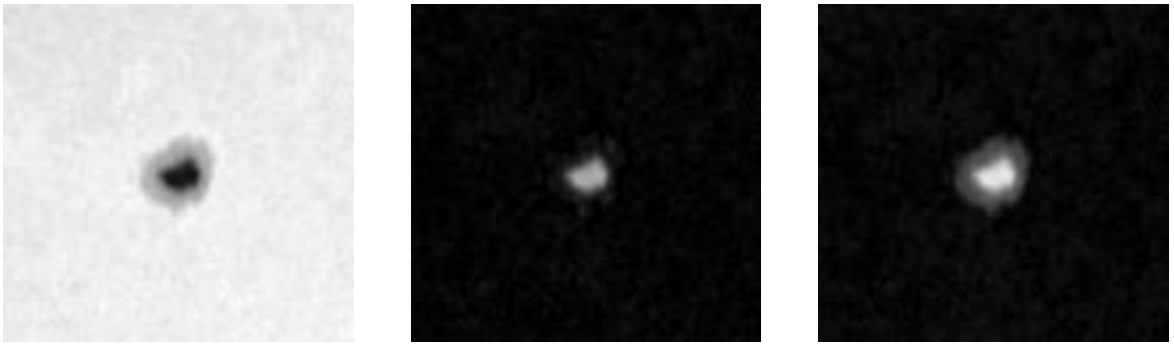
The application of the opening and the closing change less the original image than the erosion and the dilation respectively, i.e., the effects of the two first transforms are more attenuated than the other two.

### 5.2.4 Top-hat

The top-hat transforms are combinations of openings and closings. These transforms allow the extraction of valleys or peaks of a dimension less than the one of the chosen SE. The black top-hat (BTH) creates an image by subtracting the image f from its closing,

$$BTH(f) = \phi_B(f) - f.$$

The closing transform suppresses all the valleys (locally dark areas) which are smaller than the dimension of the SE. Subtracting the initial image $f$ from the image resulting of the closing allows recovering these structures by putting them at the same level, and by simultaneously filtering the ones that were not modified.



(a) Original image of a sunspot.

(b) Black top-hat by a disk of size 10.

(c) Black top-hat by a disk of size 20.

Fig. 5.6 Examples of the application of the black top-hat operator on a sunspot image.

The Figure 5.6 illustrates results of the application of a black top-hat transforms. As expected, the transformation performed extract dark areas, but, in Figure 5.6b only the umbra was extracted. This fact happens because of the size of the SE chosen (10). The increase of its dimension to 20 (Figure 5.6c), allows the extraction of the whole sunspot. It should be noted that the choice of the shape and the size of the SE, depends on the morphology of the structures to be extracted.

There is also the dual transformation of black to-hat which is called white top-hat. This transformation was not used in the sunspot detection agorithm and for that reason it is not presented here.

### 5.2.5 Geodesic reconstruction

The operator to be described belongs to a set of transformations of the MM called geodesics. Geodesics transformations resort to distances that correspond to measures carried out in a predetermined region, i.e., they are confined to that region. The difference between the geodesics and elementary transformations (also called Euclidean transformations) is, therefore, the space where they are defined, the geodesic, and the Euclidean space respectively. Geodesic transformations have as input two images: one where the transformation is applied (marker), and the other one that defines the region where the transformation is applied (mask). In these kind of transformations, the SE is not explicit, is the pair of images itself that produce new morphological primitives to use. Also the choice of the size of

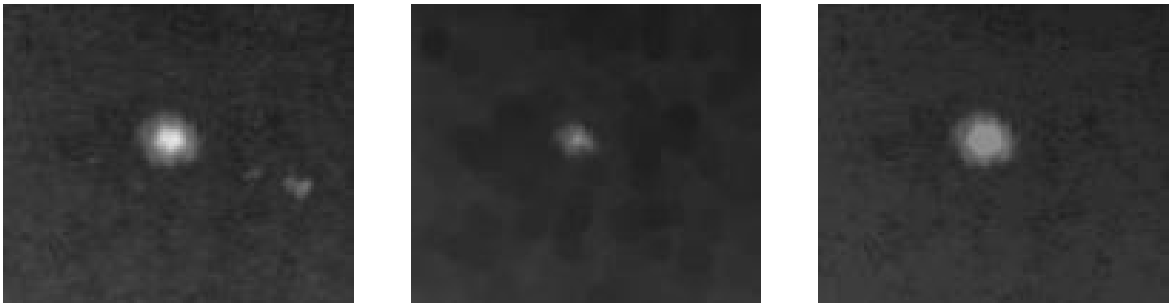
the primitives is unnecessary since these transformations are applied iteratively until idempotence is reached. The interest of geodesic transformations lies in the fact that it is possible to analyze a structure without having to analyze the entire space of the image.

In what concerns geodesic reconstruction, there are two types: reconstruction by dilation, and by erosion. In the reconstruction by dilation the marker image is subjected to successive geodesic dilations, with respect to the mask image, until idempotency (see [73] for more details). By definition the mask and marker images must have the same domain, and the mask image must be greater or equal than the marker image. Therefore, the reconstruction by dilation, denoted by $R_g^{\delta}$, of a mask image $g$, from a marker image $f$ is given by,

$$R_g^{\delta}(f) = \delta_g^{(i)}(f),$$

where $i$ is such that $\delta_g^{(i)}(f) = \delta_g^{(i+1)}(f)$. The reconstruction by erosion ($R_g^{\varepsilon}$) was not used in this work, but is defined analogously.

How the reconstruction is applied depends on the intended purpose. When the restore of peaks that are marked is desired a reconstruction by dilation should be used. By the other hand, to restore valleys that are marked, a reconstruction by erosion should be chosen.



(a) Mask image. (b) Marker image. (c) Reconstructed image.

Fig. 5.7 Example of the application of the geodesic reconstruction by dilation operator.

As an example of the application of the reconstruction by dilation operator, an image containing a sunspot and 3 pores, Figure 5.7a, is used as mask image. Figure 5.7b, which was obtained eroding the mask image by a disk of size 5, is used as marker image. The reconstruction by dilation of the mask image from the marker image is shown in Figure 5.7c. As can be seen the sunspot was reconstructed, but the pores, once were not marked, were not reconstructed. Note that this operator reconstruct the shape, but not digital values.

The pair of transformations erosion and reconstruction is often used for filtering, eliminating unwanted structures (erosion) and restoring the shape of the structures that matter (reconstruction).

### 5.2.6 Thinning

The thinning is part of a set of transformations of the MM called hit-or-miss transformations. Generally speaking, this type of operations consists of first analyzing the image so that it is modified or not, according to neighborhood criteria. In what concerns the thinning, when applied to grey level images, it replaces the attribute of the central point by the attribute of one of the points of the

chosen neighborhood configuration, or it maintains that value.The formal definition of thinning is not presented in this work, once it requires more in-depth knowledge of MM. For more details, see [73].

## 5.3 Database

The Geophysical and Astronomical Observatory of the University of Coimbra has a collection of solar observations on a daily basis that spans near nine decades until today. This extensive collection acquired with the same instrumental apparatus is presently entirely available in digital format. The image acquisition instrument, is a spectroheliograph based on Deslandres principles [48]. It consists of a coelostat, with a primary flat mirror of 0.4 m of diameter, and a secondary one of 0.3 m which sends the sunlight into an optical system. This spectrograph system is made of a converging lens that focuses the solar disk into the slit, then the light passes through the pre-filters, after through a collimator, and finally by a diffraction grating before being collected by the recording device. The projected solar disk is entirely swept, across a slit, by mechanically moving the first focal lens of the system. That process takes around 80 seconds to do the full scan. Consequently, the solar disk is not recorded instantaneously but recorded in "slices" onto a CCD (onto a photographic plate before 2007). Those slices are combined later by a specific data reduction software allowing to record the full solar disk at a specific wavelength in one image. These records are called spectroheliograms and they correspond to the intensity observed at a specific wavelength. Regular observations of the full solar disk in the spectral line of $Ca\,\textsc{ii}\,K$ started in 1926, where $K_1$ (393.23 nm) and $K_3$ (393.37 nm) have been recorded. These $K_1$ and $K_3$ corresponds to fine structures inside the $Ca\,\textsc{ii}\,K$ line. Observations in the spectral line of $H_\alpha$ (655.87 nm) started in 1989. In 2007 the $H_\alpha$ *continuum* wavelength (656.28 nm) was added [31]. Those different wavelengths have the property that each one probe different layers of the sun and consequently allow observing a variety of solar phenomena. Observations in the $Ca\,\textsc{ii}\,K$, and in the $H_\alpha$ core, allow observing mainly the solar chromosphere due to their absorption and emission line formation behaviour. On the other hand, when looking at the $H_\alpha$ *continuum*, one looks at the white light of the Sun, meaning the general look of the photosphere when looked at naked eye. That is caused by the absence of significant spectral lines at that specific wavelength allowing to look directly at the photosphere and the structures that are present there. One of the most prominent structures that can observed on the *continuum* are the sunspots. Sunspots look as darker regions on the bright background. Their dark appearance originates from the interaction between the intense solar magnetic field, and the photospheric plasma where the convection that brings hot material from the bottom layers is suppressed by the so-called "frozen flux" ([77]). That leads to the cooling of the plasma inside the sunspot, making them irradiate less, and by consequence making them darker when observed on the *continuum*. An example of a $H_\alpha$ *continuum* spectroheliogram acquired at the Geophysical and Astronomical Observatory of the University of Coimbra is shown in Fig. (5.8). The image consists of the solar disk and some overwritten information related with the acquisition: orientation, place, spectral line, and date. The data set used to develop the algorithm presented in this chapter consists of 144 spectroheliograms, which are 8 bits digital images with 1200 x 1000 pixels, taken at $H_\alpha$ *continuum*. The set comprises images of the solar cycle 24. In a simplified way, a solar cycle can be described as a periodic variation in the solar activity that can be seen, for example, in the number of sunspots [37]. The cycles have a periodicity of approximately 11 years, in which the solar

