Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Rigorous Numerics for Dissipative PDEs III. An effective algorithm for rigorous computation of trajectories and Poincaré maps

by

Piotr Zgliczynski

Preprint no.: 39 2005

Rigorous Numerics for Dissipative PDEs III. An effective algorithm for rigorous computation of trajectories and Poincaré maps

Piotr Zgliczyński[1 2]

Jagiellonian University, Institute of Computer Science,
Nawojki 11, 30–072 Kraków, Poland
e-mail: umzglicz@cyf-kr.edu.pl

and

Max Planck Institute for Mathematics in the Sciences,
Inselstrasse 22, D-04103 Leipzig, Germany

April 13, 2005

Abstract

We describe a Lohner-type algorithm for rigorous computation of dissipative PDEs with periodic boundary conditions. The algorithm have been implemented for the Kuramoto-Sivashinsky PDE with odd and periodic boundary conditions and was used to prove the existence of multiple periodic orbits, both stable and unstable ones.

Keywords: periodic orbits, dissipative PDEs, Galerkin projection, rigorous numerics

AMS classification: 35B40, 35B45, 65G30, 65N30

1 Introduction

The goal of this paper is to present the algorithm for a rigorous numerical integration of a certain class of dissipative PDEs. To be more specific we consider PDEs of the following type

$$u_t = Lu + N(u, Dx, \ldots, D^r x), \tag{1}$$

where $u \in \mathbb{R}^n$, $x \in \mathbb{T}^d$, ($\mathbb{T}^d = (\mathbb{R}/2\pi)^d$ is an d-dimensional torus), L is a linear operator, N - a polynomial and by $D^s u$ we denote s-th order derivative of u, i.e. the collection of all partial derivatives of u of order s. We require that L is diagonal in the Fourier basis $\{e^{kx}\}_{k \in \mathbb{Z}^d}$, namely

$$Le^{ikx} = \lambda_k e^{ikx}, \tag{2}$$

[1]Research supported in part by Polish State Committee for Scientific Research grant 2 P03A 019 22

[2]most of this work was done during authors stay at Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

and the eigenvalues λ_k satisfy

$$\lambda_k = -v(|k|)|k|^p \quad (3)$$

$$0 < v_0 \leq v(|k|) \leq v_1, \qquad \text{for } |k| > K_- \quad (4)$$

$$p > r. \quad (5)$$

The fact that we are considering functions on the torus means that we impose periodic boundary conditions. We may eventually seek odd or even solutions or impose some other conditions.

Our approach starts with replacing (1) by an infinite ladder of ordinary differential equations for Fourier coefficients of $u(t,x) = \sum_k u_k(t)e^{ikx}$. We obtain

$$\frac{du_k}{dt} = \lambda_k u_k + N_k(u), \qquad \text{for all } k \in \mathbb{Z}^d. \quad (6)$$

The next step is to split the phase space for (6) into two parts: the finite dimensional part, X, containing the Fourier modes most relevant for the dynamics of (1) and the tail $T \subset X^\perp$. After this splitting the problem (6) is replaced by two problems (7) and (8). The first part consist of a finite dimensional differential inclusion for $p \in X$, given by

$$\frac{dp}{dt} \in P(Lp + N(p+T)), \qquad p \in X \quad (7)$$

where P is a projection onto X. The second part is concerned with the evolution of T, which is governed by an infinite set of inequalities of the form

$$\lambda_k u_{k,j} + N^-_{k,j} < \frac{du_{k,j}}{dt} < \lambda_k u_{k,j} + N^+_{k,j}, \qquad j = 1,\ldots,n \text{ and for } k \text{ not in } X \quad (8)$$

where $N^\pm_{k,j}$ are suitably chosen constants. Obviously, to infer from (7) and (8) any information on the behavior of solutions of the full system (6) one needs some consistency conditions. A systematic treatment of this issue is at the heart of our *method of self-consistent bounds*, which was introduced in [ZM] and later developed in [ZAKS, ZGal, ZNS, Z2].

The main example treated in this paper is the Kuramoto-Sivashinsky PDE [KT, S] (in the sequel we will refer to it as the KS equation)

$$u_t = -\nu u_{xxxx} - u_{xx} + (u^2)_x, \qquad \nu > 0 \quad (9)$$

where $x \in \mathbb{R}$, $u(t,x) \in \mathbb{R}$ and we impose odd and periodic boundary conditions

$$u(t,x) = -u(t,-x), \qquad u(t,x) = u(t, x+2\pi). \quad (10)$$

The choice of the KS equation for this study is motivated by the following facts

- the existence theory and asymptotic properties of solutions of (9) are well established, see for example [CEES, FT, FNST] and the literature cited there. It should be stressed that we are not using these results in our work, but they assure us that all interesting dynamics is 'finite dimensional' and should be accessible using the method of self-consistent bounds combined with topological tools.

- there exists a lot of numerical studies of the dynamics of the KS equation (see for example [CCP, HN, JKT, JJK]), where it was shown that the dynamics of the KS equation is highly nontrivial and it is well represented by relatively small number of modes.
- we believe that the experience gained and new tools developed in the study of the KS equation may help in the rigorous study of the dynamics of the Navier-Stokes equations or the Ginzburg-Landau equation [Si].

We implemented the proposed algorithm for the KS equation (9) with the odd and periodic boundary conditions (10). Using it we proved the existence of several periodic orbits, both attracting and unstable ones, for various parameter values of ν in the interval $[0.02991, 0.128]$. Proofs are topological and are based on the Brouwer Theorem in case of attracting orbits and on the Miranda Theorem[Mi] in case of unstable ones. The main difference between this paper and [Z2] is the generality and the efficiency of the algorithm for a rigorous integration of (9). The algorithm described in [Z2] required some preparatory work to construct the a-priori bounds, which have to be verified during the computation, moreover the tail was fixed in the computation. The present algorithm allows for the tail evolution and do not require any a-priori bounds to start the computation, hence it could be used also to obtain rigorous bounds for the forward orbit of any initial condition with a finite description, this was not possible using the previous algorithm. Other improvements, while rather technical, are also of great importance for the performance of the algorithm. They include a new function for the generation of the rough enclosure for differential inclusions, which allows to use much larger time steps. All these improvements taken together result in more than 6 fold speed up of the proof of the existence of periodic orbit for $\nu = 0.127$. This orbit has the reflectional symmetry and this fact was essential in the proof because it allowed us to consider the half-Poincaré map instead of the full Poincaré map. Our attempts to compute the full Poincaré map along this periodic orbit using the previous algorithm failed due to blow-up of rigorous enclosures produced. Using the current algorithm we were able to overcome this problem and also treat smaller values of ν, which is more difficult computationally and more interesting from the dynamics standpoint.

The choice of odd boundary conditions was motivated by earlier numerical studies [CCP, JKT], but the basic mathematical reason is: equation (9) with periodic boundary conditions has the translational symmetry, which implies that for fixed value of ν periodic orbits (fixed points, etc) are members of one-parameter families of periodic orbits (fixed points, etc). The restriction to the subspace of odd functions breaks this symmetry and gives a hope that the dynamically interesting objects are topologically isolated, which is later confirmed in proofs.

The content of this paper can be described as follows: in Sections 2 and 3 we outline the method of self-consistent bounds and discuss how it can be used for the study of dynamics of dissipative PDEs. This material is based on [ZM, ZNS, ZGal], but some new theorems about the general applicability to other dissipative PDEs are added in Section 3. In Section 4 we present a Lohner-type algorithm for the integration of ordinary differential inclusions, which in the context of the rigorous integration of PDEs is used to provide enclosures for (7). In Section 5 we present a new effective enclosure theorem for ordinary differential inclusions and a function based on it. In Sections 6, 7 and 8 we discuss the algorithm with an evolving tail for the rigorous integration of dissipative PDEs with periodic boundary conditions. In Section 9 we treat the issue of Poincaré maps. In the remaining sections we report on the computer assisted proofs of the existence of periodic orbits for the KS equation, both stable and unstable ones.

