

# Shadowing of non-transversal heteroclinic chains

Amadeu Delshams<sup>1</sup>, Adrià Simon<sup>2</sup>

Departament de Matemàtiques, Universitat Politècnica de Catalunya  
Av. Diagonal 647, 08028 Barcelona  
Amadeu.Delshams@upc.edu, adria.simon@gmail.com

Piotr Zgliczyński<sup>3</sup>

Jagiellonian University, Institute of Computer Science and Computational  
Mathematics,  
Lojasiewicza 6, 30-348 Kraków, Poland  
e-mail: umzglicz@cyf-kr.edu.pl

January 27, 2017

## Abstract

We present a new result about the shadowing of nontransversal chain of heteroclinic connections based on the idea of dropping dimensions. We illustrate this new mechanism with several examples. As an application we discuss this mechanism in a simplification of a toy model system derived by Colliander *et al.* in the context of cubic defocusing nonlinear Schrödinger equation.

**keywords:** shadowing, heteroclinic chain, non-transversal intersection, covering relations

## 1 Introduction

In the present paper we deal with the problem of shadowing a nontransversal chain of heteroclinic connections between invariant sets (fixed points, periodic orbits, etc). The motivation for us is the work [CKS+] (see also [GK]) on the transfer of energy to high frequencies in the nonlinear Schrodinger equation (just NLS from now on). From the dynamical systems viewpoint there is one remarkable feature of the construction in [CKS+], namely that the authors were able to shadow a non-transversal highly degenerated chain of heteroclinic connections between some periodic orbits. The length of the chain is arbitrary, but finite. Neither in [CKS+] nor in [GK] we were able to find a clear geometric picture showing how this is achieved, so it could be easily applicable to other systems.

---

<sup>1</sup>Partially supported by the MINECO-FEDER Grant MTM2015-65715-P and the Russian Scientific Foundation Grant 14-41-00044 at the Lobachevsky University of Nizhny Novgorod.

<sup>2</sup>Partially supported by the MINECO-FEDER Grant MTM2012-31714 and the Catalan Grant 2014SGR504.

<sup>3</sup>Research has been supported by Polish National Science Centre grant 2011/03B/ST1/04780

In this work we present a mechanism, which we believe gives a geometric explanation of what is happening. Moreover, we strive to establish an abstract framework, which will make it easier to apply this technique to other systems, both PDEs and ODEs, in questions related to the existence of diffusing orbits. The term *diffusing orbit* relates to the Arnold's diffusion [Ar] for the perturbation of integrable Hamiltonian systems. Throughout the paper we will often call diffusing orbit an orbit shadowing a chain of heteroclinic connections, and occasionally the existence of such an orbit will be referred to as the diffusion.

In our picture we think of evolving a disk of dimension  $k$  along a heteroclinic transition chain and when a given transition is not transversal, then we 'drop' one or more dimensions of our disk, i.e., we select a subdisk of lower dimension "parallel to expanding directions in future transitions". After at most  $k$  transitions, our disk is a single point and we cannot continue further. We will refer to this phenomenon as the *dropping dimensions* mechanism. Since this is a new mechanism, we have found it convenient to include several figures to illustrate the main differences between transversal and non-transversal heteroclinic chains. While thinking about disks has some geometric appeal, we consider instead in our construction a thickened disk called h-set in the terminology of [ZGi] and our approach is purely topological (just as the one presented in [CKS+]).

The main technical tool used in our work is the notion of *covering relations* as introduced in [ZGi], which differs from the notion used under the same name in [CKS+]. Similar ideas about the dropping exit dimensions appear implicitly also in the works [BM+, WBS].

In our work we present an abstract topological theorem about shadowing chains of covering relations with dropping dimensions, and we show how such chains of coverings can be obtained in the presence of chains of heteroclinic connections in two examples, a linear model and a simplified Toy Model that the one in [CKS+], which however contains all the difficulties present in its prototype. We intend to treat more complicated examples, in particular NLS from [CKS+, GK] in subsequent papers.

The content of this paper can be described as follows. In Section 2 we first describe the consequences of the difference between transversal and non-transversal intersection of invariant manifolds of fixed points and we present the model problem with a non-transversal heteroclinic chain and state our conjecture about the existence of shadowing orbits arbitrarily close to such chain. We also introduce an example formed by a triangular system, where the existence of the diffusion is quite obvious. In Section 3 we explain the basic geometric idea of our dropping dimensions mechanism. In Section 4 we recall from [ZGi] the notions of h-sets and the covering relation. In Section 5 we prove the main topological result on shadowing of chains of covering relations with dropping dimensions. Using this new mechanism, in the next two sections we rigorously analyze two simple models, a linear model in Section 6 and a simplified Toy Model in Section 7.

## 1.1 Notation

By  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$  we denote the set of natural, integer, rational, real and complex numbers, respectively. We assume that  $0 \in \mathbb{N}$ .  $\mathbb{Z}_-$  and  $\mathbb{Z}_+$  are nonpositive and nonnegative integers, respectively. By  $S^1$  we will denote the unit circle on the complex plane.

In  $\mathbb{R}^n$  by  $e_i$  for  $i = 1, \dots, n$  we will denote the  $i$ -th vector from the canonical basis in  $\mathbb{R}^n$ , i.e. the  $j$ -th coordinate of  $e_i$  is equal to 1, when  $j = i$  and 0 otherwise.

For  $\mathbb{R}^n$  we will denote the norm of  $x$  by  $\|x\|$  and when in some context the formula for the norm is not specified, then it means that any norm can be used. For  $x_0 \in \mathbb{R}^s$ ,  $B_s(x_0, r) = \{z \in \mathbb{R}^s : \|x_0 - z\| < r\}$  and  $B_s = B_s(0, 1)$ .

Sometimes, if  $V$  is a vector space with a norm, then  $B_V(a, r)$  will denote an open ball in  $V$  centered at  $a$  with radius  $r$ .

For  $z \in \mathbb{R}^u \times \mathbb{R}^s$  we will call usually  $x$  the first coordinate and  $y$  the second one. Hence  $z = (x, y)$ , where  $x \in \mathbb{R}^u$  and  $y \in \mathbb{R}^s$ . We will use the projection maps  $\pi_x(z) = x(z) = x$  and  $\pi_y(z) = y(z) = y$ . For functions  $f : \mathbb{R}^u \times \mathbb{R}^s \rightarrow \mathbb{R}^u \times \mathbb{R}^s$  we will use the shortcuts  $f_x = \pi_x f$  and  $f_y = \pi_y f$ .

Let  $z \in \mathbb{R}^n$  and  $U \subset \mathbb{R}^n$  be a compact set and  $f : U \rightarrow \mathbb{R}^n$  be continuous map, such that  $z \notin f(\partial U)$ . Then the local Brouwer degree [S] of  $f$  on  $U$  at  $z$  is defined and will be denoted by  $\deg(f, U, z)$ . See for example the Appendix in [ZGi] and references given there for the properties of  $\deg(f, U, z)$ .

If  $V, W$  are two vector spaces, then by  $\text{Lin}(V, W)$  we will denote the set of all linear maps from  $V$  to  $W$ . When  $V = \mathbb{R}^k$  and  $W = \mathbb{R}^m$ , we will identify  $\text{Lin}(\mathbb{R}^k, \mathbb{R}^m)$  with the set of matrices with  $m$  columns and  $k$  rows, denoted by  $\mathbb{R}^{k \times m}$ .

## 2 Non-transverse diffusion, the statement of the problem, some examples

In this section we introduce the geometric assumptions under which we expect to construct the orbits shadowing a non-transversal heteroclinic chain. Our approach is motivated by the work [CKS+] on the NLS.

One of the key ingredients of the constructions of the energy transfer in [CKS+] consists on finding an orbit which visits the neighborhoods of  $N$  invariant 1-dimensional objects in a  $N$ -dimensional complex system. Each object is connected with the previous and the following one with heteroclinic connections, so the authors look for a solution that 'concatenates' these connections. This kind of scheme seems similar to the Arnold diffusion[Ar], but we plan to explain that it is a different phenomenon since we do not have a transverse intersection between the invariant manifolds. In addition, the proposed mechanism could be applied to integrable systems in contrast to the Arnold diffusion, which is a phenomenon that only takes place in non integrable systems.

## 2.1 Transverse versus Non-Transverse

In this subsection we will explain the difference between the transverse and the non-transverse situation. The hint about the idea of dropping dimensions will be given.

To do so, we are going to consider a two dimensional map with four fixed points:

$$p_0 = (0,0) \quad p_1 = (1,0) \quad p_2 = (1,1) \quad p_3 = (2,1).$$

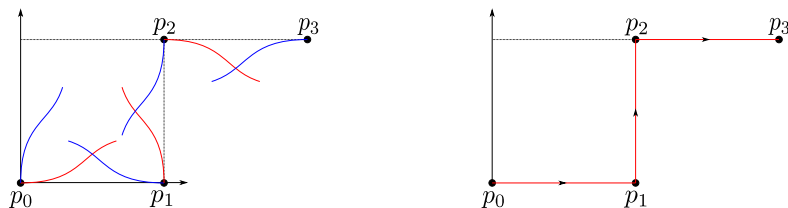
We are going to assume also that each point  $p_i$  has a one dimensional stable manifold,  $\mathcal{T}^s(p_i)$ , and a one dimensional unstable manifold,  $\mathcal{T}^u(p_i)$ , both tangent to some linear subspaces. That is

- $\mathcal{T}^s(p_0)$  is tangent to the subspace generated by  $e_2$  at  $p_0$  and  $\mathcal{T}^u(p_0)$  is tangent to the subspace generated by  $e_1$  at  $p_0$ .
- $\mathcal{T}^s(p_1)$  is tangent to the subspace generated by  $e_1$  at  $p_1$  and  $\mathcal{T}^u(p_1)$  is tangent to the subspace generated by  $e_2$  at  $p_1$ .
- $\mathcal{T}^s(p_2)$  is tangent to the subspace generated by  $e_2$  at  $p_2$  and  $\mathcal{T}^u(p_2)$  is tangent to the subspace generated by  $e_1$  at  $p_2$ .
- $\mathcal{T}^s(p_3)$  is tangent to the subspace generated by  $e_1$  at  $p_3$  and  $\mathcal{T}^u(p_3)$  is tangent to the subspace generated by  $e_2$  at  $p_3$ .

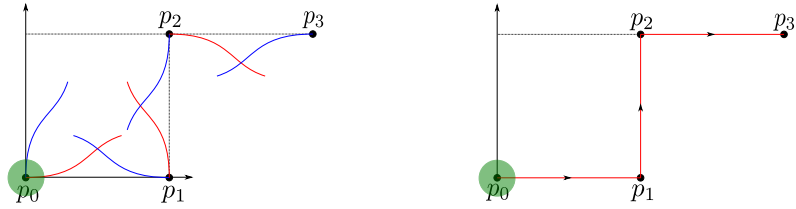
We are now going to consider two different scenarios. The first one consists on assuming that the unstable manifold of a point  $p_i$  intersects transversally the stable manifold of the following point  $p_{i+1}$ .

The second will be given by a non-transverse intersection of the manifolds, with one of the branches of  $\mathcal{T}^u(p_i)$  coinciding with one of the branches of  $\mathcal{T}^s(p_{i+1})$ .

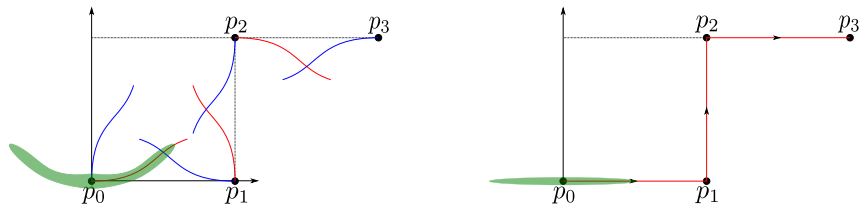
The schematic situation is the following (unstable manifolds are in red and stable manifolds are in blue):



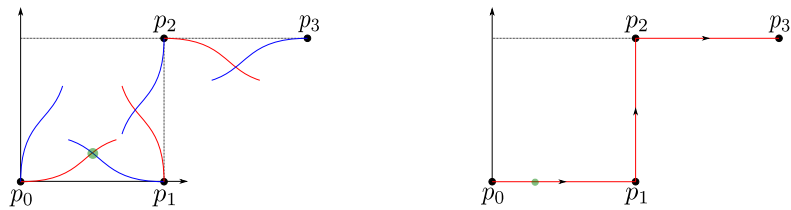
We wonder if it is possible to connect  $p_0$  with  $p_3$  through the map, in both situations. To do so, we consider a ball containing the first fixed point  $p_0$ :



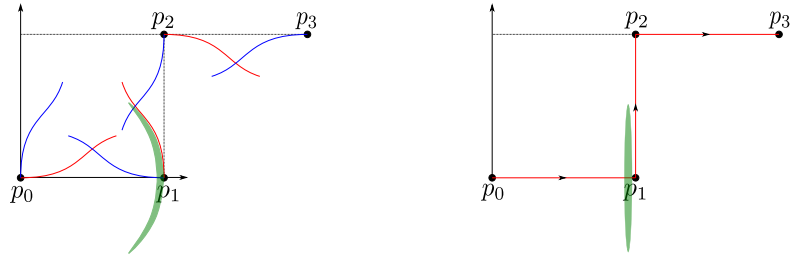
If we compute iterates of the ball through the maps we can expect that it is expanded in the unstable direction and contracted in the stable direction:



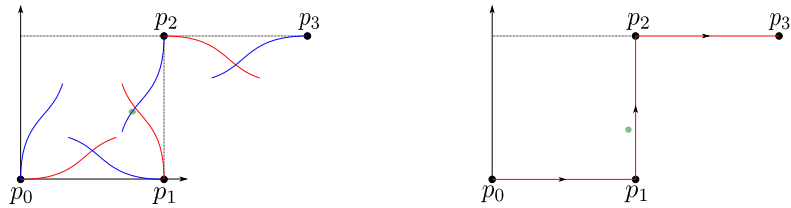
Notice that in both cases the domain intersects the stable manifold of the following fixed point,  $p_1$ . In the transverse case the domain contains a heteroclinic point. In the non-transverse situation this is obvious because the manifolds are coincident. We can now restrict our domain precisely around that intersection point for the transverse case and at some place in the right-hand side of the fixed point  $p_0$  for the non-transverse case.



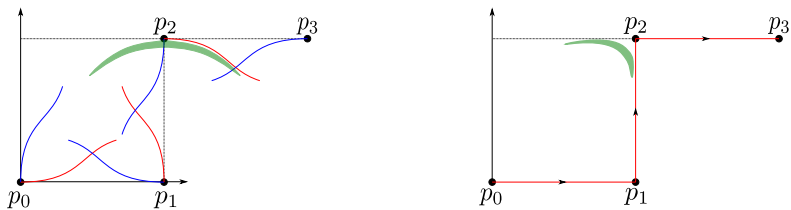
The forward iterates of the restricted domain approach the fixed point  $p_1$  since, in both cases, our domains contain heteroclinic points. The domains will not only approach  $p_1$  but also, after some iterates, will spread along the unstable manifold of  $p_1$ :



While our domain spreads along the unstable manifold of  $p_1$ , here we find the first big difference between the transverse and the non-transverse case. In the transverse case, it is clear that our domain will intersect the stable manifold of  $p_2$ . In the non-transverse situation, our domain will never cross the stable manifold of  $p_2$ . Then, we restrict our domain in the intersection for the transverse case and in the upper part of  $p_1$  since we want to reach  $p_3$ :



Using, in the transverse case, the same argument as before, since our domain contains a heteroclinic point in the transverse situation, forward iterates will spread our domain along the unstable manifold of  $p_2$ . For the non transverse case we will reach the proximity of  $p_2$  after some iterates, but our domain will be trapped and will not visit the following fixed point,  $p_3$ :



In the transverse case, we could continue and see that the domain will visit  $p_3$ .

The above analysis, shows that, on the one hand, in the transverse situation there are no geometric obstructions in shadowing the heteroclinic chain. On the other hand, in the non-transverse case, we can see that, in general, we cannot visit as many invariant objects as we want. So, now, we wonder why the authors of [CKS+] can connect  $N$  periodic orbits in the Toy Model System. The main reason is the large dimension of the system and the fact that each connection takes place in a direction that has not been used in the past.

