Instituto Nacional de Matemática Pura e Aplicada

Optimization Methods for Locating Heteroclinic Orbits

Pedro Henrique Borges de Melo

Advisor: Dan Marchesin
Co-advisor: Grigori Chapiro - UFJF

May 3, 2016

Abstract

Assume we are given a system of ordinary differential equations $x' = f(x, p)$ depending on a parameter $p \in \mathbb{R}^{\widetilde{p}}$. In this dissertation we consider the problem of locating a parameter p and an initial condition ξ that give rise to a heteroclinic orbit. In the case that such p and ξ are not unique we can introduce a cost function in the parameter space and ask for the parameters generating heteroclinic orbits that also minimize the cost function.

We provide optimization problems to solve the mentioned issues. We also provide formulas to solve the optimization problems accurately. The advantage of approaching these problems using techniques from optimization is the conceptual simplicity of the formulations that appear combined with the robustness of the solvers available in standard packages. This approach can be easily extended to situations in which the properties of the orbits can be seen as optimums of optimization problems.

All technical difficulties are concentrated in the accurate estimation of first and second order derivatives of solutions and equilibria of ordinary differential equations. Our methods can be used if we obtain parametrizations of equilibria. The equilibria can be non-hyperbolic and the multiplicities of eigenvalues at equilibria do not matter. The dimensions of the stable and unstable spaces can change.

As an application of our main method we considered the problem of finding travelling waves in a system of partial differential equations modelling combustion studied in [13]. The corresponding heteroclinic orbit approaches a non-hyperbolic equilibrium and the ordinary differential equation has solutions very sensitive to initial conditions and parameters. The orbits are very slow in the positive direction and very fast in the negative direction. The analytical velocity of the travelling wave is found in [13]. The analytical and numerical results are shown.

Keywords: heteroclinic orbits, optimization, travelling waves

Acknowledgements

I was born in a small town in the countryside of Brazil. In my family, my mother was also the father. My grandfather, Agostinho, was generous and all smiles. My older sister carried her studies in the state capital. I always stayed with her during vacations. My mother worked in my grandfather's supermarket. She always took me there. There was a day in which I and my twin sister set fire to the supermarket.

When I got older, I realized that my mother worked much harder than other people. She said she needed to make the supermarket succeed so that she could take care of the family. When I asked her the reason she did not drink or smoke, she just said she did not want to set bad examples to her children. As she had not enough time to participate in my homework from school, she decided to hire a tutor. On the weekends she had to clean the house because my twin sister could not breathe dust. She had to wash clothes, cook, take care of the kitchen garden as well as take spinach to the neighbour with anemia.

I never saw Mrs. Emília complain or relent for a moment. She always seemed to be a strong rock. She always had that strong will that never weakened and that good advice. She always dreamed for a bright future for her children. Always fought to fix her mistakes and those of her children. Always was a person with values, sincere and solidary. People say that difficulties make simple people write their names in history and my mother did it. She is the greatest example of heroism I know.

I want to thank her for all values she taught me, for all times she said no, for all times she helped me, was patient, lovely and strong. Everything I got followed from what she taught me. Thank you Mother. I am extremely grateful for everything you have done.

I should thank Brazil's National Institute for Pure and Applied Mathematics for all courses and opportunities. I want to thank my advisor D. Marchesin for choosing a fruitful theme for my dissertation and for posing the question that led to our numerical method to search optimum parameters giving rise to heteroclinic orbits. Besides, I want to acknowledge my co-advisor G. Chapiro for his helpful comments on the structure of the present text and for posing applications with non-hyperbolic equilibria. I am also grateful to A. Mailybaev for his explanations on his method and for his availability. I want to thank Brazil's Petroleum Agency for the financial support.

Agradecimentos

Nasci em uma cidade pequena do interior de Minas Gerais, Brasil. Na minha família, minha mãe era também pai. Meu avô, Agostinho, generoso e soridente. Minha irmã mais velha estudava na capital. Sempre ia passar férias com ela. Minha mãe trabalhava no armazém do Vô Tim. Ela sempre me levava pro mercado. Teve um dia que eu e minha irmã gêmea colocamos fogo lá.

Quando, aos poucos, desenvolvi consciência, fui percebendo que minha mãe trabalhava muito mais que todas as outras pessoas. Ela dizia que tinha que fazer o mercado dar lucro e que ela dependia dele pra sustentar a família. Quando eu a perguntei porque ela não bebia ou fumava, ela dizia que era porque não queria dar mau exemplo. Como ela não tinha tempo pra estudar comigo, ela resolveu pagar uma professora particular. No final de semana ela tinha que arrumar casa porque minha irmã gêmea não podia respirar poeira. Tinha que lavar roupa, cozinhar, cuidar da horta e levar espinafre pra vizinha com anemia.

Nunca vi a Dona Emília reclamar ou desistir por um momento. Ela sempre pareceu aquela rocha forte. Sempre teve aquela vontade que não enfraquece. Aquele conselho amigo. Sempre sonhou que os filhos chegassem longe. Sempre lutou pra corrigir os defeitos dela e os dos filhos. Sempre foi aquela pessoa de valores, sincera e solidária. Dizem que as dificuldades fazem pessoas simples escreverem seus nomes na história e minha mãe escreveu o dela. Ela é o maior exemplo de heroísmo que eu tive.

Quero agradecê-la por todos os valores que ela sempre me ensinou, por todas as vezes que ela me disse não, por todas as vezes que ela me ajudou, que foi paciente, que foi amorosa e forte. Tudo que eu consegui até hoje foi devido ao que ela me ensinou. Obrigado Mãe. Sou extremamente grato por tudo o que você fez.

Agradeço ao Instituto Nacional de Matemática Pura e Aplicada por todos os cursos e oportunidades. Quero agradecer ao meu orientador D. Marchesin por ter escolhido um tema frutífero para minha dissertação e por ter me feito a pergunta que levou a formulação do método que encontra parâmetros ótimos que ainda geram órbitas heteroclínicas. Quero também agradecer ao meu coorientador G. Chapiro pelos seus comentários sobre a estrutura do presente texto e por ter introduzido situações de interesse com equilíbrios não-hiperbólicos. Também sou grato a A. Mailybaev por ter pacientemente sanado dúvidas e colaborado. Agradeço à Agência Nacional de Petróleo pelo suporte financeiro.

Contents

Chapter 1

Introduction

Assume we are given a system of ordinary differential equations $x' = f(x, p)$ depending on a parameter $p \in \mathbb{R}^{\widetilde{p}}$. Fix $p_0 \in \mathbb{R}^{\widetilde{p}}$. A solution $x(t)$ of $x' = f(x, p_0)$ such that $\lim_{t \longrightarrow \infty} x(t) = x_0$ and $\lim_{t \longrightarrow -\infty} x(t) = y_0$ with $f(x_0, p_0) = 0$, $f(y_0, p_0) = 0$ and $x_0 \neq y_0$ is called a heteroclinic orbit. The points x_0 and y_0 are called equilibria.

Heteroclinic orbits arise, for instance, in the study of travelling waves in reaction-diffusion partial differential equations and diffusive conservation laws as well as in the study of chaos in ordinary differential equations. According to [10], they have been shown to underlie chaotic vibrations of structures [19] and chaotic behaviour of electronic circuits [18, 9, 8] and chemical reactions [14]. Heteroclinic orbits are also important to the study of bifurcations of phase portraits [17, 24].

We consider the problem of locating a parameter p and an initial condition ξ that give rise to a heteroclinic orbit in the ordinary differential equation above. We also consider the problem of locating such parameters p and ξ that also minimize a cost function over the space of initial conditions and parameters. Note that the second problem is more general than the first one.

Assume we are given a viscous conservation law $u_t + g(u)_x = u_{xx}$. To find a travelling wave solution of this conservation law is equivalent to find a heteroclinic orbit in a certain ordinary differential equation with a real parameter σ called velocity. For instance, solving the second problem we can identify, if they exist, the minimum and maximum velocities such that the mentioned viscous conservation law has a travelling wave.

The method developed in [7] to compute heteroclinic orbits considered only hyperbolic equilibria. They derived equations involving heteroclinic orbits and solved them by continuation techniques through discretizations and a phase condition to deal with the non uniqueness of the initial time of the orbits. Their method was extended to the case of center manifolds in [11, 12] and to the case of non-hyperbolic equilibria in [6, 3]. Both hyperbolic and non-hyperbolic equilibria were considered in [5] using multiple shooting for discretizing a certain approximate truncated problem. The majority of papers on computing heteroclinic orbits use approximations of invariant manifolds and continuation techniques to find orbits connecting points in invariant manifolds.

Assume we have two hyperbolic equilibria $u_0 \in \mathbb{R}^n$ and $u_1 \in \mathbb{R}^n$ of a certain ordinary differential equation depending on a parameter $p \in \mathbb{R}^{\widetilde{p}}$. The algorithm

proposed in [10] for locating heteroclinic orbits identifies an orbit $u : [0,1] \longrightarrow \mathbb{R}^n$ such that $u(0) - u_0$ is in the tangent space of the unstable manifold of u_0. Using a successive continuation procedure, the orbit and the equilibria are adjusted so that $u(1) - u_1$ (resp. $u(0) - u_0$) is in the tangent space of the stable (resp. unstable) manifold of u_1 (resp. u_0).

The equations derived in [10] to put $u(0) - u_0$ and $u(1) - u_1$ in tangent spaces of invariant manifolds are different depending on the type of the eigenvalues of the hyperbolic equilibria. Their equations depend on the multiplicity of the eigenvalues as well as in the sign of the real part of a certain eigenvalue.

The method for locating heteroclinic orbits presentd in [2] considers hyperbolic equilibria and two orbits, one with initial condition in an unstable manifold and the other one with initial condition in a stable manifold. Then it tries to annihilate the difference vector between the intersection points of these orbits with a hyperplane. This is done using the sensitivity analysis of the invariant manifolds presented there. Continuation steps are not used.

The plane used in [2] is left to the choice of the user as well as the identification of the initial orbits in the invariant manifolds and their contact times with the plane. Also, the iterative process derived there to get a difference vector equal to zero possess small basin of convergence and presents numerical instability due to the existence of vectors of arbitrarily small norm satisfying a certain non linear equation. It also requires the dimensions of the stable and unstable spaces of the equilibria to remain constant.

In [15] the authors refer to variational methods that can be applied to certain Hamiltonian ordinary differential equations. Variational methods try to give variational formulations for which heteroclinic orbits are critical points. According to them, such techniques are useful to prove existence of heteroclinic orbits. The difficulties lies in finding formulations for which minimization techniques, mountain-pass theorems or other global methods can be applied. For more details see [21].

Our main contribution is to introduce the use of optimization techniques for locating heteroclinic orbits in systems of ordinary differential equations. The ideas introduced here can be used whenever the kind of orbit that we are searching for is a solution of an optimization problem.

For instance, they can be applied to search for periodic orbits, homoclinic orbits and orbits satisfying boundary value problems. Our approach allows locating heteroclinic orbits minimizing fairly general cost functions. The cost functions can be the velocity of a travelling wave, the length of an orbit or its energy.

The best way to show our contributions is to introduce briefly our main method, explain how it works and compare it with the methods proprosed in [10, 2]. Assume we have a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$. Let $x(t, \xi, p)$ be the orbit beginning at ξ, associated with the parameter p and evaluated at time t:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{1.1}$$

By the Differentiable Dependency Theorem, see [17], the function $x(t, \xi, p)$ is of class C^k and is defined in an open set $V = \{(t, \xi, p) : t \in I_{(\xi,p)}, \quad (\xi, p) \in$

$\mathbb{R}^n \times \mathbb{R}^{\widetilde{p}}\}$. The set $I_{(\xi,p)} \subset \mathbb{R}$ is the maximal interval of $x(t,\xi,p)$.

Assume also that we have an open set S and functions $\hat{x}, \hat{y} : S \subset \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n$ with $\hat{x}, \hat{y} \in C^k$ satisfying $f(\hat{x}(p), p) = 0$, $f(\hat{y}(p), p) = 0$ and $\hat{x}(p) \neq \hat{y}(p)$ for all $p \in S$. The functions $\hat{x}$ and $\hat{y}$ are parametrizations of equilibria. Such parametrizations can be obtained via the Implicit Function Theorem or analytically.

The idea of our main method is to find times $\tau, t \in \mathbb{R}$ such that the distance between the point $x(t,\xi,p)$ and the equilibria $\hat{x}(p)$ is very small. The same holds for the distance between $x(\tau,\xi,p)$ and $\hat{y}(p)$. To keep initial conditions in a known place we require the values of ξ to be in a certain plane separating the equilibria. The times τ and t of (1.2) have opposite signs. These ideas can be formalized by (1.2):

$$\begin{aligned} \underset{(\tau,t,\xi,p)}{\text{minimize}} \quad & \| x(\tau,\xi,p) - \hat{x}(p) \|^2 + \| x(t,\xi,p) - \hat{y}(p) \|^2 \\ \text{subject to} \quad & \langle \hat{x}(p) - \hat{y}(p), 2\xi - \hat{x}(p) - \hat{y}(p) \rangle = 0. \end{aligned} \tag{1.2}$$

Now the only thing we should do to solve numerically (1.2) is to provide first and second order derivatives of the objective function and of the constraint. The derivatives of solutions can be found integrating certain initial value problems and the derivatives of equilibria can be found solving certain linear systems.

If the parametrizations $\hat{x}, \hat{y}$ are found via the Implicit Function Theorem, we only require the spatial derivative of f to be non-singular. In the case the equilibria are found analytically, we do not need hypotheses on the type of the equilibria to calculate their derivatives.

We are able to connect non-hyperbolic equilibria and even equilibria where the spatial derivative of the field is singular. The multiplicity of the eigenvalues and the sign of their real part do not matter as opposed to the method in [10]; changes in the dimensions of the unstable spaces of the equilibria do not break the method as opposed to the method in [2].

The feasible set of (1.2) is not compact as τ and t are in the maximal intervals of solutions, which may be unbounded. The situation of interest is when we have a sequence $(\tau_n, t_n, \xi_n, p_n)$, possibly with $|\tau_n|$ and $|t_n|$ going to ∞ and convergent (in a broader sense), where the objective function of (1.2) goes to zero. In contrast with [10, 2], we can try to develop a convergence theory connected with widely studied and used optimization methods. We prove a theorem stating that convergent sequences for which the objective function of (1.2) goes to zero generate heteroclinic orbits under certain conditions. For details see Appendix A.

To solve (1.2) we can use, for instance, the Interior Point Algorithm or trust-region algorithms if we remove the constraint. To use these methods we need to supply certain derivatives, which is a standard procedure. In [10, 2] they used particular approaches to solve their equations. Here we are concerned in giving robust formulations for which we can give accurate estimates of derivatives, but not necessarily with a particular method to minimize.

We do not need to use a phase condition or to increase the period of the approximate heteroclinic orbit to improve accuracy as opposed to [10]. In contrast with [2], we do not need to find manually a plane, initial conditions of orbits in invariant manifolds and their contact times with the plane. The times of the orbits are already included in the optimization and the plane is predefined.

Consider given a cost function $\chi : \mathbb{R}^{n+\widetilde{p}} \longrightarrow \mathbb{R}$ such that $\chi \in C^k$. Functions $x, \hat{x}, \hat{y}$ and f are as before. To find parameters minimizing the cost function and generating heteroclinic orbits, we minimize χ over the set of vectors (ξ, p) such that $x(t, \xi, p)$ is approximately heteroclinic in the sense that it has points close to certain equilibria. For small values of $\epsilon > 0$ and $\widehat{\epsilon} > 0$, we hope that the minimum of (1.3) approaches the minimum of χ over the set of heteroclinic orbits:

$$\begin{array}{ll} \underset{(\tau,t,\xi,p)}{\text{minimize}} & \chi(\xi, p) \\ \text{subject to} & \| \, x(\tau, \xi, p) - \hat{x}(p) \, \|^2 + \| \, x(t, \xi, p) - \hat{y}(p) \, \|^2 \leq \epsilon, \\ & \| \, \hat{x}(p) - \hat{y}(p) \, \|^2 \geq \widehat{\epsilon}, \\ & \langle \hat{x}(p) - \hat{y}(p), 2\xi - \hat{x}(p) - \hat{y}(p) \rangle = 0. \end{array} \tag{1.3}$$

It is harder to solve (1.3) due to the possible bifurcation behaviour of the equilibria. In practice, we need strategies to avoid numerical instabilities. One can also include a constraint in (1.3) requiring the determinant of the Jacobian of the field at the equilibria to be uniformly apart from zero. Certainly, formulation (1.3) can be greatly improved considering special cases of bifurcations.

In Chapter 2 we provide formulae to calculate first and second order derivatives of solutions and equilibria of ordinary differential equations. Such formulae are useful to calculate numerically derivatives of the objective functions and constraints of our optimization problems. Second order derivatives of solutions and equilibria are needed in second order optimization methods to solve the formulations.

Chapter 3 contains detailed explanations of the formulations described here and formulae to solve the optimization problems, i.e., gradients and Hessians of the objective functions and constraints in terms of the values, gradients and Hessians of solutions and equilibria. In Section 3.1 we give details about (1.2) and Section 3.2 about (1.3). In Appendix B we reformulate the method proposed in [2]. Appendix D proposes, as a future work, an optimization problem for locating heteroclinic orbits that do not use information about the equilibria.

In Section 4.1 we use (1.2) for locating a heteroclinic orbit found numerically in [2] and another one in a simple system for which we know the analytical parameter that gives rise to the heteroclinic orbit. In Section 4.2 we use (1.3) for locating optimum parameters generating heteroclinic orbits in a system of ordinary differential equations for which we know the analytical optimum parameter. We provide several convergence tables. All numerical results agree with the analytical ones or with the results from the literature.

In Section 4.3 we considered the problem of finding travelling waves in a system of partial differential equations modelling combustion studied in [13]. The corresponding heteroclinic orbit approaches a non-hyperbolic equilibrium and the ordinary differential equation has solutions very sensible to initial conditions and parameters. Besides, the orbits are very slow in the positive direction and very fast in the negative direction. The analytical velocity of the travelling wave is found in [13]. The analytical and numerical results are shown.

Chapter 2

Preliminaries

In this chapter we state results employed to build and solve optimization problems for locating heteroclinic orbits. Since our optimization problems involve solutions of ordinary differential equations and parametrizations of equilibria of a field, we show here how to verify differentiability of solutions and equilibria as well as how to calculate their first and second order derivatives.

In Section 2.1 we prove a corollary of the Differentiable Dependency Theorem that guarantees differentiability of solutions with respect to time, initial conditions and parameters when the initial conditions are restricted to hypersurfaces depending on parameters.

Section 2.2 shows how to calculate first and second order derivatives of solutions. The second order derivatives of solutions appear in the Hessian of the objective functions and constraints of the optimization problems in Chapter 3.

In Section 2.3 we show how to obtain first and second order derivatives of equilibria when the spatial derivative of the field is non singular; we also explain how to obtain the derivatives of equilibria in special singular cases. The singular case is important for locating travelling waves in a system modelling combustion studied in [13].

In Appendix C we study the dependency of invariant manifolds on field parameters and derive a function to be used as initial condition in systems of ordinary differential equations. We use the ideas of Appendix C in Appendix B.

2.1 Differentiable Dependency Theorem

To solve optimization problems involving solutions of autonomous systems of ordinary differential equations we need to understand the hypothesis under which derivatives of such solutions exist as well as how to calculate them. The Differentiable Dependency Theorem provides this information under reasonable hypotheses about the parameter dependent field of the ordinary differential equation. For similar results on the non autonomous case see [17].

The purpose of this section is to prove a corollary of the Differentiable Dependency Theorem to justify differentiability of solutions with initial conditions contained in hypersurfaces that depend on parameters. Before stating our corollary, we reproduce a form of the Differentiable Dependency Theorem, which can

be found in [17]:

Theorem 2.1.1. *Assume that $U \subset \mathbb{R}^{n+\widetilde{p}}$ is an open set, $f : U \longrightarrow \mathbb{R}^n$ and $f \in C^k$. Let $(\xi, p) \in U, \quad \xi \in \mathbb{R}^n, \quad p \in \mathbb{R}^{\widetilde{p}}$ and $x(t, \xi, p)$ be the unique function of t that satisfies the following initial value problem:*

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{2.1}$$

Consider also the set $V = \{(t, \xi, p) : t \in I_{(\xi,p)}, \quad (\xi, p) \in U\}$, where $I_{(\xi,p)}$ is the maximal interval of $x(t, \xi, p)$. Then V is an open set, the function $x(t, \xi, p) : V \longrightarrow \mathbb{R}^n$ is of class C^k and the following relations hold:

$$\begin{cases} \dfrac{\partial}{\partial t}\dfrac{\partial x}{\partial \xi}(t, \xi, p) = \dfrac{\partial f}{\partial x}(x(t, \xi, p), p)\dfrac{\partial x}{\partial \xi}(t, \xi, p) \\ \dfrac{\partial x}{\partial \xi}(0, \xi, p) = I, \end{cases} \tag{2.2}$$

$$\begin{cases} \dfrac{\partial}{\partial t}\dfrac{\partial x}{\partial p}(t, \xi, p) = \dfrac{\partial f}{\partial x}(x(t, \xi, p), p)\dfrac{\partial x}{\partial p}(t, \xi, p) + \dfrac{\partial f}{\partial p}(x(t, \xi, p), p) \\ \dfrac{\partial x}{\partial p}(0, \xi, p) = 0. \end{cases} \tag{2.3}$$

The derivative of the solution with respect to ξ is an $n \times n$ matrix, while the derivative of the solution with respect to p is an $n \times \widetilde{p}$ matrix. Note that (2.2) and (2.3) can be considered separately from (2.1) given the existence of $x(t, \xi, p)$.