1.1 Notation

Let (T,ρ) be a metric space. For a set $X \subset T$ by $\mathrm{int}\, X$, $\overline{X}$ and ∂X we denote the interior, the closure and the boundary of X, respectively. If $X \subset Y \subset T$, then by $\mathrm{int}_Y X$ and by $\partial_Y X$ we will denote respectively the interior and the boundary of X with respect to the metric space (Y,ρ). By $B(c,r) = \{x \mid \rho(c,x) < r\}$ we will denote the ball of radius r. For a point $p \in T$ put $\rho(p,X) = \inf\{\rho(p,q) \mid q \in X\}$. We define $B(X,\epsilon) = \{y \mid \rho(y,X) < \epsilon\}$. The Hausdorff distance, $\mathrm{dist}\,(G,H)$, between two closed sets G and H is defined by the formula

$$\mathrm{dist}\,(G,H) = \max\{\sup_{q\in G} \rho(q,H), \sup_{h\in H} \rho(h,G)\}.$$

For $-\infty \leq t_0 < t_1 \leq \infty$ by $C([t_0,t_1],\mathbb{R}^s)$ we will denote the set of all continuous functions defined on $[t_0,t_1]$ with the values in $\mathbb{R}^s$ and by $C_b([t_0,t_1],\mathbb{R}^s)$ we will denote the set of all bounded and continuous functions defined on $[t_0,t_1]$ with the values in $\mathbb{R}^s$.

Let $cf(\mathbb{R}^n)$ denotes the set of all nonempty, convex and compact subsets of $\mathbb{R}^n$. A multivalued map $f : \mathbb{R}^n \to cf(\mathbb{R}^n)$ is said to be *continuous* if it is continuous with respect to the Hausdorff distance.

For an ordinary differential equation

$$x' = f(x), \qquad x \in \mathbb{R}^n \tag{11}$$

where $f \in \mathcal{C}^1$, by φ we will denote the local flow induced by (11). We set $\varphi(t,x_0) = x(t)$ where $x(t)$ is the unique solution of (11) with the initial condition $x(0) = x_0$.

Let $f : \mathbb{R}^n \to cf(\mathbb{R}^n)$ be continuous. Consider a differential inclusion

$$x' \in f(x), \tag{12}$$

By a solution of (12) through x_0 we will understand a C^1 function $x : (t_0,t_1) \in \mathbb{R}^n$, such that $0 \in (t_0,t_1)$, $x(0) = x_0$ and (12) holds for $t \in (t_0,t_1)$. Moreover we will always assume that the solution is defined on the maximal existence interval.

We define the local flow, φ, induced by (12) as follows: $(t,x_0) \in \mathbb{R} \times \mathbb{R}^n$ is in the domain of φ if for all solutions x through x_0 the value of $x(t)$ is defined and then

$$\varphi(t,x_0) = \{x(t) \mid x(t) \text{ is a solution through } x_0\}. \tag{13}$$

While we will use the same symbol $\varphi(t,x)$ to indicate the local flow induced both by an ODE or an inclusion it will be always clear from the context what type of the flow we are considering.

In the sequel we will use an expression of the form

$$\varphi([0,h],x_0) \subset Z. \tag{14}$$

Such expression means that $\varphi([0,h],x_0)$ is defined for $t \in [0,h]$ and the stated inclusion holds, i.e. $\varphi(t,x_0) \subset Z$ for $t \in [0,h]$.

2 The method of self-consistent bounds

We begin with an abstract nonlinear evolution equation in a real Hilbert space H (L^2 or some its subspaces in our treatment of dissipative PDEs) of the form

$$\frac{du}{dt} = F(u) \tag{15}$$

where the domain of F is dense in H. By a solution of (15) we understand a function $u : [0,t_{max}) \to \mathrm{dom}\,(F)$, such that u is differentiable and (15) is satisfied for all $t \in [0,t_{max})$.

The scalar product in H will be denoted by $(u|v)$. Throughout the paper we assume that there is a set $I \subset \mathbb{Z}^d$ and a sequence of subspaces $H_k \subset H$ for $k \in I$, such that $\dim H_k = d_1 < \infty$ and H_k and $H_{k'}$ are mutually orthogonal for $k \neq k'$. Let $A_k : H \to H_k$ be the orthogonal projection onto H_k. We assume that for each $u \in H$ holds

$$u = \sum_{k \in I} u_k = \sum_{k \in I} A_k u \tag{16}$$

The above equality for a given $u \in H$ and $k \in I$ defines u_k. Analogously if B is a function with the range in H, then $B_k(u) = A_k B(u)$.

For $k \in \mathbb{Z}^d$ we define

$$|k| = \sqrt{\sum_{i=1}^{d} k_i^2}$$

For $n > 0$ we set

$$X_n = \bigoplus_{|k| \leq n, k \in I} H_k$$
$$Y_n = X_n^{\perp},$$

by $P_n : H \to X_n$ and $Q_n : H \to Y_n$ we will denote the orthogonal projections onto X_n and onto Y_n, respectively.

Definition 1 *We say that $F : H \supset \mathrm{dom}\,(F) \to H$ is admissible if the following conditions are satisfied for any $i \in \mathbb{R}$, such that $\dim X_i > 0$*

- $X_i \subset \mathrm{dom}\,(F)$
- $P_i F : X_i \to X_i$ *is a* C^1 *function*

Definition 2 *Assume* F *is admissible. For a given number* $n > 0$ *the ordinary differential equation*

$$x' = P_n F(x), \qquad x \in X_n \tag{17}$$

will be called the n-th Galerkin projection *of (15).*

By $\varphi^n(t, x)$ *we denote the local flow on* X_n *induced by (17).*

Definition 3 *Assume* F *is an admissible function. Let* $m, M \in \mathbb{R}$ *with* $m \leq M$*. Consider an object consisting of: a compact set* $W \subset X_m$ *and a sequence of compact sets* $B_k \subset H_k$ *for* $|k| > m$*. We define the conditions* ***C1****,* ***C2****,* ***C3****,* ***C4a*** *as follows:*

C1 *For* $|k| > M$*,* $0 \in B_k$ *.*

C2 *Let* $\hat{a}_k := \max |a_k^{\pm}|$ *for* $|k| > m$*,* $k \in I$ *and then* $\sum_{|k|>m, k\in I} \hat{a}_k^2 < \infty$*. In particular*

$$W \oplus \Pi_{|k|>m} B_k \subset H \tag{18}$$

and for every $u \in W \oplus \Pi_{|k|>m} B_k$ *holds,* $\|Q_n u\| \leq \sum_{|k|>n} \hat{a}_k^2$*.*

C3 *The function* $u \mapsto F(u)$ *is continuous on* $W \oplus \prod_{|k|>m} B_k \subset H$*.*

Moreover, if we define for $k \in I$*,* $f_k = \max_{u \in W \oplus \prod_{|k|>m} B_k} |F_k(u)|$*, then* $\sum f_k^2 < \infty$*.*

C4a *For* $|k| > m$ B_k *is given by (19) or (20)*

$$B_k = \overline{B(c_k, r_k)}, \quad r_k > 0 \tag{19}$$

$$B_k = \Pi_{s=1}^{d} [a_s^-, a_s^+], \qquad a_s^- < a_s^+,\ s = 1, \dots, d_1 \tag{20}$$

Let $u \in W \oplus \Pi_{|k|>m} B_k$*. Then for* $|k| > m$ *holds:*

- *if* B_k *is given by (19) then*

$$u_k \in \partial_{H_k} B_k \quad \Rightarrow \quad (u_k - c_k | F_k(u)) < 0. \tag{21}$$

- *if* B_k *is given by (20) then*

$$u_{k,s} = a_{k,s}^- \quad \Rightarrow \quad F_{k,s}(u)) > 0, \tag{22}$$

$$u_{k,s} = a_{k,s}^+ \quad \Rightarrow \quad F_{k,s}(u)) < 0. \tag{23}$$

In the sequel we will refer to equations (21) and (22–23) as *isolation equations.*

Definition 4 *[ZM, Def. 2.1, 2.11] Assume F is an admissible function. Let $m, M \in \mathbb{R}$ with $m \leq M$. Consider an object consisting of: a compact set $W \subset X_m$ and a sequence of compacts $B_k \subset H_k$ for $|k| > m, k \in I$. We say that set $W \oplus \Pi_{|k|>m} B_k$ forms* self-consistent bounds *for F if conditions C1, C2, C3 are satisfied.*

If additionally condition C4a holds, then we say that $W \oplus \Pi_{|k|>m} B_k$ forms topologically self-consistent bounds for F

If F is clear from the context, then we will often drop F, and we will speak simply about *self-consistent bounds* or *topologically self-consistent bounds*.

In our previous works on the KS equation [ZM, ZAKS, Z2], we had $I = \mathbb{Z}_+$, $H_k = \mathbb{R}$ and $B_k = [a_k^-, a_k^+]$. The conditions from Def. 3 are generalizations of the conditions given there to a more general setting.