Indeed notice that, if in the non-transverse example the last fixed point  $p_3$  is located in a new dimension (that means that the system is three dimensional) it might be possible continue with the argument and visit  $p_3$ .

In the next subsection we are going to generate an example for which it is clear that one can shadow a non-transverse heteroclinic chain.

## 2.2 Examples with diffusion in a non-transverse situation - the triangular system

The triangular system is

$$\dot{x} = F(x) \tag{1}$$

with

$$\begin{cases} F_1(x) &= \lambda_1 x_1 - \lambda_1 x_1^2 \\ F_i(x) &= (\lambda_i - \mu_i)x_i - \lambda_i x_i^2 + \mu_i x_i x_{i-1} \end{cases} \quad \text{for } 1 < i \leq n \tag{2}$$

with  $\lambda_i > 0$  for  $1 \leq i \leq n$  and  $\mu_i \in \mathbb{R}$  for  $1 < i < n$ .

Note that we can integrate (1) recursively, since

$$\dot{x}_i = f_i(t)x_i - \lambda_i x_i^2 \Rightarrow x_i(t) = \frac{x_i(0) \exp\left(\int_0^t f_i(s) ds\right)}{1 + \lambda_i \int_0^t \exp\left(\int_0^s f_i(r) dr\right) ds},$$

with  $f_i(t) = \lambda_i - \mu_i + \mu_i x_{i-1}(t)$  and  $x_{-1}(t) = 1$ .

For  $i = 0, 1, \dots, n$  we denote by  $p_i$  the following fixed point of (1)

$$p_i = (p_{i,1}, \dots, p_{i,n}), \quad p_{i,j} = 1, j \leq i-1, \quad p_{i,j} = 0, \quad j \geq i. \tag{3}$$

Therefore we have

$$p_0 = (0, \dots, 0), \quad p_1 = (1, 0, 0, \dots, 0), \quad p_2 = (1, 1, 0, \dots, 0), \dots, p_n = (1, 1, \dots, 1).$$

Observe that the interval  $C_i = (1, \dots, 1, x_i, 0, \dots, 0)$ , where the length of the initial sequence of 1's is equal  $i-1$ , for  $x_i \in (0, 1)$  is the heteroclinic orbit connecting  $p_i$  and  $p_{i+1}$ . The question is whether we can follow this heteroclinic chain, i.e., whether there exists an orbit which starts in the close vicinity of  $p_0$  then it visits consecutively the neighborhoods of  $p_1, p_2, \dots, p_n$ .

We can check now the linear behavior around the equilibrium points computing the derivative of the vector field:

$$DF(p_0) = \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \lambda_2 - \mu_2 & & & & & & & \\ & & \ddots & & & & & & \\ & & & \lambda_i - \mu_i & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_n - \mu_n & & & \end{pmatrix}$$

$$DF(p_i) = \begin{pmatrix} -\lambda_1 & & & & & & & & \\ \mu_2 & -\lambda_i & & & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & \mu_i & -\lambda_i & & & & \\ & & & & & \lambda_{i+1} & & & \\ & & & & & & \lambda_{i+2} - \mu_{i+2} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \lambda_n - \mu_n \end{pmatrix}$$

so we have different possible choices for the parameters.

- If  $\mu_i > \lambda_i$  for all  $i = 1, \dots, n$ , each point  $p_j$  has only one unstable direction defined by  $e_{j+1}$ , while the rest of the directions are stable.
- If  $\mu_i = \lambda_i$  for all  $i = 1, \dots, n$ , each point  $p_j$  has only one unstable direction, defined by  $e_{j+1}$ . All the “past” directions, defined by  $\{e_1, \dots, e_j\}$  are stable while all the “future” directions, defined by  $\{e_{j+2}, \dots, e_n\}$ , are linearly neutral.
- If  $\mu_i < \lambda_i$  for all  $i = 1, \dots, n$ , at each point,  $p_j$ , all the “past” directions, defined by  $\{e_1, \dots, e_j\}$ , are stable while all the “future” directions, defined by  $\{e_{j+2}, \dots, e_n\}$ , are linearly unstable.

Let us now display some numerical integration of the system for these three possibilities for  $n = 4$ . For the sake of concreteness we are going to assume  $\lambda = \lambda_i$  and  $\mu = \mu_i$  for all  $i = 1, \dots, n$ . In Figures 1, 2 and 3 we display the numerical evidence for the existence of diffusive orbits.



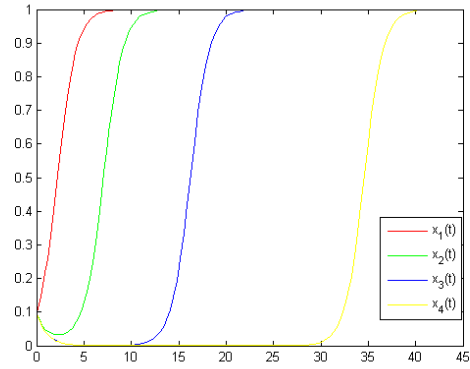


Figure 1: Solution of system (1) for  $\lambda = 1$  and  $\mu = 2$ . The initial condition is:  $x_1(0) = x_2(0) = x_3(0) = x_4(0) = \frac{1}{10}$ .

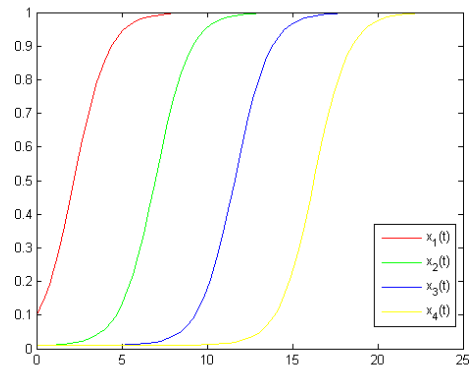


Figure 2: Solution of system (1) for  $\lambda = 1$  and  $\mu = 1$ . The initial condition is:  $x_1(0) = \frac{1}{10}$ ,  $x_2(0) = x_3(0) = x_4(0) = \frac{1}{100}$ .

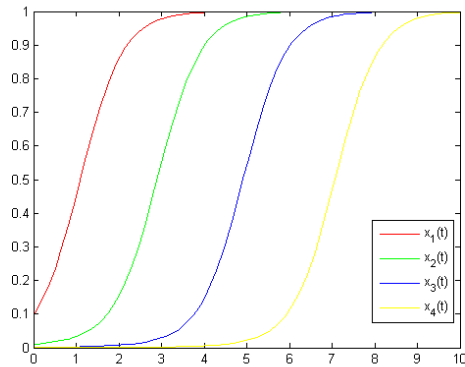


Figure 3: Solution of system (1) for  $\lambda = 2$  and  $\mu = 1$ . The initial condition is:  $x_1(0) = 10^{-1}$ ,  $x_2(0) = 10^{-2}$ ,  $x_3(0) = 10^{-3}$ ,  $x_4(0) = 10^{-4}$ .

Of course, we could expect these numerical evidences just by looking at the equations. It is clear that

$$\lim_{t \rightarrow \infty} x_1(t) = 1.$$

By induction, assuming

$$l_{i-1} = \lim_{t \rightarrow \infty} x_{i-1}(t)$$

and looking for equilibria of  $x_i(t)$  for  $t \rightarrow \infty$ , it has to satisfy

$$\lim_{t \rightarrow \infty} \dot{x}_i(t) = 0,$$

i.e.

$$0 = \lambda_i l_i (1 - l_i) - \mu_i l_i (1 - l_{i-1}).$$

So, we get  $l_i = 0$  or  $l_i = 1$  but since  $\dot{x}_i(t) > 0$  for  $t$  large enough we can conclude  $l_i = 1$ . This reasoning with some small effort can be turned into the rigorous proof.

The main conclusion of this part is that we have obtained an easy example for which we can ensure the transition chain even if the intersection between the invariant manifolds is not transverse, regarding the high dimension and the disposition of the equilibrium points and the heteroclinic connections: each one in a new direction not used before. In addition, the system is integrable by quadratures which goes against the notion of the Arnold's diffusion. However, we have detected the reason why the connection could be possible. The geometric mechanism relies on the fact that we are dealing with a high dimensional system and that each new connection is defined by a direction that *has not been used before*.

### 2.3 The conjecture

We are now in condition to present a conjecture, which can be proved by our method discussed in the following sections under some additional assumptions.

Let  $n_i > 0$  for  $i = 1, \dots, L$  and let  $n_1 + n_2 + \dots + n_L = n$ ,  $d_i = n_1 + \dots + n_i$ .

For  $i = 1, \dots, L$  the subspaces  $V_i$  which are spanned by  $\{e_{k_{i-1}+j}\}_{j=1, \dots, n_i}$ .

In this notation  $\mathbb{R}^n = \bigoplus_{i=1}^L V_i$ .

We will use the following notation: for  $l \in \mathbb{N}$ ,  $0^l, 1^l$  will denote the sequences of length  $l$  consisting of  $l$  0's or  $l$  1's, respectively.

Assume that we have a diffeomorphism  $f : \mathbb{R}^n \times \mathbb{R}^{w_u} \times \mathbb{R}^{w_s} \rightarrow \mathbb{R}^n \times \mathbb{R}^{w_u} \times \mathbb{R}^{w_s}$ ,  $w = w_u + w_s$  with the following properties:

- there exists a sequence of fixed points

$$\begin{aligned}
 p_0 &= (0^n, 0^w), \\
 p_1 &= (1, 0^{n_1-1}, 0, \dots, 0^w), \\
 p_2 &= (1, 0^{n_1-1}, 1, 0^{n_2-1}, 0, \dots, 0^w), \\
 &\vdots \\
 p_k &= (1, 0^{n_1-1}, 1, 0^{n_2-1}, \dots, 1, 0^{n_k-1}, 0^{n_{k+1}+\dots+n_L}, 0^w), \\
 &\vdots \\
 p_L &= (1, 0^{n_1-1}, 1, 0^{n_2-1}, \dots, 1, 0^{n_L-1}, 0^w)
 \end{aligned}$$

- for any  $i = 1, \dots, L$ , the interval connecting  $p_{i-1}$  and  $p_i$  denoted by  $C_i$

$$C_i = \{z = (1, 0^{n_1-1}, 1, 0^{n_2-1}, \dots, 1, 0^{n_{i-1}-1}, t, 0^{-1+n_i+\dots+n_L}, 0^w), \quad t \in [0, 1]\}$$

is invariant under  $f$  and for any  $z \in C_i$

$$\lim_{k \rightarrow \infty} f^k(z) = p_i, \quad \lim_{k \rightarrow -\infty} f^k(z) = p_{i-1} \quad (4)$$

- at  $p_i$  the subspace  $V_i$  defined the exit directions and these directions are 'dominating' for the scattering when passing by  $p_i$ . This statement is vague, because the scenario we are going to present most likely can be realized under various sets of assumptions.

Only the first  $n$ -directions really count, the others will be treated as the entry directions (in the sense of covering relations, see Section 4).

Our result is is

**Conjecture 1** *Under the above assumptions for any  $\epsilon > 0$  there exists a point  $z$  and a sequence of integers  $k_1 < k_2 < \dots < k_L$ , such that*

$$\begin{aligned}
 \|z - p_0\| &< \epsilon, \\
 \|f^{k_i}(z) - p_i\| &< \epsilon, \quad i = 1, \dots, L
 \end{aligned}$$

Our idea of the proof of this result requires a construction of covering relations (see (25), and then Conjecture 1 follows directly from Theorem 8.

We will show how a construction of suitable h-sets and coverings can be done for a linear model in Section 6 and for the simplified version of the toy model from [CKS+] in Section 7.

### 3 The geometric idea of dropping dimensions

In this section we will explain our idea of dropping dimensions along a non-transversal heteroclinic chain.

Let us start with a simplified version of the example from Section 2.3. Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a diffeomorphism with the following properties:

1. The points  $p_i = (1, \dots, 1, 0, \dots, 0)$  are fixed under  $f$  for  $i = 0 \dots n$ .
2. The segments  $C_i$  that connect the points  $p_{i-1}$  and  $p_i$ ,

$$C_i = \{(1, \dots, 1, t, 0, \dots, 0), 0 \leq t \leq 1\}$$

for  $1 \leq i \leq n$  are invariant under  $f$  and, for all  $x \in C_i$ :

$$\lim_{k \rightarrow \infty} f^k(x) = p_i \quad \lim_{k \rightarrow -\infty} f^k(x) = p_{i-1}.$$

3. At each point  $p_i$  the  $i$ -th direction is stable and the  $(i+1)$ -th is unstable. This means:

$$\begin{aligned} Df(p_i)e_i &= \mu_i e_i, & |\mu_i| &< 1 \\ Df(p_i)e_{i+1} &= \lambda_i e_{i+1}, & |\lambda_i| &> 1 \end{aligned}$$

4. The past directions, defined by  $e_1, \dots, e_{i-1}$ , are contracting directions around the fixed point  $p_i$  but with a rate greater than  $\mu_i$ . The future directions, defined by  $e_{i+2}, \dots, e_n$ , are expanding directions around the fixed point  $p_i$  but with a rate lower than  $\lambda_i$ .

**Proposition 2** *Under the previous assumptions, for all  $\epsilon > 0$  there exists a point  $x$  and a sequence of integers  $0 = k_0 < k_1 < \dots < k_n$  such that:*

$$\|f^{k_i}(x) - p_i\| < \epsilon \quad i = 0, \dots, n.$$

**Remark 3** *Notice that we connect  $n+1$  points in a  $n$  dimensional space. We cannot guarantee that the result is valid for more points.*

**Remark 4** *Observe that we are assuming that there are only two dominant coordinates around each fixed point. That means that this could not be applied to the Toy Model System in NLS [CKS+, GK], where we have four dominant directions. In Section 7 we treat such system.*

Proposition 2 is proved in Section 6 for the linear model.

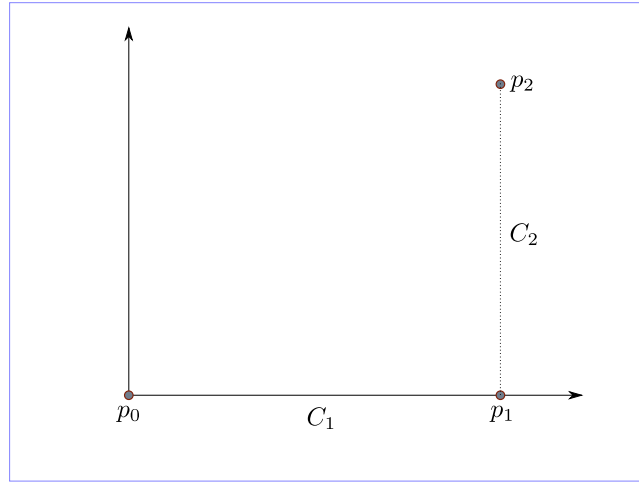
### 3.1 The idea of the proof: dropping dimensions

Here we sketch the basic idea of the proof of Proposition 2 with some pictures. We are going to consider only a two dimensional map. So, consider  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  with three fixed points:

$$p_0 = (0, 0), \quad p_1 = (1, 0), \quad p_2 = (1, 1),$$

with invariant segments  $C_1$  and  $C_2$  defined as

$$C_1 = \{(x_1, x_2) : 0 \leq x_1 \leq 1, x_2 = 0\} \quad C_2 = \{(x_1, x_2) : x_1 = 1, 0 \leq x_2 \leq 1\}$$

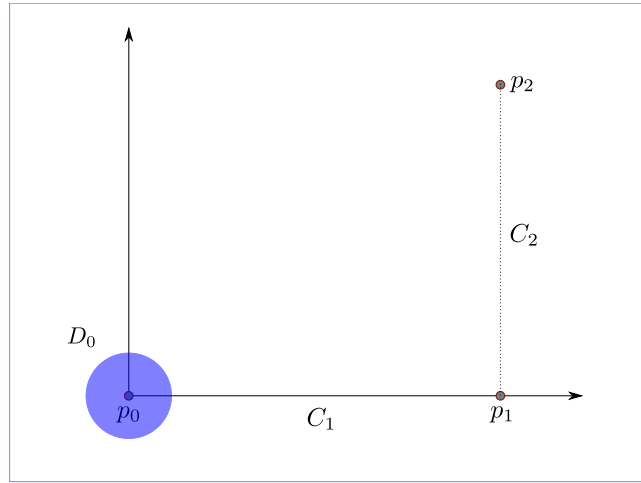


Assume also that the derivatives of the map around the fixed points have the following structure:

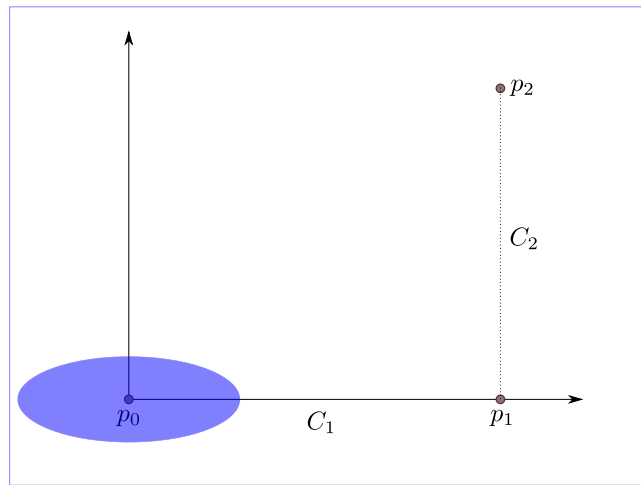
$$Df(p_0) = \begin{pmatrix} \lambda_{0,1} & 0 \\ 0 & \lambda_{0,2} \end{pmatrix}, \quad Df(p_1) = \begin{pmatrix} \mu_{1,1} & 0 \\ 0 & \lambda_{1,2} \end{pmatrix}, \quad Df(p_2) = \begin{pmatrix} \mu_{2,1} & 0 \\ 0 & \mu_{2,2} \end{pmatrix}, \quad (5)$$

where  $\lambda_{0,1}, \lambda_{0,2}, \lambda_{1,2} > 1$  and  $0 < \mu_{1,1}, \mu_{2,1}, \mu_{2,2} < 1$ .