One sees that if (2.2) holds, then we can commute the derivatives of the solution with respect to t and ξ. Conversely, if we can commute these derivatives, then (2.2) holds. Analogously, (2.3) is equivalent to commuting the derivatives with respect to t and p of the solution.

In (2.2) and (2.3) we wrote explicitly where the derivatives of the field f should be evaluated. One way to calculate numerically solutions of (2.2) and (2.3) is to first calculate the solution of (2.1) and then interpolate this solution to solve (2.2) and (2.3). To solve the initial value problems (2.2) and (2.3) one can solve each of their columns separately. The accurate way to compute the derivatives of solutions is to solve (2.1), (2.2) and (2.3) in the coupled form.

The difference between Cor. 2.1.2 and Th. 2.1.1 is essentially the initial conditions of solutions. In Cor. 2.1.2 we consider a function η as given. For instance, in Appendix B we use Cor. 2.1.2 with η equal to a certain parametrization of an unstable manifold; in Section 3.1 we consider $\eta(p, \xi) = \xi \in \mathbb{R}^n$.

Corollary 2.1.2. *Let $S \subset \mathbb{R}^{d+\widetilde{p}}$ be an open set and consider $\eta : S \longrightarrow \mathbb{R}^n$ of class C^k as given. Assume $f : \mathbb{R}^{n+\widetilde{p}} \longrightarrow \mathbb{R}^n$ and $f \in C^k$. Let $(\xi, p) \in U, \quad \xi \in \mathbb{R}^n, \quad p \in \mathbb{R}^{\widetilde{p}}$ and $x(t, \xi, p)$ be the unique function of t satisfying the following initial value problem:*

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \eta(p, \xi). \end{cases} \tag{2.4}$$

Consider also the set $V = \{(t,\xi,p) : t \in I_{(\xi,p)}, \quad (\xi,p) \in S\}$, *where* $I_{(\xi,p)}$ *is the maximal interval of* $x(t,\xi,p)$. *Then* V *is an open set, the function* $x(t,\xi,p) : V \longrightarrow \mathbb{R}^n$ *is of class* C^k *and the following relations hold:*

$$\begin{cases} \dfrac{\partial}{\partial t}\dfrac{\partial x}{\partial \xi}(t,\xi,p) = \dfrac{\partial f}{\partial x}(x(t,\xi,p),p)\dfrac{\partial x}{\partial \xi}(t,\xi,p) \\ \quad \dfrac{\partial x}{\partial \xi}(0,\xi,p) = \dfrac{\partial \eta}{\partial \xi}(p,\xi), \end{cases} \tag{2.5}$$

$$\begin{cases} \dfrac{\partial}{\partial t}\dfrac{\partial x}{\partial p}(t,\xi,p) = \dfrac{\partial f}{\partial x}(x(t,\xi,p),p)\dfrac{\partial x}{\partial p}(t,\xi,p) + \dfrac{\partial f}{\partial p}(x(t,\xi,p),p) \\ \quad \dfrac{\partial x}{\partial p}(0,\xi,p) = \dfrac{\partial \eta}{\partial p}(p,\xi). \end{cases} \tag{2.6}$$

Proof. Take $p \in \mathbb{R}^{\widetilde{p}}$ and $\xi \in \mathbb{R}^d$ and consider the vector $\hat{p} = (p,\xi) \in \mathbb{R}^{\widetilde{p}+d}$. Define $\hat{f}(x,\hat{p}) = f(x+\eta(p,\xi),p)$. As long as $f, \eta \in C^k$, we also have $\hat{f} \in C^k$. The domain of $\hat{f}$ is $\hat{U} = \{(x,\hat{p}) : \hat{p} = (p,\xi) \in S\}$, which is open. By Theorem 2.1.1, the function $\hat{x}(t,\hat{p})$ satisfying $\hat{x}(0,\hat{p}) = 0$ and $\frac{d}{dt}\hat{x}(t,\hat{p}) = \hat{f}(\hat{x}(t,\hat{p}),\hat{p})$ belongs to C^k. To conclude, by uniqueness we have $x(t,\xi,p) = \hat{x}(t,\hat{p}) + \eta(p,\xi)$, which implies $x(t,\xi,p) \in C^k$. Using the last equality, it is simple to obtain (2.5) and (2.6). □

2.2 Derivatives of Solutions

We could try to estimate derivatives of solutions using finite differences. However, this approach does not give reasonable results as can be seen using experiments. To solve properly initial value problems (2.5) and (2.6) gives better estimates of the derivatives of the solution than to use finite differences.

We begin this section calculating the Jacobian of solutions. Later we write the Hessian of coordinates of solutions of ordinary differential equations in blocks because it is convenient for calculations. Then, we spend some subsections obtaining the ordinary differential equations that can be used to calculate certain blocks of the mentioned Hessian.

Throughout this section we consider given a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field. Assume also we have an open set $S \subset \mathbb{R}^{d+\widetilde{p}}$ and a function $\eta : S \longrightarrow \mathbb{R}^n$ of class C^k called parametrization of initial conditions. We denote by $x : V \subset \mathbb{R}^{1+d+\widetilde{p}} \longrightarrow \mathbb{R}^n$ the unique solution (of class C^k) satisfying the initial value problem (2.4). The set V is the same one given by Cor. 2.1.2.

The Jacobian of $x(t,\xi,p)$ at a point (t,ξ,p) is a matrix of dimension $n \times 1 + d + \widetilde{p}$. The derivative of the solution with respect to ξ and p are, respectively, matrices of dimensions $n \times d$ and $n \times \widetilde{p}$. The time derivative is a matrix of dimension $n \times 1$. The mentioned derivatives can be found by solving initial value problems (2.4), (2.5) and (2.6):

$$Dx(t,\xi,p) = \begin{bmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial p} \end{bmatrix}(t,\xi,p). \tag{2.7}$$

2.2.1 Hessian of Coordinates of Solutions

Assume $k \geq 2$. The second order derivatives of the coordinates of the solution of (2.4) are square matrices of sizes $1+d+\widetilde{p}$ and in what follows we write them in blocks for convenience. The derivatives in (2.8) involving time are obtained in this section, while the other ones are found in subsections 2.2.2, 2.2.3 and 2.2.4:

$$\nabla^2 x_i(t,\xi,p) = \begin{bmatrix} \frac{\partial^2 x_i}{\partial t^2} & \frac{\partial^2 x_i}{\partial t \partial \xi}^T & \frac{\partial^2 x_i}{\partial t \partial p}^T \\ \frac{\partial^2 x_i}{\partial t \partial \xi} & \frac{\partial^2 x_i}{\partial \xi^2} & \frac{\partial^2 x_i}{\partial \xi \partial p}^T \\ \frac{\partial^2 x_i}{\partial t \partial p} & \frac{\partial^2 x_i}{\partial \xi \partial p} & \frac{\partial^2 x_i}{\partial p^2} \end{bmatrix} (t,\xi,p). \tag{2.8}$$

For clarity, we write down the dimensions of the blocks in (2.8). The double derivatives with respect to t, ξ and p are, respectively, square matrices of sizes 1, d and $\widetilde{p}$. The mixed derivative with respect to t and ξ has dimension $d \times 1$, while the mixed derivative with respect to t and p has dimension $\widetilde{p} \times 1$. The derivative with respect to ξ and p, taken in this order, has dimension $\widetilde{p} \times d$.

Assume we know $x(t,\xi,p)$ and $Dx(t,\xi,p)$. In order to calculate the mixed derivative with respect to t and ξ at (t,ξ,p) we need to use (2.5); to obtain the mixed derivative with respect to t and p at (t,ξ,p) we need to use (2.6). The double derivative with respect to t of solutions is given by:

$$\frac{\partial^2 x}{\partial t^2}(t,\xi,p) = \frac{\partial f}{\partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial t}(t,\xi,p). \tag{2.9}$$

2.2.2 Double Derivative Relative to Parameters

Assume $k \geq 2$. In this subsection we obtain an initial value problem whose solution is the double derivative with respect to p of the coordinates of the solution of (2.4). By (2.6) we know that the derivative of the solution with respect to p_i $(i = 1, ..., \widetilde{p})$ satisfies the following:

$$\begin{aligned} \frac{\partial}{\partial t}\frac{\partial x}{\partial p_i}(t,\xi,p) &= \frac{\partial f}{\partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial p_i}(t,\xi,p) + \frac{\partial f}{\partial p_i}(x(t,\xi,p),p) \\ &= \frac{\partial f}{\partial p_i}(x(t,\xi,p),p) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial p_i}(t,\xi,p). \end{aligned} \tag{2.10}$$

If we differentiate (2.10) with respect to p_k and in the resulting equation we commute the derivatives with respect to t and p_k we obtain (2.11) for $i,k \in \{1, ..., \widetilde{p}\}$. We can commute the derivatives due to Cor. 2.1.2. Note that we need to apply the chain rule to the terms appearing on the right side of (2.10):

$$
\begin{aligned}
\frac{\partial}{\partial t}\frac{\partial}{\partial p_k}\frac{\partial x}{\partial p_i}(t,\xi,p) &= \frac{\partial}{\partial p_k}\frac{\partial}{\partial t}\frac{\partial x}{\partial p_i}(t,\xi,p) \\
&= \frac{\partial}{\partial p_k}\frac{\partial f}{\partial p_i}(x(t,\xi,p),p) + \frac{\partial}{\partial p_k}\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial p_i}(t,\xi,p) \\
&= \frac{\partial^2 f}{\partial x \partial p_i}(x(t,\xi,p),p)\frac{\partial x}{\partial p_k}(t,\xi,p) \\
&+ \frac{\partial^2 f}{\partial p_k \partial p_i}(x(t,\xi,p),p) \\
&+ \sum_{j,l=1}^{n}\frac{\partial^2 f}{\partial x_l \partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial p_i}(t,\xi,p)\frac{\partial x_l}{\partial p_k}(t,\xi,p) \\
&+ \sum_{j=1}^{n}\frac{\partial^2 f}{\partial p_k \partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial p_i}(t,\xi,p) \\
&+ \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial^2 x_j}{\partial p_k \partial p_i}(t,\xi,p).
\end{aligned}
\tag{2.11}
$$

Note that (2.11) is an ordinary differential equation. We now write it using matrix notation. Define the functions $\Phi^m(t,\xi,p)$ for $m \in \{1,...,n\}$ by relation (2.12). Observe that $\Phi^m(t,\xi,p)$ are symmetric matrices of size $\widetilde{p}$:

$$
\begin{aligned}
\Phi^m(t,\xi,p) &= \frac{\partial x}{\partial p}(t,\xi,p)^T \frac{\partial^2 f_m}{\partial x \partial p}(x(t,\xi,p),p)^T \\
&+ \frac{\partial^2 f_m}{\partial p^2}(x(t,\xi,p),p) \\
&+ \frac{\partial x}{\partial p}(t,\xi,p)^T \frac{\partial^2 f_m}{\partial x \partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial p}(t,\xi,p) \\
&+ \frac{\partial^2 f_m}{\partial x \partial p}(x(t,\xi,p),p)\frac{\partial x}{\partial p}(t,\xi,p).
\end{aligned}
\tag{2.12}
$$

The entries $(k,i) \in \{1,...,\widetilde{p}\} \times \{1,...,\widetilde{p}\}$ of $\Phi^1(t,\xi,p)$, ..., $\Phi^n(t,\xi,p)$ are equal to the coordinates of the non-homogeneous term of (2.11). Then, (2.11) is equivalent to (2.13). As usual, we are denoting by Φ^m_{ki} the entry (k,i) of the matrix Φ^m:

$$
\frac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial p_k \partial p_i} \\ \frac{\partial^2 x_2}{\partial p_k \partial p_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial p_k \partial p_i} \end{bmatrix} = \frac{\partial f}{\partial x}(x(t,\xi,p),p)\begin{bmatrix} \frac{\partial^2 x_1}{\partial p_k \partial p_i} \\ \frac{\partial^2 x_2}{\partial p_k \partial p_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial p_k \partial p_i} \end{bmatrix} + \begin{bmatrix} \Phi^1_{ki} \\ \Phi^2_{ki} \\ \vdots \\ \Phi^n_{ki} \end{bmatrix}. \tag{2.13}
$$

We write where the spatial derivative of the parameter dependent field should be evaluated to avoid misunderstanding. By (2.6), to obtain the derivative of

x with respect to p_k and p_i at (t, ξ, p) we need to integrate (2.13) using initial conditions (2.14):

$$\frac{\partial^2 x}{\partial p_k \partial p_i}(0, \xi, p) = \frac{\partial^2 \eta}{\partial p_k \partial p_i}(p, \xi). \tag{2.14}$$

To summarize we introduce the symbolic initial value problem (2.15). Note that if we take the (k, i) entries in the double parameter derivatives, non-homogeneous term and initial conditions in (2.15), then (2.13) and (2.14) are obtained. Moreover, if we pretend that the derivatives of coordinates of solutions below are numbers and perform the matrix multiplication and examine the entries (k, i) of the resulting matrix equation we also obtain (2.13):

$$\begin{cases} \dfrac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial p^2} \\ \frac{\partial^2 x_2}{\partial p^2} \\ \vdots \\ \frac{\partial^2 x_n}{\partial p^2} \end{bmatrix} = \dfrac{\partial f}{\partial x}(x, p) \begin{bmatrix} \frac{\partial^2 x_1}{\partial p^2} \\ \frac{\partial^2 x_2}{\partial p^2} \\ \vdots \\ \frac{\partial^2 x_n}{\partial p^2} \end{bmatrix} + \begin{bmatrix} \Phi^1 \\ \Phi^2 \\ \vdots \\ \Phi^n \end{bmatrix} \\ \dfrac{\partial^2 x_i}{\partial p^2}(0, \xi, p) = \dfrac{\partial^2 \eta_i}{\partial p^2}(p, \xi) \quad i = 1, ..., n. \end{cases} \tag{2.15}$$

2.2.3 Derivative Relative to Parameters and Initial Conditions

Assume $k \geq 2$. In this subsection we obtain initial value problems whose solutions are mixed derivatives of coordinates of the solution of (2.4) with respect to ξ and p. We apologize for the similarities of this subsection with the previous one, however the steps in the derivations are similar.

As before, we know that the derivative of the solution with respect to p_i ($i = 1, ..., \widetilde{p}$) satisfies (2.10). Differentiating (2.10) with respect to ξ_k ($k = 1, ..., d$) and using Cor. 2.1.2 to commute the derivatives with respect to t and ξ_k, we obtain (2.16):

$$\begin{aligned} \frac{\partial}{\partial t}\frac{\partial}{\partial \xi_k}\frac{\partial x}{\partial p_i}(t, \xi, p) &= \frac{\partial}{\partial \xi_k}\frac{\partial}{\partial t}\frac{\partial x}{\partial p_i}(t, \xi, p) \\ &= \frac{\partial}{\partial \xi_k}\frac{\partial f}{\partial p_i}(x(t, \xi, p), p) + \frac{\partial}{\partial \xi_k}\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t, \xi, p), p)\frac{\partial x_j}{\partial p_i} \\ &= \frac{\partial^2 f}{\partial x \partial p_i}(x(t, \xi, p), p)\frac{\partial x}{\partial \xi_k}(t, \xi, p) \\ &+ \sum_{j,l=1}^{n}\frac{\partial f}{\partial x_l \partial x_j}(x(t, \xi, p), p)\frac{\partial x_l}{\partial \xi_k}(t, \xi, p)\frac{\partial x_j}{\partial p_i}(t, \xi, p) \\ &+ \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t, \xi, p), p)\frac{\partial^2 x_j}{\partial \xi_k \partial p_i}(t, \xi, p). \end{aligned} \tag{2.16}$$

We now recognize that (2.16) is an ordinary differential equation. Let us write it using matrix notation. Define for $m \in \{1, ..., n\}$ the functions $\Psi^m(t, \xi, p)$ by relation (2.17). Note that $\Psi^m(t, \xi, p)$ are matrices of dimension $d \times \widetilde{p}$:

$$\begin{aligned}\Psi^m(t,\xi,p) = &\frac{\partial x}{\partial \xi}(t,\xi,p)^T \frac{\partial^2 f_m}{\partial x \partial p}(x(t,\xi,p),p)^T \\ &+ \frac{\partial x}{\partial \xi}(t,\xi,p)^T \frac{\partial^2 f_m}{\partial x \partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial p}(t,\xi,p).\end{aligned} \tag{2.17}$$

The entries $(k, i) \in \{1, ..., d\} \times \{1, ..., \widetilde{p}\}$ of the matrices $\Psi^1(t, \xi, p)$, ...,$\Psi^n(t, \xi, p)$ are equal to the coordinates of the non-homogeneous term of (2.16). Then, (2.16) is equivalent to (2.18). We denote by ${\Psi^m}_{ki}$ the (k, i) entry of Ψ^m:

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi_k \partial p_i} \\ \frac{\partial^2 x_2}{\partial \xi_k \partial p_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi_k \partial p_i} \end{bmatrix} = \frac{\partial f}{\partial x}(x(t,\xi,p),p)\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi_k \partial p_i} \\ \frac{\partial^2 x_2}{\partial \xi_k \partial p_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi_k \partial p_i} \end{bmatrix} + \begin{bmatrix} {\Psi^1}_{ki} \\ {\Psi^2}_{ki} \\ \vdots \\ {\Psi^n}_{ki} \end{bmatrix}. \tag{2.18}$$

For clarity we write explicitly where the spatial derivative of the parameter dependent field should be evaluated. By (2.6), to calculate the derivative of the solution with respect to ξ_k and p_i at a point (t, ξ, p) we need to integrate (2.18) with the initial condition (2.19):

$$\frac{\partial^2 x}{\partial \xi_k \partial p_i}(0,\xi,p) = \frac{\partial^2 \eta}{\partial \xi_k \partial p_i}(p,\xi). \tag{2.19}$$

Motived by the same reasons as in (2.15), we summarize the result obtained here introducing the symbolic initial value problem below. The word symbolic is being used here because the matrix multiplication in (2.20) is not defined in general:

$$\begin{cases} \dfrac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi \partial p} \\ \frac{\partial^2 x_2}{\partial \xi \partial p} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi \partial p} \end{bmatrix} = \dfrac{\partial f}{\partial x}(x,p)\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi \partial p} \\ \frac{\partial^2 x_2}{\partial \xi \partial p} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi \partial p} \end{bmatrix} + \begin{bmatrix} {\Psi^1}^T \\ {\Psi^2}^T \\ \vdots \\ {\Psi^n}^T \end{bmatrix} \\ \dfrac{\partial^2 x_i}{\partial \xi \partial p}(0,\xi,p) = \dfrac{\partial^2 \eta_i}{\partial \xi \partial p}(p,\xi) \quad i = 1, ..., n. \end{cases} \tag{2.20}$$

2.2.4 Double Derivative Relative to Initial Conditions

Assume $k \geq 2$. We now show how to obtain the double derivative with respect to ξ of the coordinates of the solution of (2.4). As before, we apologize for the similarities of this subsection with the last two. By (2.5), the derivative of the solution with respect to ξ_i $(i = 1, ..., d)$ satisfies the following:

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial x}{\partial \xi_i}(t,\xi,p) &= \frac{\partial f}{\partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial \xi_i}(t,\xi,p) \\ &= \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial \xi_i}(t,\xi,p).\end{aligned} \tag{2.21}$$

Differentiating (2.21) with respect to ξ_k ($k = 1, ..., d$) and using Cor. 2.1.2 to change the order of differentiation between t and ξ_k, we have (2.22). Note that we need to apply the chain rule to terms appearing in the right side of (2.21):