Reader familiar with our earlier works should be also warned in the terminology of [Z2, Def. 2] conditions C1,C2,C3,C4a defined *self-consistent a-priori bounds*. In this paper we returned to the terminology used in [ZM] and we dropped the phrase *a-priori*.

Given self-consistent bounds W and $\{B_k\}_{k \in I, |k|>m}$, by T (the tail) we will denote

$$T := \prod_{|k|>m} B_k \subset Y_m.$$

Here are some useful lemmas illustrating the implications of conditions C1, C2, C3.

From condition C2 it follows immediately that

Lemma 1 *If $W \oplus T$ forms self-consistent bounds, then $W \oplus T$ is a compact subset of H.*

The following lemma is an immediate consequence of conditions C2 and C3.

Lemma 2 *Given self-consistent bounds $W \oplus T$, then*

$$\lim_{n \to \infty} P_n(F(u)) = F(u), \quad \textit{uniformly for } u \in W \oplus T$$

The lemma below was proved in [Z2, Lemma 5], where the definition of self-consistent bounds required conditions C1,C2, C3 and C4a and $\dim H_k = 1$, but the condition C4a and the dimension of H_k were not used in the proof. Hence we can write this lemma as follows

Lemma 3 *Let $W \oplus T$ forms self-consistent bounds for (15). Let $\{d_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ be a sequence, such that $\lim_{n \to \infty} d_n = \infty$. Assume that, for all n, $x_n : [t_1, t_2] \to W \oplus T$ is a solution of*

$$\frac{dp}{dt} = P_{d_n}(F(p)), \qquad p(t) \in X_{d_n}. \tag{24}$$

Then there exists a convergent subsequence $\{d_{n_l}\}_{l \in \mathbb{N}}$ such that, $\lim_{l \to \infty} x_{n_l} = x^*$, where $x^* : [t_1, t_2] \to W \oplus T$ and the convergence is uniform on $[t_1, t_2]$. Moreover, x^* satisfies (15).

Later we will need a slightly stronger version of the above lemma, which we state without a proof, because the proof of Lemma 3 works for this version.

Lemma 4 *Let $W_i \oplus T_i$, $i = 1, \dots, k$ forms self-consistent bounds for (15). Let $\{d_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence, such that $\lim_{n\to\infty} d_n = \infty$. Assume that, for all n, $x_n : [t_1, t_2] \to \bigcup_{i=1}^k W_i \oplus T_i$ is a solution of*

$$\frac{dp}{dt} = P_{d_n}(F(p)), \qquad p(t) \in X_{d_n}. \tag{25}$$

Then there exists a convergent subsequence $\{d_{n_l}\}_{l\in\mathbb{N}}$ such that, $\lim_{l\to\infty} x_{n_l} = x^*$, where $x^* : [t_1, t_2] \to \bigcup_{i=1}^k W_i \oplus T_i$ and the convergence is uniform on $[t_1, t_2]$. Moreover, x^* satisfies (15).

3 The existence of uniform bounds for all Galerkin projections for short time steps

Consider equation (1). We assume that conditions (3),(4) and (5) are satisfied. If $a(t, x)$ is a sufficiently regular solution of (1), then we can expand it in Fourier series $a(t, x) = \sum_{k\in\mathbb{Z}^d} a_k(t)e^{ik\cdot x}$ to obtain an infinite ladder of ordinary differential equations for the coefficients a_k

$$\frac{da_k}{dt} = \lambda_k a_k + N_k(a), \quad k \in \mathbb{Z}^d, \tag{26}$$

where $N_k(a)$ is k-th Fourier coefficient of function $N(a, Da, \dots, D^r a)$.

Observe that $a_k \in \mathbb{C}^n$ and (26) are not independent, because the reality of a imposes the following condition

$$a_{-k} = \overline{a}_k. \tag{27}$$

To put (26) in the context of the previous sections we define

$$I = \mathbb{Z}^d, \quad H = \{(a_k)_{k\in I} \mid \sum_{k\in I} |a_k|^2 < \infty\}$$

and consider the subspace defined by condition (27). This subspace is invariant for all Galerkin projections onto X_n. Other constraints like oddness or evenness of $a(t, x)$ may cause the change of I, moreover also the basis in our Hilbert space may change accordingly, for example for the KS equation (9) with odd and periodic boundary conditions (10) we have $I = \mathbb{Z}_+$ and $u(t, x) = \sum_{k\in I} -2a_k(t)\sin(kx)$, where $a_k \in \mathbb{R}$ and equation (26) becomes [CCP, ZM]

$$\frac{da_k}{dt} = k^2(1 - \nu k^2)a_k - k\sum_{n=1}^{k-1} a_n a_{k-n} + 2k\sum_{n=1}^{\infty} a_n a_{n+k}, \quad k = 1, 2, 3\dots \tag{28}$$

Observe that conditions (3),(4) and (5) are satisfied for the KS equation. Namely we have $v(|k|) = \nu - \frac{1}{|k|^2}$, $p = 4$, $r = 1$. For the Navier-Stokes equations with periodic boundary conditions $p = 2$, $r = 1$ and $v(k) = \nu$, where ν is the viscosity.

3.1 Estimates

In this subsection our goal is to prove the following

Lemma 5 *Let $s > s_0 = d + r + 1$. If $|a_k| \leq C/|k^s|$, $|a_0| \leq C$, then there exists $D = D(C, s)$*

$$|N_k| \leq \frac{D}{|k|^{s-r}}, \qquad |N_0| \leq D \tag{29}$$

Before we proceed with the proof we need several lemmas. To make expression some formulas less cumbersome in this subsection for $0 = \{0\}^d \in \mathbb{Z}^d$ we redefine its norm by setting $|0| = 1$.

The following lemma was proved in [Sa]

Lemma 6 *Assume that $\gamma > d$. Then there exists $C(d, \gamma) \in \mathbb{R}$ such that for any $k \in \mathbb{Z}^d \setminus \{0\}$ holds*

$$\sum_{k_1 \in \mathbb{Z}^d \setminus \{0,k\}} \frac{1}{|k_1|^\gamma |k - k_1|^\gamma} \leq \frac{C_Q(d, \gamma)}{|k|^\gamma}. \tag{30}$$

From the above lemma one easily obtains

Lemma 7 *Assume that $\gamma > d$. Then there exists $C_2(d, \gamma) \in \mathbb{R}$ such that for any $k \in \mathbb{Z}^d$ holds*

$$\sum_{k_1, k_2 \in \mathbb{Z}^d, k_1 + k_2 = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma} \leq \frac{C_2(d, \gamma)}{|k|^\gamma}. \tag{31}$$

Proof: Consider two cases $k = 0$ and $k \neq 0$.

If $k = 0$, then there exists $\widetilde{C}(d, \gamma) \in \mathbb{R}$ such that

$$\sum_{k_1, k_2 \in \mathbb{Z}^d, k_1 + k_2 = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma} = 1 + \sum_{k_1 \in \mathbb{Z}^d \setminus \{0\}} \frac{1}{|k_1|^{2\gamma}} = \widetilde{C}(d, \gamma).$$

If $k \neq 0$, then from Lemma 6 it follows that

$$\sum_{k_1, k_2 \in \mathbb{Z}^d, k_1 + k_2 = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma} = \frac{2}{|k|^\gamma} + \sum_{k_1, k_2 \in \mathbb{Z}^d \setminus \{0\}, k_1 + k_2 = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma} \leq \frac{C_Q(d, \gamma) + 2}{|k|^\gamma}.$$

Hence the assertion holds for $C_2(d, \gamma) = \max(\widetilde{C}(d, \gamma), C_Q(d, \gamma) + 2)$. □

Lemma 8 *Assume $\gamma > d$. For any $n \in \mathbb{Z}_+$, $n > 1$ there exists $C_n(d, \gamma) \in \mathbb{R}$ such that for any $k \in \mathbb{Z}^d$ holds*

$$\sum_{k_1, k_2, \ldots, k_n \in \mathbb{Z}^d, \sum_{i=1}^n k_i = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma \cdot \ldots \cdot |k_n|^\gamma} \leq \frac{C_n(d, \gamma)}{|k|^\gamma}. \tag{32}$$