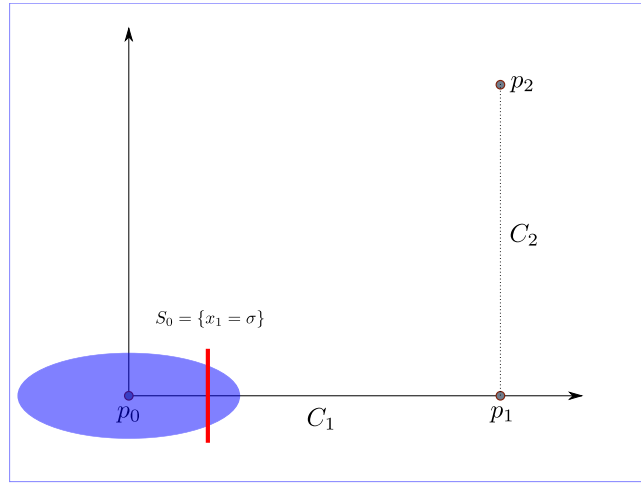
We start by considering a domain (ball)  $D_0$  of full dimension centered around  $p_0$ .



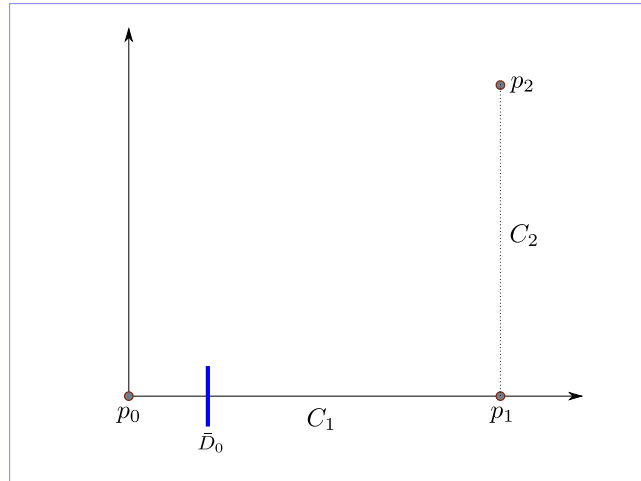
Given the linear stability from (5), we can assume that after one iteration of the map, our initial ball  $D_0$  will be expanded in both directions:



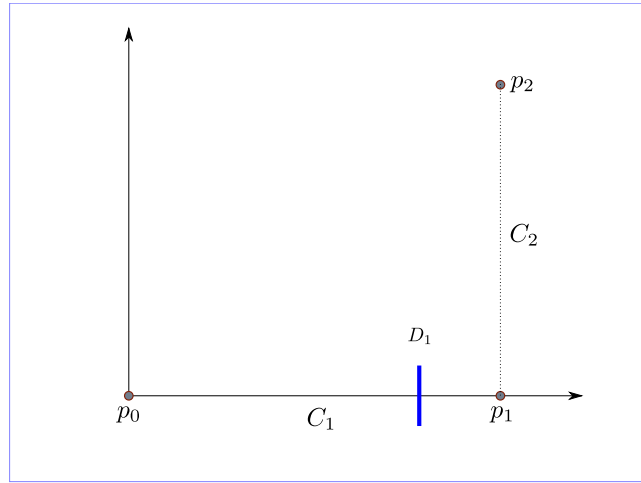
It is now time to make a decision: from all the possible directions, we are only interested in the one defined by the outgoing heteroclinic connection, that is, the segment  $C_1$ . Then we consider a section  $S_0 = \{x_1 = \sigma\}$  where  $\sigma$  is some small parameter:



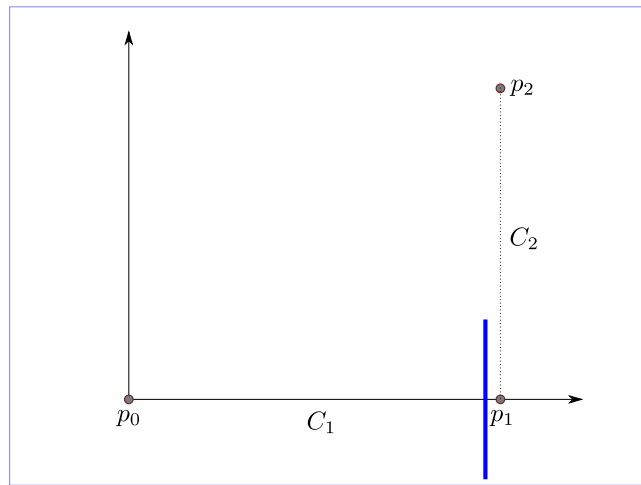
Since we are only interested in the points of our ball close to the heteroclinic connection, we intersect the domain with the section  $S_0$ . We say that we have dropped the  $x_1$  direction. We will not use this direction in future steps. Our domain,  $\bar{D}_0$ , has one dimension less than  $D_0$ , that is, it has dimension one.



Then we continue with this domain. After several iterations of the map since the domain is close to the heteroclinic connection, we can ensure that  $\bar{D}_0$  will approach  $p_1$  and we obtain a domain  $D_1$ :

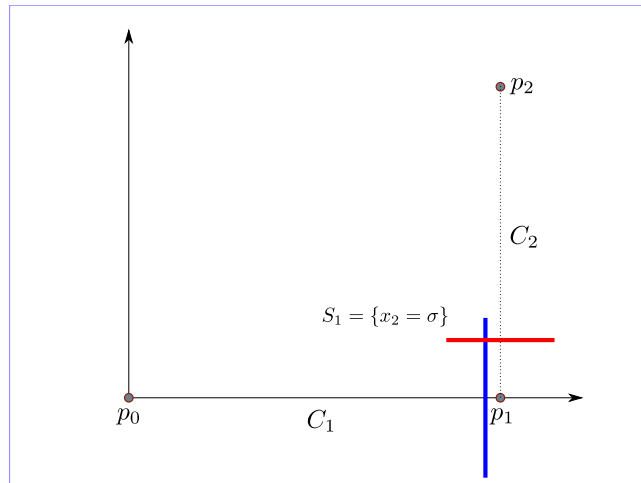


We can use, again, the linear prediction of the map, (5), to be sure that our domain is expanded in the  $x_2$  direction.

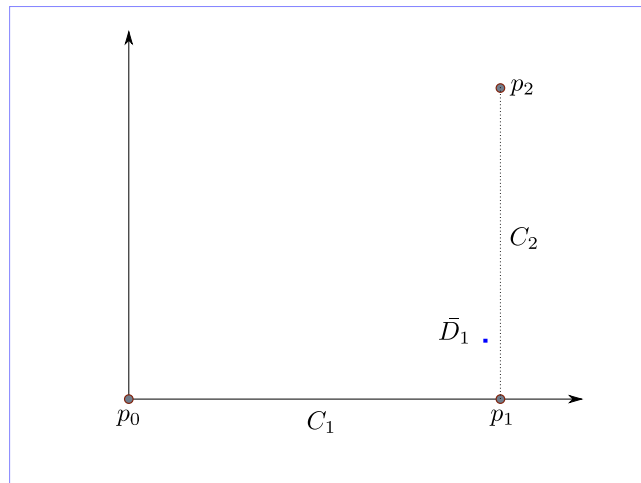


We use now the same argument. From all the possible directions that  $f(D_1)$  covers, we want to escape through the one defined by the heteroclinic connection to  $p_2$ . So we put a section defined in the same spirit as before:  $S_1 = \{x_2 = \sigma\}$ .

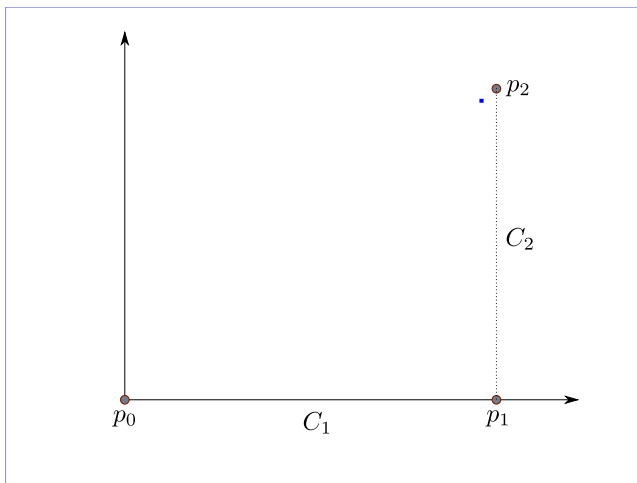




We restrict now our domain in its intersection with the section  $S_1$ . The resulting domain  $\bar{D}_1$  will have, then, one dimension less than  $D_1$ , which means that it will have dimension zero.



We have no more dimensions to drop, since our initial domain becomes a single point. This point is close to the heteroclinic defined in  $C_2$ , so that we are sure that after some iterates, it will approach the final fixed point  $p_2$ :



## 4 h-sets, covering relations

The goal of this section is to recall from [ZGi] the notions of h-sets and covering relations, and to state the theorem about the existence of point realizing the chain of covering relations. This will be the main technical tool in proving the existence of the orbits shadowing the heteroclinic chain in the next sections.

### 4.1 h-sets and covering relations

**Definition 1** [ZGi, Definition 1] *An h-set  $N$  is a quadruple  $(|N|, u(N), s(N), c_N)$  such that*

- $|N|$  is a compact subset of  $\mathbb{R}^n$
- $u(N), s(N) \in \{0, 1, 2, \dots, n\}$  are such that  $u(N) + s(N) = n$
- $c_N : \mathbb{R}^n \rightarrow \mathbb{R}^n = \mathbb{R}^{u(N)} \times \mathbb{R}^{s(N)}$  is a homeomorphism such that

$$c_N(|N|) = \overline{B_{u(N)}} \times \overline{B_{s(N)}}.$$

We set

$$\begin{aligned} \dim(N) &:= n, \\ N_c &:= \overline{B_{u(N)}} \times \overline{B_{s(N)}}, \\ N_c^- &:= \partial B_{u(N)} \times \overline{B_{s(N)}}, \\ N_c^+ &:= \overline{B_{u(N)}} \times \partial B_{s(N)}, \\ N^- &:= c_N^{-1}(N_c^-), \quad N^+ = c_N^{-1}(N_c^+). \end{aligned}$$

Hence a  $h$ -set  $N$  is a product of two closed balls in some coordinate system. The numbers  $u(N)$  and  $s(N)$  are called the exit and entry dimensions, respectively. The subscript  $c$  refers to the new coordinates given by the homeomorphism  $c_N$ . Observe that if  $u(N) = 0$ , then  $N^- = \emptyset$  and if  $s(N) = 0$ , then  $N^+ = \emptyset$ . In the sequel to make the notation less cumbersome we will often drop the bars in the symbol  $|N|$  and we will use  $N$  to denote both the  $h$ -sets and its support.

We will call  $N^-$  the exit set of  $N$  and  $N^+$  the entry set of  $N$ . These names are motivated by the Conley index theory [C, MM] and the role that these sets will play in the context of covering relations.

**Definition 2** [ZGi, Definition 6] *Assume that  $N, M$  are  $h$ -sets, such that  $u(N) = u(M) = u$  and  $s(N) = s(M) = s$ . Let  $f : N \rightarrow \mathbb{R}^n$  be a continuous map. Let  $f_c = c_M \circ f \circ c_N^{-1} : N_c \rightarrow \mathbb{R}^u \times \mathbb{R}^s$ . Let  $w$  be a nonzero integer. We say that*

$$N \xrightarrow{f,w} M$$

( $N$   $f$ -covers  $M$  with degree  $w$ ) iff the following conditions are satisfied

1. *There exists a continuous homotopy  $h : [0, 1] \times N_c \rightarrow \mathbb{R}^u \times \mathbb{R}^s$ , such that the following conditions hold true*

$$h_0 = f_c, \tag{6}$$

$$h([0, 1], N_c^-) \cap M_c = \emptyset, \tag{7}$$

$$h([0, 1], N_c) \cap M_c^+ = \emptyset. \tag{8}$$

2. *If  $u > 0$ , then there exists a map  $A : \mathbb{R}^u \rightarrow \mathbb{R}^u$  such that*

$$h_1(p, q) = (A(p), 0), \text{ for } p \in \overline{B_u}(0, 1) \text{ and } q \in \overline{B_s}(0, 1), \tag{9}$$

$$A(\partial B_u(0, 1)) \subset \mathbb{R}^u \setminus \overline{B_u}(0, 1). \tag{10}$$

Moreover, we require that

$$\deg(A, \overline{B_u}(0, 1), 0) = w, \tag{11}$$

We will call condition (7) the exit condition and condition (8) the entry condition.

Note that in the case  $u = 0$ , if  $N \xrightarrow{f,w} M$ , then  $f(N) \subset \text{int}M$  and  $w = 1$ .

In fact in the above definition  $s(N)$  and  $s(M)$  can be different, see [W2, Def. 2.2].

**Remark 5** *If the map  $A$  in condition 2 of Def. 2 is a linear map, then condition (10) implies that*

$$\deg(A, \overline{B_u}(0, 1), 0) = \pm 1.$$

Hence condition (11) is fulfilled with  $w = \pm 1$ .

In fact, this is the most common situation in the applications of covering relations.

Most of the time we will not be interested in the value of  $w$  in the symbol  $N \xrightarrow{f,w} M$  and we will often drop it and write  $N \xrightarrow{f} M$ , instead. Sometimes we may even drop the symbol  $f$ , if known from the context, and write  $N \implies M$ .

To handle the inverse maps in the context of covering relations the following two definitions are useful.

**Definition 3** [ZGi, Definition 3] *Let  $N$  be a  $h$ -set. We define a  $h$ -set  $N^T$  as follows*

- $|N^T| = |N|$
- $u(N^T) = s(N)$ ,  $s(N^T) = u(N)$
- We define a homeomorphism  $c_{N^T} : \mathbb{R}^n \rightarrow \mathbb{R}^n = \mathbb{R}^{u(N^T)} \times \mathbb{R}^{s(N^T)}$ , by

$$c_{N^T}(x) = j(c_N(x)),$$

where  $j : \mathbb{R}^{u(N)} \times \mathbb{R}^{s(N)} \rightarrow \mathbb{R}^{s(N)} \times \mathbb{R}^{u(N)}$  is given by  $j(p, q) = (q, p)$ .

■

Observe that  $N^{T,+} = N^-$  and  $N^{T,-} = N^+$ . This operation is useful in the context of inverse maps.