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial}{\partial \xi_k}\frac{\partial x}{\partial \xi_i}(t,\xi,p) &= \frac{\partial}{\partial \xi_k}\frac{\partial}{\partial t}\frac{\partial x}{\partial \xi_i}(t,\xi,p) \\ &= \frac{\partial}{\partial \xi_k}\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial \xi_i}(t,\xi,p) \\ &= \sum_{j,l=1}^{n}\frac{\partial f}{\partial x_l \partial x_j}(x(t,\xi,p),p)\frac{\partial x_j}{\partial \xi_i}(t,\xi,p)\frac{\partial x_l}{\partial \xi_k}(t,\xi,p) \\ &+ \sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x(t,\xi,p),p)\frac{\partial^2 x_j}{\partial \xi_k \partial \xi_i}(t,\xi,p).\end{aligned} \tag{2.22}$$

We now write the non-homogeneous term appearing in (2.22) using matrix notation. To do so, define for $m \in \{1, ..., n\}$ the functions $\Theta^m(t,\xi,p)$ by relation (2.23), which are square matrices of size d:

$$\Theta^m(t,\xi,p) = \frac{\partial x}{\partial \xi}(t,\xi,p)^T \frac{\partial^2 f_m}{\partial x \partial x}(x(t,\xi,p),p)\frac{\partial x}{\partial \xi}(t,\xi,p). \tag{2.23}$$

The entries $(k,i) \in \{1, ..., d\} \times \{1, ..., d\}$ of the matrices $\Theta^1(t,\xi,p)$, ..., $\Theta^n(t,\xi,p)$ are equal to the coordinates of the non-homogeneous term of (2.22). Then (2.22) is equivalent to (2.24). As before, we denote by ${\Theta^m}_{ki}$ the entry (k,i) of Θ^m:

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi_k \partial \xi_i} \\ \frac{\partial^2 x_2}{\partial \xi_k \partial \xi_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi_k \partial \xi_i} \end{bmatrix} = \frac{\partial f}{\partial x}(x(t,\xi,p),p)\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi_k \partial \xi_i} \\ \frac{\partial^2 x_2}{\partial \xi_k \partial \xi_i} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi_k \partial \xi_i} \end{bmatrix} + \begin{bmatrix} {\Theta^1}_{ki} \\ {\Theta^2}_{ki} \\ \vdots \\ {\Theta^n}_{ki} \end{bmatrix}. \tag{2.24}$$

By the same reasons as before, we summarize the result obtained here introducing the symbolic initial value problem (2.25). Note that we provide the initial conditions to integrate (2.24) below:

$$\begin{cases} \dfrac{d}{dt}\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi^2} \\ \frac{\partial^2 x_2}{\partial \xi^2} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi^2} \end{bmatrix} = \dfrac{\partial f}{\partial x}(x,p)\begin{bmatrix} \frac{\partial^2 x_1}{\partial \xi^2} \\ \frac{\partial^2 x_2}{\partial \xi^2} \\ \vdots \\ \frac{\partial^2 x_n}{\partial \xi^2} \end{bmatrix} + \begin{bmatrix} \Theta^1 \\ \Theta^2 \\ \vdots \\ \Theta^n \end{bmatrix} \\ \dfrac{\partial^2 x_i}{\partial \xi^2}(0,\xi,p) = \dfrac{\partial^2 \eta_i}{\partial \xi^2}(p,\xi) \quad i = 1,...,n. \end{cases} \tag{2.25}$$

2.3 Derivatives of Equilibria

We begin explaining how to calculate the value of a equilibrium of a parameter dependent field. For our purposes we need methods to find such equilibrium with basins of attraction as large as possible. Trust-region methods are a good choice because they have larger basins of attraction than Newton's Method.

Derivatives of equilibria appear when we compute gradients and Hessians of the objective functions and constraints appearing in the optimization problems we proposed for locating heteroclinic orbits. In most cases we cannot have explicit formulae for the derivatives of the equilibrium. Here we give implicit formulae for such derivatives in terms of the value of the equilibria.

In this section we consider given $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field. Assume the existence of a parameter $p_0 \in \mathbb{R}^{\widetilde{p}}$ and a point $x_0 \in \mathbb{R}^n$ such that $f(x_0, p_0) = 0$ and that the spatial derivative of the parameter dependent field evaluated at (x_0, p_0) is non singular. Using the Implicit Function Theorem (see [22]), we obtain a function $\hat{x} : B(p_0, \epsilon) \longrightarrow \mathbb{R}^n, \quad \hat{x} \in C^k$ such that $\hat{x}(p_0) = x_0$ and $f(\hat{x}(p), p) = 0 \quad \forall p \in B(p_0, \epsilon)$.

Using the chain rule to differentiate equation $f(\hat{x}(p), p) = 0$ and inverting the spatial derivative of the parameter dependent field in the resulting equation we obtain the following:

$$\frac{\partial \hat{x}}{\partial p}(p) = -\frac{\partial f}{\partial x}(\hat{x}(p), p)^{-1}\frac{\partial f}{\partial p}(\hat{x}(p), p). \tag{2.26}$$

Let us now calculate the Hessian of the coordinates of the equilibrium $\hat{x}$. Differentiating the m-th coordinate of equation $f(\hat{x}(p), p) = 0$ with respect to p_i we obtain (2.27) for $i \in \{1, ..., \widetilde{p}\}$ and $m \in \{1, ..., n\}$:

$$\sum_{j=1}^{n} \frac{\partial \hat{x}_j}{\partial p_i}(p)\frac{\partial f_m}{\partial x_j}(\hat{x}(p), p) = -\frac{\partial f_m}{\partial p_i}(\hat{x}(p), p). \tag{2.27}$$

Assume $k \geq 2$. Differentiating (2.27) with respect to p_k for $k \in \{1, ..., \widetilde{p}\}$ we obtain (2.28) that we couple for $m \in \{1, ..., n\}$ to build a linear system through which we find the Hessian of the coordinates of the equilibrium:

$$\begin{aligned}\sum_{j=1}^{n}\frac{\partial f_m}{\partial x_j}(\hat{x}(p),p)\frac{\partial^2\hat{x}_j}{\partial p_k\partial p_i}(p) = &-\sum_{j,l=1}^{n}\frac{\partial\hat{x}_j}{\partial p_i}(p)\frac{\partial^2 f_m}{\partial x_l\partial x_j}(\hat{x}(p),p)\frac{\partial x_l}{\partial p_k}(p)\\ &-\sum_{j=1}^{n}\frac{\partial\hat{x}_j}{\partial p_i}(p)\frac{\partial f_m}{\partial p_k\partial x_j}(\hat{x}(p),p)\\ &-\sum_{l=1}^{n}\frac{\partial^2 f_m}{\partial x_l\partial p_i}(\hat{x}(p),p)\frac{\partial\hat{x}_l}{\partial p_k}(p)\\ &-\frac{\partial f_m}{\partial p_k\partial p_i}(\hat{x}(p),p).\end{aligned} \tag{2.28}$$

To solve (2.28) it is convenient to introduce matrix notation. Define for $m = 1, ..., n$ the functions $\widehat{\Phi}^m(p)$ by (2.29). Note that $\widehat{\Phi}^m(p)$ are symmetric matrices of size $\widetilde{p}$:

$$\begin{aligned}\widehat{\Phi}^m(p) = &\frac{\partial^2 f_m}{\partial x\partial p}(\hat{x}(p),p)\frac{\partial\hat{x}}{\partial p}(p)\\ &+\frac{\partial\hat{x}}{\partial p}(p)^T\frac{\partial^2 f_m}{\partial x\partial p}(\hat{x}(p),p)^T\\ &+\frac{\partial\hat{x}}{\partial p}(p)^T\frac{\partial^2 f_m}{\partial x\partial x}(\hat{x}(p),p)\frac{\partial\hat{x}}{\partial p}(p)\\ &+\frac{\partial^2 f_m}{\partial p\partial p}(\hat{x}(p),p).\end{aligned} \tag{2.29}$$

Coupling equations (2.28) for $m = 1, ..., n$ and using functions $\widehat{\Phi}^m(p)$, we obtain the linear system below. Observe the similarity of functions $\widehat{\Phi}^m(p)$ and $\Phi^m(p)$. As before, we are denoting by $\widehat{\Phi}^m_{ki}$ the entry (k, i) of the matrix $\widehat{\Phi}^m$:

$$\frac{\partial f}{\partial x}(\hat{x}(p),p)\begin{bmatrix}\frac{\partial^2\hat{x}_1}{\partial p_k\partial p_i}\\ \frac{\partial^2\hat{x}_2}{\partial p_k\partial p_i}\\ \vdots\\ \frac{\partial^2\hat{x}_n}{\partial p_k\partial p_i}\end{bmatrix} = -\begin{bmatrix}\widehat{\Phi}^1_{ki}\\ \widehat{\Phi}^2_{ki}\\ \vdots\\ \widehat{\Phi}^n_{ki}\end{bmatrix}. \tag{2.30}$$

As we are assuming that the spatial derivative of the parameter dependent field is non singular, (2.30) can be solved. Given the value and the Jacobian of the equilibrium $\hat{x}$ at a parameter p, we can find its second order derivatives. We summarize the results obtained here introducing the symbolic equation (2.31) by the same reasons as in (2.15):

$$\frac{\partial f}{\partial x}(\hat{x}(p),p)\begin{bmatrix}\frac{\partial^2\hat{x}_1}{\partial p\partial p}\\ \frac{\partial^2\hat{x}_2}{\partial p\partial p}\\ \vdots\\ \frac{\partial^2\hat{x}_n}{\partial p\partial p}\end{bmatrix} = -\begin{bmatrix}\widehat{\Phi}^1\\ \widehat{\Phi}^2\\ \vdots\\ \widehat{\Phi}^n\end{bmatrix}. \tag{2.31}$$

2.4 Conclusion

The Differentiable Dependency Theorem is our main tool to state differentiability of solutions of ordinary differential equations with respect time, initial conditions and parameters and also tells how to find good numerical estimates of derivatives of solutions solving certain initial value problems. As we already mentioned, to use finite differences to approximate derivatives of solutions is not accurate.

To state differentiability of equilibria we can try to solve analytically a non linear equation or we can try to use the Implicit Function Theorem. If we have a formula to the equilibria, we can check if they are differentiable. In the case we are trying to use the Implicit Function Theorem, we showed that when the spatial derivative of the parameter dependent field is non singular the derivatives of equilibria are solutions of linear systems.

Appendix C was intended to explore issues related to the dependency of invariant manifolds on field parameters. We showed there that the hypothesis about the equilibria needed to consider the function ϕ are strong and also that certain second order partial derivatives of ϕ are difficult to find because the calculations involve derivatives of eigenvectors.

The same theory developed here to find derivatives of solutions of autonomous ordinary differential equations can be done to the non autonomous case. Numerical estimates of derivatives of solutions of non autonomous ordinary differential equations may help us to solve non linear equations or optimization problems involving such solutions. Therefore, most of the content in this chapter is of general interest.

Chapter 3

Locating Heteroclinic Orbits

In this chapter we propose optimization problems for locating heteroclinic orbits of ordinary differential equations. To solve these optimization problems, we provide formulae for the gradients and Hessians of objective functions and constraints. This is sufficient because solvers available in standard packages need only the value, gradient and Hessian of the objective function and constraints for each point in the search space.

Section 3.1 introduces a new numerical method for locating heteroclinic orbits that can be used whenever we provide local parametrizations of equilibria. For instance, we can use this method to connect non-hyperbolic equilibria using the Implicit Function Theorem to obtain parametrizations. We provide a theorem that justifies the method.

In the case we do not have uniqueness for the initial conditions and parameters that generate a heteroclinic orbit we can ask for the ones that also minimize a cost function. In Section 3.2 we propose a numerical method that solves this issue by a sequence of optimization problems.

In Appendix B we reformulate the method presented in [2] for locating heteroclinic orbits. The methods introduced here make weaker hypotheses than the ones used in the methods of Appendix B and of [10]. They can also be used in more general contexts.

Appendix D introduces a method for locating heteroclinic orbits that do not need information on the localization of the equilibria. Such method shows small basin of attraction, but from a theoretical point of view it may be interesting. We left a more complete analysis of its properties as a future work.

3.1 Global Method

The idea of the Global Method is to take initial conditions in a predefined plane. Then, using a general optimization procedure, it iteratively chooses a positive time t_+, a negative time t_-, an initial condition ξ and a parameter p that minimizes the distance between the equilibria and the orbit beginning at ξ with parameter p evaluated at t_+ and t_-.

Assume we have a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field, an open set S and also functions $\hat{x}, \hat{y} : S \subset \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n$ with $\hat{x}, \hat{y} \in C^k$ satisfying $f(\hat{x}(p), p) = 0$, $f(\hat{y}(p), p) = 0$ and $\hat{x}(p) \neq \hat{y}(p)$ for all $p \in S$. The set S is contained in the parameter space of f. The functions $\hat{x}$ and $\hat{y}$ are parametrizations of equilibria. They can be obtained via the Implicit Function Theorem or analytically.

We assume that the parametrizations of the equilibria are given. This assumption allows our formulation to remove the hypothesis about the nature of the equilibria. However, such hypothesis should be used, for instance, to obtain the parametrizations of equilibria. It also tells how to calculate derivatives of equilibria. More details about this subject are found in Section 2.3.

Let $x(t, \xi, p)$ be the unique solution satisfying the initial value problem (3.1) below. Note that the initial condition $\xi \in \mathbb{R}^n$ can be everywhere in $\mathbb{R}^n$. The domain of x is the one given by Cor. 2.1.2:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{3.1}$$

Now we introduce (3.2). The objective function measures the distance between the orbit at times $\tau, t \in \mathbb{R}$ to $\hat{x}$ and $\hat{y}$. So that orbits can reach different equilibria, τ and t must have opposite signs. We require the initial condition $\xi \in \mathbb{R}^n$ of the orbit $x(t, \xi, p)$ to be in the plane whose normal is $\hat{x}(p) - \hat{y}(p)$ and that passes through the middle point of $\hat{x}(p)$ and $\hat{y}(p)$. The choice of the plane is arbitrary. Note that the optimization is over $\tau \in \mathbb{R}$, $t \in \mathbb{R}$, $\xi \in \mathbb{R}^n$ and $p \in \mathbb{R}^{\widetilde{p}}$:

$$\begin{array}{ll} \underset{(\tau,t,\xi,p)}{\text{minimize}} & \| x(\tau, \xi, p) - \hat{x}(p) \|^2 + \| x(t, \xi, p) - \hat{y}(p) \|^2 \\ \text{subject to} & \langle \hat{x}(p) - \hat{y}(p), 2\xi - \hat{x}(p) - \hat{y}(p) \rangle = 0. \end{array} \tag{3.2}$$

In general, the feasible set of (3.2) is not compact. Assume there is an heteroclinic orbit, which we represent by the vector (ξ^*, p^*). If the equilibria $\hat{x}$ and $\hat{y}$ are different, the orbit $x(t, \xi^*, p^*)$ cannot reach them for finite times. This means that there are not times τ^* and t^* such that the objective function is zero at $(\tau^*, t^*, \xi^*, p^*)$.

The situation of interest is when we have a convergent sequence $(\tau_n, t_n, \xi_n, p_n)$ for which the objective function approaches zero. If the equilibria are different, we can conclude that the absolute values of τ_n and t_n converge to infinity. Note that we are using the word convergent in a broader sense for τ_n and t_n. Assume that the sequence (ξ_n, p_n) converges to a certain (ξ_f, p_f). We would like to know under which conditions $x(t, \xi_f, p_f)$ is heteroclinic.

Let us forget for a moment the non compactness of the feasible set of (3.2). Imagine that to solve (3.2) we use standard optimization methods such as the Interior Point Algorithm with $(\tau_0, t_0, \xi_0, p_0)$ as a starting point. As usual, we hope that the last two coordinates of the sequence generated by the optimization method converge to a point (ξ^*, p^*) that gives rise to a heteroclinic orbit if the starting point is close enough to (ξ^*, p^*). We expect the convergence to be less sensitive to (τ_0, t_0).

We now introduce a procedure needed to calculate the values of the objective function of (3.2), which we call Equilibrium Non-existence Procedure (ENP).

The procedure consists to assign the value ∞ to the objective function of (3.2) at every point (τ, t, ξ, p) such that we are not able to locate numerically a equilibrium associated with the parameter p. The procedure ENP does not change the minimum of (3.2) and is useful when the equilibria disappear for a certain parameter p.

To calculate the value of the objective function and of the constraint, we need to find the values of $\hat{x}(p)$ and $\hat{y}(p)$ and to integrate (3.1) to find the values of $x(\tau, \xi, p)$ and $x(t, \xi, p)$. To see how to calculate the Jacobian of x read Section 2.2 using $\eta(p, \xi) = \xi$ and for information about derivatives of equilibria see Section 2.3. Note that the initial conditions of the derivatives of the solution are simple:

$$\frac{\partial x}{\partial \xi}(0, \xi, p) = I \in M(n \times n), \tag{3.3}$$

$$\frac{\partial x}{\partial p}(0, \xi, p) = 0 \in M(n \times \widetilde{p}). \tag{3.4}$$

Assume $g, h \in C^1$. Formulae (3.5) and (3.6) are used to calculate the gradient of the objective function and constraint of (3.2). In Appendix E we obtain them:

$$\nabla < h, g >= Dh^T g + Dg^T h, \tag{3.5}$$

$$\nabla \parallel g \parallel^2 = 2Dg^T g. \tag{3.6}$$

We now give formulae for the derivatives of functions appearing inside squared norms and inner products in (3.2). As the formulae that appear are large, we introduce some notation to organize the results. Let Δ, Γ, Ω and Λ be the following functions:

$$\begin{aligned} \Delta(\tau, t, \xi, p) &= x(\tau, \xi, p) - \hat{x}(p), \\ \Gamma(\tau, t, \xi, p) &= x(t, \xi, p) - \hat{y}(p), \\ \Omega(\tau, t, \xi, p) &= \hat{x}(p) - \hat{y}(p), \\ \Lambda(\tau, t, \xi, p) &= 2\xi - \hat{x}(p) - \hat{y}(p). \end{aligned} \tag{3.7}$$

Note that we should take the gradient with respect to τ, t, ξ and p. The first order derivatives of Δ, Γ, Ω and Λ are given by the formulae below, which together with (3.5) and (3.6) give the gradients of the objective function and constraint:

$$\begin{aligned} D\Delta &= \left[\begin{array}{cccc} \frac{\partial x}{\partial \tau} & 0 & \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial p} - \frac{\partial \hat{x}}{\partial p} \end{array} \right], \\ D\Gamma &= \left[\begin{array}{cccc} 0 & \frac{\partial x}{\partial t} & \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial p} - \frac{\partial \hat{y}}{\partial p} \end{array} \right], \\ D\Omega &= \left[\begin{array}{cccc} 0 & 0 & 0 & \frac{\partial \hat{x}}{\partial p} - \frac{\partial \hat{y}}{\partial p} \end{array} \right], \\ D\Lambda &= \left[\begin{array}{cccc} 0 & 0 & 2I & -\frac{\partial \hat{x}}{\partial p} - \frac{\partial \hat{y}}{\partial p} \end{array} \right]. \end{aligned} \tag{3.8}$$

The derivatives of x building $D\Delta$ at (τ, t, ξ, p) should be evaluated at (τ, ξ, p); the derivatives of x building $D\Gamma$ at the point (τ, t, ξ, p) should be evaluated at (t, ξ, p). The non constant blocks building $D\Omega$ and $D\Lambda$ depend only on p.