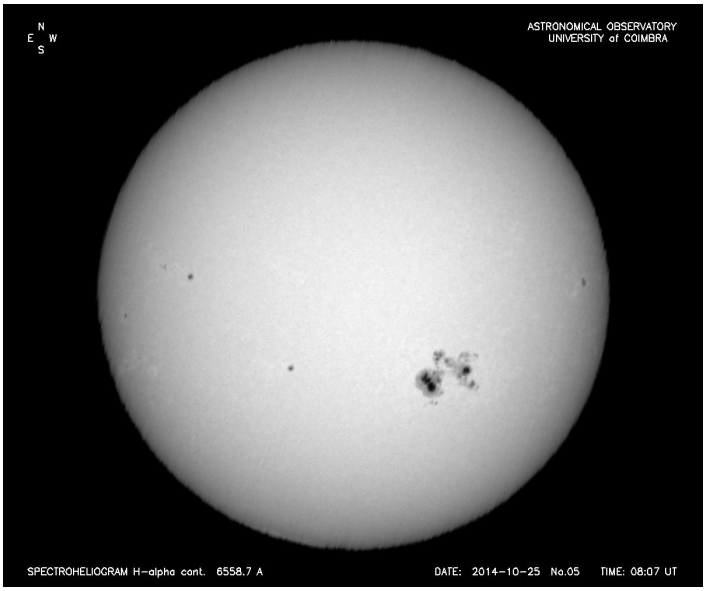



Fig. 5.8 Example of a $H_\alpha$ *continuum* spectroheliogram with sunspots (some of them with umbra and penumbra), acquired on the $25^{th}$ of October 2014.

activity waxes and wanes. As can be seen in Figure 5.9 the solar cycle 24 started around 2009 and is expected to end in 2020. The images of the data set used were chosen to represent the whole cycle:

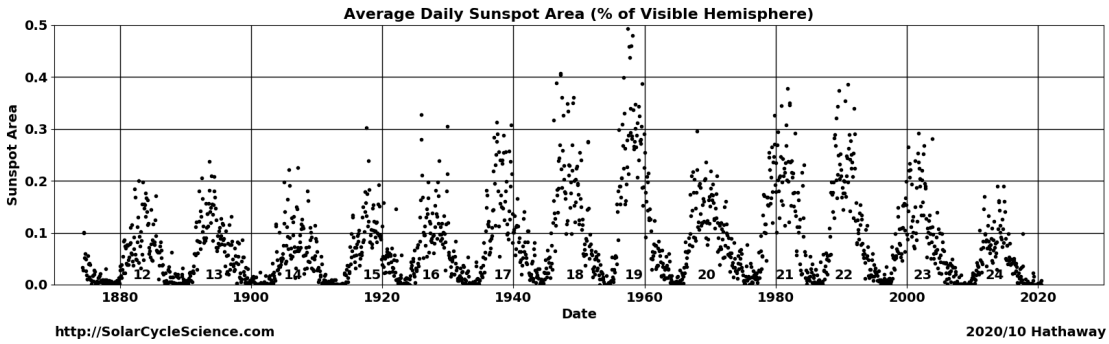



Fig. 5.9 Solar Cycles: Average daily sunspot area record. (Courtesy of D. Hathaway).

taken in different years and in different seasons. Additionally, for each image, an observer, with about 40 years of experience, delineated, manually, the umbra and penumbra regions in order to build the ground-truth data set used to validate the results obtained.

## 5.4 Automatic detection and segmentation of sunspots

Before the development of the sunspots detection and segmentation algorithm, some important aspects were enumerate and have been taken into account posteriorly. Observing Figure 5.8, that shows an example of a spectroheliogram acquired in Coimbra, can be seen that there are some text around the image providing some information as the place, date, and time of the acquisition. This informational text hampers any automatic processing algorithm. It can be also seen in Figure 5.8 that the solar disk

appears to have a solid black background (except the informational text), but actually it is not true once not all pixels have digital levels equal to zero. This characteristic of spectroheliograms results from not only the instrumental noise but also from the characteristics of the acquisition process (one spectroheliogram results from the juxtaposition of multiple "slices" of the solar disk). As it may lead to false sunspots detection outside the solar disk, this issue must be taken into account. It is also true that the solar disk is not a perfect circle, being slightly flatted over its rotation axis, but, relatively to the hundreds of pixels of the solar disk diameter, and not thousands, the flattening effect can be neglected. Therefore, the solar disk can be taken as perfectly circular, facilitating the construction of the algorithm once the irregular boundary could be a problem in the detection of sunspots in the limb. Due to the issues raised previously, an automatic preprocessing was built and should be performed before the detection and segmentation algorithm.

The programming language used to develop the automatic method to detect and segment sunspots was MATLAB using the image processing toolbox. Throughout this section, it is shown the results of the main steps of the algorithm using Figure 5.10 as starting point.

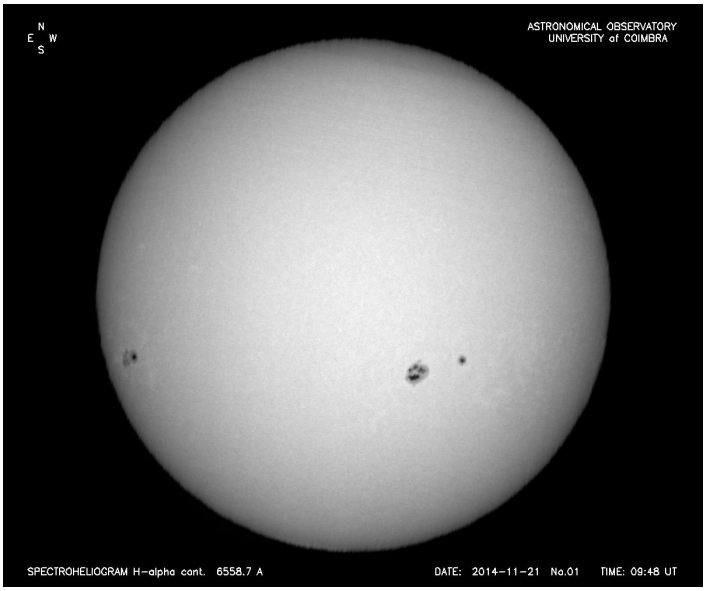



Fig. 5.10 Spectroheliogram of 21$^{st}$ of November 2014, used to exemplify the application of the sunspots detection and segmentation algorithm.

### 5.4.1 Preprocessing

The aims of the preprocessing are the identification of the solar disk, removal of informational text, and the homogenization of the background. The preprocessing starts by applying a closing over the original image using as structuring element a disk of 10 pixels of diameter, which essentially removes small holes since it unites some objects/shapes. The choice of the structuring element was done in order to preserve the circular nature of the solar disk. The resulting image, Figure 5.11a, became more homogeneous, but the text was not removed. Therefore, a morphological operation opening, using as structuring element a disk of size 20, is performed. When applying this transformation the objective is essentially remove small objects/shapes. Figure 5.11b is the result of the application of the opening

operation. Analyzing it, can be observed that the text disappeared, but the digital levels inside the solar disk were not preserved. Therefore, another set of operations must be done until reaching the desired final image. First, a morphological reconstruction is performed using the original image as a mask, and the subtraction of Figure 5.11c by the original image as a marker, originating Figure 5.11d. Hereupon an adaptive threshold filter is applied with a lower cutting-value of 30 and an upper cutting-value of 124. This last operation allows to recover the solar disk as a binary image. The result is shown in Figure 5.11e. Analyzing the image, was detected the existence of a hole (a black spot) inside the solar disk. The presence of holes at this stage of the preprocessing happens for some images of the data set. In order to suppressed it, a fill hole operation is performed (Figure 5.11f). The next objective is to make the solar disk of the image obtained in the previous step, a perfect circle. For that it is necessary to determine the center $(x_c, y_c)$, and the radius $r$ of the solar disk. Recall that the spectroheliograms used in this work have 1000 rows and 1200 columns. Let start with the calculation of x-coordinate of the center, $x_c$. Given the geometric center $(x_{gc}, y_{gc}) = (\frac{1200}{2}, \frac{1000}{2})$ of the image, it is determined for all lines of the image which are 20% above and 20% below that point, the x-coordinate of the first ($x_{first}$) and last ($x_{last}$) pixels belonging to the solar disk, in the x-direction. That is, for each line $i \in \{y_{gc} - 0.2 * 1000, ..., y_{gc} + 0.2 * 1000\} = \{300, ..., 700\}$, the pixels $x_{first_i}$ and $x_{last_i}$ are determined, and $x_c$ is calculated as follows

$$x_c = \frac{1}{1000 * 0.4} \sum_{i=300}^{700} \frac{x_{first_i} + x_{last_i}}{2}.$$

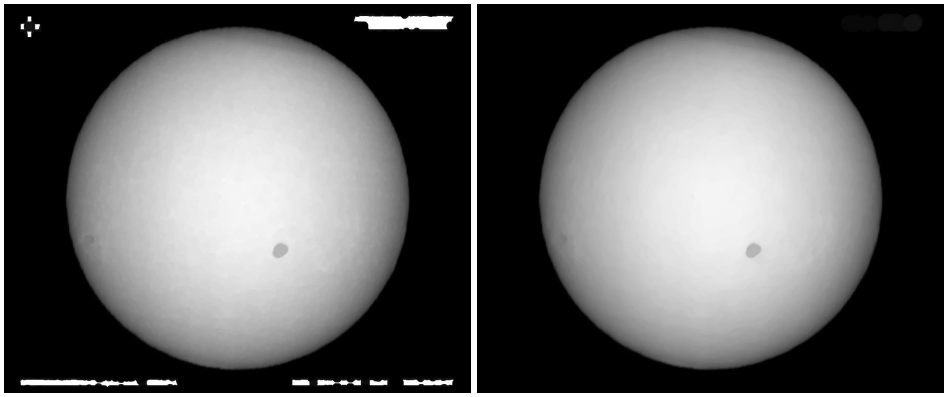
The y-coordinate of the center, $y_c$, is calculated analogously, by the formula that follows,

$$y_c = \frac{1}{1200 * 0.4} \sum_{i=360}^{840} \frac{y_{first_i} + y_{last_i}}{2}.$$