**Definition 4** [ZGi, Definition 7] *Assume  $N, M$  are  $h$ -sets, such that  $u(N) = u(M) = u$  and  $s(N) = s(M) = s$ . Let  $g : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^n$ . Assume that  $g^{-1} : |M| \rightarrow \mathbb{R}^n$  is well defined and continuous. We say that  $N \xleftarrow{g} M$  ( $N$   $g$ -backcovers  $M$ ) iff  $M^T \xrightarrow{g^{-1}} N^T$ .*

## 4.2 Main theorem about chains of covering relations

**Theorem 6 (Thm. 9)** [ZGi] *Assume  $N_i$ ,  $i = 0, \dots, k$ ,  $N_k = N_0$  are  $h$ -sets and for each  $i = 1, \dots, k$  we have either*

$$N_{i-1} \xrightarrow{f_i, w_i} N_i \tag{12}$$

or

$$N_i \subset \text{dom}(f_i^{-1}) \quad \text{and} \quad N_{i-1} \xleftarrow{f_i, w_i} N_i. \tag{13}$$

*Then there exists a point  $x \in \text{int}N_0$ , such that*

$$f_i \circ f_{i-1} \circ \dots \circ f_1(x) \in \text{int}N_i, \quad i = 1, \dots, k \tag{14}$$

$$f_k \circ f_{k-1} \circ \dots \circ f_1(x) = x \tag{15}$$

The reader is referred to [ZGi] for a proof. The basic idea of the proof of this theorem is the homotopy and the local Brouwer degree.

The following corollary is an immediate consequence of Theorem 6.

**Collorary 7** Let  $N_i, i \in \mathbb{Z}_+$  be h-sets. Assume that for each  $i \in \mathbb{Z}_+$  we have either

$$N_{i-1} \xrightarrow{f_i, w_i} N_i \quad (16)$$

or

$$N_i \subset \text{dom}(f_i^{-1}) \quad \text{and} \quad N_{i-1} \xleftarrow{f_i, w_i} N_i. \quad (17)$$

Then there exists a point  $x \in \text{int}N_0$ , such that

$$f_i \circ f_{i-1} \circ \cdots \circ f_1(x) \in \text{int}N_i, \quad i \in \mathbb{Z}_+. \quad (18)$$

Moreover, if  $N_{i+k} = N_i$  for some  $k > 0$  and all  $i$ , then the point  $x$  can be chosen so that

$$f_k \circ f_{k-1} \circ \cdots \circ f_1(x) = x. \quad (19)$$

### 4.3 Natural structure of a h-set

Observe that all the conditions appearing in the definition of the covering relation are expressed in ‘internal’ coordinates  $c_N$  and  $c_M$ . Also the homotopy is defined in terms of these coordinates. Sometimes this makes statements and notation seem a bit cumbersome. With this in mind we introduce the notion of a natural structure on a h-set.

**Definition 5** We will say that  $N = \{(x_0, y_0)\} + \overline{B}_u(0, r_1) \times \overline{B}_s(0, r_1) \subset \mathbb{R}^u \times \mathbb{R}^s$  is an h-set with a natural structure if:

$$u(N) = u, \quad s(N) = s, \quad c_N(x, y) = \left( \frac{x-x_0}{r_1}, \frac{y-y_0}{r_2} \right).$$

### 4.4 The operation of dropping exit dimensions

**Definition 6** Assume that we have a decomposition  $\mathbb{R}^n = \mathbb{R}^{u_1} \oplus \mathbb{R}^t \oplus \mathbb{R}^{s_1}$  and the norm for  $(x_1, x_2, x_3) \in \mathbb{R}^{u_1} \oplus \mathbb{R}^t \oplus \mathbb{R}^{s_1}$  is  $\|(x_1, x_2, x_3)\| = \max(\|x_1\|, \|x_2\|, \|x_3\|)$ .

Assume that  $N$  is an h-set, with  $u(N) = u_1 + t$  and  $s(N) = s_1$ . In view of the norm on  $\mathbb{R}^n$  we have

$$c_{|N|} = (\overline{B}_{u_1} \oplus \overline{B}_t) \oplus \overline{B}_{s_1} \quad (20)$$

where the parentheses enclose the exit directions.

Let us denote by  $V$  the subspace  $\{0\} \times \mathbb{R}^t \times \{0\}$ . We define a new h-set  $R_V(N)$  by setting

- $|R_V(N)| = |N|$
- $u(R_V(N)) = u_1, \quad s(R_V(N)) = s_1 + t$
- $c_{R_V(N)} = c_N$

Roughly speaking,  $R_V(N)$  is obtained from  $N$  by relabeling some exit coordinates in  $N$  as the entry directions.

## 5 The mechanism of dropping dimensions—the main topological theorem

### 5.1 h-sets $M_i$ and $\widetilde{M}_i$

Our setting is motivated by Proposition 1. For the sake of completeness we recall here from Section 2 some assumptions.

Let  $n_i > 0$  for  $i = 1, \dots, L$  and let  $n_1 + n_2 + \dots + n_L = n$  and  $w = w_u + w_s$ , where  $w_u, w_s \in \mathbb{N}$ .

For  $i = 1, \dots, L$  let  $V_i$  be a subspace with  $\dim V_i = n_i$ . Let  $W_u$  and  $W_s$  be two subspaces of dimensions  $w_u$  and  $w_s$ , respectively.

In  $\mathbb{R}^n \times \mathbb{R}^{w_u} \times \mathbb{R}^{w_s}$  we will represent points as  $(z_1, \dots, z_L, z_{L+1}, z_{L+2})$ , where  $z_i \in V_i$  for  $i = 1, \dots, L$ ,  $z_{L+1} \in W_u$  and  $z_{L+2} \in W_s$ . In each of the spaces  $V_i$ ,  $W_j$  we have some fixed basis and an isomorphism with some  $\mathbb{R}^d$  equipped with the metric, so we can define balls in this subspace.

First, we put sections (hyperplanes of codimension  $n_i$ ) in the vicinity of each point  $p_i$  (not to be confused with the Poincaré sections for ODEs): the exit section  $S_i$  is given by conditions

$$z_i = \delta_i = (\Delta_i, 0^{n_i-1}) \in V_i, \quad i = 1, \dots, L, \quad (21)$$

where  $\Delta_i > 0$ .

We also define

$$\delta_{L+1} = 0. \quad (22)$$

For  $1 \leq i \leq L+1$  we define the set  $M_i$  (centered on the section  $S_i$  for  $i \leq L$  and on  $p_{L+1}$  for  $M_{L+1}$ )

$$M_i = p_i + \delta_i + \left( \prod_{j=1}^L \overline{B}_{n_j}(0, t_{i,j}) \right) \times \overline{B}_{w_u}(0, t_{i,L+1}) \times \overline{B}_{w_s}(0, t_{i,L+2}), \quad (23)$$

where  $t_{i,j}$  are positive real numbers.

We equip the set  $M_i$  with two different h-set structures. To define the first one, denoted by  $M_i$  for  $i = 1, \dots, L, L+1$ , we declare  $V_i \oplus V_{i+1} \oplus \dots \oplus V_L \oplus W_u$  as the exit directions and  $V_1 \oplus V_2 \oplus \dots \oplus V_{i-1} \oplus W_s$  as the entry directions. For the second one, we set

$$\widetilde{M}_i = R_{V_i}(M_i), \quad i = 1, \dots, L. \quad (24)$$

This means that (see Definition 6) we declare  $V_{i+1} \oplus \dots \oplus V_L \oplus W_u$  as the exit directions (i.e. when compared to  $M_i$  we drop the subspace  $V_i$  of exit directions).

We will not need  $\widetilde{M}_{L+1}$ .

Observe that  $\widetilde{M}_L$  and  $M_{L+1}$  have  $W_u$  as the exit directions and  $M_L$  has  $V_L \oplus W_u$  as the exit directions.

Now we assume that

$$\widetilde{M}_i \xrightarrow{f^i} M_{i+1} \quad i = 1, \dots, L. \quad (25)$$

The above covering relations are expected by combining the transition where we drop the connection direction (plus some others we decide to treat from that point on as the entry ones) with the local hyperbolic behavior near  $p_{i+1}$ , whereas in other directions for both covering relations where the dynamics might not help us we just adjust the sizes to obtain the correct inequalities. For this purpose we need to increase the sizes during the transition if these are treated as the entry directions or to decrease the sizes if they are treated as the exit ones.

## 5.2 The main topological shadowing theorem

About the same time as this work was under development a theorem about shadowing a chain covering relations with decreasing number of exit directions appeared in works [BM+, WBS], where a slightly different technique of proof was used, but it still is based on the same covering relations we are using.

**Theorem 8** *Assume that the following covering relations are satisfied*

$$\widetilde{M}_i = R_{V_i}(M_i) \xrightarrow{f^{l_i}} M_{i+1} \quad i = 1, \dots, L. \quad (26)$$

Let  $k_i = \sum_{j=1}^i l_j$ .

Then there exists  $q$  such that

$$\begin{aligned} q &\in M_1, \\ f^{k_i}(q) &\in M_{i+1} \quad i = 1, \dots, L. \end{aligned}$$

**Proof:** Equation (23) allows us to introduce the coordinates on  $M_i$  through the map

$$\begin{aligned} \mathcal{C}_i : M_i &\rightarrow \prod_{i=1}^L \overline{B_{V_i}}(0, 1) \times \overline{B_{W_u}}(0, 1) \times \overline{B_{W_s}}(0, 1), \quad (27) \\ \mathcal{C}_i(z_1, \dots, z_L, z_{L+1}, z_{L+2}) &= \left( \frac{z_1 - p_{i,1} - \delta_{i,1}}{t_{i,1}}, \dots, \frac{z_{L+2} - p_{i,L+2} - \delta_{i,L+2}}{t_{i,L+2}} \right). \end{aligned}$$

Observe that the above coordinates  $\mathcal{C}_i$ , up to a permutation required to put the exit direction first, are the ones from the natural structure of h-set.

From now on we will use these coordinates. Without any loss of generality we will assume that  $M_i = M_{i,c} = \prod_{i=1}^L \overline{B_{V_i}}(0, 1) \times \overline{B_{W_u}}(0, 1) \times \overline{B_{W_s}}(0, 1)$ .

We will prove the following statement, which implies the assertion of our theorem.

For any  $\bar{y} \in B_{W_s}(0, 1)$ ,  $\bar{x} \in B_{W_u}(0, 1)$ ,  $\eta_i \in B_{V_i}(0, 1)$ ,  $i = 1, \dots, L$ , there exists  $q$  such that

$$q \in M_1, \quad \pi_{W_s}(q) = \bar{y}, \quad \pi_{V_1}(q) = \eta_1, \quad (28)$$

$$f^{k_i}(q) \in M_{i+1}, \quad \pi_{V_{i+1}} f^{k_i}(q) = \eta_{i+1} \quad i = 1, \dots, L, \quad (29)$$

$$\pi_{W_u}(f^{k_{L+1}}(q)) = \bar{x}. \quad (30)$$

In the sequel we will denote  $f^{l_i}$  by  $f_i$ .

To obtain  $q_1 \in M_1$  satisfying (28–30) it is enough to find a sequence  $\{q_i\}_{i=1}^{L+1}$  satisfying the following conditions

$$y(q_1) - \bar{y} = 0, \quad (31)$$

$$z_i(q_i) - \eta_i = 0, \quad i = 1, \dots, L, \quad (32)$$

$$f_i(q_i) - q_{i+1} = 0, \quad i = 1, \dots, L, \quad (33)$$

$$z_{L+1}(q_{L+1}) - \bar{x} = 0. \quad (34)$$

which we will consider in the set

$$D = \prod_{i=1}^{L+1} M_i.$$

Let us remind the reader that the supports of  $M_i$  and  $\widetilde{M}_i$  coincide, but  $M_i^\pm$  and  $\widetilde{M}_i^\pm$  differ.

Observe that the number of equations in system (31–34) coincides with the number of variables in  $D$ . Indeed the equation count goes as follows:

- (31) gives  $w_s$  equations
- (32) consists of  $n_1 + n_2 + \dots + n_L = n$  equations
- (33) consists of  $L \cdot (n + w_u + w_s)$  equations
- (34) gives  $w_u$  equations,

which gives  $(L+1)(n + w_u + w_s)$  equations in the system, which coincides with the dimension of the set  $D$ .

If  $w_u = 0$ , then  $\bar{x} = 0$  and equation (34) is dropped from further considerations when defining maps  $F$ ,  $H_t$ . Analogously, when  $w_s = 0$  then  $\bar{y} = 0$  and we drop equation (31).

Let us denote by  $F$  the map given by the left hand side of system (31–34). We have for  $q = (q_1, \dots, q_{L+1}) \in D$

$$F(q) = \begin{pmatrix} y(q_1) - \bar{y} \\ z_i(q_i) - \eta_i & i = 1, \dots, L \\ f_i(q_i) - q_{i+1} & i = 1, \dots, L, \\ z_{L+1}(q_{L+1}) - \bar{x} \end{pmatrix}. \quad (35)$$

We will prove that system (31–34) has a solution in  $D$ , by using the homotopy argument to show that the local Brouwer degree  $\deg(F, \text{int}D, 0)$  is nonzero.

Let  $h_i$  for  $i = 1, 2, \dots, L$  be the homotopies from the covering relations (26).

We imbed  $F$  into a one-parameter family of maps (a homotopy)  $H_t$  as follows

$$H_t(q) = \begin{pmatrix} y(q_1) - (1-t)\bar{y} \\ z_i(q_i) - (1-t)\eta_i & i = 1, \dots, L \\ h_{t,i}(q_i) - q_{i+1} & i = 1, \dots, L, \\ z_{L+1}(q_{L+1}) - (1-t)\bar{x} \end{pmatrix} \quad (36)$$

It is easy to see that  $H_0(q) = F(q)$ .



We show that if  $q \in \partial D$  then  $H_t(q) \neq 0$  holds for all  $t \in [0, 1]$ . This will imply that  $\deg(H_t, D, 0)$  is defined for all  $t \in [0, 1]$  and does not depend on  $t$ .

Let  $q \in \partial D$ . Then  $q_i \in \partial M_i$  for some  $i = 1, \dots, L+1$ . We will use the following decomposition of  $\partial M_i$  for  $i = 1, \dots, L$ :  $\partial M_i = M_i^+ \cup (\widetilde{M}_i^+ \cap M_i^-) \cup \widetilde{M}_i^-$ , whereas for  $i = L+1$ , since  $\widetilde{M}_{L+1}$  is not defined,  $\partial M_i = M_i^+ \cup M_i^-$ . It may happen that  $M_{L+1}^- = \emptyset$ .

- the case  $q_i \in M_i^+$ .

If  $i > 1$ , then we consider  $\widetilde{M}_{i-1} \xrightarrow{f_{i-1}} M_i$  and we see from (8) that  $q_i \notin h_{t,i}(q_{i-1})(M_{i-1})$ . Therefore in this case  $h_{t,i}(q_{i-1}) - q_i \neq 0$ .

If  $i = 1$ , then  $y(q_1) \neq (1-t)\bar{y}$ , because in this case  $y(q_1) \in \partial B_{w_s}(0, 1)$ , hence  $\|y(q_1)\| = 1 > \|\bar{y}\|$ .

- the case  $i \leq L$  and  $q_i \in \widetilde{M}_i^+ \cap M_i^-$ . Then  $q_i \in \partial B_{V_i}(0, 1)$ , hence  $\|z_i(q_i)\| = 1 > \|\eta_i\|$ .

- the case  $i \leq L$  and  $q_i \in \widetilde{M}_i^-$ .

From the exit condition (7), in the covering relation  $\widetilde{M}_i \xrightarrow{f_i} M_{i+1}$  it follows  $h_{t,i}(q_i) \notin M_{i+1}$ , therefore  $h_{t,i}(q_i) - q_{i+1} \neq 0$ .

- the case  $i = L+1$  and  $q_{L+1} \in M_{L+1}^-$ .

We have  $\|z_{L+1}(q_{L+1})\| = 1 > \|\bar{x}\|$ .

We have proved that  $\deg(H_t, \text{int}D, 0)$  is defined. By the homotopy invariance we have

$$\deg(F, \text{int}D, 0) = \deg(H_1, \text{int}D, 0). \quad (37)$$

In the sequel the points in  $M_i$  (and  $\widetilde{M}_i$ ) will be denoted by  $(z_{i,1}, \dots, z_{i,L}, z_{i,L+1}, y_i)$ , where  $z_{i,k} \in V_k$  for  $k = 1, \dots, L$ ,  $z_{i,L+1} \in W_u$  and  $y_i \in W_s$ .