Proof: By induction. Case $n = 2$ is contained in Lemma 7. Assume now that the assertion holds for n. We have

$$\sum_{k_1,k_2,\ldots,k_{n+1}\in\mathbb{Z}^d, \sum_{i=1}^{n+1} k_i = k} \frac{1}{|k_1|^\gamma |k_2|^\gamma \cdot \ldots \cdot |k_{n+1}|^\gamma} =$$

$$\sum_{k_{n+1}\in\mathbb{Z}^d} \left(\frac{1}{|k_{n+1}|^\gamma} \sum_{k_1,k_2,\ldots,k_n\in\mathbb{Z}^d, \sum_{i=1}^{n} k_i = k-k_{n+1}} \frac{1}{|k_1|^\gamma |k_2|^\gamma \cdot \ldots \cdot |k_n|^\gamma} \right) \leq$$

$$\sum_{k_{n+1}\in\mathbb{Z}^d} \frac{1}{|k_{n+1}|^\gamma} \cdot \frac{C_n(d,\gamma)}{|k-k_{n+1}|^\gamma} \leq \frac{C_2(d,\gamma)C_n(d,\gamma)}{|k|^\gamma}.$$

□

Proof of Lemma 5: For the proof it is enough to assume that N is a monomial. After plugging-in the Fourier expansion for $u, Du, \ldots, D^r u$ (observe that because $s > d + r + 2$, then partial derivatives of u are obtained as a sum of partial derivatives of its Fourier expansion) we obtain the expression of the following type

$$N_k(u) = \sum_{k_1+\cdots+k_l=k} v_{k_1} \cdot v_{k_2} \cdot \ldots \cdot v_{k_l}, \tag{33}$$

where each of the variables v_{k_i}, $i = 1, \ldots, l$ is some Fourier coefficient of one the components of u or its partial derivatives of the order less than or equal to r.

Observe that for the Fourier coefficients of partial derivatives up to order r we have the following estimates

$$\left| \frac{\partial^{\beta_1+\cdots+\beta_l} u}{\partial x_1^{\beta_1} \ldots \partial x_d^{\beta_l}} \right| \leq \frac{C}{|k|^{s-(\beta_1+\cdots+\beta_l)}} \leq \frac{C}{|k|^{s-r}}. \tag{34}$$

From conditions (33) and (34), and Lemma 8 we obtain

$$|N_k(u)| \leq \sum_{k_1+\cdots+k_n=k} \frac{C^n}{|k_1|^{s-r} \cdot \ldots \cdot |k_n|^{s-r}} \leq \frac{C^n C_n(d, s-r)}{|k|^{s-r}} \tag{35}$$

□

3.2 Existence theorems

The main result in this section is Theorem 10, which states that equation (26) satisfying conditions (3), (4), (5) has solutions within self-consistent bounds for a sufficiently short time.

Theorem 9 *Consider (26). Assume that conditions (3), (4) and (5) hold. Let be $s_0 = p + d + 1$ and $m \in \mathbb{R}$.*

Consider compact set $W \subset X_m$ and a sequence of compact sets $B_k \subset H_k$ for $|k| > m$, such that there exist $s \geq s_0$ and $C \in \mathbb{R}$ such that the following condition hold

$$|B_k| \leq \frac{C}{|k|^s}, \qquad |k| > m,\ k \in I. \tag{36}$$

Then $W \oplus \Pi_{k\in I,|k|>m}$ satisfies conditions C2,C3.

Proof: Condition C2 is obvious.

It remains to prove C3. Let $T = \Pi_{k\in I,|k|>m} B_k$.

The first question is, whether $W \oplus T \subset \operatorname{dom} F$. Consider $u \in W \oplus T$. From Lemma 5 it follows that $F_k(u)$ is defined and for $|k| > m$ holds

$$|F_k(u)| \leq v_1 C|k|^{p-s} + D|k|^{r-s} \leq \frac{D_2}{|k|^{s-p}} \tag{37}$$

for some constants D and D_2. Hence

$$f_k = \max_{u\in W\oplus T} |F_k(u)| \leq \frac{D_2}{|k|^{s-p}}, \qquad |k| > m, k \in I \tag{38}$$

For $s \geq s_0$ we have $\sum_{|k|>m,k\in I} f_k^2 < \infty$. From this it follows that $W \oplus T \subset \operatorname{dom}(F)$.

It remains to prove the continuity of $F : W \oplus T \to H$. From condition (38) it follows that

$$\lim_{n\to\infty} \sum_{|k|>n} |Q_n F(x)|^2 = 0 \tag{39}$$

uniformly on $W \oplus T$. Hence it is enough to prove that $F_k : W \oplus T \to H_k$ is continuous.

Let us fix $k \in I$ and assume $u^n \to u^*$, $n \in \mathbb{Z}_+$, $u^n, u^* \in W \oplus T$. We have (compare the proof of Lemma 5)

$$F_k(u) = \lambda_k u_k + N_k(u) = \lambda_k u_k + \sum_{i\in I'} N_{k,i}(u), \tag{40}$$

where for each $i \in I'$ $N_{k,i}$ is polynomial depending on the finite number of u_l.

The term $\lambda_k u_k$ is continuous, hence it is enough to consider N_k, only. Let us fix $\epsilon > 0$. From Lemma 5 it follows that there exists a finite set $S \subset I'$, such that

$$\sum_{i\in I'\setminus S} |N_{k,i}(u)| < \epsilon/3, \qquad \text{for all } u \in W \oplus T. \tag{41}$$

There exists L, such that for all $i \in S$ the polynomials $N_{k,i}(u)$ depend in fact on the variables u_l for $|l| \leq L$, hence $\sum_{i\in S} N_{k,i}(u)$ is continuous on $W \oplus T$. Hence there exists n_0, such that

$$\left| \sum_{i\in S} N_{k,i}(u^n) - \sum_{i\in S} N_{k,i}(u^*) \right| < \epsilon/3. \tag{42}$$

From (42) and (41) we obtain for $n > n_0$

$$|N_k(u^n) - N_k(u^*)| \leq \left| \sum_{i\in S} N_{k,i}(u^n) - \sum_{i\in S} N_{k,i}(u^*) \right| + \sum_{i\in I'\setminus S} |N_{k,i}(u^n)| + \sum_{i\in I'\setminus S} |N_{k,i}(u^*)| < \epsilon.$$

Hence $\lim_{n\to\infty} N_k(u^n) = N_k(u^*)$. $\square$

Our goal is to prove the following theorem.

Theorem 10 *Consider (26). Assume that conditions (3), (4) and (5) hold. Let* $s_0 = p + d + 1$

Let $Z \oplus T_0$ *form self-consistent bounds for (26), such that for some* C_0 *and* $s \geq s_0$ *holds*

$$|T_{0,k}| \leq \frac{C_0}{|k|^s}, \qquad |k| > m,\ k \in I,\ s > s_0 \tag{43}$$

Then there exist $h > 0$, $W \oplus T_1$ *- self-consistent bounds for (26) and* $L > 0$, *such that for all* $l > L$ *and* $u \in P_l(Z \oplus T_0)$

$$\varphi^l([0,h],u) \subset W \oplus T_1. \tag{44}$$

and

$$|T_{1,k}| \leq \frac{C_1}{|k|^s}, \qquad |k| > m, k \in I. \tag{45}$$

Proof: Let $W \subset X_m$, be a compact set, such that $Z \subset \operatorname{int}_{X_m} W$.

By eventually increasing C_0 we can assume that

$$|u_k| \leq \frac{C_0}{|k|^s}, \qquad u \in W \oplus T_0. \tag{46}$$

We set $C_1 = 2C_0$ and define the tail T_1 by

$$T_1 = \Pi_{|k|>m, k\in I} \overline{B}\left(0, \frac{C_1}{|k|^s}\right). \tag{47}$$

From Lemma 5 it follows that there exists $D = D(C_1, s)$, such that

$$|N_k(u)| < \frac{D}{|k|^{s-r}}, \quad \text{for all } u, \text{ such that } |u_k| \leq \tfrac{C_1}{|k|^s} \tag{48}$$

Let $u \in W \oplus T_1$ and $|u_{k_0}| = \frac{C_1}{|k_0|^s}$ for some $|k_0| > K_-$.

$$\frac{1}{2}\frac{d}{dt}(u_{k_0}|u_{k_0}) < -v_0|k_0|^p|u_{k_0}|^2 + |u_{k_0}||N_{k_0}(u)| \leq \tag{49}$$

$$\left(-v_0 C_1 |k_0|^{p-s} + D|k_0|^{r-s}\right)|u_{k_0}| \tag{50}$$

$$\frac{d|u_{k_0}|^2}{dt} < 0, \qquad |k_0| > L, \tag{51}$$

for L sufficiently large.