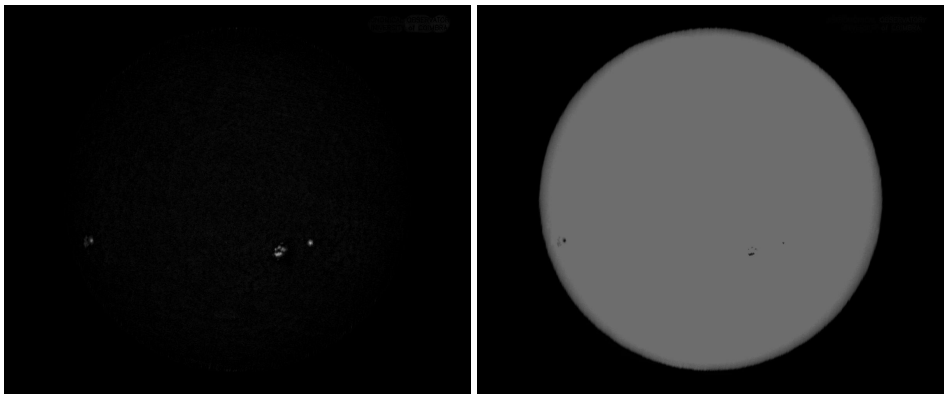
Determining the center of the solar disk allows to calculate the radius, which is given by the average between the distances of every pixel used in the calculation of the center, and the center itself. Having the center and radius of Figure 5.11f determined, the solar disk of Figure 5.11f was transformed into a perfect circle. The result is shown in Figure 5.11g. Finally, this image is multiplied by the original one and the result is shown in Figure 5.11h. This last image is the final result of the preprocessing, with the original digital levels inside the solar disk, and with digital levels equal to zero (corresponding to the black color) outside of it.

### 5.4.2 Morphological detection of sunspots

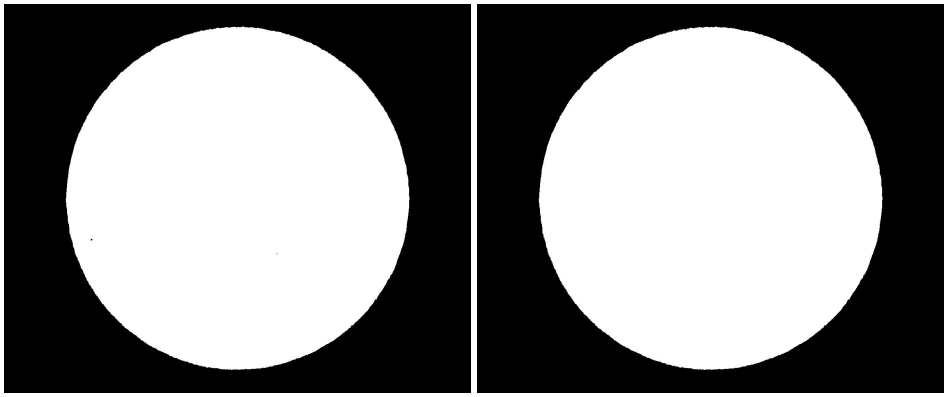
The main goal of the algorithm, is the automatic detection of sunspots in Coimbra's spectroheliograms. The initial image fed into the algorithm is the final image of the pre-processing (Figure 5.11h). To enhance the (possible) sunspots on the image, a black top-hat is applied once it allows extracts the small elements and details seen in the image. This transform consists of the difference between the closing by a disk of size 20 of the initial image (Figure 5.11h), and that very same initial image. Then, the image resulted from the black top-hat (Figure 5.12a) is used to extract the contour of sunspots. For that, an adaptive threshold is applied with limits 20 and 255, originating the image shown in Figure 5.12b. As can be seen, not only the sunspots were identified but also a lot of noise. An erosion by a



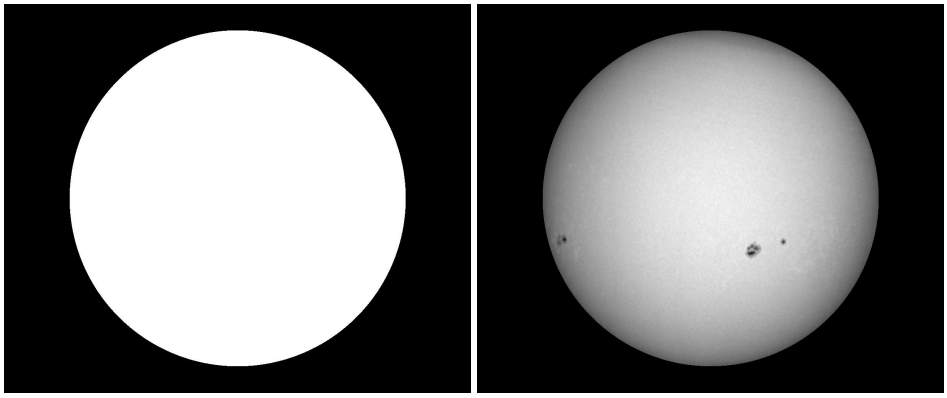
(a) Closing. (b) Opening.



(c) Subtraction. (d) Reconstruction.



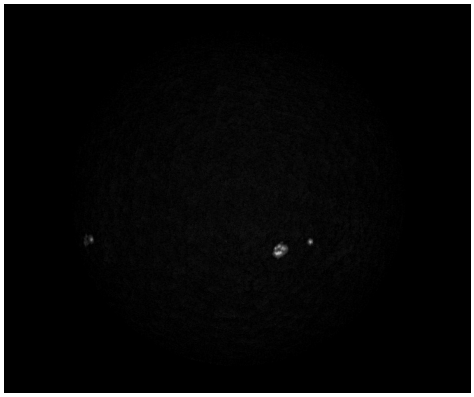
(e) Adaptive threshold. (f) Filling holes.



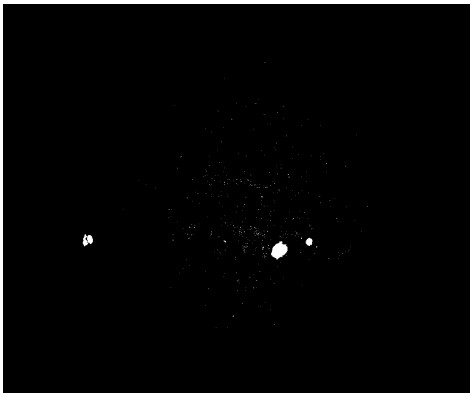
(g) Solar disk as a perfect circle. (h) Multiplication.

Fig. 5.11 Preprocessing.

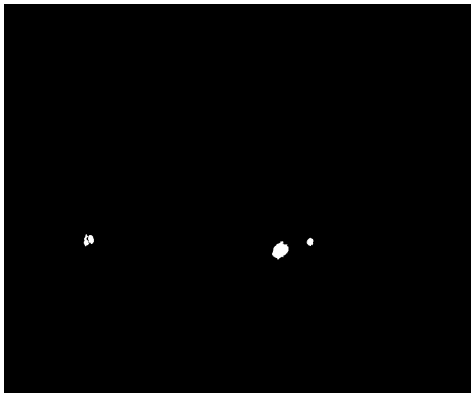
disk of diameter 1 is applied, which eliminates noisy specs. Since that also erodes real features, an application of reconstruction is required to recover their original shape. The result is shown in Figure 5.12c. The sunspots are now correctly identified, and the next step is the extraction of its contours which is done through the morphological gradient operation (Figure 5.12d) followed by a thinning operation (Figure 5.12e). While gradient enhances the contours of the sunspots, the thinning allows to reduce that contour to one pixel only, preserving the relationship between structures and holes. The final result of this stage of the algorithm is shown in Figure 5.12f, where the sunspots' contours were superimposed over the original image.



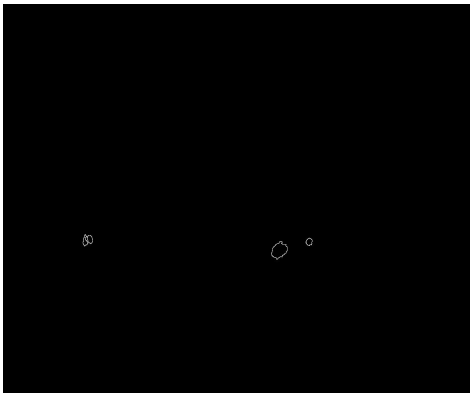
(a) Black top-hat.



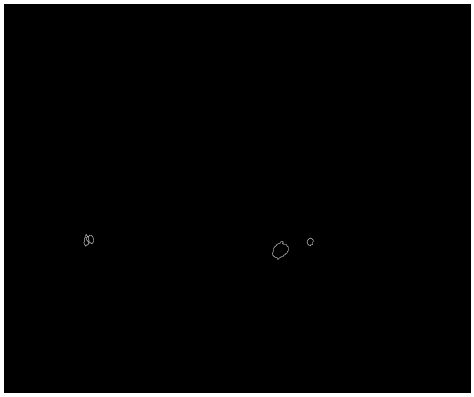
(b) Adaptive threshold.



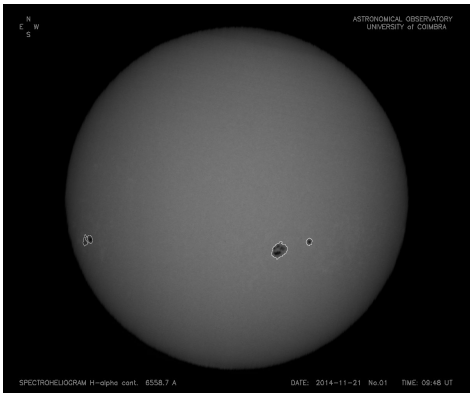
(c) Reconstruction.



(d) Morphological gradient.



(e) Sunspots contours.




(f) Sunspots contours over Figure 5.10.

Fig. 5.12 Sunspots detection step by step.