Observe that  $H_1(q) = 0$  is the following system of *linear equations*

$$\begin{aligned} y_1 &= 0, \\ z_{1,1} &= 0, \\ (0, A_1(z_{1,2}, \dots, z_{1,L+1}), 0) - (z_{2,1}, z_{2,2}, \dots, z_{2,L+1}, y_2) &= 0, \\ z_{2,2} &= 0, \\ (0, 0, A_2(z_{2,3}, \dots, z_{2,L+1}), 0) - (z_{3,1}, z_{3,2}, \dots, z_{3,L+1}, y_3) &= 0, \\ &\dots \\ (0, \dots, A_L(z_{L,L+1}), 0) - (z_{L+1,1}, z_{L+1,2}, \dots, z_{L+1,L+1}, y_{L+1}) &= 0, \\ z_{L+1,L+1} &= 0, \end{aligned}$$

where  $A_i$  is a linear map which appears at the end of the homotopy  $h_i$ .

It is not hard to see that  $q = 0$  is the only solution of this system. For the proof observe that  $y_i = 0$  for  $i = 1, 2, \dots$  because the first term in each equation involving  $A_i$  has zero on the last ( $y$ ) coordinate. To prove that  $z_{i,j} = 0$  for

$i, j = 1, \dots, L, L + 1$ , we should start from the two bottom equations to infer that  $z_{L+1,i} = 0$  for  $i = 1, \dots, L + 1$ , and since  $A_L$  is an isomorphism then also  $z_{L,L+1} = 0$ . Now we consider  $z_{L,i}$  from the next two equations from the bottom and so on.

Therefore  $\deg(H_1, \text{int}D, 0) = \pm 1$ . This and (37) implies that

$$\deg(F, \text{int}D, 0) = \pm 1, \quad (38)$$

hence there exists a solution of equation  $F(q) = 0$  in  $D$ . This finishes the proof.  $\blacksquare$

### 5.3 Generalization

In the theorem below we allow chains of coverings relations combined with dropping some directions.

**Theorem 9** *Assume that we have h-sets  $N_i$  and  $M_j$  (and  $\tilde{M}_j$  when some exit dimensions have been dropped) and the following covering relations are satisfied*

$$\begin{aligned} N_{0,0} &\xrightarrow{f_{0,0}} N_{0,1} \xrightarrow{f_{0,1}} \cdots \xrightarrow{f_{0,i_0}} N_{0,i_0+1} = M_0, \\ \tilde{M}_0 = N_{1,0} &\xrightarrow{f_{1,0}} N_{1,1} \xrightarrow{f_{1,1}} \cdots \xrightarrow{f_{1,i_1}} N_{1,i_1+1} = M_1, \\ \tilde{M}_1 = N_{2,0} &\xrightarrow{f_{2,0}} N_{2,1} \xrightarrow{f_{2,1}} \cdots \xrightarrow{f_{2,i_2}} N_{2,i_2+1} = M_2, \\ &\cdots \\ \tilde{M}_L = N_{L,0} &\xrightarrow{f_{L,0}} N_{L,1} \xrightarrow{f_{L,1}} \cdots \xrightarrow{f_{L,i_L}} N_{L,i_L+1} = M_L. \end{aligned}$$

Then there exists  $q_0, \dots, q_L$ , such that

$$\begin{aligned} q_k &\in N_{k,0}, \quad f_{k,j} \circ \cdots \circ f_{k,1} \circ f_{k,0}(q_k) \in N_{k,i_j+1}, \quad j = 0, \dots, i_k, \quad k = 0, \dots, L \\ q_{k+1} &= f_{k,i_k} \circ \cdots \circ f_{k,1} \circ f_{k,0}(q_k), \quad k = 0, \dots, L - 1. \end{aligned}$$

**Proof:** Conceptually the same as the proof of Theorem 8.  $\blacksquare$

## 6 Diffusion in the linear model

We prove now Proposition 2 for  $f$  being a linear model. To formulate the precise assumptions about our linear model we need first to introduce some notations.

Let  $z = (x_1, \dots, x_n)$ . For  $i = 0, \dots, n$ , define  $z_i = (z_{i,p}, z_{i,\text{inc}}, z_{i,\text{out}}, z_{i,f})$  where

- $z_{i,p} = (x_1, \dots, x_{i-1})$  are the past coordinates
- $z_{i,\text{inc}} = x_i$  is the incoming coordinate
- $z_{i,\text{out}} = x_{i+1}$  is the outgoing coordinate

- $z_{i,f} = (x_{i+2}, \dots, x_n)$  are the future coordinates.

These are the local coordinates around each fixed point  $p_i$ .

We assume that we have a sequence of linear maps:  $f_i$  for  $i = 0, \dots, n$  and affine maps  $f_{i-1,i}$  for  $i = 1, \dots, n$ .

We assume that the map  $f$  is equal to the linear map  $f_i$  around the fixed point  $p_i$  and defined as  $f_{i-1,i}$  close to the heteroclinic connection. Therefore will refer to  $f_i$ ' as *the local maps* and  $f_{i-1,i}$ 's will be called *the transition maps*.

## 6.1 Local maps

Let  $f_i(z_i) = (f_{i,p}(z_i), f_{i,inc}(z_i), f_{i,out}(z_i), f_{i,f}(z_i))$  the decomposition of the map  $f_i$  in terms of the previous splitting of the coordinates  $z_i$ . We assume that:

$$\begin{aligned} f_{i,p}(z_i) &= A_{i,p}z_{i,p} \\ f_{i,inc}(z_i) &= \mu_i z_{i,inc} \\ f_{i,out}(z_i) &= \lambda_i z_{i,out} \\ f_{i,f}(z_i) &= A_{i,f}z_{i,f}, \end{aligned} \tag{39}$$

where  $A_{i,p}$  and  $A_{i,f}$  are matrices that satisfy

$$|A_{i,p}z_{i,p}| \leq \mu_i |z_{i,p}| \quad |A_{i,f}z_{i,f}| \geq \lambda_i |z_{i,f}|,$$

with  $|\mu_i|, \mu_i < 1$  and  $1 < \lambda_i, |\lambda_i|$ . The norm that we are using here and for the rest of the proof is the maximum norm,  $|\cdot| = \|\cdot\|_\infty$ .

Let us fix  $\epsilon > 0$ ,  $0 < \sigma < \epsilon$  and  $0 < \eta < 1$  for all  $i = 0, \dots, n$ .

For each  $i = 0, \dots, n$ , we want to define  $h$ -sets that will be centered in the following points  $q_{i,inc}$  and  $q_{i,out}$ :

- $q_{i,inc} = (0, \sigma, 0, 0)$
- $q_{i,out} = (0, 0, \sigma, 0)$ .

Notice that  $q_{i,inc}$  is located close to the fixed point  $p_i$  in the direction of the incoming heteroclinic connection, defined by the segment  $C_i$ . The point  $q_{i,out}$  is also located close to the fixed point  $p_i$ , but in the direction of the outgoing heteroclinic, defined by the segment  $C_{i+1}$ . Define the sets:

$$N_i^{inc} = \{z_i \in \mathbb{R}^n : |z_i - q_{i,inc}| \leq \epsilon\} \tag{40}$$

$$\begin{aligned} N_i^{out} &= \{z_i \in \mathbb{R}^n : |z_{i,p}| \leq (1 - \eta)\epsilon, |z_{i,inc}| \leq (1 - \eta)\epsilon, \\ &\quad |z_{i,out} - \sigma| \leq (1 + \eta)\epsilon, |z_{i,f}| \leq (1 + \eta)\epsilon\}. \end{aligned} \tag{41}$$

We equip these sets with the natural  $h$ -set structure (see Definition 5). We declare the directions  $(z_{i,p}, z_{i,inc})$  as the entry directions and  $(z_{i,out}, z_{i,f})$  as the exit directions in both cases.

**Lemma 10** *For every  $i = 0, \dots, n - 1$ , there exists an integer  $k_i$  such that the following covering relation holds:*

$$N_i^{inc} \xrightarrow{f_i^{k_i}} N_i^{out}.$$

**Proof:** Since the map is linear we only have to prove that the entry (stable) components of  $N_i^{\text{inc}}$  are mapped inside  $N_i^{\text{out}}$  and that the exit (unstable) directions of  $N_i^{\text{inc}}$  cover the exit components of  $N_i^{\text{out}}$ . This is, the boundary of the exit directions of  $N_i^{\text{inc}}$  is mapped outside  $N_i^{\text{out}}$ .

Let us start with the past components. We have to show that  $\left|f_{i,p}^{k_i}(z_i)\right| < (1 - \eta)\epsilon$  for  $z_i \in N_i^{\text{inc}}$ . But

$$\left|f_{i,p}^{k_i}(z_i)\right| = \left|A_{i,p}^{k_i} z_{i,p}\right| \leq \mu_{i,p}^{k_i} |z_{i,p}| \leq \mu_{i,p}^{k_i} \epsilon,$$

and the requested inequality holds for

$$k_i > \frac{\ln(1 - \eta)}{\ln \mu_{i,p}}. \quad (42)$$

Consider the incoming component,  $z_{i,\text{inc}}$ . We want  $k_i$  such that  $f_{i,\text{inc}}^{k_i}(z_i) < (1 - \eta)\epsilon$  for  $z_i \in N_i^{\text{inc}}$ . But

$$\left|f_{i,\text{inc}}^{k_i}(z_i)\right| = \mu_i^{k_i} |z_{i,\text{inc}}| \leq \mu_i^{k_i} (\sigma + \epsilon).$$

If we take

$$k_i > \frac{\ln \frac{\sigma + \epsilon}{(1 - \eta)\epsilon}}{\ln \mu_i^{-1}}, \quad (43)$$

we obtain the desired inequality.

Now we study the exit components. Take  $z_i \in N_i^{\text{inc}}$  such that its outgoing component  $z_{i,\text{out}}$  satisfies  $|z_{i,\text{out}}| = \epsilon$ . We want to see that

$$\left|f_{i,\text{out}}^{k_i}(z_i)\right| > \sigma + (1 + \eta)\epsilon.$$

Notice that we have:

$$\left|f_{i,\text{out}}^{k_i}(z_i)\right| = \lambda_i^{k_i} |z_{i,\text{out}}| = \lambda_i^{k_i} \epsilon.$$

If we take  $k_i$  such that

$$k_i > \frac{\ln \frac{\sigma + (1 + \eta)\epsilon}{\epsilon}}{\ln \lambda_i}, \quad (44)$$

we obtain the desired inequality.

Finally, for the future components we proceed in the same way. Take  $z_i \in N_i^{\text{inc}}$  such that its future component  $z_{i,f}$  satisfies  $|z_{i,f}| = \epsilon$ . We want to see that

$$\left|f_{i,f}^{k_i}(z_i)\right| > (1 + \eta)\epsilon.$$

But

$$\left|f_{i,f}^{k_i}(z_i)\right| = \left|A_{i,f}^{k_i} z_{i,f}\right| \geq \lambda_{i,f}^{k_i} |z_{i,f}| = \lambda_{i,f}^{k_i} \epsilon,$$

and the requested inequality holds for when

$$k_i > \frac{\ln(1 + \eta)}{\ln \lambda_{i,f}}. \quad (45)$$

To finish the proof of Lemma 10, we take  $k_i$  large enough to satisfy the derived above lower bounds (42)-(45) for  $k_i$ . ■

## 6.2 Dropping of one direction

Now we are going to equip  $N_i^{\text{out}}$  with another  $h$ -set structure,  $\widetilde{N}_i^{\text{out}}$ . We are going to put the outgoing coordinate  $z_{i,\text{out}}$  in the set of entry directions. Notice that  $\widetilde{N}_i^{\text{out}}$  is the same as  $N_i^{\text{out}}$  as sets. We are only changing the declaration of entry and exit coordinates, that is, the  $h$ -set structure.

Notice that it is precisely at this moment where we drop the outgoing direction. This argument is equivalent to the one in Section 3.1 where we intersected some domain with a section of co-dimension one located in the desired outgoing direction. Notice that  $\widetilde{N}_i^{\text{out}}$  have the same number of entry (and exit) components than  $N_{i+1}^{\text{inc}}$ .

## 6.3 Transition along the heteroclinic connection

We define the map close to the heteroclinic segment just as a translation,  $f_{i,i+1}$ . For points in  $\widetilde{N}_i^{\text{out}}$ , that is for points of the form  $q_{i,\text{out}} + z_i$  the map  $f_{i,i+1}$  is defined as:

$$f_{i,i+1}(q_{i,\text{out}} + z_i) = q_{i+1,\text{inc}} + z_i. \quad (46)$$

Notice that, with the transition written in this way we do not have to perform a change of variables that would locate the fixed point  $p_{i+1}$  at the origin. The change is included in the transition.

Our goal is to prove that  $\widetilde{N}_i^{\text{out}}$  covers  $N_{i+1}^{\text{inc}}$ . If we write the transition map in terms of  $z_i$  and  $z_{i+1}$  we have:

$$\begin{aligned} (z_{i,p}, z_{i,\text{inc}}) &= z_{i+1,p} \\ z_{i,\text{out}} &= z_{i+1,\text{inc}} \\ z_{i,f} &= (z_{i+1,\text{out}}, z_{i+1,f}). \end{aligned}$$

With this relation, taking into account the relative sizes of the entry and exit directions in  $\widetilde{N}_i^{\text{out}}$  and  $N_{i+1}^{\text{inc}}$  we can conclude that:

**Lemma 11** *The following covering relation hold:*

$$\widetilde{N}_i^{\text{out}} \xrightarrow{f_{i,i+1}} N_{i+1}^{\text{inc}},$$

for all  $i = 0, \dots, n - 1$ .

## 6.4 The conclusion

By combining Lemmas 10,11 with Theorem 9 we obtain the following

**Theorem 12** *Under the previous assumptions (39), (46), for all  $\epsilon > 0$  there exists a point  $x$  and a sequence of integers  $0 = k_0 < k_1 < \dots < k_n$  such that:*

$$\|f^{k_i}(x) - p_i\| < \epsilon \quad i = 0, \dots, n.$$

## 7 Diffusion in a simplified Toy Model from [CKS+]

### 7.1 The toy model from [CKS+]

The toy model system from [CKS+, page 59, eq. (31)] is given by

$$\frac{d}{dt}b_j = -i|b_j|^2b_j + 2i\bar{b}_j(b_{j-1}^2 + b_{j+1}^2), \quad j \in \mathbb{Z} \quad (47)$$

where  $b_j \in \mathbb{C}$ ,  $j \in \mathbb{Z}$ ,  $i = \sqrt{-1}$ , and  $\bar{z}$  denote the complex conjugation of  $z$ .

System (47) is Hamiltonian with

$$\partial_t b_j = -2i \frac{\partial H}{\partial \bar{b}_j}, \quad \partial_t \bar{b}_j = 2i \frac{\partial H}{\partial b_j}, \quad (48)$$

where

$$H(b) = \sum_j \left( \frac{1}{4}|b_j|^4 - \operatorname{Re}(\bar{b}_j^2 b_{j-1}^2) \right) = \sum_j \left( \frac{1}{4}|b_j|^4 - \frac{1}{2}\bar{b}_j^2 b_{j-1}^2 - \frac{1}{2}b_j^2 \bar{b}_{j-1}^2 \right) \quad (49)$$

Another conserved quantity for (47) called *the mass* is given by

$$M(b) = \sum_j |b_j|^2. \quad (50)$$

Following [CKS+, GK] we are interested in the dynamics (47) on the hypersurface  $M(b) = 1$ , and more specifically in the diffusing solutions, which transport the mass from modes with  $j$  small to modes with large  $j$ . This phenomenon is one of the main ingredients in the transfer of energy to high frequencies in the cubic defocusing NLS on a 2D torus established in [CKS+, GK].

Observe from (47) that on the hypersurface  $M(b) = 1$  there exists the following family of periodic solutions indexed by  $j \in \mathbb{Z}$ :  $b_j(t) = e^{it}$  and  $b_k(t) = 0$  for  $k \neq j$ . Following [CKS+, GK] we denote the  $j$ -th orbit in this family by  $\mathbb{T}_j$ . It turns out there exist heteroclinic connections from  $\mathbb{T}_j$  to  $\mathbb{T}_{j\pm 1}$ . The diffusing orbits constructed in [CKS+, GK] follow the chain of heteroclinic connections  $\mathbb{T}_0 \rightarrow \mathbb{T}_1 \rightarrow \dots \rightarrow \mathbb{T}_N$  for arbitrary  $N \in \mathbb{N}$ .