Assume $k \geq 2$. In order to use second order methods to solve (3.2), we need to calculate second order derivatives of solutions and equilibria. For more details on these subjects see sections 2.2 and 2.3. The initial conditions needed to find second order derivatives of solutions are given by the following equations:

$$\frac{\partial^2 x_i}{\partial p \partial p}(0, \xi, p) = 0 \in M(\widetilde{p} \times \widetilde{p}), \tag{3.9}$$

$$\frac{\partial^2 x_i}{\partial \xi \partial \xi}(0, \xi, p) = 0 \in M(n \times n), \tag{3.10}$$

$$\frac{\partial^2 x_i}{\partial \xi \partial p}(0, \xi, p) = 0 \in M(\widetilde{p} \times n). \tag{3.11}$$

Assume $g, h \in C^2$. We use formulae (3.12) and (3.13), which are derived in Appendix E, to compute the Hessian of the objective function and constraint of (3.2):

$$\nabla^2 < g, h >= Dg^T Dh + Dh^T Dg + \sum g_i \nabla^2 h_i + \sum h_i \nabla^2 g_i, \tag{3.12}$$

$$\nabla^2 \parallel g \parallel^2 = 2Dg^T Dg + 2\sum g_i \nabla^2 g_i. \tag{3.13}$$

Formulae (3.12) and (3.13) are used with functions g, h equal to functions in (3.7). Then, we need to give the Hessians of the coordinates of such functions. The reader should be careful to evaluate the derivatives below at the right places. The subscript stands for the coordinates of the functions:

$$\begin{aligned}
\nabla^2 \Delta_i &= \begin{bmatrix} \frac{\partial^2 x_i}{\partial \tau \partial \tau} & 0 & \frac{\partial^2 x_i}{\partial \tau \partial \xi}^T & \frac{\partial^2 x_i}{\partial \tau \partial p}^T \\ 0 & 0 & 0 & 0 \\ \frac{\partial^2 x_i}{\partial \tau \partial \xi} & 0 & \frac{\partial^2 x_i}{\partial \xi \partial \xi} & \frac{\partial^2 x_i}{\partial \xi \partial p}^T \\ \frac{\partial^2 x_i}{\partial \tau \partial p} & 0 & \frac{\partial^2 x_i}{\partial \xi \partial p} & \frac{\partial^2 x_i}{\partial p \partial p} - \frac{\partial^2 \hat{x}_i}{\partial p \partial p} \end{bmatrix}, \\
\nabla^2 \Gamma_i &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial^2 x_i}{\partial t \partial t} & \frac{\partial^2 x_i}{\partial t \partial \xi}^T & \frac{\partial^2 x_i}{\partial t \partial p}^T \\ 0 & \frac{\partial^2 x_i}{\partial t \partial \xi} & \frac{\partial^2 x_i}{\partial \xi \partial \xi} & \frac{\partial^2 x_i}{\partial \xi \partial p}^T \\ 0 & \frac{\partial^2 x_i}{\partial t \partial p} & \frac{\partial^2 x_i}{\partial \xi \partial p} & \frac{\partial^2 x_i}{\partial p \partial p} - \frac{\partial^2 \hat{y}_i}{\partial p \partial p} \end{bmatrix}, \\
\nabla^2 \Omega_i &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +\frac{\partial^2 \hat{x}_i}{\partial p \partial p} - \frac{\partial^2 \hat{y}_i}{\partial p \partial p} \end{bmatrix}, \\
\nabla^2 \Lambda_i &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\partial^2 \hat{x}_i}{\partial p \partial p} - \frac{\partial^2 \hat{y}_i}{\partial p \partial p} \end{bmatrix}.
\end{aligned} \tag{3.14}$$

We can remove the non linear constraint appearing in (3.2) without sacrificing its utility for locating heteroclinic orbits. However, if we keep the constraint

we know how apart the initial conditions are from the equilibria and turn more predictable the behaviour of the times τ and t. Without the constraint the solver would approximate ξ to a equilibrium letting τ or t unchanged.

If we remove the constraint we can use methods based in the trust-region algorithm (see [1]). Such methods are precise and fast to minimize quadratic objective functions and they also keep control on the size of the step of each iteration of the optimization. To solve the unconstrained problem, we need to compute derivatives only of the objective function.

When we integrate ordinary differential equations near equilibria we may face stiff integration problems. For instance consider the flux near a hyperbolic equilibria. In these cases, the time derivatives of the solutions near a certain equilibrium can be high and this abrupt variation introduces inaccuracies in the integration. For details about stiff integrators see [4]. The problem related occurs in (3.2) when we use times for which the orbits have points near equilibria.

The next theorem justifies the main idea behind the Global Method. It states that if we have a convergent sequence for which the objective function of (3.2) goes to zero, then the limit of such sequence generates a heteroclinic orbit under reasonable conditions. We prove Theorem 3.1.1 is Appendix A.

Theorem 3.1.1. *Assume we have sequences $\tau_n \longrightarrow \infty$, $t_n \longrightarrow -\infty$, $(\xi_n, p_n) \longrightarrow (\xi^*, p^*) \in \mathbb{R}^n \times S$ and that the objective function of (3.2) goes to zero over the sequence $z_n = (\tau_n, t_n, \xi_n, p_n)$. The sequence z_n is supposed to be in the feasible set of (3.2). Assume also that*

$$\begin{aligned} \| \, x(\tau_n, \xi^*, p^*) - x(\tau_n, \xi_n, p_n) \, \| &\longrightarrow 0, \\ \| \, x(t_n, \xi^*, p^*) - x(t_n, \xi_n, p_n) \, \| &\longrightarrow 0. \end{aligned} \tag{3.15}$$

If the equilibria $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are hyperbolic, then the orbit $x(t, \xi^*, p^*)$ is heteroclinic. Independently of the type of the equilibria, $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are accumulation points of the orbit $x(t, \xi^*, p^*)$.

3.2 Optimum Parameter Method

In the previous section we formulated a optimization problem to locate initial conditions and parameters that give rise to a heteroclinic orbit. In the case there is not uniqueness to such initial conditions and parameters we can ask for the ones that minimize a cost function. Here we provide a numerical method for this problem.

The non uniqueness of initial conditions that we are interested here is when the intersection between unstable and stable manifolds has dimension greater than one. When such intersection has dimension equal to one, to restrict initial conditions in a hyperplane can induce uniqueness.

In the context of travelling waves in conservation laws, the cost function could be the velocity of the wave or minus the velocity of the wave. If we have a situation of non uniqueness of the velocity that gives rise to a travelling wave, these cost functions can tell what are the minimum and maximum velocities of the possible travelling waves.

Assume we have a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field. Let S be an open set and $\hat{x}, \hat{y} : S \subset \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n$ be functions satisfying $\hat{x}, \hat{y} \in C^k$, $f(\hat{x}(p), p) = 0$ and $f(\hat{y}(p), p) = 0$ for all $p \in S$. We do

not require the equilibria to be different because optimum parameters can occur when the equilibria collide. Assume also we have a cost function $\chi : \mathbb{R}^{n+\widetilde{p}} \longrightarrow \mathbb{R}$ such that $\chi \in C^k$. Let $x(t, \xi, p)$ be the unique solution satisfying the initial value problem (3.16) below:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{3.16}$$

Consider given $\epsilon, \widehat{\epsilon} > 0$. Let us now introduce (3.17). Note that the constraints of (3.17) require the initial conditions to be in the same hyperplane used in the Global Method, the distance between certain points of the orbits to the equilibria to be less than ϵ and the equilibria to be $\widehat{\epsilon}$ apart:

$$\begin{aligned} \underset{(\tau,t,\xi,p)}{\text{minimize}} \quad & \chi(\xi, p) \\ \text{subject to} \quad & \| x(t, \xi, p) - \hat{y}(p) \|^2 + \| x(\tau, \xi, p) - \hat{x}(p) \|^2 \leq \epsilon, \\ & \| \hat{x}(p) - \hat{y}(p) \|^2 \geq \widehat{\epsilon}, \\ & \langle \hat{x}(p) - \hat{y}(p), 2\xi - \hat{x}(p) - \hat{y}(p) \rangle = 0. \end{aligned} \tag{3.17}$$

As before, to solve (3.17) we use the Equilibrium Non-existence Procedure (ENP) introduced in Section 3.1 applied to the constraint requiring the orbit to be near the equilibria. In practice, this procedure removes from the feasible set of (3.17) points for which we are not able to locate numerically the equilibria. The ENP together with the constraint that requires the equilibria to be uniformly apart helps (3.17) to deal with bifurcation situations.

We now introduce the Optimum Parameter Method. Consider sequences $\{\epsilon_n\}$ and $\{\widehat{\epsilon}_n\}$ of non-increasing numbers converging to zero. Let us call P_n the optimization problem (3.17) associated with ϵ_n and $\widehat{\epsilon}_n$. The method starts with the Global Method to find a vector $\nu_0 = (\tau_0, t_0, \xi_0, p_0)$ that identifies approximately a heteroclinic orbit. Then, for $k \geq 1$, we use ν_{k-1} as a starting point to solve P_k, which gives an analogous point ν_k. We iterate this process until a certain criteria is satisfied. Note that ν_{k-1} may be infeasible to P_k. Then, the existing optimization methods that are able to restore feasibility are useful.

In order for the minimum of (3.17) to exist, we need to assume more hypotheses on χ and on certain regions of interest of the search space. We can assume that the set of vectors (ξ, p) such that certain points of the orbit $x(t, \xi, p)$ are at most ϵ apart from the equilibria is compact. This hypothesis is sufficient for the existence of the minimum, but is not necessary. For example, consider the parameter dependent field introduced in Section 4.2.

Assume $\widehat{\epsilon}_n = 0$. Note that heteroclinic orbits are in the feasible set of (3.17). Therefore, the global minima of (3.17), when they exist, are lower bounds to the minimum value of χ over the set of heteroclinic orbits. In reasonable situations we hope that as n grows, the minima of P_n indeed converge to the minimum of χ over the set of heteroclinic orbits. Here we are abusing notation calling a vector (ξ, p) that generates a heteroclinic orbit of heteroclinic orbit itself. Observe also that the global minimum values of P_1, P_2, ... form a non-decreasing sequence as their feasible sets shrink with n.

It is natural to ask if the sequence $\{\nu_k\}_k$ mentioned above converges. We do not provide this analysis as well as we do not show here the formulae to solve (3.17) because all are similar to the ones already developed in Section 3.1. As before, we must calculate first and second order derivatives of solutions and equilibria to build the gradients and Hessians of the constraints and objective functions of (3.17).

3.3 Conclusion

The method proposed in Section 3.1 make weaker requirements on the equilibria than the ones made in [10] and in the method of Appendix B. The spatial derivative of the parameter dependent field should be non singular if we do not know how to find the equilibria analytically. The Global Method is intended to obtain vectors $(\tau, t, \xi, p) \in \mathbb{R}^{2+n+\widetilde{p}}$ that generate orbits with small distances to the equilibria.

The Optimum Parameter Method finds initial conditions and parameters minimizing a certain cost function and generating heteroclinic orbits. As in Section 3.1, we can also remove the constraint. However, the minimum may change depending on the cost function. There are important questions that we neglected. We did not discuss theoretical questions related with the topology of the set of vectors (ξ, p) that generate heteroclinic orbits as well as we did not considered the relations of Section 3.2 with Bifurcation Theory.

The method of Appendix D locates heteroclinic orbits without information on the equilibria of the parameter dependent field. The basin of attraction of the Equilibrium-Free Method is much smaller than the ones of the other methods. One can also try to figure out equilibrium-free formulations to find optimum parameters that give rise to heteroclinic orbits. We left a more careful treatment of these possible methods as a future work.

To solve numerically the optimization problems we proposed here there are several parameters we should use to do all calculations. For example, we should fix a maximum amount of iterations of the numerical methods used to solve the optimization problems, we should say if the numerical method should use Hessians, what is the time increment to solve initial value problems as well as the starting point of the numerical method. Good implementations should allow modifications of these parameters.

The ideas used here, which consists to use non linear equations or optimization problems for locating heteroclinic orbits, may be used for locating parameters of ordinary differential equations that give rise to other special orbits. For instance, if the excepted properties of the orbits can be seen as zeros of non linear equations or minima of optimization problems. As we commented in the conclusions of the previous chapter, we do not need the ordinary differential equation to be autonomous to apply such ideas.

Chapter 4

Experiments and Applications

This chapter is intended to show experiments in which we use the optimization problems introduced in Chapter 3 for locating heteroclinic orbits in ordinary differential equations. Through these experiments we give clues about the convergence rate of the optimization problems and also about their basins of attraction. The experiments are focused on the Global Method and on the Optimum Parameter Method.

In Section 4.1 we use the Global Method for locating the heteroclinic orbit found in [2] using worse starting points than the ones used in [2]. Also in Section 4.1 we use the Global Method in a simple ordinary differential equation for which we know the existence and local uniqueness of a parameter that give rise to a heteroclinic orbit. In the latter case we compare analytical and numerical results.

Section 4.2 introduces a simple ordinary differential equation for which we can calculate analytically the regions of parameters that give rise to heteroclinic orbits. We compare the analytical results with the results given by the Optimum Parameter Method.

In Section 4.3 we use the Global Method for locating travelling waves in a system modelling combustion in a porous medium. The analytical results given by [13] and the numerical ones given by the Global Method agree.

In what follows we say that a sequence $\{x_k\} \subset \mathbb{R}^n$ converges linearly to $L \in \mathbb{R}^n$ if the following holds:

$$\lim_{k \longrightarrow \infty} \frac{\| x_{k+1} - L \|}{\| x_k - L \|} = \alpha \in (0, 1) \tag{4.1}$$

4.1 Global Method Experiments

In all experiments of this section we use the *ode45* Matlab solver to integrate the initial value problems (2.4), (2.5), (2.6), (2.15), (2.20) and (2.25) in the coupled manner. To solve (3.2) we use the $fmincon$ Matlab solver with the implementation of the Interior Point Algorithm provided by the *Optimization*

Matlab Toolbox. We provide the Hessians and gradients of the objective function and constraint of (3.2). For details about the solver *fmincon* see [1].

The tolerances *TolX* and *TolFun* used in *fmincon* should be taken very small to avoid the solver to stop due to numerical criteria for the local minima. This sets the solver to stop only due to reaching the maximum amount of iterations. Here we take *TolX* and *TolFun* both equal to 10^{-30}. In the *ode45* solver we use *AbsTol*=10^{-6} and *RelTol*=10^{-3}.

First Experiment

In the first experiment we use the Global Method for locating a heteroclinic orbit in the parameter dependent field (4.2). It can be shown that for $p = 1$ the field has four heteroclinic orbits connecting hyperbolic equilibria. In Figure 4.1 we show the field and the heteroclinic orbits, while in its caption we sketch the proof of the existence of the heteroclinic orbits:

$$f_1(x, p) = \begin{bmatrix} x_1(x_1 - 2)(1 - x_2) + (1 - p) \\ x_2(x_2 - 2)(x_1 - 1) + (1 - p) \end{bmatrix}. \tag{4.2}$$

In Table 4.1 we present an experiment in which we can see the convergence rate of the parameter p side by side with the value of the objective function of (3.2). The starting and final points of the optimization are reported in the caption. Using Figure 4.1 we can see that the point $\xi_f = (1, 2)^T$ reported in the caption of Table 4.1 lies in the line whose normal is given by the difference of equilibria and passing through the middle point of the equilibria.

In Figure 4.4 we see the log of the values of the objective function of Table 4.1 and in Figure 4.6 the log of the absolute value of $p - 1$ of Table 4.1. These figures show evidence that the parameter and the objective function converge linearly to 0 and 1, respectively.

Figure 4.2 shows the coordinates of the orbits at different iterations of the Global Method. We can see that the coordinates approach equilibria. The heteroclinic orbit found in Figure 4.2 is the one connecting $(0, 2)^T$ and $(2, 2)^T$ in Figure 4.1. Note that the first coordinate of the heteroclinic orbit varies between 0 and 2 and second coordinate is equal to two. In Figure 4.2 we can also see that the absolute values of τ and t increase during iterations.

Second Experiment

In the second experiment we use the Global Method to search heteroclinic orbits in the system of ordinary differential equations associated with f_2 of (4.3). In all experiments the field f_2 is multiplied by 13. For the parameter dependent field f_2 in (4.3) we do not know how to prove analytically the existence of heteroclinic orbits. Here we reproduce the experiment presented in [2] with a worse starting point and taking advantage of the Global Method to automatically choose integration times:

$$\begin{cases} x_0 = (0.13, 0.07, 0.05, 0.02)^T, \\ g(x_1, x_2, x_3, x_4) = \begin{bmatrix} 0.01x_1 + 0.1x_2 + 0.02x_3 - 0.5x_1^2 + 0.5x_2^2 \\ -0.1x_1 + 0.01x_4 + x_1x_2 \\ -0.02x_2 - 0.12x_3 + 0.01x_4 + 0.2x_3x_4 \\ 0.01x_1 + 0.01x_3 + 0.11x_4 + 0.1x_1x_4 \end{bmatrix}, \\ f_2(x, p) = g(x) - g(x_0) - p(x - x_0). \end{cases} \tag{4.3}$$

In Table 4.2 we see the convergence of the parameter and the value of the objective function of (3.2) throughout the optimization. Note that we take almost 80 iterations to reach a value of the objective function comparable to the one obtained in Table 4.1 after 30 iterations. In Figure 4.5 we see the log of the values of the objective function of Table 4.2.

In Figure 4.5 we see the log of the values of the objective function of Table 4.2 and in Figure 4.7 the log of the absolute value of $p + 0.025955$ of Table 4.2. These figures show evidence that the parameter and the objective function converge linearly to 0 and -0.025955, respectively.

In Figure 4.3 we see the coordinates of the orbits during the optimization. More precisely, for each iteration of the optimization corresponds a point $(\tau_n, t_n, \xi_n, p_n)$. Using the same notation of Section 3.1, the coordinates shown in these figures are the ones of the orbits $x(t, \xi_n, p_n)$. The orbits are calculated between times τ_n and t_n. Note that the coordinates of the orbits approach equilibria.

Iteration	p	Obj. Function
0	-1.00000000000000	1.471138e+01
1	1.113449803983675	6.023752e-01
2	1.090525494364682	3.324847e-01
3	0.986282107733741	1.095251e-01
4	1.026289398698400	9.235322e-02
5	0.998510619452611	3.883934e-02
6	1.000141301120699	2.418643e-02
7	1.000749340057069	9.264985e-03
8	0.999702148507327	4.179318e-03
9	1.000117923890576	2.587890e-03
10	0.999914961801699	1.213904e-03
11	1.000055752435782	6.713668e-04
12	0.999983104718338	5.228764e-04
13	0.999999713131138	5.147506e-04
14	1.000000008891142	2.508180e-04
15	0.999999922487481	1.477628e-04
16	0.999999974884461	1.227807e-04
17	0.999999967175688	1.206788e-04
18	1.000000012033894	5.842547e-05
19	0.999999989277257	3.435677e-05
20	0.999999999503452	2.844560e-05
21	0.999999997925784	2.792907e-05
22	1.000000000814548	1.345917e-05
23	0.999999999304902	7.906797e-06
24	0.999999999980743	6.531921e-06
25	0.999999999874891	6.409349e-06
26	1.000000000051292	3.077943e-06
27	0.999999999957345	1.806914e-06
28	0.999999999999316	1.490408e-06
29	0.999999999992676	1.461831e-06
30	1.000000000003096	7.002595e-07

Table 4.1: In this table we can see the behaviour of the parameter and of the objective function of (3.2). The field used is (4.2). The starting point is given by $\tau_0 = -0.5$, $t_0 = 0.5$, $\xi_0 \approx (1.44, 1.96)^T$ and $p_0 = -1$. The point at iteration 30 is $\tau_f \approx -4.0$, $t_f \approx 4.0$, $\xi_f \approx (1.0, 2.0)^T$ and $p_f \approx 1$. Remember that from analytical results we know that for $p = 1$ there is a heteroclinic orbit connecting $(0, 2)^T$ and $(2, 2)^T$.

Iteration	p	Obj. Function
0	0.050000000000000	1.208019e-01
1	0.025617353419102	7.301090e-02
2	-0.054172909178207	5.190768e-02
3	-0.070845580797018	2.924967e-02
6	-0.001465554511265	1.301198e-02
9	-0.015077450368743	2.406843e-03
12	-0.020566247039409	7.416111e-04
15	-0.021526337818207	5.562179e-04
18	-0.023279407617198	2.233993e-04
21	-0.024098977041522	1.512871e-04
24	-0.025235193861268	5.290054e-05
27	-0.025307026458626	3.909746e-05
30	-0.025586522031574	1.857730e-05
33	-0.025704626673826	1.002380e-05
36	-0.025744778418354	9.723812e-06
39	-0.025816278257899	4.931695e-06
42	-0.025858402201518	2.930190e-06
45	-0.025889078280486	2.363904e-06
48	-0.025889209605272	2.361206e-06
51	-0.025910732214042	1.134499e-06
54	-0.025929540703318	7.207929e-07
57	-0.025935028465313	5.695806e-07
60	-0.025934809757395	5.690189e-07
63	-0.025941084051114	3.124348e-07
66	-0.025945059901117	2.336004e-07
69	-0.025947497810183	1.404993e-07
72	-0.025949030992367	1.373761e-07
75	-0.025949094799122	1.371522e-07
76	-0.025949094799137	1.371522e-07
78	-0.025949094799137	1.371522e-07

Table 4.2: This table shows the behaviour of the parameter side by side with the value of the objective function of (3.2) to find a heteroclinic orbit in the field f_2 of (4.3). The starting point is given by $\tau_0 = -0.2$, $t_0 = 0.2$, $\xi_0 \approx (-0.06, 0.09, 0.05, 0.08)^T$ and $p_0 = 0.05$. The point at iteration 78 is $\tau_f \approx -4.0$, $t_f \approx 4.0$, $\xi_f \approx (-0.03, 0.00, 0.07, 0.03)^T$ and $p_f \approx -0.025949$. For the same problem, the parameter for which there is a heteroclinic reported in [2] is -0.025955. The initial parameter used in [2] is -0.01, while here the initial parameter is 0.05.