### 5.4.3 Umbra-Penumbra Segmentation

Another aim of this work is the umbra-penumbra segmentation. Sunspots are, generally, constituted by umbra and penumbra, which leads to a bimodal distribution of grey levels within each spot (as an example see Figure 5.13). Nevertheless, there are sunspots constituted only by umbra (the so called pores) and, in this case, it is assumed that the distribution of grey levels is unimodal. For this reason, a representative set of sunspots was selected, and their histograms analyzed. For all practical purposes, if the difference between the maximum and minimum grey levels of the histogram is greater than 20, then the sunspot is considered to have a bimodal distribution, else the distribution should be considered unimodal. Furthermore, concerning the bimodal distributions, a threshold value $t$ is automatically estimated between the two peaks of each histogram to segment umbra and penumbra, as follows

$$t = \frac{sunspot_{max} - sunspot_{min}}{2} + sunspot_{min}, \tag{5.1}$$

where $sunspot_{max}$ and $sunspot_{min}$ are, respectively, the maximum and the minimum values of the grey level inside the sunspot.

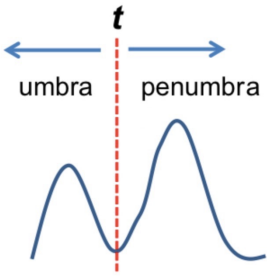
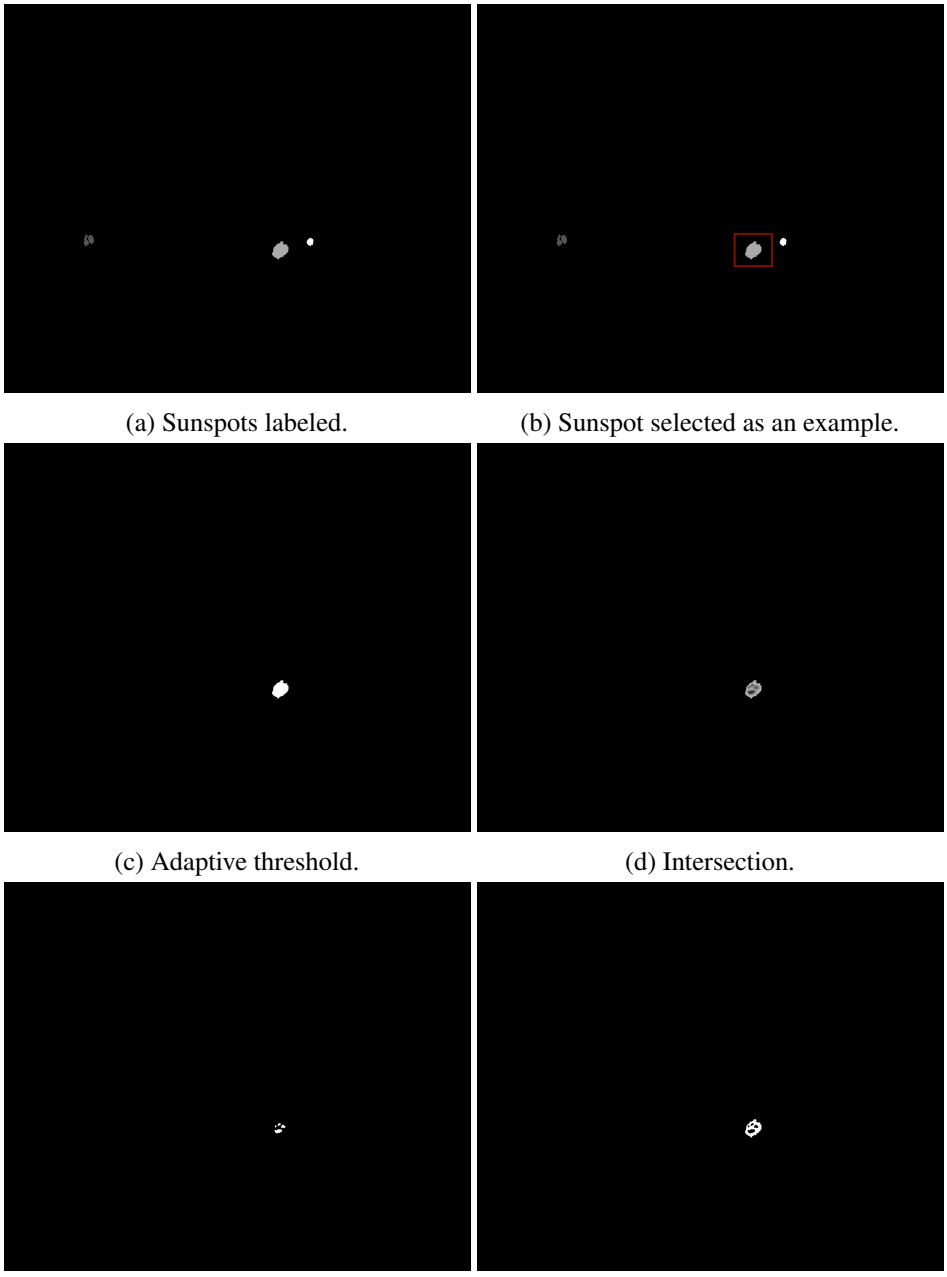



Fig. 5.13 Example of a sunspot histogram (grey level *versus* count of pixels in the image) of a bimodal distribution.

The stage of umbra and penumbra segmentation then starts by labeling each sunspot so that each one could be treated separately (Figure 5.14a). A sunspot in the image was chosen, as an example, to explain the segmentation implementation in detail. It is highlighted with a square around in Figure 5.14b. To get the original grey levels of this sunspot, two operations are necessary: first, to isolate the sunspot, an adaptive threshold is applied to the image in Figure 5.14a using the label number as limits, which results in the binary image represented in Figure 5.14c. After that, an intersection between that image and the original image is carried out resulting the image in Figure 5.14d. Hereupon, the $sunspot_{max}$ and $sunspot_{min}$ are computed in order to determine the type of grey-scale distribution. In the case of the sunspot chosen as an example, the distribution was bimodal and therefore the value $t$ was estimated following equation 5.1. However, in the case of unimodal distribution, $t$ was assumed to be 2. Thereafter, two adaptive thresholds are performed: the first one with limits 1 and $t-1$, to segment the umbra, and the second one with limits $t$ and 255 to segment the penumbra. The results of these operations are shown in Figure 5.14e and Figure 5.14f, respectively. This stage ends with the creation of four images: the first one resulting of the union of all the segmented umbrae (Figure 5.15a), the second one resulting of the union of all the segmented penumbrae (Figure 5.15b), the third one resulting of the sum of all umbrae and penumbrae together (Figure 5.15c), and the last one

resulting of the composite of the gradients of the umbrae and the penumbrae, superimposed over the original image (Figure 5.15d).



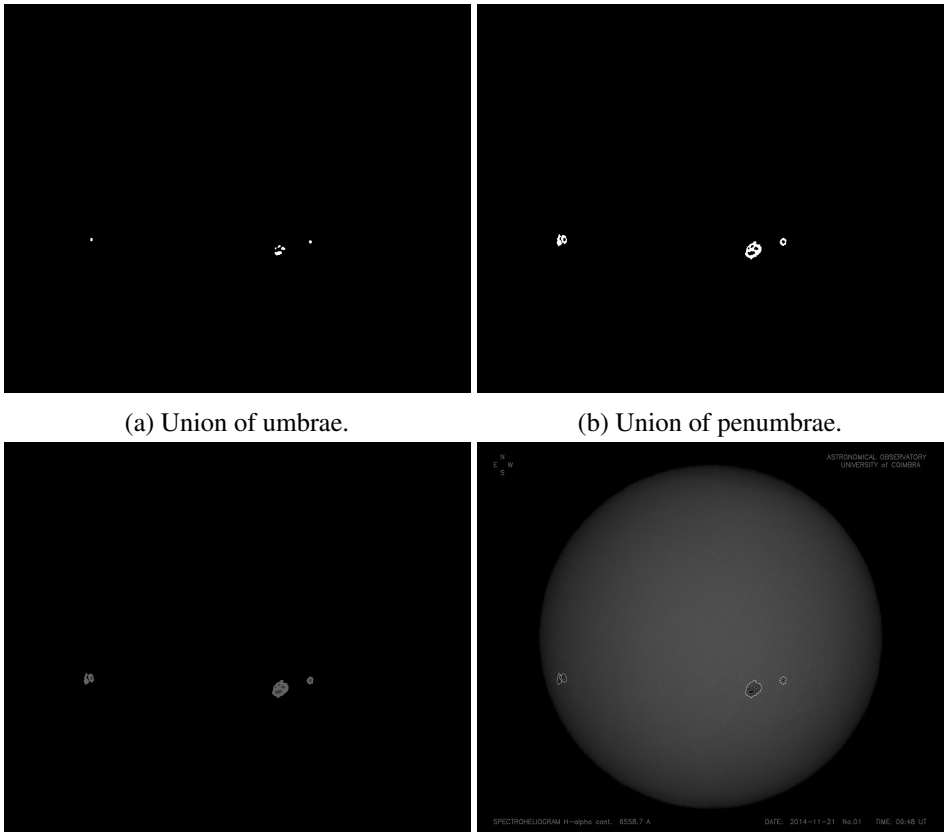
(a) Sunspots labeled. (b) Sunspot selected as an example.

(c) Adaptive threshold. (d) Intersection.

(e) Umbra segmentation. (f) Penumbra segmentation.

Fig. 5.14 Segmentation step by step.

A zoom to one of the sunspots in Figure 5.10, and the respective output of the detection and segmentation algorithm, is shown in the Figure5.16. Note that the sunspot in this example is very complex once it presents four different umbrae all surrounded by the same patch of penumbra. A possible scenario for the formation of a sunspot of this type is the branching of a magnetic flux tube

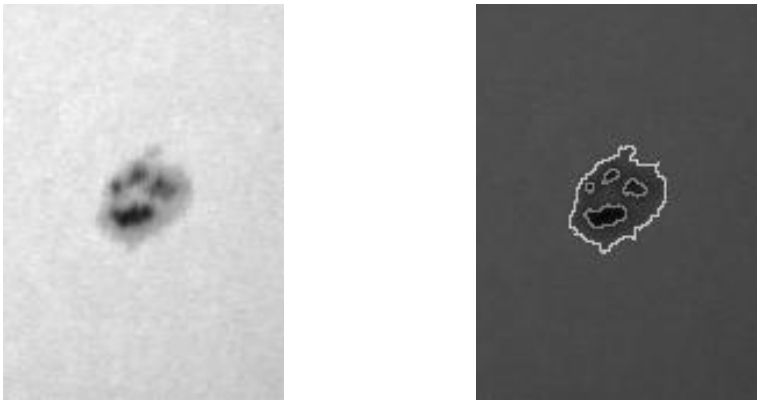



(a) Union of umbrae. (b) Union of penumbrae.