### 7.1.1 Coordinates from [CKS+] and the local form of the toy model

In the analysis of (47) in [CKS+] (see also [GK]) new coordinates were introduced, which turn each periodic orbit  $\mathbb{T}_j$  into a fixed point.

These coordinates are defined as follows.

We fix  $j$  and we introduce new coordinates  $r, \theta, c_k$  for  $k \neq j$

$$b_j = r e^{i\theta}, \quad b_k = c_k e^{i\theta}. \quad (51)$$

Let

$$\omega = e^{i2\pi/3} = \left( \frac{-1}{2}, \frac{\sqrt{3}}{2} \right). \quad (52)$$

For  $c \in \mathbb{C}$  we set

$$c = \omega c^- + \omega^2 c^+, \quad (53)$$

where  $c^-, c^+ \in \mathbb{R}$ .

The above decomposition means that we represent a complex number  $c$  in the basis  $\{\omega, \omega^2\}$  over the field  $\mathbb{R}$ .

Then we introduce  $c_{j\pm 1}^\pm \in \mathbb{R}$  by

$$c_{j\pm 1} = \omega c_{j\pm 1}^- + \omega^2 c_{j\pm 1}^+. \quad (54)$$

The form of the toy model system in these coordinate is given by the following lemma.

**Lemma 13** [CKS+, Prop. 3.1, page 69] *Consider the toy model with the constraint  $M(b) = 1$ . Let  $c = (c_k)_{k \neq j}$ ,  $c_* = (c_k)_{k \neq j-1, j, j+1}$ .*

*Then the equations for the toy model system have the following form*

$$\dot{c}_{j-1}^- = -\sqrt{3}c_{j-1}^- + O_1(c^2 c_{j-1}^-) + O_2(c_{k \neq j-1}^2 c_{j-1}^+), \quad (55)$$

$$\dot{c}_{j-1}^+ = \sqrt{3}c_{j-1}^+ + O_3(c^2 c_{j-1}^+) + O_4(c_{k \neq j-1}^2 c_{j-1}^-), \quad (56)$$

$$\dot{c}_{j+1}^- = -\sqrt{3}c_{j+1}^- + O_1(c^2 c_{j+1}^-) + O_2(c_{k \neq j+1}^2 c_{j+1}^+), \quad (57)$$

$$\dot{c}_{j+1}^+ = \sqrt{3}c_{j+1}^+ + O_3(c^2 c_{j+1}^+) + O_4(c_{k \neq j+1}^2 c_{j+1}^-), \quad (58)$$

$$\dot{c}_k = ic_k + O(c^2 c_k), \quad k \neq j \pm 1, j, \quad (59)$$

where  $c_{k \neq j-1} = (c_k)_{k \neq j, k \neq j-1}$  Moreover, all these  $O_\gamma$  are uniform with respect to  $k$  and  $j$

### 7.1.2 The transition between consecutive charts

In the coordinate systems ‘centered’ at the  $j$ -th torus we have

$$r_j^2 = 1 - \sum_{k \neq j} |b_k|^2 = 1 - \sum_{k \neq j} |c_k|^2 = 1 - \sum_{k \neq j} r_k^2, \quad (60)$$

$$b_j = r_j e^{i\theta_j}, \quad b_k = c_k e^{i\theta_j}, \quad k \neq j \quad (61)$$

$$c_{j+1} = \omega c_{j+1}^- + \omega^2 c_{j+1}^+, \quad c_{j-1} = \omega c_{j-1}^- + \omega^2 c_{j-1}^+. \quad (62)$$

We look for the relation between coordinates in the  $j$ -th and  $(j+1)$ -th charts. First we will prove the following simple lemma.

**Lemma 14** *If  $c = \omega c_- + \omega^2 c_+$ , then*

$$\frac{1}{c} = \frac{c_+}{|c|^2} \omega + \frac{c_-}{|c|^2} \omega^2.$$

**Proof:** It is easy to see that

$$\overline{c_- \omega + c_+ \omega^2} = c_- \omega^2 + c_+ \omega.$$

Since  $c^{-1} = \frac{\bar{c}}{|c|^2}$ , the assertion follows.  $\blacksquare$

We will denote by the variable with tilde the coordinates expressed in the  $(j+1)$ -th chart.

Since

$$r_{j+1} = |c_{j+1}| = ((c_{j+1}^-)^2 + (c_{j+1}^+)^2 - c_{j+1}^- c_{j+1}^+)^{1/2}, \quad (63)$$

$$e^{i\theta_{j+1}} = \frac{c_{j+1}}{|c_{j+1}|} e^{i\theta_j}, \quad (64)$$

$$e^{i(\theta_j - \theta_{j+1})} = \frac{|c_{j+1}|}{c_{j+1}} = \frac{c_{j+1}^+}{|c_{j+1}|} \omega + \frac{c_{j+1}^-}{|c_{j+1}|} \omega^2 \quad (65)$$

we therefore obtain

$$\tilde{c}_k = c_k e^{i(\theta_j - \theta_{j+1})} = c_k \cdot \frac{|c_{j+1}|}{c_{j+1}}, \quad k < j - 1 \text{ or } k > j + 2, \quad (66)$$

$$\tilde{c}_{j-1} = c_{j-1} \cdot \frac{|c_{j+1}|}{c_{j+1}} = (c_{j-1}^- \omega + c_{j-1}^+ \omega^2) \frac{|c_{j+1}|}{c_{j+1}}, \quad (67)$$

$$\tilde{c}_j = r_j e^{i(\theta_j - \theta_{j+1})} = r_j \frac{|c_{j+1}|}{c_{j+1}} = r_j \left( \frac{c_{j+1}^+}{|c_{j+1}|} \omega + \frac{c_{j+1}^-}{|c_{j+1}|} \omega^2 \right) \quad (68)$$

$$\tilde{c}_j^- = \frac{r_j c_{j+1}^+}{|c_{j+1}|} \quad \tilde{c}_j^+ = \frac{r_j c_{j+1}^-}{|c_{j+1}|} \quad (69)$$

$$\tilde{c}_{j+2} = c_{j+2} \frac{|c_{j+1}|}{c_{j+1}} = \tilde{c}_{j+2}^- \omega + \tilde{c}_{j+2}^+ \omega^2. \quad (70)$$

## 7.2 Simplified toy model

Our model is a simplification of the toy model system from [CKS+, GK]. However, while it is simpler analytically than the original, the problems and obstacles to be overcome to prove the existence of diffusive orbits are very similar.

Our phase space is defined by a sequence of coordinate charts indexed by  $j \in \mathbb{Z}$ . Each of these maps has at its center a fixed point  $\mathbb{T}_j$ .

The  $j$ -th node centered chart (we will sometimes refer to it as  $j$ -th chart) uses the following coordinates



- $c_k \in \mathbb{C}$ ,  $k \leq j - 2$  or  $k \geq j + 2$
- $y_-, x_-, y_+, x_+ \in \mathbb{R}$ .

The coordinates  $y_{\pm}$  and  $x_{\pm}$  are related to the ones from the previous subsection as follows

$$y_- = c_{j-1}^-, \quad x_- = c_{j-1}^+, \quad y_+ = c_{j+1}^-, \quad x_+ = c_{j+1}^+.$$

Let us fix  $\sigma > 0$ . For example we can take  $\sigma = 0.5$  or  $\sigma = 0.1$ . We assume that our system preserves the mass  $\sum_j |c_j|^2$ .

### 7.2.1 Local evolution close to a fixed point

The evolution in the  $j$ -th chart is given by the following ODE

$$\dot{y}_- = -y_- + O(x_-(y_+)^2), \quad (71)$$

$$\dot{x}_- = x_- + O(y_-(x_+)^2), \quad (72)$$

$$\dot{y}_+ = -y_+ + O(x_+(y_-)^2), \quad (73)$$

$$\dot{x}_+ = x_+ + O(y_+(x_-)^2), \quad (74)$$

$$\dot{c}_k = ic_k(1 + O(c^2)), \quad k \leq j - 2 \quad \text{or} \quad k \geq j + 2, \quad (75)$$

where we assume that all  $O()$  terms satisfy

$$|O(z)| \leq K|z|. \quad (76)$$

Let us stress that each  $O(\dots)$  function may depend on all the variables.

Comparing with (55–59), we make the following simplifications:

- we have left only the resonant terms in the first four equations. Removing the non-resonant terms makes perfect sense, as this can be achieved by a suitable coordinate change (see [GK]).
- $\sqrt{3}$  was replaced by 1 in the first four equations (only for simplicity of the notation).

For a given  $T > 0$  we will denote by  $\varphi_T$  a shift by time  $T$  along a trajectory of (71–75) in the domain of the  $j$ -th chart.

### 7.2.2 The dynamics near a heteroclinic connection

In the domain of the  $j$ -th chart we assume that system (71–75) governs the dynamics, and we will use  $\varphi_T$  to map the neighborhood of the entry section  $y_- = \sigma$  to the neighborhood of the exit section  $x_+ = \sigma$ . Next we compose this local map with the global map from the domain of the  $j$ -th chart to the domain of the  $(j + 1)$ -th chart.

Below we postulate the precise form of this “jump between charts” (compare with subsection 7.1.2), which should correspond to the movement along the heteroclinic connection between  $\mathbb{T}_j$  and  $\mathbb{T}_{j+1}$ , followed by the change of coordinates to the  $(j + 1)$ -th chart.

In the formulas given below the variables without tildes are the ones referring to the  $j$ -th node chart, while those with tildes denote the variables with respect to the  $(j + 1)$ -th chart.

$$\begin{aligned}
\tilde{c}_{k \leq j-2} &= c_{k \leq j-2} \\
\tilde{c}_{j-1} &= g_1(x_-, y_-) \\
\tilde{x}_- &= y_+ \\
\tilde{y}_- &= x_+ \\
(\tilde{x}_+, \tilde{y}_+) &= g_2(c_{j+2}) \\
\tilde{c}_{k \geq j+3} &= c_{k \geq j+3},
\end{aligned}$$

where  $g_1 : \mathbb{R}^2 \rightarrow \mathbb{C}$  and  $g_2 : \mathbb{C} \rightarrow \mathbb{R}^2$  are defined by

$$g_1(c^-, c^+) = \omega c^- + \omega^2 c^+, \quad g_2(c) = g_1^{-1}(c). \quad (77)$$

We will denote by  $J$  the map defined by the above equations. In principle the map  $J$  depends on  $j$ , but it will be always clear from the context what is the  $j$  used.

When comparing with the toy model in [CKS+] we basically do the following simplification:

- we assume that entry section  $y_- = \sigma$  in the  $j + 1$ -th chart and the exit section  $x_+ = \sigma$  in the  $j$ -th chart coincide. Therefore there is no need to study the evolution along the heteroclinic connection between these sections, which the authors in [CKS+] were forced to do.
- this simplification makes perfect sense as in [CKS+] this transition map was also studied using the system derived in Lemma 13, which the local system in our approach. Therefore in our model the transition along heteroclinic is achieved by the computation of the local map with the entry and exit sections located at the macroscopic distance from the  $\mathbb{T}_j$ 's.

### 7.3 Fixed points and heteroclinic connections

Observe that in our system we have fixed points  $\mathbb{T}_j$  parameterized by  $j \in \mathbb{Z}$  given in the coordinates centered in the  $j$ -th node by

$$c_{k \leq j-2} = 0, \quad c_{k \geq j+2} = 0, \quad x_- = y_- = x_+ = y_+ = 0. \quad (78)$$

For the fixed point  $\mathbb{T}_j$  the directions  $c_{k \leq j-2}, c_{k \geq j+2}$  are the center directions,  $x_-, x_+$  are the unstable directions and  $y_-, y_+$  are the stable directions.

Consecutive fixed points  $\mathbb{T}_j$  and  $\mathbb{T}_{j+1}$  are connected by a heteroclinic connection escaping the neighborhood of  $\mathbb{T}_j$  along the solution  $c_k(t) = 0$  for  $k \leq j - 2$  or  $k \geq j + 2$ ,  $x_-(t) = y_-(t) = y_+(t) = 0$ ,  $x_+(t) = e^t$  and continued later in the coordinates centered on  $\mathbb{T}_{j+1}$  as  $c_j(t) = 0$  for  $k \leq j - 1$  or  $k \geq j + 3$ ,  $x_-(t) = x_+(t) = y_+(t) = 0$ ,  $y_-(t) = e^{-t}$ .

There are more heteroclinic connections in the toy model (47) see Figure 7.3, however we will only use the connection  $\mathbb{T}_j \rightarrow \mathbb{T}_{j+1}$  described above.

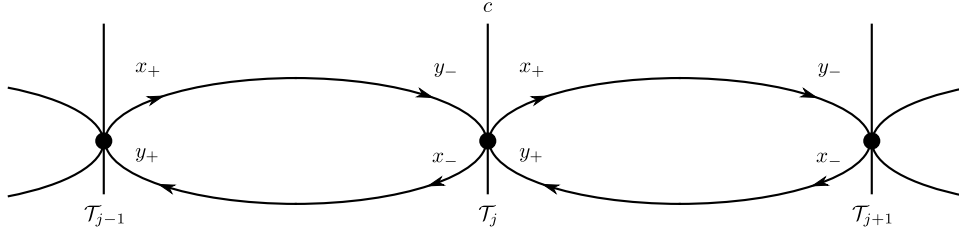


Figure 4: The heteroclinic connections in the toy model

#### 7.4 Some heuristic discussion

Following the heteroclinic chain  $\mathbb{T}_0 \rightarrow \mathbb{T}_1 \rightarrow \dots \rightarrow \mathbb{T}_N$  requires the analysis of consecutive compositions of maps  $\varphi_T$  (the shift along the trajectory by  $T$  in the  $j$ -th chart) with  $J$ , which maps from the  $j$ -th chart to the  $(j+1)$ -th chart.

After passing by the  $j$ -th torus  $\mathbb{T}_j$  we have to drop the  $x_+$ -coordinate in the  $j$ -th chart and then after the application of the  $J$  map, this coordinate becomes  $y_-$  and it is related to  $c_j$ . Therefore we decide that in the vicinity of  $\mathbb{T}_j$  we want  $c_k$  with  $k < j$  (the past modes) to be entry variables and  $c_k$  with  $k > j$  (the future modes) to be the exit directions.

The realization of the above idea must take into account the following issues.

- we declare  $x_-$  to be an entry direction, despite being unstable.
- in the center directions we can have also some growth or decay which might be non-desirable for the past modes and for the future modes, respectively. This growth or decay might turn out to be unbounded for  $T \rightarrow \infty$ .

To overcome these difficulties we have to choose carefully the relative sizes of the variables. For heuristic reasons we will deal with three different sizes: macro, micro and nano. The macro size will be  $O(1)$ , the micro size  $w(T)e^{-T}$  and the nano size  $w(T)e^{-2T}$ , where  $w$  is some polynomial. By increasing  $T$  we will be able to follow the heteroclinic chain as close as we desire, keeping the relative sizes of the variables. This idea appears also in [CKS+].

In the evolution related to passing by  $\mathbb{T}_j$  in the  $j$ -th chart on the exit section  $x_+ = \sigma$  (which is of macro size), we want all the other variables of micro size, i.e.  $w(T)e^{-T}$ , where  $w(T)$  is some polynomial. To get this sizes on the exit section  $x_+ = \sigma$ , we will need to impose on the entry section  $y_- = \sigma$  (the macro size) to the variable  $x_-$  to be of nano size  $w(T)e^{-2T}$ , whereas all the other variables should be of micro size.

Regarding the behavior of the center directions, it turns out that it is possible to maintain the micro size throughout this transition for the variables  $c_k$  for  $k \leq j-2$  or  $k \geq j+2$ , because the possible decay or growth is bounded by a constant, which does not depend on  $T$ . This is the content of Theorem 17.