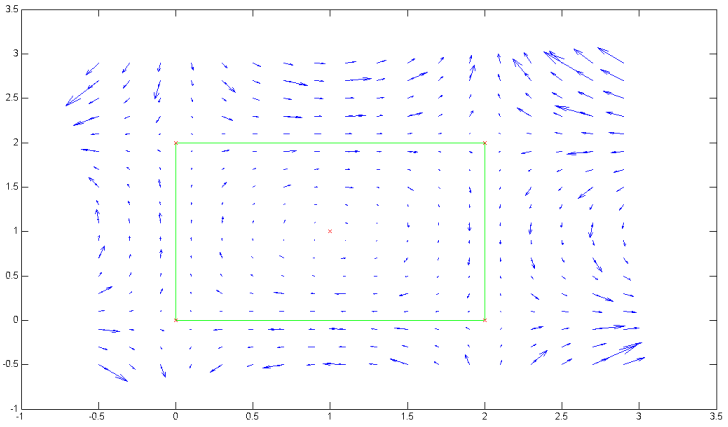

Figure 4.1: In this figure we show the field of (4.2) for $p = 1$. One sees that it possesses four saddle points at $(0,0)^T$, $(0,2)^T$, $(2,0)^T$ and $(2,2)^T$. It also has four heteroclinic orbits shown in green. One can prove the existence of the orbit connecting $(0,2)^T$ and $(2,2)^T$ taking the initial condition $(1,2)^T$ and reducing the analysis to a real ordinary differential equation. The heteroclinic orbit found in Figure 4.2 is the one the connects $(0,2)^T$ and $(2,2)^T$. We can also see a center at $(1,1)^T$. The parameter dependent field (4.2) does not have heteroclinic orbits for parameters $p \in B(1,\delta) - \{1\}$.

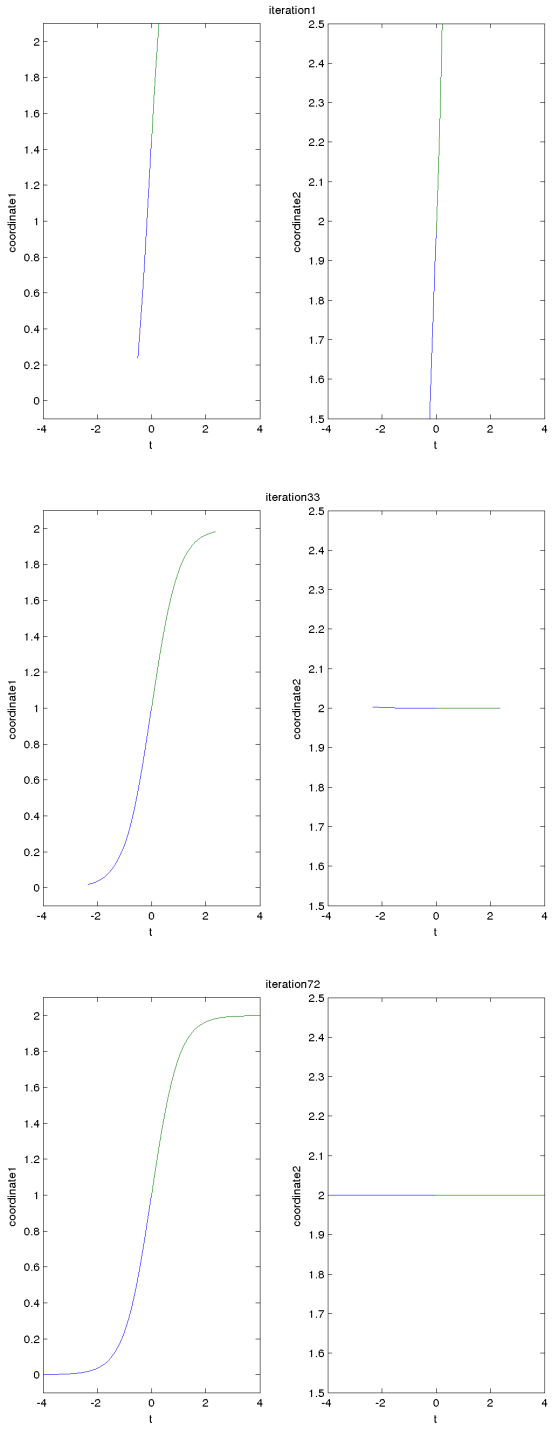

Figure 4.2: Each figure shows the coordinates of the orbits during the iterations of the Global Method to search heteroclinic orbits in (4.2). The starting point is $\tau_0 = -0.5$, $t_0 = 0.5$, $\xi_0 \approx (1.44, 1.96)^T$ and $p_0 = -1$. The final point is $\tau_f \approx -4.0$, $t_f \approx 4.0$, $\xi_f \approx (1.0, 2.0)^T$ and $p_f \approx 1$.

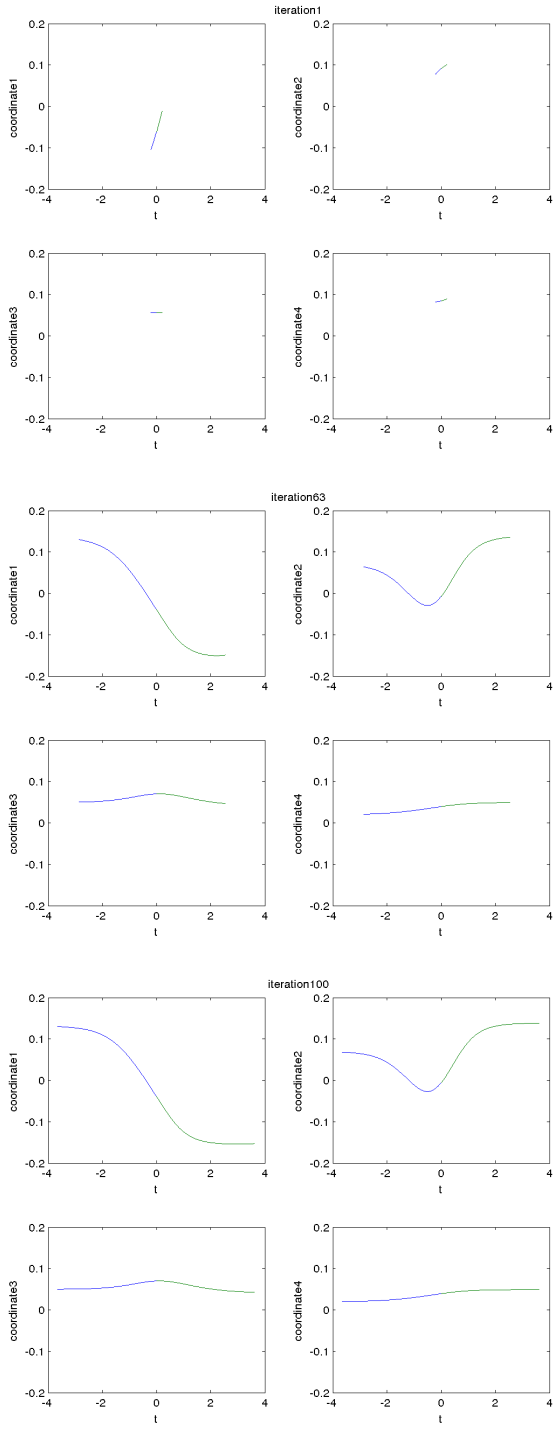

Figure 4.3: Each figure shows the coordinates of the orbits during the iterations of the Global Method to search heteroclinic orbits in the field f_2 of (4.3). Here the starting and final points are approximately the same of Table 4.2. Note that the coordinates of the orbits approach equilibria.

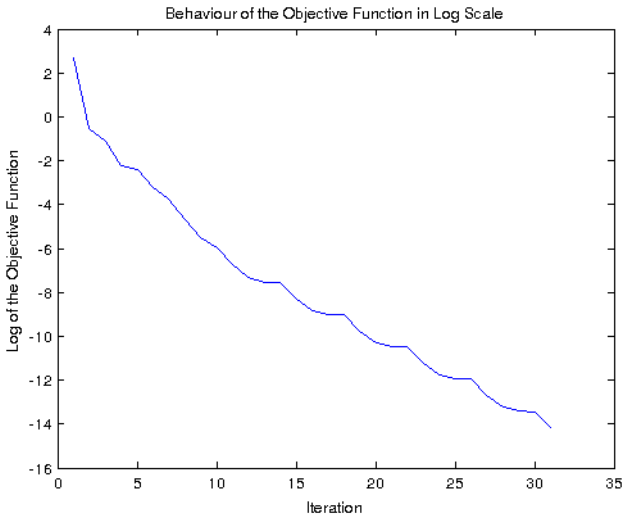

Figure 4.4: In this figure we show the values of the objective function of Table 4.1 in log scale.

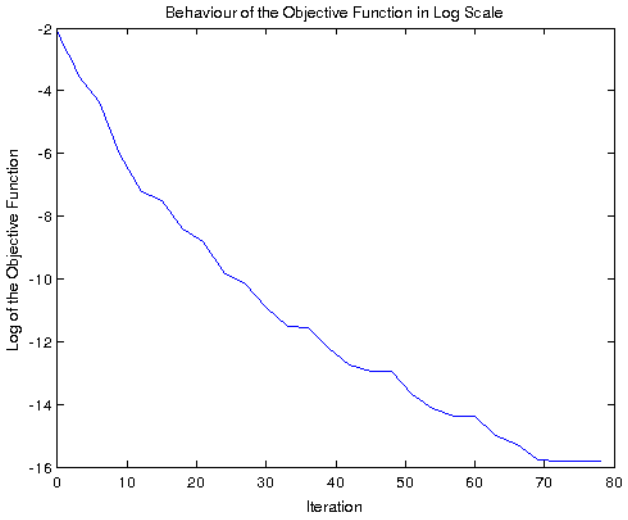

Figure 4.5: In this figure we show the values of the objective function of Table 4.2 in log scale. We think that due to lack of accuracy of the derivatives, the last 3 iterations of the optimization do not decrease the value of the objective function.

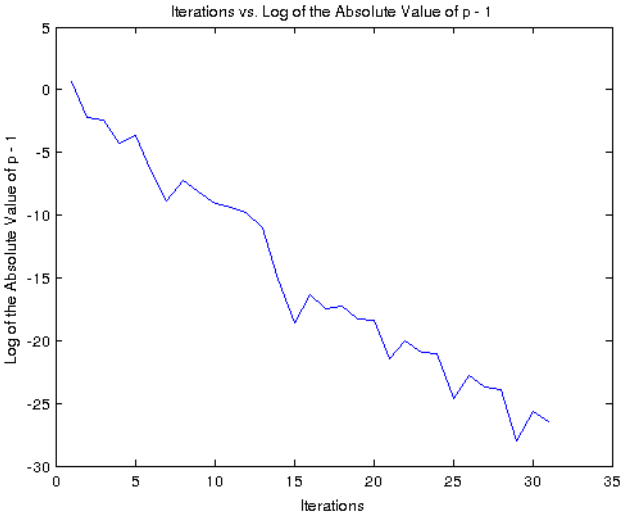

Figure 4.6: This figure shows evidence that the parameter converges linearly to 1 in the experiment shown in Table 4.1.

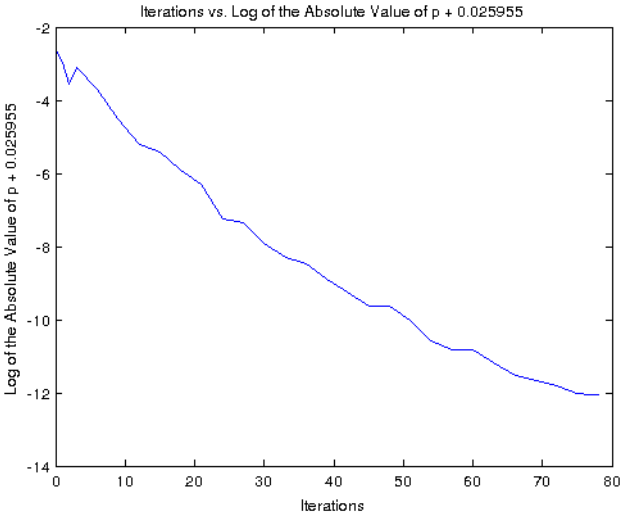

Figure 4.7: This figure shows evidence that the parameter converges linearly to -0.025955 in the experiment shown in Table 4.2.

4.2 Optimum Parameter Experiments

In this section we provide a parameter dependent field for which we can calculate analytically the regions of parameters that give rise to heteroclinic orbits. Then we compare the results given by the Optimum Parameter Method and by the analytical results. Let $f : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}^2$ be the function given by (4.4). Note that the second coordinate of (4.4) does not depend on p:

$$f(x,p) = \begin{bmatrix} x_1(x_1-2)(1-x_2) + (1-p) \\ x_2(x_2-2)(x_1-1) \end{bmatrix} \tag{4.4}$$

Consider as given $p \in \mathbb{R}$. If $\hat{x} \in \mathbb{R}^2$ satisfies $f(\hat{x},p) = 0$, then $\hat{x}_2 = 0$ or $\hat{x}_2 = 2$ or $\hat{x}_1 = 1$. Using these values of the coordinates of $\hat{x}$ in the first coordinate of (4.4), we get a quadratic equation through which we can find the undetermined coordinate of $\hat{x}$. For the undetermined coordinate of $\hat{x}$ to be real a condition in the parameter p appears. We prove that every $p \in \mathbb{R}$ satisfying this condition gives rise to a heteroclinic orbit.

Note that if $\hat{x}_1 = 1$, then $\hat{x}_2 = p$. If $\hat{x}_2 = 0$, then $\hat{x}_1 = 1+\sqrt{p}$ or $\hat{x}_1 = 1-\sqrt{p}$. If $\hat{x}_2 = 2$, then $\hat{x}_1 = 1+\sqrt{2-p}$ or $\hat{x}_1 = 1-\sqrt{2-p}$. Observe that if $p \in (0,2)$, then (4.4) has five equilibria. If $p \in [2,\infty)$ or $p \in (-\infty,0]$, then it has 3 equilibria. In this section we shall consider a different definition of heteroclinic orbit in which the equilibria can be equal. This is necessary to conclude the existence of minimum parameters.

We call $A = (1-\sqrt{2-p},2)^T, B = (1+\sqrt{2-p},2)^T, C = (1-\sqrt{p},0)^T$ and $D = (1+\sqrt{p},0)^T$. Considering the initial condition $(1,2)^T$ we can conclude that for $p \leq 2$ there is a heteroclinic orbit between A and B. The initial condition $(1,0)^T$ allows us to conclude the same for C and D for $p \geq 0$.

Note that there is a maximum (resp. minimum) parameter for which there is a heteroclinic orbit connecting A and B (resp. C and D). To identify numerically these parameters we introduce the following cost functions. Let $\chi_{A,B} : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}$, $\chi_{C,D} : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}$ be given by (4.5). Note that $\chi_{A,B}$ (resp. $\chi_{C,D}$) has a minimum for $p \leq 2$ (resp. $p \geq 0$) although this set is not compact:

$$\begin{aligned} \chi_{A,B}(\xi,p) &= -p \\ \chi_{C,D}(\xi,p) &= p \end{aligned} \tag{4.5}$$

We now report the numerical experiments. As before, we used *ode45* Matlab solver to integrate the initial value problems (2.4), (2.5), (2.6), (2.15), (2.20) and (2.25) in the coupled manner. Again, to solve (3.17) we used the *fmincon* Matlab solver with the implementation of the Interior Point Algorithm provided by the Optimization Matlab Toolbox. The tolerances *TolX*, *TolFun*, *AbsTol* and *RelTol* are equal to the ones used in Section 4.1. The time increment in integrations is 0.01.

In Table 4.3 we present the parameter and the value of the objective function of (3.17) with $\chi = \chi_{A,B}$ to search the maximum parameter for which there is a heteroclinic orbit connecting A and B. The starting and final points of the experiment are given in the caption of Table 4.3. As expected we can see numerical instabilities when the solver approaches the minimum. In Table 4.4 we present analogous results when $\chi = \chi_{C,D}$. The numerical and analytical results agree.

Iteration	p	Obj. Function
0	1.000000000000000	-1.000000e+00
1	1.994997512500000	-1.994998e+00
2	1.997171840516523	-1.997172e+00
3	1.999021079391932	-1.999021e+00
4	1.999015465925340	-1.999015e+00
5	1.997226242241598	-1.997226e+00
6	1.997226288399357	-1.997226e+00
7	1.998308528462846	-1.998309e+00
8	1.997535533378942	-1.997536e+00
9	1.999003131042513	-1.999003e+00
10	1.999170026806225	-1.999170e+00

Table 4.3: In this table we show the values of the parameter p of (3.17) with $\chi = \chi_{A,B}$ and solutions given by the field (4.4). The equilibria are $\{(1 - \sqrt{2-p}, 2)^T, (1 + \sqrt{2-p}, 2)^T\}$. We used $\epsilon = 10^{-8}$ and $\widehat{\epsilon} = 10^{-5}$. The optimum value of p is 2. The cost function $\chi_{A,B}$ sets the optimization problem (3.17) to find the **maximum** parameter p for which there is a heteroclinic orbit. The starting point is $\tau_0 = -2$, $t_0 = 2$, $\xi_0 = (1.0, 2.0)^T$ and $p_0 = 1.0$. The final point is $\tau_f \approx -7.5$, $t_f \approx 7.5$, $\xi_f \approx (1.0, 2.0)^T$ and $p_f \approx 1.999$.

Iteration	p	Obj. Function
0	1.000000000000000	1.000000e+00
1	0.005002487500000	5.002487e-03
2	0.033013915085243	3.301392e-02
3	0.001129737941614	1.129738e-03
4	7.117609946982290e-04	7.117610e-04
5	7.462635416566996e-04	7.462635e-04
6	7.462656019680963e-04	7.462656e-04
7	7.111138979676881e-04	7.111139e-04
8	7.553894403250936e-04	7.553894e-04
9	7.553896994937028e-04	7.553897e-04
10	7.300423854943774e-04	7.300424e-04

Table 4.4: In this table we show the values of the parameter p of (3.17) with $\chi = \chi_{C,D}$ and solutions given by the field (4.4). The equilibria are $\{(1 - \sqrt{p}, 0)^T, (1 + \sqrt{p}, 0)^T\}$. We used $\epsilon = 10^{-8}$ and $\widehat{\epsilon} = 10^{-5}$. The optimum value of p is 0. The cost function $\chi_{C,D}$ sets the optimization problem (3.17) to find the **minimum** parameter p for which there is a heteroclinic orbit. The starting point is $\tau_0 = 1.0$, $t_0 = -1.0$, $\xi_0 \approx (1.0, 0.0)^T$ and $p_0 = 1.0$. The final point is $\tau_f \approx 3.3$, $t_f \approx -3.3$, $\xi_f \approx (1.0, 0.0)^T$ and $p_f \approx 0.0$.

4.3 Combustion Waves in Porous Medium

Let us present a model introduced in [13] for the injection of air into a porous medium that contains solid fuel. We consider three dimensionless variables: θ, ρ and Y. The variables ρ and Y stand, respectively, for the amount of fuel and oxygen and θ for the temperature. The model for the joint behaviour of fuel, oxygen and temperature is given by (4.6):

$$\begin{cases} \partial_t\theta + a\partial_x\theta = \partial_{xx}\theta + \rho Y\Phi, \\ \partial_t\rho = -\rho Y\Phi, \\ \partial_t Y + a\partial_x Y = -\rho Y\Phi. \end{cases} \tag{4.6}$$

$$\Phi(\theta) = \begin{cases} exp(\frac{-1}{\theta}), & \theta > 0 \\ 0, & \theta \leq 0 \end{cases}. \tag{4.7}$$

The idea of the model above is that heated air is injected into a porous medium with velocity a. When the oxygen enters in contact with the fuel the combustion occurs and releases energy that is a function of the amount of fuel, oxygen and temperature. The energy coming from combustion increases the temperature. The amount of fuel and oxygen decrease in time proportionally to the energy liberated in combustion. For details see [13].

We want solutions of (4.6) in the form of travelling waves. More precisely, we want to find a velocity $c \in \mathbb{R}$ and a curve $\widehat{v} : \mathbb{R} \longrightarrow \mathbb{R}^3, \quad \widehat{v} \in C^2$ such that $\lim_{t\longrightarrow\infty} \widehat{v}(t)$ and $\lim_{t\longrightarrow-\infty} \widehat{v}(t)$ exist, are finite and $v(x,t) = \widehat{v}(x-ct)$ is a solution of (4.6).