(c) Sum of all umbrae and penumbrae. (d) umbrae and penumbrae contours.

Fig. 5.15 Results of the segmentation umbra-penumbra.

which creates nearby sunspots. The penumbrae of these sunspots end up overlapping, giving rise to a single structure.



(a) Sunspot in the original image. (b) Segmentation obtained for Figure 5.16a.

Fig. 5.16 Zoomed segmented sunspot.

## 5.5 Analysis and discussion of results

The automatic algorithm developed to detect and segment sunspots was applied to a representative set of 144 Coimbra's spectroheliograms of the solar cycle 24. These images were compared with the correspondent ground-truth images built by a solar observer expert. By ground-truth images is understood images with the information provided by direct observation (empirical evidence). To evaluate the performance of the method, two distinct evaluations stages were considered: one for sunspots detection quality, and another for the ability to separate umbra from penumbra. Note that the resulting images presented in this section as examples are with reduced contrast for better visualization.

In what concerns the evaluation of sunspots detection, the metrics used rely on pixel-based comparisons. Therefore, each pixel of each output image need to be identified as a sunspot or non-sunspot pixel, considering it positive or negative respectively. Depending on this identification and the actual one (information taken from the ground-truth image), the pixel is classified according to the confusion matrix shown in Table 5.1. True positive means that a pixel in the output image detected as part of a sunspot is a sunspot pixel in the ground-truth image. By true negative is understood a pixel that is detected as not being part of a sunspot, and in the ground-truth image it also not belongs to a sunspot. False positive means a pixel detected as part of a sunspot that is not a sunspot pixel in the ground-truth image. A false negative is a pixel not detected as belonging to a sunspot but actually is a sunspot pixel in the ground-truth image.

Table 5.1 Confusion matrix.

| | **Actual Positive** | **Actual Negative** |
|---|---|---|
| **Predicted Positive** | True Positive | False Positive |
| **Predicted Negative** | False Negative | True Negative |

The metrics chosen to evaluate the performance of the sunspots detection are the following,

$$precision = \frac{TP}{TP+FP},$$

$$recall = \frac{TP}{TP+FN},$$

$$f-score = \frac{2*(Precision*Recall)}{Precision+Recall},$$

where $TP$ is the number of true positives, $TN$ the number of true negatives, $FP$ the number of false positives, and $FN$ the number of false negatives. The *precision* measures the proportion of pixels that are actually positives out of all the pixels that are detected as positives. The *recall*, also known as sensitivity, or true positive rate ($TPR$), gives the information about the proportion of pixels that are detected as positives and are actually positives relatively to the universe of all pixels that are, in fact, positives. While the *precision* allows to evaluate the cost of having false positives in large number, the *recall* allows to select which is the best model when the number false negatives is high. The $f-score$ represents a trade-off between the two previous metrics. This set of metrics was chosen because it is universally used in evaluation of binary detection algorithms in most diverse areas, for example in [46, 49, 70].

Concerning the ability to differentiate umbra from penumbra, the overall accuracy ($OA$), which is the ratio of pixels that were correctly classified to all the classified pixels, was the metric chosen and is given by,

$$OA(\%) = \frac{UU+PP}{UU+UP+PU+PP}*100,$$

where $UU$ is the number of umbra pixels that are detected as umbra; $PP$ is the number of penumbra pixels that are detected as penumbra; $UP$ is the number of umbra pixels detected as penumbra pixels, and $PU$ is the number of penumbra pixels detected as umbra pixels. The values obtained for the metrics are summarized in Table 5.2.

The automatic method developed presents satisfactory results for most of the images as can be verified analyzing the Table 5.2. A comparison of developed algorithm in the scope of this thesis

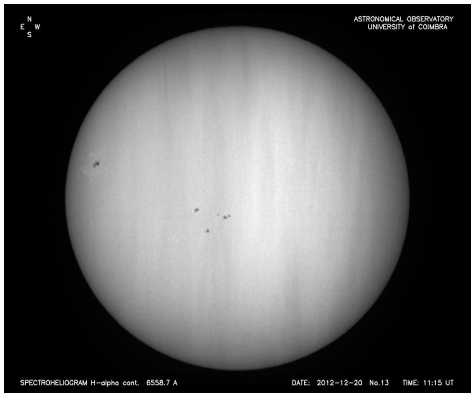
Table 5.2 Evaluation of the performance of the algorithm.

| $precision(\%)$ | $recall(\%)$ | $f-score(\%)$ | $OA(\%)$ |
|---|---|---|---|
| 81.33 | 79.42 | 78.98 | 86.25 |

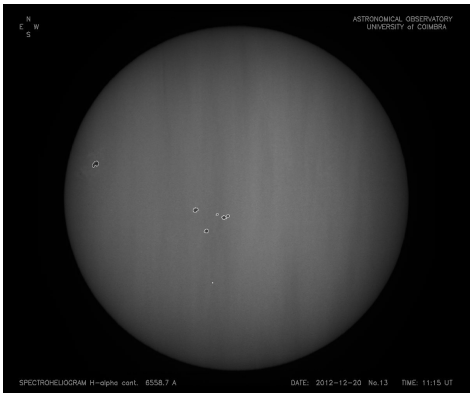
with another algorithm, based on pixel intensities, was done in [15]. This study demonstrate a better efficiency of the method based on MM.

Following are listed some artefacts that may appear in spectroheliograms which may present additional difficulties for automatic detection algorithms. Examples of outputs of the algorithm developed for spectroheliograms with these same artefacts are also presented.

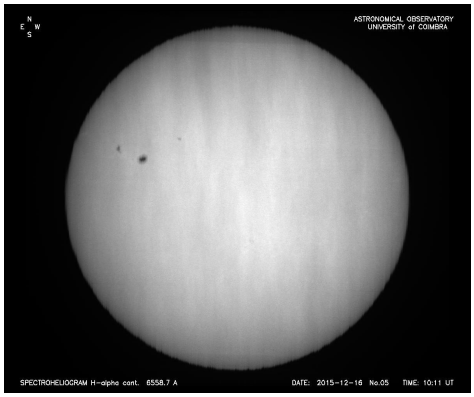
Due the Earth's atmosphere and meteorological factors, applying automatic detection methods to ground-based images present some specific hindrances. Despite this, the good performance of the algorithm developed in the scope of this thesis is essentially kept when applying to most of the images with strong atmospheric effects. Examples can be seen in Figure 5.17.



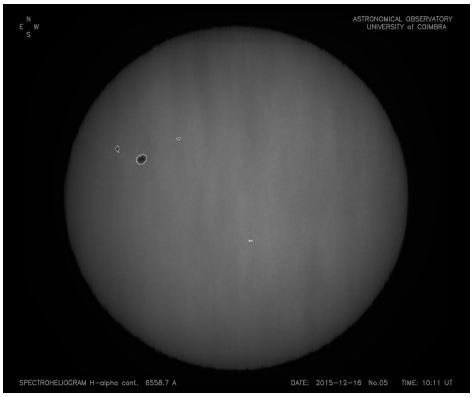

(a) Image of 20/12/2012.




(b) Output of the image of 20/12/2012.



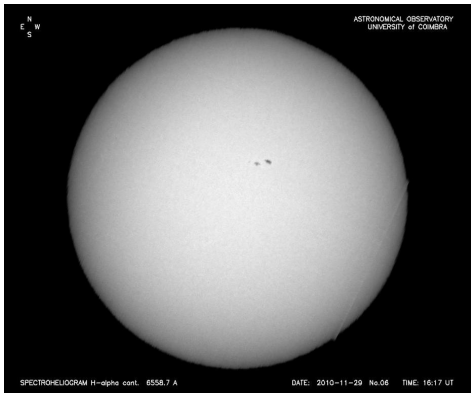

(c) Image of 16/12/2015.



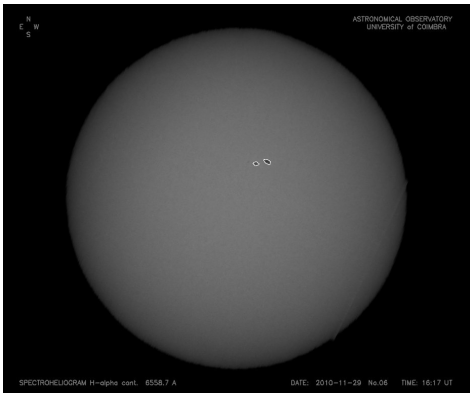

(d) Output of the image of 16/12/2015.

Fig. 5.17 Examples of detection results on spectroheliograms with clouds.

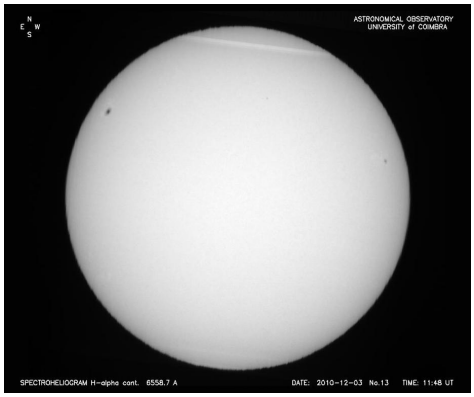
Problems during the image acquisition process is another difficulty that can appear in some spectroheliograms. Although in these cases the solar disk may present some deformations, the method performs well, as shown in Figure 5.18.