## 7.5 Evolution estimates in the hyperbolic directions

We consider the system (71–75). Let

$$\sigma' = 1.01 \cdot \sigma . \quad (79)$$

**Theorem 15** *Consider (71–75) satisfying (76) with initial conditions*

$$y_-(0) = \eta, \quad x_-(0) = e^{-2T} a_0, \quad y_+(0) = e^{-T} b_0, \quad x_+(0) = e^{-T} d_0 \quad (80)$$

where  $\eta \in (0, \sigma')$  and  $a_0 \in a$ ,  $b_0 \in b$ ,  $d_0 \in d$  where  $a, b, d$  are real intervals such that

$$a, b, d \subset [-T^k, T^k], \quad k \geq 1, \quad (81)$$

and

$$\sum_l |c_l(0)|^2 = 1. \quad (82)$$

Assume that  $T \geq T_0 > 1$  is large enough, so that the following inequalities are satisfied

$$e^T \geq (K(T^k + T2K\sigma'(T^k + \sigma')^2) (T^k + T4K\sigma'^2(T^k + \sigma')^2))^2 \quad (83)$$

$$e^T \geq K(T^k + T4K\sigma'^2(T^k + \sigma')^2) (T^k + T2K\sigma'(T^k + \sigma')^2)^2 \quad (84)$$

$$Te^{-3T} < \sigma'. \quad (85)$$

Then for  $t \in [0, T]$

$$\begin{aligned} y_-(t) &\in \eta e^{-t} + \alpha e^{-t} t e^{-3T} \\ x_-(t) &\in e^{-2T} a_0 e^t + \alpha t e^{-2T} e^t (2K\sigma' (|d| + \sigma')^2) \\ y_+(t) &\in e^{-T} b_0 e^{-t} + \alpha t e^{-T} e^{-t} (4K\sigma'^2 (|d| + \sigma')) \\ x_+(t) &\in e^{-T} d_0 e^t + \alpha t e^{-4T} e^t, \end{aligned}$$

where

$$\alpha = [-1, 1].$$

In Theorem 15 it is implicitly assumed that we are working in the  $j_0$ -th chart. This index does not appear neither in the statement of the result nor in the proof, because Theorem 15 is concerned only about the hyperbolic directions. Nevertheless, we will complete this result with the estimates in the center directions in Theorem 17 of the next subsection, and there the use of  $j_0$  will be required.

Assumption (82) implies an a-priori bound for all variables so we have a uniform estimate on all  $O()$  terms in equations (71–75).

The proof can be obtained by a direct verification based on differential inequalities and the continuation argument (as in [CKS+]). However we prefer to do it in a more constructive way to show how these bounds have been derived. The proof uses iterations, which produce the stabilizing estimates after the third iterate. The result of the  $l$ -th iteration step will be called the  $l$ -th approximation.

In the estimates we will repeatedly use the following simple lemma.

**Lemma 16** Consider the following one-dimensional ODE

$$x' = \lambda x + e^{\lambda t} D(t), \quad x(0) = x_0$$

where the continuous function  $D : \mathbb{R} \rightarrow \mathbb{R}$  satisfies

$$|D(t)| \leq D, \quad 0 \leq t \leq T$$

Then

$$x(t) \in e^{\lambda t} x_0 + [-1, 1] \cdot D t e^{\lambda t}, \quad 0 \leq t \leq T.$$

### 7.5.1 First approximation

We begin our process with

$$\begin{aligned} y_-(t) &= \eta e^{-t} \\ x_-(t) &= a_0 e^{-2T} e^t \\ y_+(t) &= b_0 e^{-T} e^{-t} \\ x_+(t) &= d_0 e^{-T} e^t. \end{aligned}$$

The estimates for the nonlinear terms (we skip the absolute value sign for  $a, b, d$  in the estimates for the nonlinear terms) are

$$K|x_-(t)(y_+(t))^2| \leq K(e^{-2T} a e^t)(e^{-T} b e^{-t})^2 \leq e^{-4T} e^{-t} (K a b^2) \leq e^{-3T} e^{-t}$$

for  $T$  large enough to satisfy

$$e^T > K a b^2. \quad (86)$$

$$K|y_-(t)(x_+(t))^2| \leq K(\eta e^{-t})(e^{-T} d e^t)^2 \leq e^{-2T} e^t (K \sigma' d^2),$$

$$K|x_+(t)(y_-(t))^2| \leq K(e^{-T} d e^t)(\eta e^{-t})^2 \leq e^{-T} e^{-t} (K d \sigma'^2),$$

$$K|y_+(t)(x_-(t))^2| \leq K(e^{-T} b e^{-t})(e^{-2T} a e^t)^2 = e^{-5T} e^t (K b a^2) \leq e^{-4T} e^t$$

for  $T$  large enough to satisfy

$$e^T > K b a^2. \quad (87)$$

We obtain the following estimates in 1-st approximation

$$\begin{aligned} y_-(t) &\in \eta e^{-t} + \alpha e^{-t} t e^{-3T} \\ x_-(t) &\in e^{-2T} a_0 e^t + \alpha t e^{-2T} e^t (K \sigma' d^2) \\ y_+(t) &\in e^{-T} b_0 e^{-t} + \alpha t e^{-T} e^{-t} (K d \sigma'^2) \\ x_+(t) &\in e^{-T} d_0 e^t + \alpha t e^{-4T} e^t. \end{aligned}$$

### 7.5.2 2-nd approximation

Using the 1-st approximation we obtain

$$\begin{aligned}
K|x_-(t)(y_+(t))^2| &\leq K(e^{-2T}e^t)(a+tK\sigma'd^2)(e^{-T}e^{-t})^2(b+tKd\sigma'^2)^2 = \\
&e^{-4T}e^{-t}(K(a+tK\sigma'd^2))(b+tKd\sigma'^2)^2 \leq e^{-3T}e^{-t} \\
K|y_-(t)(x_+(t))^2| &\leq Ke^{-t}(\sigma'+te^{-3T})(e^{-T}e^t)^2(d+te^{-3T})^2 = \\
&e^{-2T}e^t\left(K(\sigma'+te^{-3T})(d+te^{-3T})^2\right) \\
&\leq e^{-2T}e^t\left(2K\sigma'(d+\sigma')^2\right) \\
K|x_+(t)(y_-(t))^2| &\leq K(e^{-T}e^t)(d+te^{-3T})e^{-2t}(\sigma'+te^{-3T})^2 = \\
&e^{-t}e^{-T}\left(K(d+te^{-3T})(\sigma'+te^{-3T})^2\right) \\
&\leq e^{-t}e^{-T}\left(4K\sigma'^2(d+\sigma')\right) \\
K|y_+(t)(x_-(t))^2| &\leq K(e^{-T}e^{-t})(b+tKd\sigma'^2)(e^{-2T}e^t)^2(a+tK\sigma'^2d)^2 = \\
&e^{-5T}e^t\left(K(b+tKd\sigma'^2)(a+tK\sigma'^2d)^2\right) \leq e^{-4T}e^t
\end{aligned}$$

provided  $T$  is large enough for the following conditions to hold for  $t \in [0, T]$

$$\begin{aligned}
e^T &> (K(a+tK\sigma'd^2))(b+tKd\sigma'^2)^2, \\
te^{-3T} &< \sigma', \\
e^T &> \left(K(b+tKd\sigma'^2)(a+tK\sigma'^2d)^2\right).
\end{aligned}$$

These bounds give us the following estimates in 2-nd approximation

$$\begin{aligned}
y_-(t) &\in \eta e^{-t} + \alpha e^{-t}te^{-3T} \\
x_-(t) &\in e^{-2T}a_0e^t + \alpha te^{-2T}e^t\left(2K\sigma'(d+\sigma')^2\right) \\
y_+(t) &\in e^{-T}b_0e^{-t} + \alpha te^{-T}e^{-t}\left(4K\sigma'^2(d+\sigma')\right) \\
x_+(t) &\in e^{-T}d_0e^t + \alpha te^{-4T}e^t.
\end{aligned}$$

Observe that we have obtained the same formula for  $y_-(t)$  and  $x_+(t)$  in the second approximation as in the first approximation.

### 7.5.3 Next iterate, the 3-rd approximation

Since the bounds for the terms  $y_-(x_+)^2$  and  $x_+(y_-)^2$  will be the same as in the previous subsection, we just compute bounds for  $x_-(y_+)^2$  and  $y_+(x_-)^2$ .

We have

$$\begin{aligned}
& K|x_-(t)(y_+(t))^2| \leq \\
& K(e^{-2T}e^t)(a+t2K\sigma'(d+\sigma')^2)(e^{-T}e^{-t})^2(b+t4K\sigma'^2(d+\sigma'))^2 = \\
& e^{-4T}e^{-t} \left( K(a+t2K\sigma'(d+\sigma')^2)(b+t4K\sigma'^2(d+\sigma'))^2 \right) \leq e^{-3T}e^{-t} \\
& K|y_+(t)(x_-(t))^2| \leq \\
& K(e^{-T}e^{-t})(b+t4K\sigma'^2(d+\sigma'))(e^{-2T}e^t)^2(a+t2K\sigma'(d+\sigma')^2)^2 = \\
& e^{-5T}e^t \left( K(b+t4K\sigma'^2(d+\sigma'))(a+t2K\sigma'(d+\sigma')^2)^2 \right) \leq e^{-4T}e^t
\end{aligned}$$

provided  $T$  is large enough for the following conditions to hold for  $t \in [0, T]$

$$\begin{aligned}
e^T & \geq (K(a+t2K\sigma'(d+\sigma')^2)(b+t4K\sigma'^2(d+\sigma'))^2) \\
e^T & \geq K(b+t4K\sigma'^2(d+\sigma'))(a+t2K\sigma'(d+\sigma')^2)^2
\end{aligned}$$

This concludes the proof of Theorem 15.

## 7.6 The estimates in the center direction

The goal of this subsection to complete the estimates for the flow near  $\mathbb{T}_j$  given in Theorem 15 by providing bounds for  $c_j$ 's for  $j \leq j_0 - 2$  or  $j \geq j_0 + 2$ , where  $j_0$  is the index of the torus on which our chart is centered.

Let  $G = 3/2$  be such that for  $c = c_-\omega + c_+\omega^2$  the following estimate holds

$$|c|^2 \leq G(|c_-|^2 + |c_+|^2).$$

**Theorem 17** *Under the same assumptions as in Theorem 15, take*

$$d = [-2.5\sigma', 2.5\sigma'] \quad (88)$$

and assume that  $c_j = 0$  for  $j < 0$  and  $j > N$ .

Let  $j_0$  be the index of the chart used and assume that

$$c_j(0) = u_j e^{-T} \in u e^{-T}, \quad u \subset [-T^k, T^k]. \quad (89)$$

for the indexes  $j = 0, \dots, N$  such that  $j \leq j_0 - 2$  or  $j \geq j_0 + 2$

Assume also that  $T$  is large enough to satisfy

$$\begin{aligned}
& Ge^{-2T} \left( (T^k + T(2K\sigma'(3.5 \cdot \sigma')^2))^2 + (T^k + T(4K\sigma'(3.5 \cdot \sigma')))^2 \right) + \\
& 2N \exp(21 \cdot KG\sigma'^2) T^{2k+1} e^{-2T} < G\sigma'^2. \quad (90)
\end{aligned}$$

Then

$$|c_j(0)|^2 \exp(-21 \cdot KG\sigma^2) < |c_j(t)|^2 < |c_j(0)|^2 \exp(21 \cdot KG\sigma^2). \quad (91)$$

**Proof:** Since  $\dot{c}_j = ic_j(1 + g(t))$ , where  $g(t) = O(c^2(t))$ , we obtain

$$\frac{d}{dt}|c_j(t)|^2 = \dot{c}_j\bar{c}_j + c_j\dot{\bar{c}}_j = -2|c_j|^2\text{Im}g(t).$$

Therefore we have

$$-2|c_j(t)|^2 \cdot |g(t)| \leq \frac{d}{dt}|c_j(t)|^2 \leq 2|c_j(t)|^2|g(t)|,$$

hence

$$|c_j(0)|^2 \exp\left(-2 \int_0^t |g(s)|ds\right) \leq |c_j(t)|^2 \leq |c_j(0)|^2 \exp\left(2 \int_0^t |g(s)|ds\right). \quad (92)$$

Therefore, it will important to show that under our assumptions on the initial conditions we have a good bound for  $\int_0^T \sum_{j \neq j_0} |c_j|^2(s)ds$ . As the main step to achieve this goal we will show that there exists  $B > 1$  such that

$$|c_j(t)| \in \alpha B |u_j| e^{-T} \quad (93)$$

holds for all  $t \in [0, T]$ . We use a continuation argument. Since  $B > 1$  then (93) is satisfied for  $t \in [0, T']$ . We will show that  $T' \geq T$ .

We have the following estimate

$$\sum_{j \neq j_0} |c_j(t)|^2 \leq G(|x_-(t)|^2 + |y_-(t)|^2 + |x_+(t)|^2 + |y_+(t)|^2) + \sum_{j \neq j_0, j_0 \pm 1} |c_j(t)|^2$$

From (93) we have that, for  $j \neq j_0, j_0 \pm 1$  and  $t \in [0, T']$ ,

$$\int_0^t |c_j(s)|^2 ds \leq \int_0^t B^2 |u_j|^2 e^{-2T} ds \leq B^2 |u_j|^2 T e^{-2T}. \quad (94)$$

To estimate  $\int_0^t |y_-(s)|^2 ds$  we use the bounds obtained in Theorem 15. From assumption (85) of Theorem 15 we get

$$|y_-(t)| \leq 2\sigma' e^{-t}$$

and therefore

$$\int_0^{T'} |y_-(t)|^2 ds \leq 4\sigma'^2 \int_0^\infty e^{-2s} ds = 2\sigma'^2. \quad (95)$$

To estimate  $\int_0^t |x_-(s)|^2 ds$  we use again Theorem 15. From

$$|x_-(t)| \leq e^t e^{-2T} (|a| + T(2K\sigma'(|d| + \sigma')^2)),$$

we get

$$\int_0^{T'} |x_-(s)|^2 ds < 0.5 \cdot e^{-2T} (|a| + T(2K\sigma'(|d| + \sigma')^2))^2. \quad (96)$$



We now estimate  $\int_0^t |y_+(s)|^2 ds$ . From Theorem 15 we have that

$$|y_+(t)| \leq e^{-T} e^{-t} (|b| + T(4K\sigma'(|d| + \sigma'))),$$

hence

$$\int_0^{T'} |y_+(s)|^2 ds < 0.5 \cdot e^{-2T} (|b| + T(4K\sigma'(|d| + \sigma')))^2. \quad (97)$$

It remains to estimate  $\int_0^t |x_+(s)|^2 ds$ . By assumption (85) and the definition (88) of  $d$  we get

$$|x_+(t)| \leq 3.5\sigma' e^{-T} e^t,$$

so that

$$\int_0^{T'} |x_+(s)|^2 ds \leq 3.5^2 \sigma'^2 e^{-2T} \int_0^T e^{2s} ds < 8 \cdot \sigma'^2. \quad (98)$$

By combining (94,95,96,97,98) we directly obtain

$$\begin{aligned} \int_0^{T'} \sum_{j \neq j_0} |c_j(s)|^2 ds &\leq NB^2 |u|^2 T e^{-2T} + G \cdot 2 \cdot \sigma^2 + G \cdot 8 \cdot \sigma'^2 \\ &\quad + G \cdot \left( 0.5 \cdot e^{-2T} (|a| + T(2K\sigma'(|d| + \sigma')^2))^2 \right) + \\ &\quad G \cdot \left( 0.5 \cdot e^{-2T} (|b| + T(4K\sigma'(|d| + \sigma')))^2 \right) < \\ &\quad NB^2 |u|^2 T e^{-2T} + 10G\sigma'^2 + \\ &0.5 \cdot G e^{-2T} \left( (|a| + T(2K\sigma'(|d| + \sigma')^2))^2 + (|b| + T(4K\sigma'(|d| + \sigma')))^2 \right) \end{aligned}$$

Introducing

$$\begin{aligned} E &= NB^2 |u|^2 T e^{-2T} + \\ &0.5 \cdot G e^{-2T} \left( (|a| + T(2K\sigma'(|d| + \sigma')^2))^2 + (|b| + T(4K\sigma'(|d| + \sigma')))^2 \right), \end{aligned}$$

and taking into account (76) we get

$$\int_0^{T'} |g(s)| ds \leq KE + 10KG\sigma'^2.$$

For the continuation argument (93), by (89) and (92), we need that

$$|u_j|^2 e^{-2T} \exp(2KE + 20KG\sigma'^2) < |u_j|^2 B^2 e^{-2T},$$

which is equivalent to

$$\exp(2KE + 20KG\sigma'^2) < B^2.$$

It is clear that this can be achieved if we take  $B$  such that

$$\exp(21 \cdot KG\sigma'^2) = B^2 \quad (99)$$

and then take  $T$  large enough to have

$$Ge^{-2T} \left( (|a| + T(2K\sigma'(|d| + \sigma')^2))^2 + (|b| + T(4K\sigma'(|d| + \sigma')))^2 \right) + 2NB^2|u|^2Te^{-2T} < G\sigma'^2,$$

which is guaranteed by hypothesis (90) on  $T$ .