Consider now the differential inclusion

$$x' \in P_L F(x) + \Delta, \qquad x \in X_L, \Delta \subset X_L \tag{52}$$

where Δ represents the Galerkin projection errors on $W \oplus T_1$ and is given by

$$\Delta = \{P_L F(u) - P_L F(P_L u) \mid u \in W \oplus T_1\}. \tag{53}$$

It is easy to see that there exists $h > 0$, such that if $x : [0,h] \to X_L$ is a solution of (52) and $x(0) \in P_L(Z \oplus T(0)))$, then

$$x(t) \in \operatorname{int}_{X_L} P_L(W \oplus T_1), \qquad t \in [0,h]. \tag{54}$$

Let $l > L$ and let $u : [0,t_1) \to X_l$ be a solution of

$$u' = P_l F(u), \qquad u(0) = u_0 \in P_l(X \oplus T_0). \tag{55}$$

By changing the vector field in the complement of $P_l(W \oplus T_1)$ we can assume that $t_1 = \infty$.

Let

$$t_m = \sup\{t > 0 \mid u([0,t]) \subset P_l(W \oplus T_1)\}. \tag{56}$$

Obviously $t_m > 0$. It is enough to prove that $t_m \geq h$. Observe that for $t \in [0,t_m]$ $P_L u(t)$ is a solution of (52), hence from (54) we obtain

$$P_L u([0,t_m]) \subset \operatorname{int}_{X_L} P_L(W \oplus T_1). \tag{57}$$

From (51) it follows immediately that

$$Q_L u([0,t_m]) \subset \operatorname{int}_{Y_l} P_l Q_L(W \oplus T_1). \tag{58}$$

Hence

$$u(t_m) \in \operatorname{int}_{X_l} P_l(W \oplus T_1). \tag{59}$$

From above condition and the continuity of u it follows that for some $\delta > 0$ holds

$$u(t_m + t') \in \operatorname{int}_{X_l} P_l(W \oplus T_1), \qquad t' \in [0,\delta] \tag{60}$$

hence $t_m = h$. $\square$

3.3 Classical solutions from self-consistent bounds

The goal in this section is to address the question, whether the solutions of (26) are classical solutions of (1).

To formulate the answer in an abstract setting we need some assumptions about the behavior of the derivatives for functions from H_k.

For $s \geq 0$ let $C^s_{per}(n) = C^s(\mathbb{T}^d, \mathbb{R}^n)$ denote the space of functions on the d-torus of class C^s. For $u \in C^0_{per}(n)$ we set $|u|_0 = \sup_{x \in \mathbb{T}^d} |u(x)|$, where on $\mathbb{R}^n$ we use any fixed norm.

Definition 5 *Let $H = \oplus_{k \in I} H_k$. We say that* the decomposition of H into H_k *is r-smooth, when the following conditions are satisfied:*

- *there exists a partial linear map $\iota : H \to C^r_{per}(n)$, such that $H_k \subset \operatorname{dom}(\iota)$ for each $k \in I$*

- *there exists constant R, such that for each $k \in I$ holds and for $l = 1, \dots, r$ holds*

$$\left| \frac{\partial^l \iota(u)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_l}} \right|_0 \leq |k|^l R |u|, \tag{61}$$

for any $(i_1, \dots, i_l) \in \{1, \dots, d\}^l$.

Observe that the Fourier expansion, which means that H_k is the space spanned by $\exp ikxe_j$, where e_j is a canonical basis in $\mathbb{R}^n$, is obviously r-smooth decomposition of $L_2([0, 2\pi], \mathbb{R}^n)$ for any r.

Theorem 11 *Consider (26). Assume that conditions (3), (4) and (5) hold. Let $s_0 = d + p + 1$.*

Assume that $H = \oplus_{k \in I} H_k$ is an s-smooth decomposition of H, for $s \geq s_0$.

Let $u : [t_1, t_2] \to W \oplus T \subset H$, where $W \oplus T$ are self-consistent bounds for (26), such that for some constants $m, C \in \mathbb{R}$, $s \geq s_0$

$$|T_k| \leq \frac{C_0}{|k|^s}, \qquad |k| > m, \ k \in I, \tag{62}$$

holds, then u is a classical solution of (1).

Proof: We define $a(t, x)$ by

$$a(t, x) = \sum_{k \in I} \iota(u_k(t))(x). \tag{63}$$

From our assumptions it follows that the above series is converging uniformly on $[t_1, t_2] \times \mathbb{T}^d$. Also for any partial derivative of order less than or equal to p holds

$$\frac{\partial^l a}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_l}} = \sum_{k \in I} \frac{\partial^l \iota(u_k)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_l}}. \tag{64}$$

Moreover the convergence is uniform on $[t_1, t_2] \times \mathbb{T}^d$. Since $s \geq s_0$, hence also the Fourier expansions for La and $N(a)$ (see Lemma 5) are converging uniformly on $[t_1, t_2] \times \mathbb{T}^d$. This finishes the proof. □

4 The algorithm for rigorous enclosure of solutions of perturbations of ODEs and differential inclusions

4.1 The interval arithmetic and notation used in the description of algorithms

The interval arithmetic [Mo, MZ] is a suitable tool to deal with the non-rigorous computer arithmetic, because it replaces a mathematical object, r, a real number or a collection of reals composing a vector, a matrix etc, by an interval or a collection of intervals, $\mathbf{r}$, such that $r \in \mathbf{r}$. Moreover, the arithmetic operations on the interval objects can be defined so that the result of the interval computation always contains the result of the corresponding real operation.

In the description of algorithms we will use the same conventions as in [ZLo] regarding the notation of single valued and multivalued (interval) objects. In the sequel, by arabic letters we denote single valued objects like vectors, real numbers, matrices. Quite often we will use square brackets, for example $[r]$, to denote sets. Usually this will be some set constructed by some algorithm. Sets will also be denoted by single letters, for example S, when it is clear from the context that it represents a set. In situations when we want to stress (for example in the detailed description of an algorithm) that we have a set in a formula involving both single-valued objects and sets we will rather use the square bracket, hence we prefer to write $[S]$ instead of S to represent the set. From this point of view $[S]$ and S are different symbols in the alphabet used to name variables and formally speaking there is no relation between the set represented by $[S]$ and the object represented by S. If in the description of an algorithm we will have a situation that both variables, $[S]$ and S, are used simultaneously, then usually $S \in [S]$, but this is always stated explicitly.

For a set $[S]$ by $[S]_I$ we denote the interval hull of $[S]$, i.e. the smallest product of intervals containing $[S]$. The symbol $\mathrm{hull}(x_1, \dots, x_k)$ will denote the interval hull of intervals $x_1, \dots, x_k$. The set $Y \subset \mathbb{R}^m$ will be called an interval set if $Y = \Pi_{i=1}^m Y_i$, where Y_i are closed intervals (we will allow also for degenerate intervals $I = [a, a]$).

For any interval $I = [a, b]$ we define a diameter of I $\mathrm{diam}\,(I)$ and the functions $left(I)$, $right(I)$, I^+ and I^- by

$$\begin{aligned} \mathrm{diam}\,(I) &= b - a \\ I^- &= left(I) = a \\ I^+ &= right(I) = b. \end{aligned}$$

For $c > 0$ and $X = [a - \delta/2, a + \delta/2]$, where $\delta \geq 0$ we define

$$\mathrm{inflate}(X, c) = [a - c\delta/2, a + c\delta/2].$$

For any interval set (vector, matrix) $[S]$ by $\mathrm{m}([S])$ we will denote a center point of $[S]$ and by $\mathrm{diam}\,([S])$ we will denote the maximum of diameters of its components.

In the description of algorithm we will use the expression $a \in bool$ to indicate that a is a boolean variable with the possible values *false* and *true*. Sometimes integer constants 0 and 1 might used for *false* and *true*, respectively.

4.2 An outline of the algorithm

For the purpose of the rigorous integration of dissipative PDEs we will study the following nonautonomous ODE,

$$x'(t) = f(x(t), y(t)) \tag{65}$$

where $x \in \mathbb{R}^m$ and $y : \mathbb{R} \supset D \to \mathbb{R}^n$ is bounded and continuous, and f is C^1. Assume that we have some knowledge about $y(t)$, for example $|y(t)| < \epsilon$ for $0 \leq t \leq t_1$. We would like to find a rigorous enclosure for $x(t)$.

What we describe below is basically the algorithm for the rigorous enclosure of the solutions of the differential inclusion

$$\frac{dx}{dt}(t) \in f(x) + [\delta], \tag{66}$$

where $[\delta] \subset \mathbb{R}^m$. In the context of the rigorous integration of dissipative PDEs the function $y(t)$ in (65) represents the tail and $[\delta]$ in (66) is the Galerkin projection error.