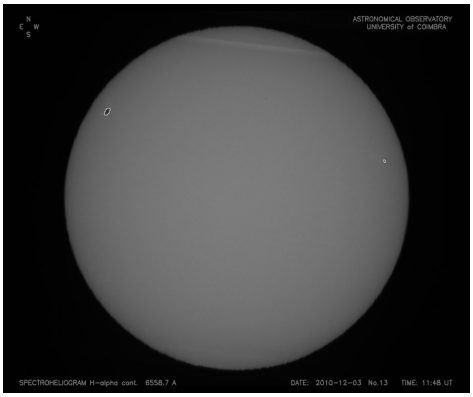

(a) Image of 29/11/2010.




(b) Output of the image of 29/11/2010.




(c) Image of 3/12/2010.



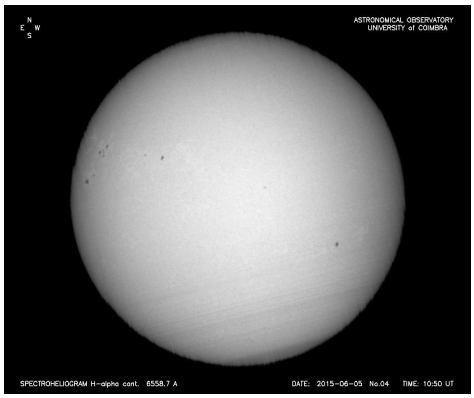

(d) Output of the image of 3/12/2010.

Fig. 5.18 Examples of detection results on spectroheliograms with acquisition errors.

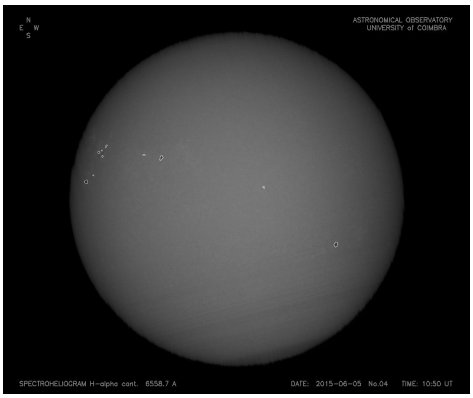
In the case of spectroheliograms with dust trapped in the slit (represented by almost horizontal lines in the image), which could not be removed or corrected by the image acquisition software, the method proved to be efficient, as can be seen in Figure 5.19.

Most methods of detecting solar phenomena take into account, in a pre-processing phase, the removing of limb-darkening so that sunspots on the solar disk's limb can be detected. The limb-darkening is an optical effect typical of starts, namely the sun, which is characterized by the gradual Since, in the essence of MM, the most important aspect is the geometric shape of the objects, the developed algorithm does not need to correct limb-darkening, keeping its efficiency. Results of the application of the method, in images with sunspots in the limb, are shown in Figure 5.20.

In fact, the method proves to be efficient in dealing with the artifacts that may appear in the spectroheliograms. However, there are cases in which the method is not efficient, presenting an over detection. This problem occurs due to the size of the structuring element in the top-hat transform. The value used was chosen in order to have a trade-off among all the images of the set considered, in order to reduce false positives. As future work is intended to solve these cases through a post processing phase. However, given the slow rotation of the sun and the slow variation of the sunspots, the analysis

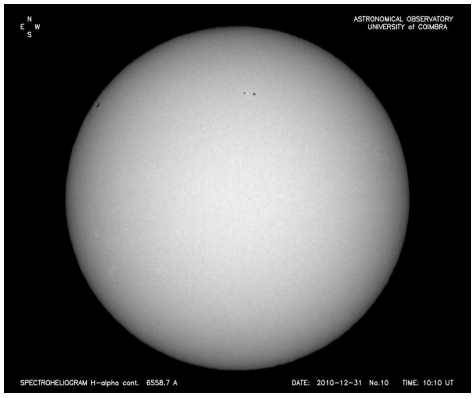



(a) Image of 5/6/2015.

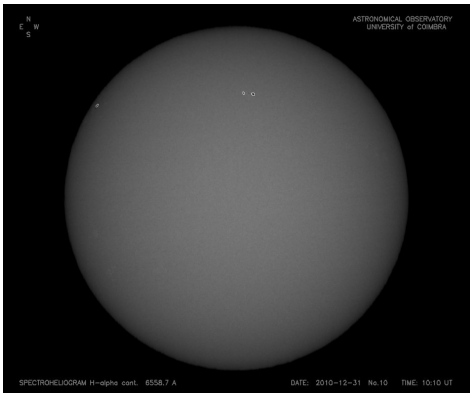



(b) Output of the image of 5/6/2015.

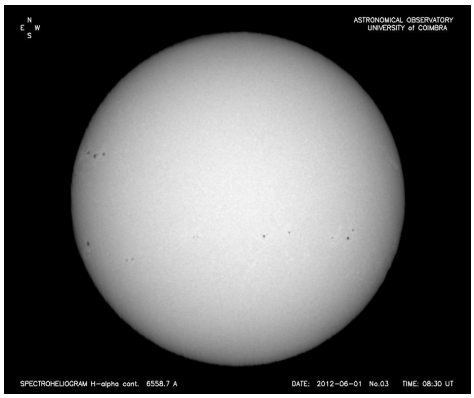
Fig. 5.19 Example of detection results on spectroheliograms with dust.



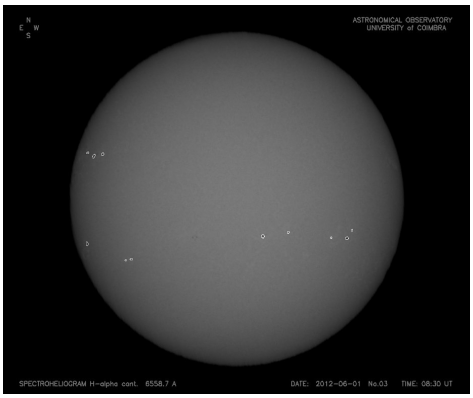

(a) Image of 31/12/2010.




(b) Output of the image of 31/12/2010.



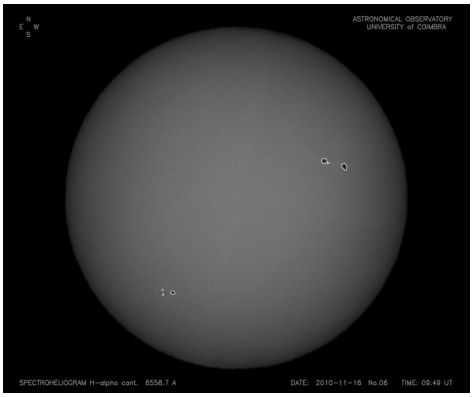

(c) Image of 3/12/2010.



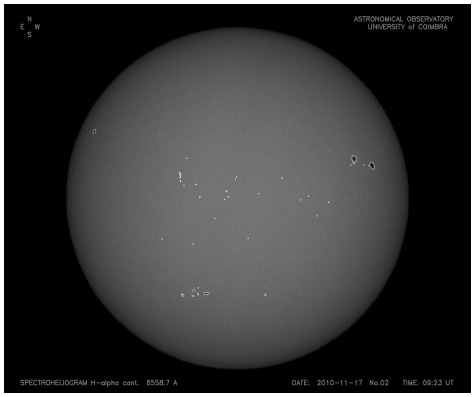

(d) Output of the image of 3/12/2010.

Fig. 5.20 Examples of detection results on spectroheliograms with sunspots in the limb.

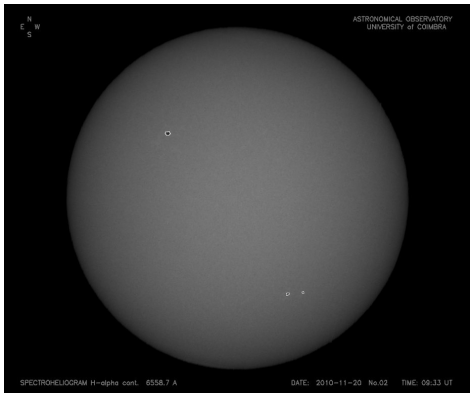
of images acquired on consecutive days, or almost consecutive, is a good way to evaluate the results, as can be seen in Figure 5.21.




(a) Output of the image of 16/11/2010.




(b) Output of the image of 17/11/2010 (over detection).




(c) Output of the image of 20/11/20106.

Fig. 5.21 Example of over detection and the results of the consecutive days.

## 5.6 Conclusions

The aim of this chapter was the detailed description of an automatic method which was developed to detect sunspots, and its separation into umbra and penumbra, in the spectroheliograms of the Geophysical and Astronomical Observatory of the University of Coimbra. The method developed was based on since it is a versatile technique which was already used with success in processing of several types of images, namely solar images. An analysis of the performance of the method was carried out, and the metrics used to evaluate it showed the great efficiency of it. As a future work it is intended to process the entire database of Geophysical and Astronomical Observatory of the University of Coimbra. Thus, it will be possible to evaluate the method's ability to deal with images obtained through other acquisition systems as well as in other formats. By the other hand, it will allow resuming the production of the Coimbra Observatory solar catalogs, which was interrupted in 1986.

# Chapter 6

# Conclusions

This final chapter starts recalling the main motivation of the present work: the computation of the magnetic field in the umbra- the darkest part of a sunspot. This motivation was the driven force for the work developed within this thesis: the study of numerical methods that can be used to simulate the magnetic field.

In Chapter 2, *Stability and convergence analysis for IBVP with Dirichlet boundary conditions*, this work started with the design and analysis of numerical methods for nonlinear parabolic IBVP and the corresponding nonlinear elliptic BVP defined in a bounded interval and Dirichlet boundary conditions. The convergence analysis is presented following two different approaches depending on the smoothness of the solution of the differential problem. For smooth solutions, a second-order error estimate is obtained, taking into account the particular structure of the spatial truncation error, while for nonsmooth solutions, the error estimates were obtained applying the Bramble-Hilbert lemma. In what concerns the stability, in a first attempt, it was observed the need for the uniform boundness of the numerical solution around which is intended to conclude the stability property. This requirement was then concluded from the established error estimates. And then the stability was easily followed. As the numerical methods proposed can be seen simultaneously as FDMs and piecewise linear FEMs, the obtained results can be seen simultaneously as supraconvergent and superconvergent results.