We have also obtained estimate (91):

$$|c_j(0)|^2 \exp(-21 \cdot KG\sigma'^2) < |c_j(t)|^2 < |c_j(0)|^2 \exp(21 \cdot KG\sigma'^2).$$

■

## 7.7 Construction of the covering relations

The goal of this section is to construct a sequence of covering relations for our model using the estimates obtained in Theorems 15 and 17.

Let us set (compare the estimate for the expansion and contraction rates in the center direction (91) in Theorem 17)

$$A = \exp(21 \cdot KG\sigma'^2). \quad (100)$$

We will have two types of  $h$ -sets in the  $j$ -th chart,  $N_{in}^j$  and  $N_{out}^j$  (to be defined later in this section), such that the following covering relations are satisfied

$$N_{in}^j \xrightarrow{\varphi_T} N_{out}^j, \quad (101)$$

$$R_{<x_+, y_+>} N_{out}^j = \tilde{N}_{out}^j \xrightarrow{J} N_{in}^{j+1}. \quad (102)$$

In relation (101) the map  $\varphi_T$  is the shift along the trajectory by the time  $T$ . The map  $J$  in relation (102) is the jump modeling the transition along the heteroclinic and the change of coordinates. Recall (see Subsection 5.1) that  $R_{<x_+, y_+>} N_{out}^j$  means that we drop the directions  $x_+$  and  $y_+$ , because we have just ‘passed’ by  $T_j$  in our heteroclinic chain.

In the derivation we will use the following conventions

- $\gamma(variable)$  - will be used for the sizes in the entry directions,
- $r(variable)$  - will be used for the sizes in the exit directions.

To be more precise, by the size will mean the radius or half-diameter of the balls or intervals used to define our  $h$ -sets.

Observe that when we are dropping some directions, the sizes in these directions become very close to zero (to set them to zero will not change anything, but it would require slight changes in Theorem 9).

By  $c_p$  (the past modes) we will denote the collection  $\{c_k\}_{k \leq j-2}$  and by  $c_f$  (the future modes) we will denote the collection  $\{c_k\}_{k \geq j+2}$ . On  $c_p$  and  $c_f$  we use the sup norm, i.e.  $\|c_p\| = \sup_{k \leq j-2} |c_k|$ .

The structure of  $h$ -sets  $N_{in}^j$  and  $N_{out}^j$  is defined as follows (we are using the  $j$ -th chart):

- the entry variables:  $c_p, x_-, y_-$
- the exit variables:  $x_+, y_+, c_f$
- parameters of  $N_{in}^j$ :

For the entry directions:

$$\begin{aligned} |c_p| &\leq \gamma_{in}^j(c_p)e^{-T}, \quad (\text{micro}) \\ y_- &\in \sigma + \alpha\gamma_{in}^j(y_-)e^{-T}, \quad \gamma_{in}^j(y_-) \approx 0, \quad (\text{macro}) \\ |x_-| &\leq \gamma_{in}^j(x_-)e^{-2T}, \quad \gamma_{in}^j(x_-) \approx 0. \quad (\text{nano}) \end{aligned}$$

For the exit directions:

$$\begin{aligned} |x_+| &\leq r_{in}^j(x_+)e^{-T}, \quad r_{in}^j(x_+) = 2.1\sigma, \quad (\text{micro}) \quad (103) \\ |y_+| &\leq r_{in}^j(y_+)e^{-T}, \quad (\text{micro}) \\ |c_f| &\leq r_{in}^j(c_f)e^{-T}. \quad (\text{micro}) \end{aligned}$$

- parameters of  $N_{out}^j$ :

For the entry directions:

$$\begin{aligned} |y_-| &\leq \gamma_{out}^j(y_-)e^{-T}, \quad (\text{micro}) \\ |x_-| &\leq \gamma_{out}^j(x_-)e^{-T}, \quad (\text{micro}) \\ |c_p| &\leq \gamma_{out}^j(c_p)e^{-T}. \quad (\text{micro}) \end{aligned}$$

For the exit directions:

$$\begin{aligned} x_+ &\in \sigma + \alpha r_{out}^j(x_+)e^{-T}, \quad r_{out}^j(x_+) \approx 0, \quad (\text{macro}) \\ |y_+| &\leq r_{out}^j(y_+)e^{-2T}, \quad r_{out}^j(y_+) \approx 0, \quad (\text{nano}) \\ |c_f| &\leq r_{out}^j(c_f)e^{-T}. \quad (\text{micro}) \end{aligned}$$

### 7.7.1 Covering $N_{in}^j \xrightarrow{\varphi^T} N_{out}^j$

We use Theorems 15 and 17 for the shift along the trajectory by time  $T$  with

$$\begin{aligned} \eta &= \sigma + \alpha\gamma_{in}^j(y_-)e^{-T}, \quad a = \alpha\gamma_{in}^j(x_-), \\ b &= \alpha r_{in}^j(y_+), \quad d = \alpha r_{in}^j(x_+). \end{aligned} \quad (104)$$

To satisfy the assumptions of Theorem 15 about the size of  $|y_-|$  we require that

$$\gamma_{in}^j(y_-)e^{-T} < \sigma' - \sigma = 0.01\sigma. \quad (105)$$

The conditions for  $N_{in}^j \xrightarrow{\varphi^T} N_{out}^j$  are

- entry conditions:

– for  $c_p$  variables:

$$A\gamma_{in}^j(c_p) \leq \gamma_{out}^j(c_p). \quad (106)$$

– for  $x_-$  from Theorem 15 and (104) we have the following condition

$$\gamma_{in}^j(x_-)e^{-T} + Te^{-T}(2K\sigma'(3.1\sigma')^2) < \gamma_{out}^j(x_-)e^{-T}$$

therefore it is enough to take

$$T(2.1K\sigma'(3.1\sigma')^2) \leq \gamma_{out}^j(x_-), \quad (107)$$

if  $\gamma_{in}^j(x_-) \approx 0$ , which will turn out to be compatible with other conditions. In fact we had replaced 2 by 2.1 and  $\sigma$  by  $\sigma'$  ( $\sigma$  appears in  $d$ ) in order to make an explicit margin for  $\gamma_{in}^j(x_-)$  given by

$$\gamma_{in}^j(x_-) < T(0.1K\sigma'(3.1\sigma')^2). \quad (108)$$

– for  $y_-$  it is enough to have

$$\sigma'e^{-T} + Te^{-4T} < \gamma_{out}^j(y_-)e^{-T},$$

and from (85) it is enough to take

$$2\sigma' \leq \gamma_{out}^j(y_-). \quad (109)$$

• exit conditions:

–  $x_+$

$$r_{in}^j(x_+) - Te^{-3T} > \sigma + r_{out}^j(x_+)e^{-T},$$

which in view of assumption (85),(79) and (104) is satisfied, if

$$r_{out}^j(x_+)e^{-T} < 0.09\sigma. \quad (110)$$

Obviously (110) is compatible with  $r_{out}^j(x_+) \approx 0$ , which is to be expected as this is the direction which will be dropped in the next covering relation.

–  $y_+$ , from Theorem 15 and (104) it follows that the following estimate is sufficient

$$r_{in}^j(y_+)e^{-2T} - Te^{-2T} \left( 4K\sigma'^2(r_{in}^j(x_+) + \sigma') \right) > r_{out}^j(y_+)e^{-2T}. \quad (111)$$

Observe that this is the direction which is dropped in the next covering relation, hence any  $r_{out}^j(y_+) > 0$ ,  $r_{out}^j(y_+) \approx 0$  is good for our construction. Therefore we can take (where we used also the known value (103) of  $r_{in}^j(x_+)$ )

$$r_{in}^j(y_+) \geq T(4.1K\sigma'^2(3.1\sigma')), \quad (112)$$

which leaves some margin for  $r_{out}^j(y_+)$ , given by

$$r_{out}^j(y_+) < T(0.1K\sigma'^2(3.1\sigma')). \quad (113)$$

–  $c_f$

$$A^{-1}r_{in}^j(c_f) \geq r_{out}^j(c_f). \quad (114)$$

### 7.7.2 Covering relation (102)

Let  $L \geq 1$  be the Lipschitz constant which holds for both functions  $g_1$  and  $g_2$  introduced in (77), where on  $\mathbb{R}^2$  the max-norm is used, whereas in  $\mathbb{C}$  we use the euclidian norm.

The conditions are as follows.

In the entry directions

$$L(\gamma_{out}^j(y_-) + \gamma_{out}^j(x_-)) \leq \gamma_{in}^{j+1}(c_p), \quad (115)$$

$$\gamma_{out}^j(c_p) < \gamma_{in}^{j+1}(c_p), \quad (116)$$

$$r_{out}^j(x_+) < \gamma_{in}^{j+1}(y_-), \quad (117)$$

$$r_{out}^j(y_+) < \gamma_{in}^{j+1}(x_-). \quad (118)$$

In the exit directions

$$r_{out}^j(c_f) \geq Lr_{in}^{j+1}(x_+), \quad (119)$$

$$r_{out}^j(c_f) \geq Lr_{in}^{j+1}(y_+), \quad (120)$$

$$r_{out}^j(c_f) > r_{in}^{j+1}(c_f). \quad (121)$$

### 7.7.3 Solving the inequalities for coverings

We have to find the following set of parameters  $\gamma_{in,out}^j(c_p, x_-, y_-)$  and  $r_{in,out}^j(x_+, y_+, c_f)$ , such that inequalities (115-121) are satisfied.

We split these parameters into two groups: the ones related to the dropped directions  $\gamma_{in}^j(y_-)$ ,  $\gamma_{in}^j(x_-)$ ,  $r_{out}^j(x_+)$ ,  $r^j(y_+)$ , and the remaining ones.

In the first group we effectively should obtain

$$\gamma_{in}^j(y_-) = 0, \quad \gamma_{in}^j(x_-) = 0, \quad r_{out}^j(x_+) = 0, \quad r_{out}^j(y_+) = 0.$$

The conditions involving these parameters are (108), (110), (113), (117) and (118). It is clear that these conditions can be easily satisfied with all these parameters being very close to zero.

Now we deal with the other directions. We already have set the value for  $r_{in}^j(x_+)$  in (103) and we now set the following parameters (compare with (107),(109), (112))

$$r_{in}^j(x_+) = 2.1\sigma, \quad (122)$$

$$\gamma_{out}^j(x_-) = TQ_1, \quad Q_1 = (2.1K\sigma'(3.1\sigma')^2), \quad (122)$$

$$\gamma_{out}^j(y_-) = 2\sigma', \quad (123)$$

$$r_{in}^j(y_+) = TQ_2, \quad Q_2 = (4.1K\sigma'^2(3.1\sigma')), \quad (124)$$

The remaining inequalities involve only the sizes for the variables  $c_p$  and  $c_f$ . These are as follows:

- for the entry directions (see (106), (115), (116))

$$\begin{aligned} A\gamma_{in}^j(c_p) &\leq \gamma_{out}^j(c_p) \\ L(2\sigma' + TQ_1) &\leq \gamma_{in}^{j+1}(c_p), \\ \gamma_{out}^j(c_p) &< \gamma_{in}^{j+1}(c_p), \end{aligned}$$

- for the exit directions (see (114), (119), (120) and (121))

$$\begin{aligned} A^{-1}r_{in}^j(c_f) &\geq r_{out}^j(c_f), \\ r_{out}^j(c_f) &\geq 2.1L\sigma, \\ r_{out}^j(c_f) &\geq TLQ_2, \\ r_{out}^j(c_f) &> r_{in}^{j+1}(c_f). \end{aligned}$$

It is clear that there exist  $\gamma_{in}^j(c_p), \gamma_{out}^j(c_p)$  satisfying

$$Q_3T < \gamma_{in}^j(c_p) \leq A^{-1}\gamma_{out}^j(c_p) < \gamma_{in}^{j+1}(c_p), \quad j = 0, \dots, N \quad (125)$$

where  $Q_3 = L\left(\frac{2\sigma'}{T} + Q_1\right)$ .

For instance we take

$$\gamma_{in}^0(c_p) = \tilde{Q}_3T, \quad \gamma_{in}^{j+1}(c_p) = \tilde{A}\gamma_{in}^j(c_p), \quad \gamma_{out}^j(c_p) = A\gamma_{in}^j(c_p), \quad j = 0, \dots, N-1$$

where  $\tilde{Q}_3 > Q_3$  and  $\tilde{A} > A$ . With such sequence we have solved the inequalities for the entry directions. Observe that as a consequence we have

$$Q_3T < \gamma_{in}^j(c_p) < \gamma_{out}^j(c_p) \leq \tilde{Q}_3\tilde{A}^N T, \quad j = 0, \dots, N. \quad (126)$$

In the exit direction the situation is similar. We just take any sequence satisfying

$$\max(TLQ_2, 2.1L\sigma) \leq r_{out}^{j+1}(c_f) \leq A^{-1}r_{in}^{j+1}(c_f) < A^{-1}r_{out}^j(c_f). \quad (127)$$

For example we can take for  $j = N, \dots, 1$

$$r_{out}^N(c_f) = \max(TLQ_2, 2.1L\sigma), \quad r_{out}^{j-1}(c_f) = \tilde{A}r_{out}^j(c_f), \quad r_{in}^j(c_f) = Ar_{out}^j(c_f). \quad (128)$$

Hence we obtain

$$r_{out}^j(c_f), r_{in}^j(c_f) \leq \tilde{A}^N \max(TLQ_2, 2.1L\sigma), \quad j = 0, \dots, N. \quad (129)$$

In (126) and (129) we introduced also an upper bound which is  $O(T)$ , so now all sizes are  $O(T)$  times a suitable weight function ( $e^{-T}$  or  $e^{-2T}$ ). Observe that this bound allows us to use Theorems 15 and 17 with  $k = 2$ , for any  $T \geq \tilde{A}^N \max(2LQ_2, 2Q_3, Q_2, Q_1, 1)$ , because then all sizes will be less than  $T^2$ .

## 7.8 The conclusion

From the chain of coverings constructed above and Theorem 9 since the distance from the heteroclinics in the chain are  $O(T)e^{-T}$  we obtain the following theorem.

**Theorem 18** *For the system discussed in this section for any  $N$ , for all  $\epsilon > 0$  there exists a point  $x_0$  close to  $\mathbb{T}_0$  whose trajectory is  $\epsilon$  close to the chain of heteroclinic connections  $\mathbb{T}_0 \rightarrow \mathbb{T}_1 \rightarrow \cdots \rightarrow \mathbb{T}_N$ .*

## References

- [Ar] V. I. Arnold, *Instability of dynamical systems with several degrees of freedom*, Soviet Mathematics 5(1964),581–585
- [BM+] R. Barrio, M.A. Martinez, S. Serrano, D. Wilczak, *When chaos meets hyperchaos: 4D Rössler model*, Physics Letters A, Vol. 379, No. 38, 2300-2305 (2015).
- [CKS+] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. math (2010) 181:39-113
- [C] C. C. Conley *Isolated Invariant Sets and the Morse Index*. 1978. CBMS vol. 38, Amer. Math. Soc., Providence
- [GK] M. Guardia, V. Kaloshin, *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation*, Journal of the European Mathematical Society, vol. 17, 1, 71–149, 2015
- [MM] K. Mischaikow, M. Mrozek, The Conley Index, in: Handbook of Dynamical Systems II: Towards Applications, (B. Fiedler, ed.) North-Holland, 2002.
- [S] J. Smoller, Shock Waves and Reaction -Diffusion Equations. Springer, 1983.
- [W2] D. Wilczak, *The existence of Shilnikov homoclinic orbits in the Michelson system: a computer assisted proof*. Foundations of Computational Mathematics, Vol.6, No.4, 495-535, (2006).
- [WBS] D. Wilczak, S. Serrano, R. Barrio, *Coexistence and dynamical connections between hyperchaos and chaos in the 4D Rössler system: a Computer-assisted proof*, SIAM Journal on Applied Dynamical Systems, 15 (2016), 356–390
- [ZGi] P. Zgliczyński and M. Gidea, *Covering relations for multidimensional dynamical systems*, Journal of Differential Equations 202/1, 32–58 (2004)