For a bounded and continuous function $y : [0, \infty) \to \mathbb{R}^n$ let $\varphi(t, x_0, y)$ denotes a solution of equation (65) with the initial condition $x(0) = x_0$. For a given $y_0 \in \mathbb{R}^n$ let $\overline{\varphi}(t, x_0, y_0)$ be a solution of the following Cauchy problem

$$x' = f(x, y_0), \quad x(0) = x_0 \tag{67}$$

with the same initial condition $x(0) = x_0$. Observe that system (67) is a particular case of (65) with $y(t) = y_0$.

We are interested in finding rigorous bounds for $\varphi(t, [x_0], [y_0])$, where $[x_0] \subset \mathbb{R}^m$ and $[y_0] \subset C_b([0, \infty), \mathbb{R}^n)$. The set $[y_0]$ might be defined be some dynamical process, in this case we may need to compute the range of $[y_0]$ during each time step or be given explicitly, for example: $y \in [y_0]$ iff y is bounded and continuous and $y(t) \in [-\epsilon, \epsilon]^n$.

To achieve the above mentioned goal we propose a modification of the original Lohner algorithm [Lo, Lo1]. Our presentation and notation follows a description of the C^0-Lohner algorithm presented in [ZLo] and almost coincide with the content of Section 6 from [Z2]. The main difference compared to [Z2] is in how the first and the fifth parts are realized in the context of dissipative PDEs. This is described in the subsequent sections.

4.3 A fundamental estimate

The following lemma is a particular case of Theorem 1 in Section 13 in [W](see subsection IV 'The Lipschitz condition'), a self-contained proof (with precisely specified assumptions) can also be found in [ZPLo].

Lemma 12 *Assume $t_0, h \in \mathbb{R}$ and $h > 0$. Let $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_1}$ be a $\mathcal{C}^1$-function. For a fixed $y_c \in \mathbb{R}^{n_2}$ and a bounded and continuous function $y : [t_0, t_0 + h] \to \mathbb{R}^{n_2}$ consider*

$$x' = f(x, y_c), \quad x(t_0) = x_0 \tag{68}$$

$$x' = f(x, y_c) + (f(x, y(t)) - f(x, y_c)), \quad x(t_0) = x_0. \tag{69}$$

Let $x_1, x_2 : [t_0, t_0 + h] \to \mathbb{R}^{n_1}$ be solutions of (68) and (69), respectively. We assume that

- *$W_y \subset \mathbb{R}^{n_2}$ is a convex set and $y([t_0, t_0+h]) \subset W_y$.*
- *Let $W_1 \subset W_2 \subset \mathbb{R}^{n_1}$ be convex and compact, such that for $s \in [t_0, t_0+h]$ holds*

$$x_1(s) \in W_1 \tag{70}$$
$$x_2(s) \in W_2 \tag{71}$$

Then the following inequality holds for $t \in [t_0, t_0+h]$ and for $i = 1, \ldots, n_1$

$$|x_{1,i}(t) - x_{2,i}(t)| \leq \left(\int_{t_0}^{t} e^{J(t-s)} C \, ds \right)_i , \tag{72}$$

provided $C \in \mathbb{R}^{n_1}$ and $J \in \mathbb{R}^{n_1 \times n_1}$ satisfy the following conditions

$$\begin{aligned} C_i &\geq \sup\{|f_i(x, y_c) - f_i(x, y)|, \quad x \in W_1, y \in W_y\}, \quad i = 1, \ldots, n_1 \\ J_{ij} &\geq \begin{cases} \sup \frac{\partial f_i}{\partial x_j}(W_2, W_y) & \text{if } i = j, \\ \sup \left| \frac{\partial f_i}{\partial x_j}(W_2, W_y) \right| & \text{if } i \neq j. \end{cases} \end{aligned}$$

Comment: It is very important for the application to dissipative PDEs, that in the above lemma the terms on the diagonal in matrix J can be negative. As a result of this fact the increasing of the dimension of the Galerkin projections does not result in a significant increase of $\|e^{Jt}\|$ for $t > 0$. This fact allows also to obtain the equicontinuity of all Galerkin projections, which can be later used to obtain an ODE-type uniqueness proof for dissipative PDEs (see [ZGal]).

4.4 One step of the algorithm

The basic outline of the algorithm in nearly the same as in [Z2]. The difference is that in the case of dissipative PDEs we have an efficient procedure for the computation of the evolution of the tail.

In the description below the objects with an index k refer to the current values and those with an index $k+1$ are the values after the next time step.

We define

$$[y_k] = \{y \in C_b([0, \infty], \mathbb{R}^n) \mid y(t) = z(t_k + t) \quad \text{for some } z \in [y_0]\}.$$

We will also use the following notation for $[y] \subset C_b([0, \infty), \mathbb{R}^s)$

$$[y]([t_1, t_2]) = \{z(t) \mid z \in [y], t \in [t_1, t_2]\}.$$

One step of the Lohner algorithm is a shift along the trajectory of system (65) with the following input and output data:

Input data:

- t_k is a current time

- h_k is a time step
- $[x_k] \subset \mathbb{R}^m$, such that $\varphi(t_k, [x_0], [y_0]) \subset [x_k]$
- bounds for $[y_k]$

Output data:

- $t_{k+1} = t_k + h_k$ is a new current time
- $[x_{k+1}] \subset \mathbb{R}^m$, such that $\varphi(t_{k+1}, [x_0], [y_0]) \subset [x_{k+1}]$
- bounds for $[y_{k+1}]$.

We do not specify here the representation of sets $[x_k]$. This issue is very important in the handling of the wrapping effect and is discussed in detail in [Lo, Lo1, Mo, MZ] (see also Section 3 in [ZLo]).

One step of the algorithm consists from the following parts:

1. Generation of a-priori bounds for φ and $[y_0]([t_k, t_{k+1}])$.

We find convex and compact set $[W_2] \subset \mathbb{R}^m$ and convex set $[W_y] \subset \mathbb{R}^n$, such that

$$\varphi([0, h_k], [x_k], [y_k]) \subset [W_2], \tag{73}$$

$$[y_k]([0, h_k]) \subset [W_y]. \tag{74}$$

2. We fix $y_c \in [W_y]$.

3. Computation of an unperturbed x-projection. We apply one step of the C^0-Lohner algorithm to (67) with the time step h_k and the initial condition given by $[x_k]$ and $y_0 = y_c$. As a result we obtain $[\overline{x}_{k+1}] \subset \mathbb{R}^m$ and convex and compact set $[W_1] \subset \mathbb{R}^m$, such that

$$\begin{aligned} \overline{\varphi}(h_k, [x_k], y_c) &\subset [\overline{x}_{k+1}], \\ \overline{\varphi}([0, h_k], [x_k], y_c) &\subset [W_1]. \end{aligned}$$

4. Computation of the influence of the perturbation. Using formulas from Lemma 12 we find set $[\Delta] \subset \mathbb{R}^m$, such that

$$\varphi(t_{k+1}, [x_0], [y_0]) \subset \overline{\varphi}(h_k, [x_k], y_c) + [\Delta]. \tag{75}$$

Hence

$$\varphi(t_{k+1}, [x_0], [y_0]) \subset [x_{k+1}] = [\overline{x}_{k+1}] + [\Delta] \tag{76}$$

5. Computation of $[y_{k+1}]$

4.5 Part 1 - comments

In the context of an ordinary differential inclusion (66) we can set $[W_y] = [\delta]$. The question of finding $[W_2]$ is treated in Section 5.

In the context of a dissipative PDE we cannot find $[W_y]$ and $[W_2]$ using independent routines, some consistency conditions are necessary. This question is treated in detail in Sections 6, 7 and 8.

4.6 Part 4 - details

We use Lemma 12.

1. We set

$$\begin{aligned} [\delta] &= [\{f(x, y_c) - f(x, y) \mid x \in [W_1], y \in [W_y]\}]_I \\ C_i &= \text{right}(|[\delta_i]|), \quad i = 1, \ldots, n \\ J_{ij} &= \begin{cases} \text{right}\left(\frac{\partial f_i}{\partial x_i}([W_2], [W_y])\right) & \text{if } i = j, \\ \text{right}\left(\left|\frac{\partial f_i}{\partial x_j}([W_2], [W_y])\right|\right). & \text{if } i \neq j. \end{cases} \end{aligned}$$

2. $D = \int_0^h e^{J(h-s)} C \, ds$
3. $[\Delta_i] = [-D_i, D_i]$, for $i = 1, \ldots, n$

For the computation of $\int_0^t e^{A(t-s)} C \, ds$, see Section 6.5 in [Z2].