Imposing such conditions, it was shown in [13] that finding such travelling waves is equivalent to find a velocity c and constants $w_1, w_2 \in \mathbb{R}$ such that (4.8) has a heteroclinic orbit. The coordinate of the travelling wave corresponding to the oxygen does not enter in (4.8) because it is automatically determined by ρ:

$$\begin{cases} \dfrac{\partial\theta}{\partial t} = (a-c)\theta - c\rho + w_1, \\ \dfrac{\partial\rho}{\partial t} = \dfrac{c\rho + w_2}{c(c-a)}\rho\Phi(\theta). \end{cases} \tag{4.8}$$

Following [13], we call the travelling waves whose velocities are greater than a fast travelling waves. When we fix that (θ^+, ρ^+, Y^+) is a steady state of the fast travelling wave, we can determine constants w_1 and w_2. Then, (4.8) becomes (4.9). If $\theta^+ \leq 0$, then the point (θ^+, ρ^+) is an equilibrium of (4.9) for every $c \neq a$:

$$\begin{cases} \dfrac{\partial\theta}{\partial t} = (a-c)(\theta-\theta^+) - c(\rho-\rho^+) \\ \dfrac{\partial\rho}{\partial t} = (\dfrac{\rho-\rho^+}{c-a} + \dfrac{Y^+}{c})\rho\Phi(\theta) \end{cases} \tag{4.9}$$

Note that for a fixed state (θ^+, ρ^+, Y^+), the ordinary differential equation (4.9) has only one real parameter. Theorem 4.3.1, which is proved in [13], answers for which velocity (4.9) has a heteroclinic orbit for a given point (θ^+, ρ^+, Y^+) satisfying certain conditions:

Theorem 4.3.1. *Let $a > 0$. Assume we are given $s^+ = (\theta^+, \rho^+, Y^+)$ satisfying $\theta^+ \leq 0, \rho^+ > 0, Y^+ > 0$ and $\theta^+ + Y^+ > 0$. Then, there is a state $s^- = (\theta^-, \rho^-, Y^-)$ such that $\theta^- > 0$ and a travelling wave with velocity c of (4.6) connecting states s^+ and s^-, where*

$$c = \frac{aY^+ - aY^-}{Y^+ - Y^- + \rho^- - \rho^+}. \tag{4.10}$$

In Figure 7 of [13] there is a travelling wave connecting states s^+ and s^- with $\theta^+ = -0.1, \rho^+ = 0.1, Y^+ = 1.5, \theta^- = 1.06, \rho^- = 0$ and $Y^- = 0.34$. The parameter a is equal to 5. The analytical velocity of this travelling wave given by Theorem 4.3.1 is 5.4716. In what follows consider fixed θ^+, ρ^+ and Y^+ equal to the numerical values informed. Our objective is to find the travelling wave of Figure 7 using the Global Method.

To use the Global Method we need to provide the values of the equilibria and their first and second order derivatives. To provide this information, we give parametrizations (4.11) and (4.12) of equilibria of (4.9) in terms of the velocity c. Note that $\hat{y}(5.4716) \approx (1.06, 0)^T$. As can be seen in Proposition 6.1 of [13], the equilibrium $\hat{x}$ is not hyperbolic:

$$\hat{x}(c) = (\theta^+, \rho^+) = (-0.1, 0.1)^T \tag{4.11}$$

$$\hat{y}(c) = (0, \theta^+ + \frac{c}{c-a}\rho^+)^T = (-0.1 + \frac{0.1c}{c-5}, 0)^T \tag{4.12}$$

To use the Global Method to identify the velocity that gives rise to a heteroclinic orbit connecting $\hat{x}$ and $\hat{y}$, we need to compute numerically the solutions of (4.9) as well as their first and second order derivatives. However, to solve initial value problems (2.4), (2.5), (2.6), (2.15), (2.20) and (2.25) in the coupled manner using our implementation results in several truncations.

The truncation occurs in our implementation because when we use the *ode45* solver to find solutions and their derivatives, the coordinates of the approximations of the solution of the corresponding ordinary differential equation are stored in the same vector and derivatives of different orders are of very different magnitudes. Then, to represent numbers of different magnitudes the *ode45* solver neglects the smaller ones.

For instance take $t_0 = -1$, $\xi_0 \approx (0.48, 0.05)^T$ and $p_0 = 5.4716$. The solution is of the order of 10^0, first order derivatives with respect to ξ and p of the solution are of the order of 10^{10} and second order derivatives with respect to ξ and p are of the order of 10^{18}. The mentioned t_0, ξ_0 and p_0 correspond to the green part of the solution shown in Figure 4.8. From this data we can see how sensible is (4.9). Derivatives are calculated numerically.

We use a simple strategy to overcome this problem. First step is to solve (2.4) to find solutions. The second step consists in solving (2.4), (2.5), (2.6) in the coupled manner. In the third step we solve the system given by (2.4), (2.5), (2.6), (2.15), (2.20) and (2.25). Then we keep the solutions of the first step, the first order derivatives of the second step and the second order derivatives given by the third step. This strategy gives reasonable precision.

In our experiments we used the *ode45* integrator with absolute tolerance equal to 10^{-16} and relative tolerance equal to 10^{-8}. See the caption of Figure 4.9 for an experiment with other tolerances. The time increment used in integrations

is 0.01. We used the Matlab solver $fmincon$ with the implementation of the Interior Point Algorithm provided by the Optimization Matlab Toolbox. The tolerances $TolX$ and $TolFun$ of $fmincon$ are equal to 10^{-30}. In all experiments we multiplied (4.9) by 50.

We now analyse numerical results of tables (4.5) and (4.6). The starting points of the experiments of these tables are equal except for the value of initial velocity. The initial velocity in Table (4.5) is 5.2 and in Table (4.6) is 5.7. The initial value of ξ_0 is infeasible for (3.2), but the $fmincon$ solver restores feasibility. In both tables, the numerical velocities approximate the analytical value given by Theorem 4.3.1 although the derivatives provided are not extremely precise.

Comparing values of the objective functions of tables 4.5 and 4.6 we see that the value of the objective function at iteration 390 in Table 4.6 is a little higher than the one at iteration 150 in Table 4.5. Pay attention to the scales of figures 4.10 and 4.11. From tables 4.5 and 4.6 we see that the method should converge somewhere between 5.465 and 5.475, which are good bounds to the analytical value of the velocity that gives rise to a travelling wave. The relative error for the velocity is approximately 0.1%.

From tables 4.5 and 4.6 we realize that the final values of the objective function are high compared to the vertical distance of the equilibria, which is 0.1. This relatively high discrepancy is partially explained by the magnitude of final times in experiments of tables 4.5 and 4.6. More than 82% of the final value of the objective function in table 4.5 comes from the distance of the orbit to the non-hyperbolic equilibria. See figures 4.10 and 4.11 for details. The orbits approach the blue equilibrium very slowly and this prevents the objective function of the Global Method to go to zero faster. See Figure 4.12.

The initial conditions of the final orbits of experiments of tables 4.5 and 4.6 reasonable approximate the one of Figure 4.8, which we found using a bisection process and is the best result possible using Matlab double precision. The norms of the difference between the initial condition used in Figure 4.8 and the final initial conditions of experiments of tables 4.5 and 4.6 are approximately 0.01. The relative errors are 10% when we compare to the vertical distance between the equilibria, which is 0.1. Comparing to the euclidean distance between the equilibria, the relative errors in the initial conditions given by the Global Method are less than 1%.

In figures 4.13 and 4.14 we show the log of the absolute value of the $c-5.4716$ of experiments of tables 4.5 and 4.6. In Figure 4.13 we see that at the final iterations the convergence rate of the parameter is slower. Remember that the influence of the non-hyperbolic equilibria on the value of the objective function of the first experiment is high. Based in numerical experiments, the opinion of the author is that this behaviour is related the slow approximation of the orbit to the non-hyperbolic equilibria and also related to the lack of accuracy of the derivatives.

We now summarize. The objective function does not go to zero faster because the orbits approach the non-hyperbolic equilibria very slowly. The equilibrium $\hat{y}$ shown in green is a saddle. This makes the corresponding green parts of the orbits much more sensible to initial conditions. The initial condition of the best result possible using Matlab double precision is 0.01 close to the ones calculated by the Global Method. The velocities approach 5.4716 starting from below and from above.

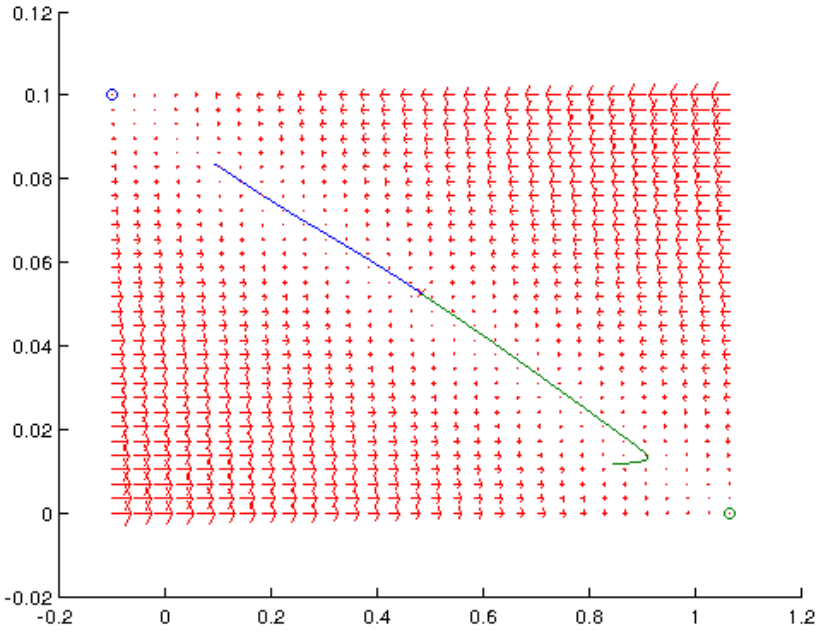

Figure 4.8: In this figure we show a solution of (4.9) for $c = 5.4716$. The equilibria are shown at $(-0.10, 0.10)^T$ and $(1.06, 0.00)^T$. The initial condition of the orbit is approximately $(0.4821, 0.0525)^T$. The blue line is the solution integrated in the positive direction until $t = 50$ and the green line is the solution integrated in the negative direction until $t = -1$. The solution approaches the blue equilibria very slowly. The absolute tolerance used in *ode45* is equal to 10^{-16} and relative tolerance is equal to 10^{-8}. We use a bisection process to generate this figure parametrizing initial conditions with a real parameter in a plane separating the equilibria.

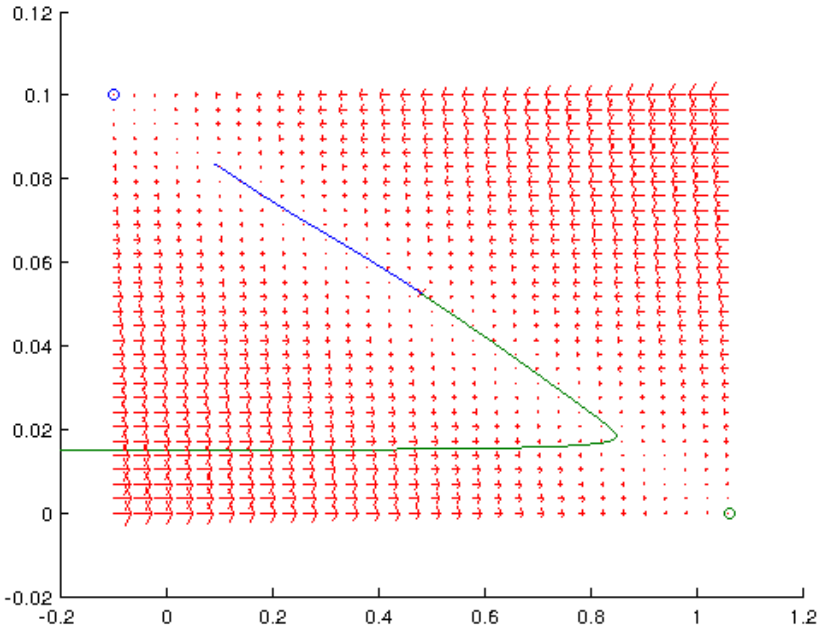

Figure 4.9: This figure shows an orbit calculated with the same initial condition and parameter of Figure 4.8. We also integrated until $t = 50$ in the positive direction (blue line) and until $t = -1$ in the negative direction (green line). The difference is that the absolute tolerance used in *ode45* is equal to 10^{-10} and the relative tolerance is equal to 10^{-7}. Comparing with Figure 4.8 we conclude that precise calculations are important to search heteroclinic orbits in (4.9).

Iteration	c	Obj. Function
0	5.200000000000000	9.933819e+00
5	5.240833797211176	1.166413e-02
10	5.252658514183594	5.905647e-02
15	5.264482918294963	5.894367e-02
20	5.276306713363986	5.591687e-02
25	5.288129779819193	5.311393e-02
30	5.299951993672620	5.080901e-02
35	5.311773216445840	4.892747e-02
40	5.323593293411717	4.738764e-02
45	5.335412051924813	4.612250e-02
50	5.347229299666272	4.507895e-02
55	5.359044822773402	4.421517e-02
60	5.370858383833522	4.349806e-02
65	5.382669719722919	4.290130e-02
70	5.394478539271828	4.240384e-02
75	5.406284520735760	4.198874e-02
80	5.418087309052199	4.164232e-02
85	5.429886512860119	4.135341e-02
90	5.441681701257576	4.111292e-02
95	5.451971199820020	3.964036e-02
100	5.458306701363148	3.877458e-02
105	5.462051087070244	3.848883e-02
110	5.464227910356054	3.840030e-02
115	5.465481094997177	3.837532e-02
120	5.466198440849154	3.836960e-02
125	5.466607713453047	3.836913e-02
130	5.466840778494809	3.836978e-02
135	5.466973356851321	3.837044e-02
140	5.467048726944483	3.837091e-02
145	5.467091559006350	3.837121e-02
150	5.467115894711891	3.837139e-02

Table 4.5: In this table we can see the behaviour of the velocity and of the objective function of (3.2). The starting point is given by $\tau_0 = 50$, $t_0 = -0.1$, $\xi_0 \approx (0.48, 0.05)^T$ and $p_0 = 5.2$. The point at iteration 150 is $\tau_f \approx 50$, $t_f \approx -0.24$, $\xi_f \approx (0.48, 0.05)^T$ and $p_f \approx 5.46$. The analytical parameter is 5.4716. We think that the difference of the analytical and numerical velocities comes from inaccuracies in the numerical calculation of derivatives of solutions.

Iteration	c	Obj. Function
0	5.700000000000000	2.745988e+01
13	5.598469643893196	4.463363e-02
26	5.531416384441518	4.316198e-02
39	5.510252692596054	3.979338e-02
52	5.504708015230803	3.959280e-02
65	5.501625515340669	3.952105e-02
78	5.499056248422352	3.946133e-02
91	5.497418517922921	3.942163e-02
104	5.496008258309351	3.938791e-02
117	5.494797947978943	3.935942e-02
130	5.493630200851647	3.933136e-02
143	5.492458630079092	3.930358e-02
156	5.486090939452496	3.915839e-02
169	5.484991217935784	3.913413e-02
182	5.484146560212103	3.911836e-02
195	5.483541312135345	3.910529e-02
208	5.482943244763810	3.909230e-02
221	5.482345180443551	3.907935e-02
234	5.481746876236924	3.906645e-02
247	5.481148367114482	3.905359e-02
260	5.480549697975461	3.904077e-02
273	5.479950912874873	3.902799e-02
286	5.479398122283584	3.901624e-02
299	5.478845056237846	3.900450e-02
312	5.478243699453054	3.899178e-02
325	5.477642043259703	3.897910e-02
338	5.477040454737247	3.896646e-02
351	5.476485246741614	3.895483e-02
364	5.475883919464934	3.894226e-02
377	5.475282780862729	3.892973e-02
390	5.474681865512331	3.891724e-02

Table 4.6: In this table we can see the behaviour of the velocity and of the objective function of (3.2). The starting point is given by $\tau_0 = 50$, $t_0 = -0.1$, $\xi_0 \approx (0.48, 0.05)^T$ and $p_0 = 5.7$. The point at iteration 390 is $\tau_f \approx 50$, $t_f \approx -0.15$, $\xi_f \approx (0.47, 0.05)^T$ and $p_f \approx 5.47$. The analytical parameter is 5.4716. We think that the difference of the analytical and numerical velocities comes from inaccuracies in the numerical calculation of derivatives of solutions.

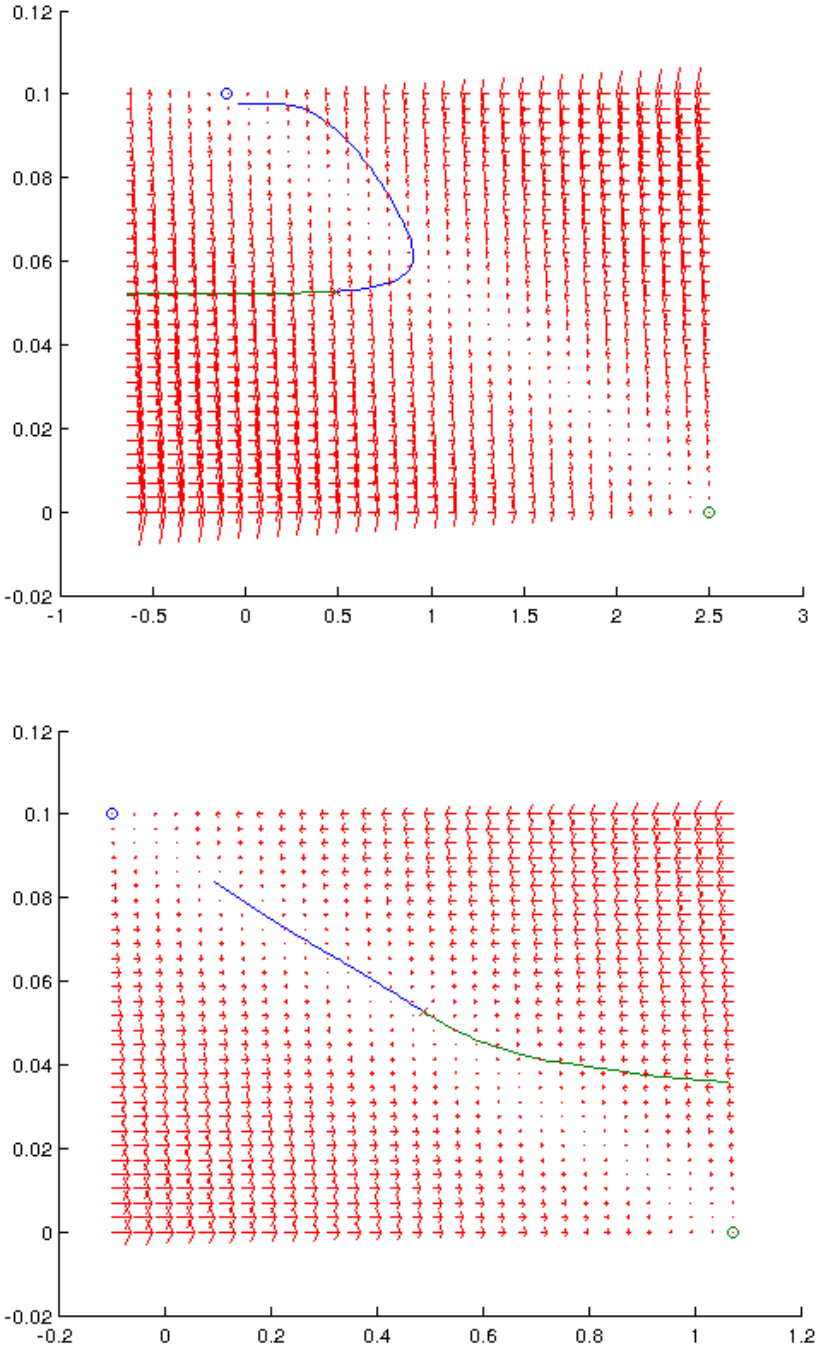

Figure 4.10: The figure in the top (resp. bottom) is the orbit in the first (resp. last) iteration of the experiment of Table 4.5. The orbits are shown in the phase space. The circles represent equilibria. Compare with Figure 4.8. We present initial and final parameters in the caption of Table 4.5.