The extension of the results presented in Chapter 2 to the IBVP defined in two-dimensional domains was the focus of Chapter 3, *An application to sunspots*. This extension was also driven by the application to the magnetic field in the umbra, where vertical magnetic fields are observed. Based on this fact, the initial motivation was then replaced by a simplified one - the numerical simulation of two of the magnetic field components being one of them, the vertical one. The initial system of 6 nonlinear parabolic equations was then replaced by a system of 4 equations, which was considered in this chapter. For a general system of nonlinear parabolic equations defined in a square, some of the results presented in Chapter 2 were rewritten for this new situation when the differential problem has smooth solutions. A numerical simulation of the studied numerical tool applied to the magnetic field (2 components) in the umbra was developed within this chapter, assuming that the velocity field is known in the space-time domain, the magnetic field is also known on the boundary of the space-time domain, as well as the initial condition. These inputs were obtained from the numerical simulation of the magnetic and velocity fields obtained with Bifrost code in a network scenario [14, 34]. This choice is based on the fact that the qualitative behaviour of the magnetic field is similar in these two

structures. A pathologic behaviour was observed in the numerical simulations, and to correct it, an artificial diffusion term was added to the method. The results obtained with the new method and the reference solution present the same qualitative behaviour at least for short times.

The numerical simulations presented in Chapter 3 were obtained considering Dirichlet boundary conditions. As such boundary conditions are not realistic in the context of the magnetic field evolution, numerical methods for IBVP with Neumann or Robin boundary conditions should be designed and studied. Chapter 4, *Convergence analysis for IBVP with Neumann boundary conditions*, is centered in the construction and convergence study of numerical methods for nonlinear parabolic IBVP with Neumann boundary conditions. Here semi-discrete approximations and fully discrete in space and time implicit-explicit approximations were analyzed, and their convergence was concluded. These results were proved, taking into account the truncation error structure and assuming that the IBVP has smooth solutions.

This thesis is concluded with the study of the automatic detection and geometric definition of sunspots, including the limits of umbra and penumbra, in solar images, done in Chapter 5, *Detection of sunspots in spectroheliograms*. Motivated by the accurate definition of the different zones of the sunspots, this research topic intends to contribute to the identification of computational domains used in the simulation of the magnetic field's behaviour in different zones of the sunspots.

To conclude, it was noticed that there are several questions that were not answered in the present work and deserve to be object of study:

- The extension of the convergence results presented in Chapter 3 for problems with nonsmooth solutions via Bramble-Hilbert lemma approach.
- The stability and convergence of implicit-explicit methods for systems of nonlinear parabolic equations with Dirichlet boundary conditions for the IBVP analized in Chapter 2 and 3.
- The stabilized methods obtained from the method studied in Chapter 3 corrected with more convenient stabilization techniques taking into account that these methods can be seen as piecewise linear finite element methods (see for instance [40, 56, 64]).
- The stability and convergence of the methods presented in Chapter 4 for nonlinear IBVP with nonsmooth solutions via Bramble-Hilbert lemma in space and time.
- The extension of the previous results for IBVPs defined in spatial domains of higher dimensions.
- Numerical simulation of the solar magnetic field and plasma velocity field in different zones of the sunspots.

# References


[1] Altschuler, M. D., Nakagawa, Y., and Lilliequist, C. G. (1968). Concerning the origin of evershed motion in sunspots. *Solar Physics*, 3(3):466–481.

[2] Barata, M. T. (2001). *Classificação do Coberto Florestal em Imagens de Detecção Remota através de uma Metodologia Baseada na Morfologia Matemática*. PhD thesis, Technical University of Lisboa.

[3] Barata, M. T., Lopes, F. C., Pina, P., Alves, E. I., and Saraiva, J. (2015). Automatic detection of wrinkle ridges in venusian magellan imagery. *Geological Society, London, Special Publications*, 401(1):357–376.

[4] Barata, T., Carvalho, S., Dorotovič, I., Pinheiro, F., Garcia, A., Fernandes, J., and Lourenço, A. (2018). Software tool for automatic detection of solar plages in the coimbra observatory spectroheliograms. *Astronomy and Computing*, 24:70 – 83.

[5] Barbeiro, S., Bardeji, S., Ferreira, J., and Pinto, L. (2018). Non-Fickian diffusion coupled models in porous media. *Numerische Mathematik*, 138:869–904.

[6] Barbeiro, S., Ferreira, J., and Grigorieff, R. (2005). Supraconvergence of a finite difference scheme for solutions in $H^s(0,L)$. *IMA Journal of Numerical Analysis*, 25:797–811.

[7] Barbeiro, S., Ferreira, J., and Pinto, L. (2011). $H^1$- second order convergent estimates for non-Fickian models. *Applied Numerical Mathematics*, 61:201–215.

[8] Bellot Rubio, L. and Orozco Suárez, D. (2019). Quiet Sun magnetic fields: an observational view. *Living Reviews in Solar Physics*, 16(1):1.

[9] Biskamp, D. (1997). *Nonlinear Magnetohydrodynamics*. Cambridge Monographs on Plasma Physics. Cambridge University Press.

[10] Borrero, J. and Ichimoto, K. (2011). Magnetic structure of sunspots. *Living Reviews in Solar Physics*, 1(8):4.

[11] Bramble, J. and Hilbert, S. (1970). Estimation of linear functionals on Sobolev spaces with application to fourier transforms and spline interpolation. *SIAM Journal on Numerical Analysis*, 7:112–124.

[12] Brandenburg, A. and Dobler, W. (2010). Pencil code: Finite-difference code for compressible hydrodynamic flows. *Astrophysics Source Code Library*, pages 10060–.

[13] Brooks, A. N. and Hughes, T. J. R. (1982). Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations. *omput. Methods Appl. Mech. Engrg.*, 32:199–259.

[14] Carlsson, Mats, Hansteen, Viggo H., Gudiksen, Boris V., Leenaarts, Jorrit, and De Pontieu, Bart (2016). A publicly available simulation of an enhanced network region of the sun. *A&A*, 585:A4.

[15] Carvalho, S., Gomes, S., Barata, T., Lourenço, A., and Peixinho, N. (2020). Comparison of automatic methods to detect sunspots in the coimbra observatory spectroheliograms. *Astronomy and Computing*, 32:100385.

[16] Carvalho, S., Pina, P., Barata, T., Gafeira, R., and Garcia, A. (2016). Ground-based Observations of Sunspots from the Observatory of Coimbra: Evaluation of Different Automated Approaches to Analyse its Datasets. In Dorotovic, I., Fischer, C. E., and Temmer, M., editors, *Coimbra Solar Physics Meeting: Ground-based Solar Observations in the Space Instrumentation Era*, volume 504 of *Astronomical Society of the Pacific Conference Series*, page 125.

[17] Cid, Consuelo, Palacios, Judith, Saiz, Elena, Guerrero, Antonio, and Cerrato, Yolanda (2014). On extreme geomagnetic storms. *J. Space Weather Space Clim.*, 4:A28.

[18] Colak, T. and Qahwaji, R. (2008). Automated mcintosh-based classification of sunspot groups using mdi images. *Solar Physics*, 248.

[19] Dasgupta, U., Singh, S., and Jewalikar, V. (2011). Sunspot number calculation using clustering. In *2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, pages 171–174.

[20] Davidson, P. A. (2001). *An Introduction to Magnetohydrodynamics*. Cambridge Texts in Applied Mathematics. Cambridge University Press.

[21] Deepa, V., Beaulah, H., and Shanmugapriya, P. (2016). Evaluating the performance of automatic sunspots detection algorithms using full-disk solar images. *IJSMC*, 5:199–205.

[22] Dorotovič, I., Shahamatnia, E., Lorenc, M., Rybanský, M., Ribeiro, R., and Fonseca, J. (2014). Sunspots and coronal bright points tracking using a hybrid algorithm of pso and active contour model. *Sun and Geosphere*, 9:81–84.

[23] Ferreira, J. and Grigorieff, R. (2006). Supraconvergence and supercloseness of a scheme for elliptic equations on non-uniform grids. *Numerical Functional Analysis and Optimization*, 27:539–564.

[24] Ferreira, J., Jordão, D., and Pinto, L. (2017). Second order approximations for kinetic and potential energies in Maxwell's wave equations. *Applied Numerical Mathematics*, 120:125–140.

[25] Ferreira, J., Jordão, D., and Pinto, L. (2018). Approximating coupled hyperbolic-parabolic systems arising in enhanced drug delivery. *Computers & Mathematics with Applications*, 76:81–97.

[26] Ferreira, J., Oliveira, P., and Silveira, E. (2020). Drug release enhanced by temperature: An accurate discrete model for solutions in $H^3$. *Computers & Mathematics with Applications*, 79:852–875.

[27] Ferreira, J. and Pinto, L. (2013). Supraconvergence and supercloseness in quasilinear coupled problems. *Journal of Computational and Applied Mathematics*, 252:120–131.

[28] Ferreira, J. A. (1994). *Métodos Adaptativos para Problemas Parabólicos: Estudo da Convergência*. PhD thesis, University of Coimbra.

[29] Fonte, C. C. and Fernandes, J. (2009). Application of fuzzy sets to the determination of sunspot areas. *Solar Physics*, 260(1):21–41.

[30] Gafeira, R., Fonte, C. C., Pais, M. A., and Fernandes, J. (2014). Temporal evolution of sunspot areas and estimation of related plasma flows. *Solar Physics*, 289(5):1531–1542.

[31] Garcia, A., Sobotka, M., Klvaňa, M., and Bumba, V. (2011). Synoptic observations with the coimbra spectroheliograph. *Contributions of the Astronomical Observatory Skalnate Pleso*, 41:69–72.

[32] Gary, G. A. (2001). Plasma beta above a solar active region: Rethinking the paradigm. *Solar Physics*, 203(1):71–86.

[33] Gastón-Romeo, M., Leon, T., Mallor, F., and Ramírez-Santigosa, L. (2011). A morphological clustering method for daily solar radiation curves. *Solar Energy*, 85(9):1824 – 1836.

[34] Gudiksen, B. V., Carlsson, M., Hansteen, V. H., Hayek, W., Leenaarts, J., and Martínez-Sykora, J. (2011). The stellar atmosphere simulation code bifrost - code description and validation. *A&A*, 531:A154.