After we compute Δ to avoid the wrapping effect we perform a rearrangement, see Section 6.6 in [Z2].

4.7 Part 5 - comments

For ordinary differential inclusions (66) we don't have to do anything. In the context of PDEs this is a very important issue and it is treated in Section 7.4.

5 Generation of a-priori bounds for ordinary differential inclusions

The goal of this section is to present an algorithm for the generation of a-priori bounds for ordinary differential inclusions. We will frequently refer to such a-priori bounds as *the rough enclosure*. The main result is Theorem 17 and the algorithm based on it is presented in Section 5.3. These developments realize for differential inclusions Part 1 of the algorithm outlined in Section 4.

5.1 A naive rough enclosure function

We start with the following easy theorem.

Theorem 13 *Consider a differential equation*

$$x' = f(x), \quad x \in \mathbb{R}^n \tag{77}$$

where $f \in C^1$. Let φ be a local flow induced by (77), $h \in \mathbb{R}$ and X, Z be interval sets, $X \subset \operatorname{int} Z$. Suppose that

$$Y := \textit{interval hull}(X + [0,h]f(Z)) \subset \operatorname{int} Z \tag{78}$$

then

$$\varphi([0,h],X) \subset Y \tag{79}$$

An easy proof is left to the reader. Above theorem can be also derived from Theorem 17 with $D = \emptyset$.

Let Y be as in the above theorem, we will refer to it as the *first order enclosure*, because it is based on the first order Taylor formula. Analogous theorems using higher order Taylor formulas are possible, but our experience show that they are not much better.

Remark 14 *From Theorem 13 it follows immediately that, if we take h sufficiently small, then there exists the first order enclosure. In fact any interval set Z, such that $X \subset \operatorname{int} Z$ is good for sufficiently small h.*

Observe that condition (78) imposes severe restrictions on the size of h even in the situation, when it is obvious that the enclosure should exists for any $h > 0$. As an example we consider a single linear equation

$$x' = f(x) = -Lx, \qquad L > 0, \quad x \in \mathbb{R}. \tag{80}$$

Assume that (78) holds for some intervals X, Y, Z and $f(x) = -Lx$. By taking diameters of both sides of (78) we obtain

$$hL \cdot \operatorname{diam}(Z) \quad < \quad \operatorname{diam}(Z), \tag{81}$$

$$hL \quad < \quad 1. \tag{82}$$

On the other side is easy to see that the interval $Y = [-\max|X|, \max|X|]$ is the enclosure for any $h > 0$.

An natural generalization of (82) to multidimensional nonlinear system is

$$h|df| < 1, \tag{83}$$

where $|df|$ is the maximum of the norm of $df(x)$ for x over the region of interest.

In the context of the Galerkin projection of dissipative PDE from condition (83) it follows that in order to obtain the first rough enclosure for n-th Galerkin projection the time step must satisfy

$$h|\lambda_k| < 1, \qquad \text{for } |k| \leq n. \tag{84}$$

This usually leads to unreasonably small time steps, which is dictated not by the dynamics of the system under the consideration, but by the inclusion in the integration of highly damped variables of little relevance for the dynamics. In the next section we will present an enclosure theorem and an algorithm based on it, which allows to use considerably larger time steps.

5.2 The rough enclosure algorithm based on isolation

Our goal is to devise a rough enclosure routine, which will take into account the strong damping for some variables and will overcome the restriction on h given by (83).

Before we proceed further we need a few easy lemmas.

Lemma 15 *Let N be a constant. Let $x(t)$ be a C^1 function. If*

$$\frac{dx}{dt} < \lambda x + N, \tag{85}$$

then for $t > 0$ holds

$$x(t) < \left(x(0) - \frac{N}{-\lambda}\right) e^{\lambda t} + \frac{N}{-\lambda} \tag{86}$$

Similarly, if

$$\frac{dx}{dt} > \lambda x + N, \tag{87}$$

then for $t > 0$ holds

$$x(t) > \left(x(0) - \frac{N}{-\lambda}\right) e^{\lambda t} + \frac{N}{-\lambda} \tag{88}$$

Lemma 16 *Let N be a constant. Let $x(t)$ be a C^1-function. If*

$$\frac{dx}{dt} < \lambda x + N, \tag{89}$$

then for $t > 0$ holds

$$x(t) < \frac{N}{-\lambda}, \quad \text{if } x(0) < \tfrac{N}{-\lambda} \tag{90}$$

$$x(t) < x(0), \quad \text{if } x(0) \geq \tfrac{N}{-\lambda} \tag{91}$$

Similarly, if

$$\frac{dx}{dt} > \lambda x + N, \tag{92}$$

then for $t > 0$ holds

$$x(t) > \frac{N}{-\lambda}, \quad \text{if } x(0) > \tfrac{N}{-\lambda} \tag{93}$$

$$x(t) > x(0), \quad \text{if } x(0) \leq \tfrac{N}{-\lambda} \tag{94}$$

We assume that our problem can be written as

$$\frac{dx_i}{dt} \in f_i(x) = \lambda_i x_i + N_i(x), \quad i = 1, \ldots, n \tag{95}$$

where $N_i : \mathbb{R}^n \to cf(R)$ is a multivalued continuous function and by φ we will denote *the (local) flow* induced by (95) (see Section 1.1 for the definition).

Now we state a theorem which is a basis of our improved enclosure function.

Theorem 17 *Consider (95). Let $h > 0$ and $X \subset Z \subset \mathbb{R}^n$ be interval sets. Let $D \subset \{1, \ldots, n\}$ (the set of dissipative(damped) directions), such that if $k \in D$, then*

$$\lambda_k < 0 \tag{96}$$

$$\lambda_k a_k + N_k^- < \frac{da_k}{dt} < \lambda_k a_k + N_k^+ \tag{97}$$

where $N_k(Z) \subset (N_k^-, N_k^+)$.

For $k \in D$ we set

$$b_k^\pm = \frac{N_k^\pm}{-\lambda_k} \tag{98}$$

$$g_k^\pm = \left(X_k^\pm - b_k^\pm\right) e^{\lambda_k h} + b_k^\pm. \tag{99}$$

Let $Y = \Pi_{i=1}^n Y_i$, be such that

$$Y_i = X_i + [0,h] f_i(Z), \qquad i \notin D \tag{100}$$

$$Y_i = Z_i, \qquad i \in D. \tag{101}$$

Then

$$\varphi([0,h], X) \subset Y, \tag{102}$$

provided the following conditions are satisfied for $i = 1, \ldots, n$

1. *if $i \notin D$, then*

$$Y_i \subset \operatorname{int} Z_i \tag{103}$$

2. upper bounds: *for $i \in D$*

$$\textit{if} \quad Z_i^+ < b_i^+, \quad \textit{then} \quad Z_i^+ \geq g_k^+ \tag{104}$$

3. lower bounds: *for $i \in D$*

$$\textit{if} \quad Z_i^- > b_i^-, \quad \textit{then} \quad Z_i^- \leq g_k^- \tag{105}$$

Proof: After a modification of the right-hand side of (95) outside a sufficiently large ball we can assume that all solutions of (95) are defined for $t \in \mathbb{R}_+$.

By a small stretching of Z in dissipative directions we can construct a new interval set $\widetilde{Z}$, such that

$$\widetilde{Z}_i = Z_i, \quad i \notin D \tag{106}$$

$$Z_i \subset \operatorname{int} \widetilde{Z}_i, \qquad i \in D \tag{107}$$

$$N_k(\widetilde{Z}) \subset (N_k^-, N_k^+), \qquad k \in D \tag{108}$$

$$X_i + [0,h] f_i(\widetilde{Z}) \subset \operatorname{int} Z_i = \operatorname{int} \widetilde{Z}_i, \quad i \notin D. \tag{109}$$

Obviously

$$Y \subset \operatorname{int} \widetilde{Z}. \tag{110}$$

Let us fix $x_0 \in X$ and $x(t)$ be a solution of (95) through x_0 and let

$$T = \sup\{t \in [0,h] \quad | \quad \forall s \in [0,t] \;\; x(s) \in Y\}. \tag{111}$$

To finish the proof it is enough to show that $T = h$.