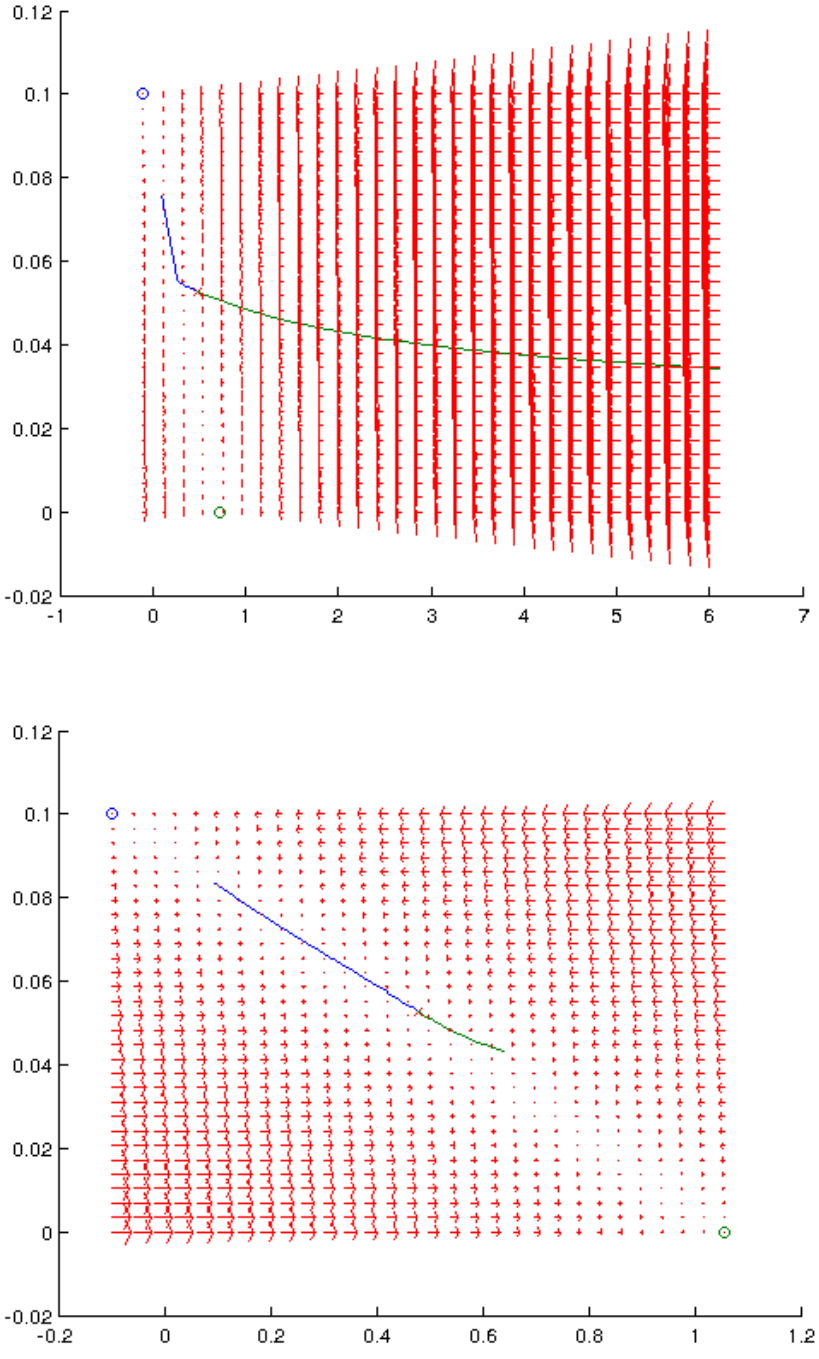

Figure 4.11: The figure in the top (resp. bottom) is the orbit in the first (resp. last) iteration of the experiment of Table 4.6. The orbits are shown in the phase space. The circles represent equilibria. The final orbit is close to the one in Figure 4.8. We present initial and final parameters in the caption of Table 4.6.

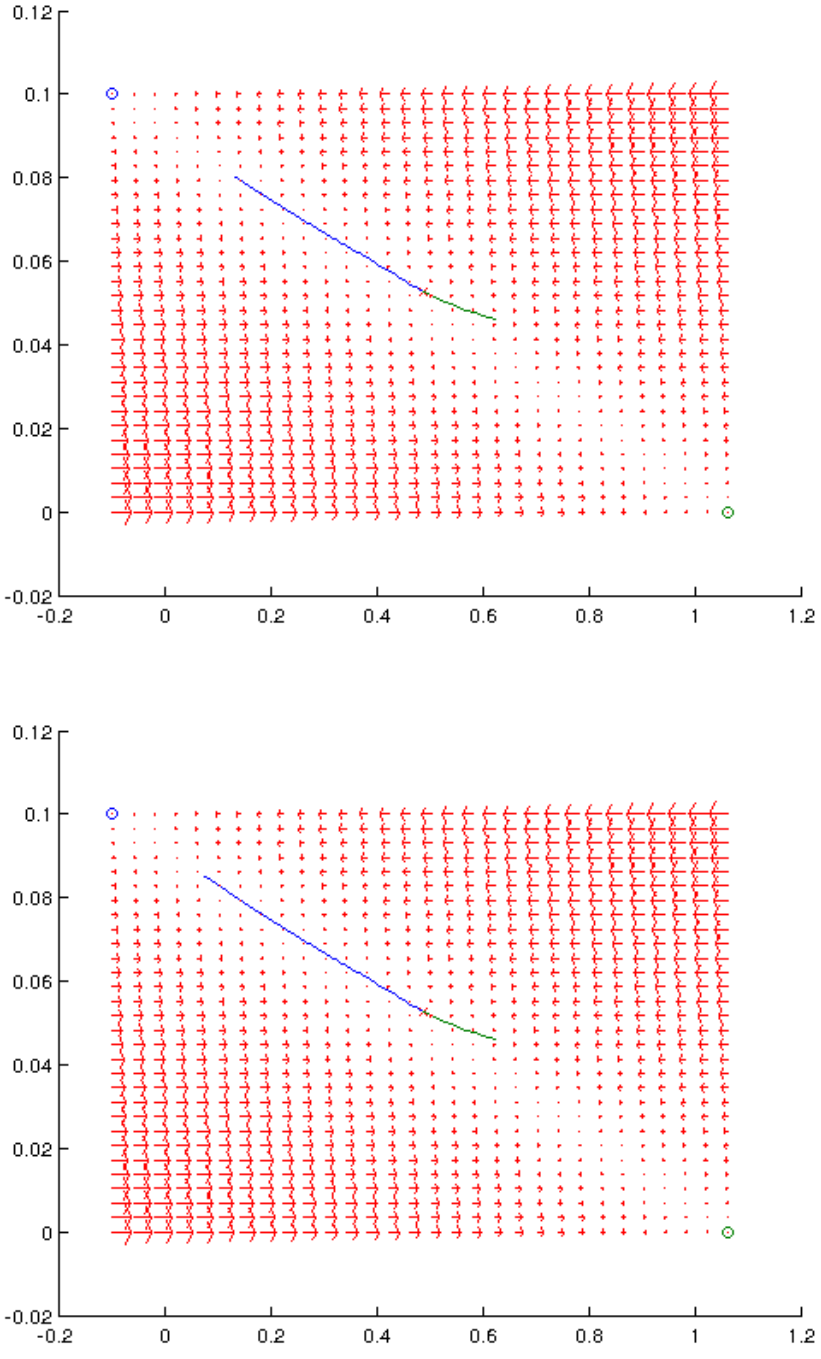

Figure 4.12: The initial condition are $(0.48, 0.05)^T$ and the parameter is equal to 5.4716. The blue part of the orbit in the top (resp. bottom) is obtained integrating until 5 (resp. 500). The orbit approaches the blue equilibrium very slowly and this prevents the objective fuction to go to zero faster.

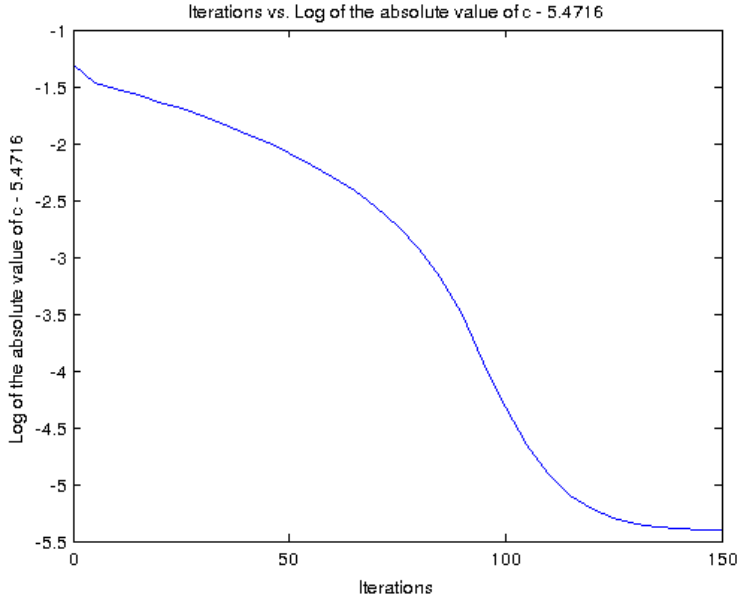

Figure 4.13: This figure shows the convergence behaviour of the velocity c in the experiment of Table 4.5.

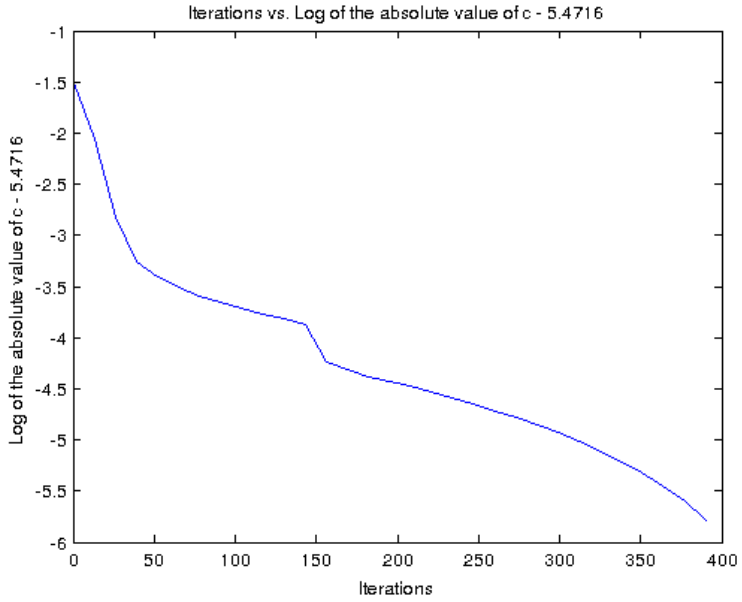

Figure 4.14: This figure shows the convergence behaviour of the velocity c in the experiment of Table 4.6.

4.4 Conclusion

In Section 4.1 we use the Global Method for locating heteroclinic orbits in two systems of ordinary differential equations. The numerical results agree with the analytical ones for the case of (4.2) and with the numerical ones available in [2] for (4.3). The simultaneous optimization over times, initial conditions and parameters reduce the need of human interference when locating heteroclinic orbits.

Section 4.2 shows the use of the Optimum Parameter Method to find minimum and maximum parameters that give rise to heteroclinic orbits. The numerical and analytical results agree. The method experienced numerical instabilities near the optimums. We left as a future work to improve the performance of the method. Probably the Optimum Parameter Method can be largely improved considering simpler bifurcation behaviour of the equilibria. We also left as a future work to test it in more complex systems.

In Section 4.3 we successfully use the Global Method to search heteroclinic orbits in a system modelling combustion in a porous medium. This experiment shows the robustness and usefulness of the Global Method to deal with non-hyperbolic equilibria. We left as a future work to improve the calculations of derivatives of solutions and to run again the experiments to see how precise the method can be.

Appendices

Appendix A

Convergence Theorem for the Global Method

Assume we have a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field, an open set S and also functions $\hat{x}, \hat{y} : S \subset \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n$ with $\hat{x}, \hat{y} \in C^k$ satisfying $f(\hat{x}(p), p) = 0$, $f(\hat{y}(p), p) = 0$ and $\hat{x}(p) \neq \hat{y}(p)$ for all $p \in S$. The set S is contained in the parameter space of f. The functions $\hat{x}$ and $\hat{y}$ are parametrizations of equilibria. Let $x(t, \xi, p)$ be the unique solution given by Cor. 2.1.2:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{A.1}$$

The next lemma uses the Hartman-Grobman Theorem to state that if a certain orbit converges to a hyperbolic equilibria in a sequence of times going to infinity, then it converges to the mentioned equilibria in every sequence of times going to infinity.

Lemma A.0.1. *Let $p^* \in S$. Assume that the equilibria $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are hyperbolic. Assume that there are sequences $\tau_n \longrightarrow \infty$ and $t_n \longrightarrow -\infty$ such that $x(\tau_n, \xi^*, p^*) \longrightarrow \hat{x}(p^*)$ and $x(t_n, \xi^*, p^*) \longrightarrow \hat{y}(p^*)$ for a certain $\xi^* \in \mathbb{R}^n$. Then, $\lim_{\tau \longrightarrow \infty} x(\tau, \xi^*, p^*) = \hat{x}(p^*)$ and $\lim_{t \longrightarrow -\infty} x(t, \xi^*, p^*) = \hat{y}(p^*)$.*

Proof. Let us denote by A the spatial derivative of f at (ξ^*, p^*). By the Hartman-Grobman Theorem (see [17]), there is an open set $N \subset \mathbb{R}^n, \quad \hat{x}(p^*) \in N$ and an homeomorphism $h : N \longrightarrow \mathbb{R}^n, \quad h(\hat{x}(p^*)) = 0$ such that $h(x(t, \xi, p^*)) = e^{tA} h(\xi) \quad \forall \xi \in N$. Take n_0 such that $x(\tau_{n_0}, \xi^*, p^*) \in N$ and define $\xi_0 = x(\tau_{n_0}, \xi^*, p^*)$. As the field f is autonomous, $x(t, \xi_0, p^*) = x(t + \tau_{n_0}, \xi^*, p^*)$. Therefore, without loss of generality we can assume that $x(\tau_n, \xi_0, p^*) \longrightarrow \hat{x}(p^*)$. Note that $h(x(\tau_n, \xi_0, p^*)) = e^{\tau_n A} h(\xi_0)$. Taking limits in the last equality we conclude that $\lim_n e^{\tau_n A} h(\xi_0) = 0$. Therefore, $h(\xi_0)$ belongs to the stable space of $\hat{x}(p^*)$. Take a sequence $\widehat{\tau}_n \longrightarrow \infty$. Note that $x(\widehat{\tau}_n, \xi_0, p^*) = h^{-1}(e^{\widehat{\tau}_n A} h(\xi_0))$. Taking limits in the last expression we have $\lim_n x(\widehat{\tau}_n, \xi_0, p^*) = h^{-1}(0) = \hat{x}(p^*)$. Therefore $\lim_{\tau \longrightarrow \infty} x(\tau, \xi^*, p^*) = \hat{x}(p^*)$. The proof for $\hat{y}$ is analogous. $\square$

Theorem A.0.2. *Assume we have sequences $\tau_n \longrightarrow \infty$, $t_n \longrightarrow -\infty$, $(\xi_n, p_n) \longrightarrow (\xi^*, p^*) \in \mathbb{R}^n \times S$ and that the objective function of (3.2) goes to zero over the sequence $z_n = (\tau_n, t_n, \xi_n, p_n)$. The sequence z_n is supposed to be in the feasible set of (3.2). Assume also that*

$$\begin{aligned} \parallel x(\tau_n, \xi^*, p^*) - x(\tau_n, \xi_n, p_n) \parallel &\longrightarrow 0, \\ \parallel x(t_n, \xi^*, p^*) - x(t_n, \xi_n, p_n) \parallel &\longrightarrow 0. \end{aligned} \tag{A.2}$$

If the equilibria $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are hyperbolic, then the orbit $x(t, \xi^*, p^*)$ is heteroclinic. Independently of the type of the equilibria, $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are accumulation points of the orbit $x(t, \xi^*, p^*)$.

Proof. Using the triangle inequality we have:

$$\begin{aligned} \parallel x(\tau_n, \xi^*, p^*) - \hat{x}(p^*) \parallel \leq & \parallel x(\tau_n, \xi^*, p^*) - x(\tau_n, \xi_n, p_n) \parallel \\ + & \parallel x(\tau_n, \xi_n, p_n) - \hat{x}(p_n) \parallel \\ + & \parallel \hat{x}(p_n) - \hat{x}(p^*) \parallel . \end{aligned} \tag{A.3}$$

The objective function of (3.2) dominates $\parallel x(\tau_n, \xi_n, p_n) - \hat{x}(p_n) \parallel^2$ and goes to zero over the sequence z_n, then $\parallel x(\tau_n, \xi_n, p_n) - \hat{x}(p_n) \parallel \longrightarrow 0$. Taking limits in the inequality above we have $\parallel x(\tau_n, \xi^*, p^*) - \hat{x}(p^*) \parallel \longrightarrow 0$. Analogously, $\parallel x(t_n, \xi^*, p^*) - \hat{y}(p^*) \parallel \longrightarrow 0$. If the equilibria $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are hyperbolic, by Lemma A.0.1 we conclude that $\lim_{\tau \longrightarrow \infty} x(\tau, \xi^*, p^*) = \hat{x}(p^*)$ and $\lim_{t \longrightarrow -\infty} x(t, \xi^*, p^*) = \hat{y}(p^*)$. As $\hat{x}(p^*) \neq \hat{y}(p^*)$, the orbit $x(t, \xi^*, p^*)$ is heteroclinic. □

Theorem A.0.2 shows that the equilibria $\hat{x}(p^*)$ and $\hat{y}(p^*)$ are accumulation points of the orbit $x(t, \xi^*, p^*)$. We can use the Krasovskii-LaSalle Theorem to generalize Theorem A.0.2 for non-hyperbolic equilibria . We left this analysis as a future work. We think that Theorem A.0.2 can provide useful information to the development of special minimization methods to solve (3.2). We are missing conditions under which (A.2) holds. We also do not know if it is possible to obtain the sequences τ_n, t_n, ξ_n and p_n when the heteroclinic orbit exists.

Appendix B

Reformulated Method of Mailybaev et al

During our studies of [2] we noticed that there was room for improvements, which we present here. Throughout this section assume we are given a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field, an open set $S \subset \mathbb{R}^{\widetilde{p}}$ and two functions $\hat{x}, \hat{y} : S \longrightarrow \mathbb{R}^n$ with $\hat{x}, \hat{y} \in C^k$ satisfying $f(\hat{x}(p), p) = 0$, $f(\hat{y}(p), p) = 0$ and $\hat{x}(p) \neq \hat{y}(p)$ for all $p \in S$. In other words, the functions $\hat{x}$ and $\hat{y}$ are parametrizations of equilibria.

We also assume that for each $p \in S$ the equilibria $\hat{x}, \hat{y}$ are hyperbolic, meaning that the real part of the eigenvalues of the spatial derivative of the parameter dependent field are different of zero. Assume that the dimension of the unstable space of $\hat{x}$ remains equal to d_u and the dimension of the stable space of $\hat{y}$ remains constant and equal to d_s for all $p \in S$.

We now fix the notation for the parametrizations of the invariant manifolds. For details see Appendix C. Let $\phi_u : U_u \subset \mathbb{R}^{\widetilde{p}+d_u} \longrightarrow \mathbb{R}^n, \quad \phi_u \in C^k$ be a parametrization of certain regions of the unstable manifolds $M_u(\hat{x}(p))$ containing the equilibrium $\hat{x}(p)$ and $\phi_s : U_s \subset \mathbb{R}^{\widetilde{p}+d_s} \longrightarrow \mathbb{R}^n, \quad \phi_s \in C^k$ be a parametrization of certain regions of the stable manifolds $M_s(\hat{y}(p))$ containing $\hat{y}(p)$. Also, we take the parametrizations satisfying $\phi_u(p, 0) = \hat{x}(p)$ and $\phi_s(p, 0) = \hat{y}(p)$.

To consider orbits in unstable and stable manifolds, we need to consider the initial value problems below. The orbit $x(t, \xi_u, p)$ has initial condition $\phi_u(p, \xi_u)$ in the unstable manifold $M_u(\hat{x}(p))$, while $y(t, \xi_s, p)$ has initial condition $\phi_s(p, \xi_s)$ in the stable manifold $M_s(\hat{y}(p))$. The domain of definition of solutions x and y are the ones given by Cor. 2.1.2 used with η equal to ϕ_u and ϕ_s. Note that t can be taken everywhere in the maximal interval of the solutions:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi_u, p) = f(x(t, \xi_u, p), p) \\ x(0, \xi_u, p) = \phi_u(p, \xi_u), \end{cases} \tag{B.1}$$

$$\begin{cases} \dfrac{dy}{dt}(t, \xi_s, p) = f(y(t, \xi_s, p), p) \\ y(0, \xi_s, p) = \phi_s(p, \xi_s). \end{cases} \tag{B.2}$$

The main idea of the method proposed in [2] is to minimize the distance between the orbits in the unstable manifold and orbits in the stable manifold. To minimize this distance, we consider a plane separating the equilibria $\hat{x}$ and $\hat{y}$ and we try to minimize the distance between the intersection of the orbits with the mentioned plane. If the plane does not separate equilibria for all $p \in S$, the ideas presented here may not work.