[35] Hale, G. E. (1908). On the Probable Existence of a Magnetic Field in Sun-Spots. *The Astrophysical Journal*, 28:315.

[36] Hapgood, M. (2017). *Space Weather*. 2399-2891. IOP Publishing.

[37] Hathaway, D. H. (2015). The solar cycle. *Living Reviews in Solar Physics*, 12(1):4.

[38] J. J. Curto, M. B. and Martínez, E. (2008). Automatic sunspots detection on full-disk solar images using mathematical morphology. *Solar Physics*, 250:411–429.

[39] Jewalikar, V. and Singh, S. (2010). Automated sunspot extraction, analysis and classification. In *International Conference on Image and Video Processing and Computer Vision (IVPCV-10)*.

[40] John, V., Knobloch, P., and Novo, J. (2018). Finite elements for scalar convection-dominated equations and incompressible flow problems: a never ending story? *Computing and Visualization in Science*, 19(5):47–63.

[41] Jordão, D. (2020). *Coupling hyperbolic and parabolic IBVP: applications to drug delivery*. PhD thesis, University of Coimbra.

[42] Krizek, M. and Neittaanmäki, P. (1984). Superconvergence phenomenon in the finite element method arising from averaging gradients. *Numerische Mathematik*, 45:105–116.

[43] Lax, P. D. and Richtmyer, R. D. (1956). Survey of the stability of linear finite difference equations. *Communications on Pure and Applied Mathematics*, 9(2):267–293.

[44] Levine, N. (1985). Superconvergent Recovery of the Gradient from Piecewise Linear Finite-element Approximations. *IMA Journal of Numerical Analysis*, 5(4):407–427.

[45] Lin, G. H., Wang, X. F., Liu, S., Yang, X., Zhu, G. F., Deng, Y. Y., Ji, H. S., Zhou, T. H., Sun, L. N., Feng, Y. L., Liu, Z. Z., Tao, J. P., Ben, M. X., Lin, J., Ding, M. D., Li, Z., Zheng, S., Zeng, S. G., He, H. L., Zeng, X. Y., Shu, Y., and Sun, X. B. (2019). Chinese sunspot drawings and their digitization – (i) parameter archives. *Solar Physics*, 294.

[46] Liu, S., Zhang, L., Zhang, Z., Wang, C., and Xiao, B. (2015). Automatic cloud detection for all-sky images using superpixel segmentation. *IEEE Geosci. Remote Sensing Lett.*, 12(2):354–358.

[47] Lopez-Marcos, J. and Sanz-Serna, J. (1988). Stability and convergence in numerical analysis iii: linear investigation of nonlinear stability. *IMA Journal of Numerical Analysis*, 8:71–84.

[48] Lourenço, A., Carvalho, S., Barata, T., Garcia, A., Carrasco, V., and Peixinho, N. (01 Jan. 2019). Solar observations at the coimbra astronomical observatory. *Open Astronomy*, 28(1):165 – 179.

[49] Lu, H., Gargesha, M., Wang, Z., Chamié, D., Attizzani, G. F., Kanaya, T., Ray, S., Costa, M. A., Rollins, A. M., Bezerra, H., and Wilson, D. L. (2012). Automatic stent detection in intravascular oct images using bagged decision trees. *Biomedical optics express*, 3:2809–24.

[50] Manish, T. I., Murugan, D., and Kumar, T. G. (2014). Automatic detection of sunspot activities using advanced detection model. *IOSR Journal of Computer Engineering*, 16:83–87.

[51] Mao, X., Zhang, B., Deng, H., Zou, Y., and Chen, J. (2016). Three-dimensional morphological analysis method for geologic bodies and its parallel implementation. *Computers & Geosciences*, 96:11 – 22.

[52] Matheron, G. (1967). *Éléments pour une théorie des milieus poreux*. Masson, Paris.

[53] Matheron, G. (1975). *Random sets and integral geometry*. Wiley series in probability and mathematical statistics. John Wiley & Sons Inc.

[54] Mehena, J. (2013). Medical images edge detection based on mathematical morphology. *International Journal of Computer and Communication Technology*, 4.

[55] Mohammed, E. and Akbar, H. (2016). Automated identification of sunspots area size using sobel edge detection and image histogram. *American Journal of Science and Technology*, 2:329–334.

[56] Nadukandi, P. and García Espinosa, J. (2012). *Stabilized finite element methods for convection-diffusion-reaction, Helmholtz and Stokes problems*. Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE).

[57] Nguyen, T. T., Willis, C. P., Derek, J. P., and Nguyen, H. S. (2006). Ios press learning sunspot classification. *Fundam. Inform.*, 72:295–309.

[58] Opgenoorth, H., Pulkkinen, A., Buchert, S., Monstein, C., Klein, K. L., Marqué, C., and Krucker, S. (2016). Solar activity during the space weather incident of Nov 4., 2015 - Complex data and lessons learned. In *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, pages EPSC2016–12017.

[59] Parker, E. N. (1982). The Dynamics of Fibril Magnetic Fields - Part Two - the Mean Field Equations. *The Astrophysical Journal*, 256:302.

[60] Pina, P. (1999). *Caracterização, Modelação e Simulação de Estruturas com base na Morfologia Matemática*. PhD thesis, Technical University of Lisboa.

[61] Pina, P., Barata, T., and Bandeira, L. (2006). Morphological recognition of the spatial patterns of olive trees. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 845–848.

[62] Pinto, L. (2013). *Parabolic Partial Integro-Differential Equations: Superconvergence Estimates and Applications*. PhD thesis, University of Coimbra.

[63] Qahwaji, R. and Colak, T. (2006). Hybrid imaging and neural networks techniques for processing solar images. *International journal of computers and their applications*, 13:9–16.

[64] Quarteroni, A. (2009). *Numerical models for differential problems*. MS&A. Springer, Milano.

[65] Reinhardt, H.-J. (1985). *Analysis of Approximation Methods for Differential and Integral Equations*. Applied Mathematics Sciences. Springer, New York.

[66] Román, J. C. M., Escobar, R., Martínez, F., Vázquez Noguera, J. L., Legal-Ayala, H., and Pinto-Roa, D. P. (2020). Medical image enhancement with brightness and detail preserving using multiscale top-hat transform by reconstruction. *Electronic Notes in Theoretical Computer Science*, 349:69 – 80. Proceedings of CLEI 19, the XLV Latin American Computing Conference.

[67] Schrijver, C. J., Kauristie, K., Aylward, A. D., Denardini, C. M., Gibson, S. E., Glover, A., Gopalswamy, N., Grande, M., Hapgood, M., Heynderickx, D., Jakowski, N., Kalegaev, V. V., Lapenta, G., Linker, J. A., Liu, S., Mandrini, C. H., Mann, I. R., Nagatsuma, T., Nandy, D., Obara, T., Paul O'Brien, T., Onsager, T., Opgenoorth, H. J., Terkildsen, M., Valladares, C. E., and Vilmer, N. (2015). Understanding space weather to shield society: A global road map for 2015–2025 commissioned by cospar and ilws. *Advances in Space Research*, 55(12):2745 – 2807.

[68] Serra, J. (1969). Introduction à la morphologie mathématique booklet no. 3. *Cahiers du Centre de Morphologie Mathématique, Fontainebleau, France*, page 24.

[69] Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, New York.

[70] Shahamatnia, E., Dorotovič, I., Mora, A., Fonseca, J., and Ribeiro, R. (2015). *Data Inconsistency in Sunspot Detection*, pages 567–577. Springer-Verlag.

[71] Silveira, M. E. (2020). *Drug transport enhanced by temperature: mathematical analysis and numerical simulation*. PhD thesis, University of Coimbra.

[72] Soille, P. (2002). *Morphological Image Analysis - Principles and Applications*. Berlin, Springer–Verlag, 2nd edition.

[73] Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, Berlin, Heidelberg, 2 edition.

[74] Solanki, S. K. (2003). Sunspots: An overview. *The Astronomy and Astrophysics Review*, 11(2-3):153–286.

[75] Solanki, S. K., Inhester, B., and Schüssler, M. (2006). The solar magnetic field. *Reports on Progress in Physics*, 69(3):563–668.

[76] Somasundaram, K. and Kalaiselvi, T. (2011). Automatic brain extraction methods for t1 magnetic resonance images using region labeling and morphological operations. *Computers in Biology and Medicine*, 41(8):716 – 725.

[77] Stix, M. (2004). *The Sun: An Introduction*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg.

[78] Vögler, A., Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T., and Linde, T. (2005). Simulations of magneto-convection in the solar photosphere: Equations, methods and results of the muram code. *Astronomy and Astrophysics, v.429, 335-351 (2005)*, 429.

[79] Wahlbin, L. (1995). *Superconvergence in Galerkin Finite Element Methods*, volume 1605 of *Lecture Notes in Mathematics*. Springer, Berlin.

[80] Warnecke, J. (2013). *Combining Models of Coronal Mass Ejections and Solar Dynamos*. PhD thesis, Stockholm University, Faculty of Science, Department of Astronomy.

[81] Wedemeyer-Böhm, S., Lagg, A., and Nordlund, A. (2008). Coupling from the photosphere to the chromosphere and the corona. *Space Science Reviews*, 144(1-4):317–350.

[82] Yu, L., Deng, L., and Feng, S. (2014). Automated sunspot detection using morphological reconstruction and adaptive region growing techniques. In *Proceedings of the 33rd Chinese Control Conference*, pages 7168–7172.

[83] Zhao, C., Lin, G., Deng, Y., and Yang, X. (2016). Automatic recognition of sunspots in hsos full-disk solar images. *Publications of Astronomical Society of Australia (PASA)*, 33:e018.

[84] Zharkov, S., Zharkova, V., Ipson, S., and Benkhalil, A. (2005). Technique for automated recognition of sunspots on full-disk solar images. *EURASIP Journal on Advances in Signal Processing*, 2005(15):318462.

[85] Zharkova, V., Ipson, S., Benkhalil, A., and Zharkov, S. (2005). Feature recognition in solar images. *Artificial Intelligence Review*, 23(3):209–266.