Let $p_0 \in S$ be our initial guess of the parameter that gives rise to a heteroclinic orbit. Assume now that we have a hyperplane $P = \{x \in \mathbb{R}^n :< x, v > +b = 0\}$ identified by a vector $v \in \mathbb{R}^n$ and a constant $b \in \mathbb{R}$, such that P separates the points $\hat{x}(p)$ and $\hat{y}(p)$ for all p in certain open subset of S. Assume we have a parameter $\xi_u^0 \in \mathbb{R}^{d_u}$ such that the orbit $x(t, \xi_u^0, p_0)$ intersects P at time $t_u > 0$ and another parameter $\xi_s^0 \in \mathbb{R}^{d_s}$ such that the orbit $y(t, \xi_s^0, p_0)$ intersects P at time $t_s < 0$. To study heteroclinic orbits we introduce the following function:

$$\zeta(\xi_u, \xi_s, p) = \begin{bmatrix} x(t_u, \xi_u, p) - y(t_s, \xi_s, p) \\ < v, x(t_u, \xi_u, p) > +b \end{bmatrix}. \tag{B.3}$$

Note that if we have a vector $(\xi_u^*, \xi_s^*, p^*) \in \mathbb{R}^{d_u+d_s+\widetilde{p}}$ such that $\zeta(\xi_u^*, \xi_s^*, p^*) = 0$, then the orbit $x(t, \xi_u^*, p^*)$ is heteroclinic and intersects the plane P at time t_u. This observation shows that to find a root of ζ is equivalent to find a heteroclinic orbit. Observe also that to find a root of ζ, we could also minimize its squared norm.

As we already explained in Appendix C, it is difficult to use second order methods to minimize ζ because it is both computationally and intellectually expensive to give estimates to the derivatives of ϕ_u and ϕ_s. Therefore, in what follows we focus into presenting one first order method to find a root of ζ.

Remember that we need to integrate certain initial value problems to find derivatives of solutions of ordinary differential equations and here we need to approximate the initial conditions in those initial value problems. The errors in the approximation of the initial condition can be significantly increased if we integrate over long times.

To calculate the value of $\zeta(\xi_u, \xi_s, p)$, we need to integrate (B.1) and (B.2). In Appendix C we explained how to give the initial conditions. To calculate the Jacobian of ζ at (ξ_u, ξ_s, p) we need to calculate the derivatives of x with respect to ξ_u and p evaluated at (t_u, ξ_u, p) and the derivatives of y with respect to ξ_s and p evaluated at (t_s, ξ_s, p). For more details on these calculations see Section 2.2. The Jacobian of ζ is given by (B.4). For simplicity we omitted the points where the derivatives are evaluated:

$$D\zeta = \begin{bmatrix} \frac{\partial x}{\partial \xi_u} & -\frac{\partial y}{\partial \xi_s} & \frac{\partial x}{\partial p} - \frac{\partial y}{\partial p} \\ v^T \frac{\partial x}{\partial \xi_u} & 0 & v^T \frac{\partial x}{\partial p} \end{bmatrix}. \tag{B.4}$$

We now rewrite the method provided in the preprint [2]. The idea is to use first order Taylor expansion and then to solve a linear system of equations, which is almost the same idea of Newton's Method. The difference is that the system appearing here is not necessarily square. Assume we have a point $w_k = (\xi_u^k, \xi_s^k, p^k)^T$. Solving by least squares (B.5) we find the next point $w_{k+1} = (\xi_u^{k+1}, \xi_s^{k+1}, p^{k+1})^T$. Equation (B.5) is equivalent to equations (5.3) and (5.4) of [2]:

$$D\zeta(w_k)(w_{k+1} - w_k) = -\zeta(w_k). \tag{B.5}$$

The values of ξ_u and ξ_s do not determine the orbits uniquely and we can have infinitely many roots of ζ. To solve (B.5) using least squares is important to assure that the norms of ξ_u and ξ_s are small so that the approximations of the initial conditions given by (C.1), (C.2) and (C.3) are good enough, but induces a numerical problem due to the fact that the resulting iterates of ξ_u and ξ_s are in practice of the order of the machine precision. Moreover, as there is not control on the sizes of ξ_u and ξ_s, standard minimization methods may not work properly to find a root of ζ as they can make the length of vectors ξ_u and ξ_s increase.

The choices of the plane and of the times $t_u, t_s \in \mathbb{R}$ are left to the user of the method. In addition, the method was designed to work in small neighbourhoods of the initial parameter $p_0 \in \mathbb{R}^{\widetilde{p}}$, however we are interested to have as large as possible basins of attraction. The assumption that the dimensions of the stable and unstable spaces of the equilibria do not change may not hold as well the hypotheses that the equilibria are hyperbolic.

Appendix C

Dependency of Invariant Manifolds on Field Parameters

Here we introduce parametrizations of invariant manifolds that also depend on field parameters. We also address their local approximations. For simplicity we develop the theory for unstable manifolds. Throughout this section assume we are given a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field, an open set $S \subset \mathbb{R}^{\widetilde{p}}$ and a function $\hat{x} : S \longrightarrow \mathbb{R}^n, \quad \hat{x} \in C^k$ satisfying $f(\hat{x}(p), p) = 0 \quad \forall p \in S$.

Given a certain field $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad g \in C^k$ and an equilibrium $x_0 \in \mathbb{R}^n$ of g, the unstable manifold of x_0 associated with the field g is the subset $M(x_0) \subset \mathbb{R}^n$ such that $\forall y_0 \in M(x_0)$ the orbit $y(t)$ of the field g beginning at y_0 is defined in a certain interval of the type $(-\infty, a)$ and $y(t)$ converges to x_0 as the time goes to $-\infty$. We denote by $M(\hat{x}(p))$ the unstable manifold associated with the equilibrium $\hat{x}(p)$ of the field $g(x) = f(x, p)$.

It is proved in [17] that for each $p \in S$ the unstable manifold $M(\hat{x}(p))$ is indeed a manifold. That means there are local diffeomorphisms between $M(\hat{x}(p))$ and open sets of a certain euclidean space. As p changes in S the dimension of $M(\hat{x}(p))$ and therefore the dimension of the mentioned euclidean spaces can change. We avoid this by supposing that the equilibria $\hat{x}(p)$ are hyperbolic and that for all $p \in S$ the dimension of the associated unstable spaces of $\hat{x}(p)$ are all equal to a certain d.

Denote by $\zeta_p : U_p \subset \mathbb{R}^d \longrightarrow V_p \subset M(\hat{x}(p))$ a diffeomorphism between an open set U_p and other set V_p such that $0 \in U_p$ and $\zeta_p(0) = \hat{x}(p)$. In other words, ζ_p is a parametrization of a small region of $M(\hat{x}(p))$ containing $\hat{x}(p)$. Fixing our choice of ζ_p for each $p \in S$, we define $\phi(p, \xi) = \zeta_p(\xi)$. We call $\phi(p, \xi)$ parametrization of the unstable manifold.

Depending on our choices of ζ_p, the function ϕ is neither C^k nor its domain open. However, it is shown in [16] that there is a choice of parametrizations so that ϕ has the same class of differentiability of the parameter dependent field and the set $U = \{(p, \xi) : p \in \hat{S}, \xi \in U_p\}$ is open, where the set $\hat{S}$ is open and contained in S.

Here, the main use of the function $\phi(p, \xi)$ is as initial condition to ordinary

differential equations. If we look at the initial value problems (2.4), (2.5) and (2.6) with $\eta = \phi$, we see that we need the value of η and of its derivatives with respect to p and ξ. As we do not have $\phi(p, \xi)$ explicitly, we try to find approximations.

If we integrate two initial value problems with close initial conditions over long times the results can be very different. Therefore, the estimates of the derivatives of solutions using approximated initial conditions can be inaccurate.

Remember that ξ is considered in the vicinity of 0. We also have $\phi(p, 0) = \hat{x}(p)$. Doing first order Taylor approximation of ϕ in ξ we have estimate (C.1) to the value of $\phi(p, \xi)$. Using the continuity of the of the derivatives of ϕ and that $\phi(p, 0) = \hat{x}(p)$ we obtain approximations (C.2) and (C.3):

$$\phi(p, \xi) = \hat{x}(p) + \frac{\partial \phi}{\partial \xi}(p, 0)\xi + O(\| \xi \|), \tag{C.1}$$

$$\frac{\partial \phi}{\partial \xi}(p, \xi) = \frac{\partial \phi}{\partial \xi}(p, 0) + O(1), \tag{C.2}$$

$$\frac{\partial \phi}{\partial p}(p, \xi) = \frac{\partial \hat{x}}{\partial p}(p) + O(1). \tag{C.3}$$

Due to the hypothesis that for fixed p the application $\zeta(\xi) = \phi(p, \xi)$ is a diffeomorfism over its range, we know that the columns of the derivative of ϕ with respect to ξ evaluated at $(p, 0)$ should span the tangent space of the unstable manifold at $\hat{x}(p)$.

It is well known (see [17]) that the tangent space of the unstable manifold is generated by the eigenvectors associated with eigenvalues with positive real part of the spatial derivative of the parameter dependent field evaluated at $(\hat{x}(p), p)$. Therefore, our choice of the eigenvectors to use as columns of the derivative of ϕ implicitly determine the parametrization ϕ used in practice. Note that ϕ is not unique.

As we can calculate equilibria, derivatives of equilibria and eigenvectors numerically, formulae (C.1)-(C.3) provide numerical approximations to ϕ and its derivatives. Note that as our numerical approximation of the derivative of ϕ with respect to ξ in (C.1) is only $O(1)$, in practice we have only $O(1)$ approximation of $\phi(p, \xi)$.

Assume $k \geq 2$. In Appendix B we use ϕ to build a non linear equation involving solutions of ordinary differential equations that summarizes the content of [2]. To use second order methods to solve that equation we need approximations to second order partial derivatives of ϕ. Using the continuity of the derivatives of ϕ and $\phi(p, 0) = \hat{x}(p)$, we obtain estimates below for $i = 1, ..., n$:

$$\frac{\partial^2 \phi_i}{\partial p \partial p}(p, \xi) = \frac{\partial^2 \hat{x}_i}{\partial p \partial p}(p) + O(1), \tag{C.4}$$

$$\frac{\partial^2 \phi_i}{\partial p \partial \xi}(p, \xi) = \frac{\partial^2 \phi_i}{\partial p \partial \xi}(p, 0) + O(1), \tag{C.5}$$

$$\frac{\partial^2 \phi_i}{\partial \xi \partial \xi}(p, \xi) = \frac{\partial^2 \phi_i}{\partial \xi \partial \xi}(p, 0) + O(1). \tag{C.6}$$

As columns of the derivative of ϕ with respect to ξ evaluated at $(p, 0)$ are eigenvectors, to calculate the mixed derivative of ϕ with respect to p and ξ we

run into differentiation of eigenvectors. For more details see [20, 23]. We do not spend time on this subject because we show other numerical methods for locating heteroclinic orbits that do not use derivatives of eigenvectors. Using (C.4) we can approximate the double derivative of ϕ with respect to p by the Hessian of the coordiantes of the equilibrium $\hat{x}$ derived in Section 2.3. For details on how to calculate numerically the right side of (C.6) see [2].

Appendix D

Equilibrium-free Method

In this section we present an equilibrium-free optimization problem for locating heteroclinic orbits. Such optimization problem minimizes the squared norm of time derivatives of the solution evaluated at times with opposite sign.

This strategy relies in the fact that if there are points $x_0, y_0 \in \mathbb{R}^n, \quad x_0 \neq y_0$ such that a certain orbit $x(t)$ satisfies $\lim_{t \longrightarrow \infty} x(t) = x_0$ and $\lim_{t \longrightarrow -\infty} x(t) = y_0$ and the time derivatives go to zero when $t \longrightarrow \pm\infty$, then $x(t)$ is heteroclinic.

Assume we have a function $f : \mathbb{R}^n \times \mathbb{R}^{\widetilde{p}} \longrightarrow \mathbb{R}^n, \quad f \in C^k$ called parameter dependent field and consider given constants $\tau_0 < 0$ and $t_0 > 0$. Let $x(t, \xi, p)$ be the unique solution satisfying the initial value problem (D.1) below:

$$\begin{cases} \dfrac{dx}{dt}(t, \xi, p) = f(x(t, \xi, p), p) \\ x(0, \xi, p) = \xi. \end{cases} \tag{D.1}$$

We now introduce the Equilibrium-free Method. The presence of the box constraints in (D.2) is to avoid times t and τ with the same signs. Without these constraints, a vector (τ, t, ξ, p) producing small values of the objective function of (D.2) would stand for an orbit approaching an attractor but without limit in the past:

$$\begin{aligned} &\underset{(\tau,t,\xi,p)}{\text{minimize}} \quad \| \frac{\partial x}{\partial t}(t, \xi, p) \|^2 + \| \frac{\partial x}{\partial t}(\tau, \xi, p) \|^2 \\ &\text{subject to} \quad \tau \leq \tau_0, \quad t \geq t_0. \end{aligned} \tag{D.2}$$

The discussion about convergence presented in Section 3.1 can be applied here with little modifications. To solve (D.2) we can use trust-region methods, which are fast and accurate. We do not show formulas of the gradient and Hessian to solve (D.2) because all are similar to the ones already developed.

Formulation (D.2) does not assure that the absolute values of times diverge. For this problem one may try to successively increase the absolute values of τ_0 and t_0. We expect the basin of attraction of the Equilibrium-free Method to be smaller than the one of the Global Method. The possibly oscillatory behaviour of the field can lead to local minima very easily, which may not be heteroclinic orbits. We left a more complete analysis of this method as a future work.

Appendix E

Derivative of the Inner Product

Assume we have two function $g, h : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ such that $g, h \in C^2$. In this section we develop formulae to the first and second order derivatives of the inner product between g and h. This is useful to solve the optimization problems of Chapter 3. The inner product between g and h is given by:

$$g^T h = \sum_{i=1}^{n} g_i h_i. \tag{E.1}$$

Differentiating equation E.1 with respect to x_k for $k \in \{1, ..., m\}$, we obtain equation E.2. It is useful to have the gradient of the inner product written in terms of the derivatives of g and h and their values, which is given in formula E.3. Using formula E.3 with $g = h$, we obtain the gradient of the squared norm of g in formula E.4:

$$\frac{\partial}{\partial x_k} g^T h = \sum_{i=1}^{n} \frac{\partial g_i}{\partial x_k} h_i + \sum_{i=1}^{n} g_i \frac{\partial h_i}{\partial x_k} = \frac{\partial g}{\partial x_k}^T h + \frac{\partial h}{\partial x_k}^T g, \tag{E.2}$$

$$\frac{\partial}{\partial x} g^T h = \frac{\partial g}{\partial x}^T h + \frac{\partial h}{\partial x}^T g, \tag{E.3}$$

$$\frac{\partial}{\partial x} g^T g = 2 \frac{\partial g}{\partial x}^T g. \tag{E.4}$$

Differentiating equation E.2 with respect to x_j for $j \in \{1, ..., m\}$, we obtain formula E.5. As before, it is useful to have the Hessian of the inner product written in terms of first order derivatives of g and h and their values, which is done in formula E.6. Using formula E.6 with $g = h$, we obtain the Hessian of squared norm of g in formula E.7:

$$\frac{\partial^2}{\partial x_l \partial x_k} g^T h = \left(\sum_{i=1}^{n} \frac{\partial^2 g_i}{\partial x_l \partial x_k} h_i \right) + \left(\sum_{i=1}^{n} \frac{\partial g_i}{\partial x_k} \frac{\partial h_i}{\partial x_l} \right) + \left(\sum_{i=1}^{n} g_i \frac{\partial^2 h_i}{\partial x_l \partial x_k} \right) + \left(\sum_{i=1}^{n} \frac{\partial g_i}{\partial x_l} \frac{\partial h_i}{\partial x_k} \right), \tag{E.5}$$

$$\nabla^2 g^T h = \sum_{i=1}^{n} h_i \nabla^2 g_i + Dg^T Dh + \sum_{i=1}^{n} g_i \nabla^2 h_i + Dh^T Dg, \tag{E.6}$$

$$\nabla^2 g^T g = 2Dg^T Dg + 2 \sum_{i=1}^{n} g_i \nabla^2 g_i. \tag{E.7}$$

Bibliography

[1] *Optimization Toolbox User's Guide.* The MathWorks, Inc, 2015.

[2] D. Marchesin A. A. Mailybaev and M. H. de Sá Vera. *Sensitivity Analysis of Stable and Unstable Manifolds: Theory and Applications.* preprint, 2007.

[3] Yu. A. Kuznetsov A. R. Champneys and B. Sandstede. *A numerical toolbox for homoclinic bifurcation analysis.* Applied Nonlinear Mathematics Research Report, University of Bristol, Bristol, UK, 1995.

[4] J. Douglas Burden, Richard L.; Faires. *Numerical Analysis.* Boston Prindle, Weber and Schmidt, 1993.

[5] V. Canale. *The computation of paths of homoclinic orbits, Ph.D. Thesis, Department of Computer Science.* University of Toronto, Toronto, Canada, 1994.

[6] A. R. Champneys and Yu. A. Kuznetsov. *Numerical detection and continuation of codimension-two homoclinic orbits.* Int. J. Bif. and Chaos, 1994.

[7] E. J. Doedel and M. J. Friedman. *Numerical computation of heteroclinic orbits.* J. Comput. and Appl. Math., 1989.

[8] E. Gamero E. Freire, A. J. Rodr ıguez-Luis and E. Ponce. *A case study for homoclinic chaos in an autonomous electronic circuit.* Physica, 1993.

[9] L. G. Franquelo E. Freire and J. Aracil. *Periodicity and chaos in an autonomous electronic circuit.* IEEE Transactions on Circuits and Systems, 1984.

[10] M. J. Friedman E. J. Doedel and A. C. Monteiro. *On locating connecting orbits.* Appl. Math. and Comput., 1994.

[11] M. J. Friedman. *Numerical analysis and accurate computation of heteroclinic orbits in the case of center manifolds.* J. Dynam. Diff. Eq., 1993.

[12] M. J. Friedman and E. J. Doedel. *Computational methods for global analysis of homoclinic and heteroclinic orbits: a case study.* J. Dynam. Diff. Eq., 1993.

[13] S. Schecter G. Chapiro, D. Marchesin. *Combustion Waves and Riemann Solutions in Light Porous Foam.* Journal of Hyperbolic Differential Equations, 2014.

[14] P. Gray and S. K. Scott. *Chemical Oscillations and Instabilities: Nonlinear Chemical Kinetics*. Oxford, 1990.

[15] A. J. Homburg and B. Sandstede. *Homoclinic and Heteroclinic Bifurcations in Vector Fields, Chapter 8, Handbook of Dynamical Systems, Volume 3*.

[16] Floris Takens Jacob Palis. *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations: Fractal Dimensions and Infinitely Many Attractors in Dynamics*. Cambridge University Press, 1995.

[17] Philip Holmes John Guckenheimer. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer New York, 1995.

[18] A. Huang L. O. Chua, C. W. Wu and C. Q. Zhong. *A universal circuit for studying and generating chaos, Part I: Routes to chaos, IEEE Trans. on Circuits and Systems-I: Fundamental Theory an Applications*. Special Issue on Chaos in Electronic Circuits, Part A, 1993.

[19] F. C. Moon. *Chaotic Vibrations: An Introduction for Applied Scientists and Engineers*. Wiley, New York, 1987.

[20] H.G. Ter Morsche N.P. Van Der AA and R.R.M. Mattheij. *Computation of Eigenvalue and Eigenvector Derivatives for a General Complex-Valued Eigensystem*. Electronic Journal of Linear Algebra, 2007.

[21] P. H. Rabinowitz. *Variational methods for Hamiltonian systems. In: Handbook of dynamical systems, Vol. 1A*.

[22] Water Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Education, 1976.

[23] A. P. Seyranian and A. A. Mailybaev. *Multiparameter Stability Theory with Mechanical Applications*.

[24] S. Wiggins. *Global Bifurcations and Chaos. Analytical Methods*. Springer, New York, 1988.