If $T < h$, then there exists $\delta > 0$, such that $T + \delta < h$ and $x(T+t) \in \widetilde{Z}$. Hence from Lemma 15 it follows that

$$x_i(T+t) \in \operatorname{int} Y_i \subset Z_i, \quad \text{for } i \in D \text{ and } t \in (0, \delta]. \tag{112}$$

Hence

$$x(s) \in Z, \qquad \text{for } s \in [0, T+\delta]. \tag{113}$$

By applying the Mean Value Theorem to x_i for $i \notin D$ for $t \in [0,\delta]$ we obtain

$$x_i(T+t) \in x_{0,i} + (T+t) \cdot f_i(Z) \subset x_{0,i} + [0,h] f_i(Z) \subset Y_i. \tag{114}$$

From this and (112) it follows that

$$x([0, T+\delta]) \subset Y.$$

This is in a contradiction with the definition of T, hence $T = h$. □

5.3 The algorithm for rough enclosure for differential inclusions

The initial guess: We define

$$Z_i = X_i + [0,h] f(X), \quad \text{if } \lambda_i \geq 0 \tag{115}$$

$$Z_i = X_i, \quad \text{if } \lambda_i < 0 \tag{116}$$

$$N_i = N_i(Z) \tag{117}$$

$$b_i = \frac{N_i}{-\lambda_i}. \tag{118}$$

We define the set D, the set of dissipative coordinates, as follows: $i \in D$ iff $\lambda_i < 0$ and at least one of the following two conditions holds

$$h|\lambda_i| \geq 1/2 \tag{119}$$

$$b_i \cap X_i \neq \emptyset. \tag{120}$$

Remark 18 *Condition (119) makes sure that the coordinates for which* $h|\lambda_i| < 1$ *(compare (82)and (83) in Section 5.1) are always 'dissipative'.*

Remark 19 *The intention behind condition (120) is to treat as non-damped the direction if we are outside the range where the linear term dominates. On the other hand this condition is somewhat arbitrary and it sometimes happens that for some variable* i *the following holds: we have two sets* $X \subset X'$ *and* $i \in D$ *for* X'*, but* $i \notin D$ *for* X*. Due to this fact the enclosure for* X' *might be smaller than the one obtained for* X*.*

We choose two real constants $c > 1$ and $0 < c_d < 1$ (we use $c = 1.1$, $c_d = 0.1$) and we redefine Z_i by setting:

$$Z_i = \mathit{inflate}(Z_i, c), \qquad i \notin D \tag{121}$$

For $i \in D$ we set

$$\begin{aligned} g_i^{\pm} &= (X_i^{\pm} - b_i^{\pm})e^{\lambda_i h} + b_i^{\pm} \\ Z_i^{+} &= \begin{cases} X_i^{+}, & \text{if } X_i^{+} \geq b_i^{+}, \\ (1-c_d)g_i^{+} + c_d b_i^{+}, & \text{if } X_i^{+} < b_i^{+}, \end{cases} \\ Z_i^{-} &= \begin{cases} X_i^{-}, & \text{if } X_i^{-} \leq b_i^{-}, \\ (1-c_d)g_i^{-} + c_d b_i^{-}, & \text{if } X_i^{-} > b_i^{-}. \end{cases} \end{aligned}$$

Validation and producing a new guess. For each i we initialize the array *validated*, by $\mathit{validated}[i] = \mathit{true}$.

For each $i \notin D$, we set

$$Y_i = X_i + [0, h] f_i(Z). \tag{122}$$

If not $Y_i \subset \operatorname{int} Z_i$, then we set $\mathit{validated}[i] = \mathit{false}$ and produce a new guess

$$Z_i = \mathit{inflate}(Y_i \cup Z_i, c). \tag{123}$$

For each $i \in D$ we do the following:

- we compute N_i and b_i.
- if not $Z_i^{+} \geq b_i^{+}$, then we compute g_i^{+}.
  If $Z_i^{+} < g_i^{+}$, then we set $\mathit{validated}[i] = \mathit{false}$ and we produce a new guess by setting
  $$Z_i^{+} = (1-c_d)g_i^{+} + c_d b_i^{+}. \tag{124}$$

- With Z_i^- we proceed symmetrically, i.e.
  if not $Z_i^- \leq b_i^-$, then we compute g_i^-.
  If $Z_i^- > g_i^-$, then we set $validated[i] = false$ and we produce a new guess by setting
$$Z_i^- = (1 - c_d)g_i^- + c_d b_i^-. \tag{125}$$

Finally the enclosure is *validated* if $validated[i] = true$ for all i.

We iterate the above step until we achieve the validation or the number of steps is larger than some limit (equal to $\max(5, n/2)$ in my program, where n is the dimension of the phase space).

If we achieved the validation then we set

$$Y_i = X_i + [0,h]f_i(Z), \qquad i \notin D \tag{126}$$
$$Y_i = Z_i, \qquad i \in D \tag{127}$$

If we have obtained the validated enclosure, then we perform the *trimming* operation.

Trimming of the enclosure via isolation. The idea is to make the isolation argument for the directions with $\lambda_i < 0$ and $i \notin D$ and in the hope of some new gains, because for these coordinates this argument, while applicable, was not used in the above algorithm.

For each $i \notin D$ and $\lambda_i < 0$ we compute

$$\begin{aligned} N_i &= N_i(Y) \\ b_i &= \frac{N_i}{-\lambda_i}. \end{aligned}$$

If $X_i^+ \geq b_i^+$, then we set $Y_i^+ = X_i^+$.
If $X_i^- \leq b_i^-$, then we set $Y_i^- = X_i^-$.

Lemma 20 . *If the above algorithm applied to (95), with the time step $h > 0$ and the set of initial conditions X as input parameters, returns* true *and $Y \subset \mathbb{R}^n$, then*
$$\varphi([0,h], X) \subset Y.$$

Proof: For the proof of the correctness of the *validation part* use Theorem 17. The *trimming part* of the algorithm is justified by Lemma 16. □

Remark 21 *For very small h the first order enclosure (see Theorem 13), which treats all variables as non dissipative ($D = \emptyset$), can be better than the one produced by the above algorithm (this happens for the KS equation).*

6 Treatment of the tail for dissipative PDEs

In this section we discuss how to realize Parts 1 and 5 of the algorithm for the rigorous integration of dissipative PDEs. The algorithm itself is presented in the next section.

We consider the problem (26) derived from (1) and we adopt the notation used in Sections 2 and 3. Let us stress that we do not assume the local existence of solutions of (1), it is a byproduct of the algorithm.

During the computation we want the bounds for solutions to be given by self-consistent bounds. These bounds will be valid for sufficiently high dimensional Galerkin projections of (26), so we can use Lemma 3 to obtain the existence of solutions of (26).

Notations: For self-consistent bounds $W \oplus T$ we will denote by $m(T)$ and $M(T)$ the numbers m and M from Definition 3, respectively. In the sequel we will often use variables T, $T(h)$, $T([0,h])$ to indicate the tail. By $T(0)$ we will usually denote the initial tail (or a candidate for such set), by $T(h)$ the tail at time $t = h$ (or a candidate) and by $T([0,h])$ the tail for $t \in [0,h]$ (or a candidate). For tail T, by $\varphi(t, x_0, T)$, where $t \in \mathbb{R}$ and $x_0 \in X_{m(T)}$, we will denote the set all possible values of $x(t)$, where x is a solution of the following differential inclusion defined on the maximum interval of the existence.

$$x' \in P_{m(T)} F(x + T), \qquad x(0) = x_0. \tag{128}$$

If $x(t)$ does not exists for some t, then also $\varphi(t, x_0, T)$ is undefined. In the sequel we will use expression of the form

$$\varphi([0,h], x_0, T) \subset Z. \tag{129}$$

It means that $\varphi([0,h], x_0, T)$ defined, hence any solution of (128) is defined for $t \in [0,h]$, and the stated inclusion holds.

Standing assumptions: In this section we assume that $I = \mathbb{Z}_+$ and $H_k = \mathbb{R}$, hence the sets B_k in self-consistent bounds can be represented as $[a_k^-, a_k^+]$, where $a_k^- \leq a_k^+$, $a_k^\pm \in \mathbb{R}$. In this situation we can also assume that $m, M \in \mathbb{N}$. The generalization to a more general situation is straightforward, usually it is enough to take $B_k = \Pi_{i=1}^{d_1} [a_{k,i}^-, a_{k,i}^+]$ for $m < |k| \leq M$ and $B_k = \overline{B_{H^k}}(0, r_k)$ for $|k| > M$.

Moreover we assume that conditions (3), (4) and (5) are satisfied.

Lemma 22 *Assume that $W_2 \subset \mathbb{R}^m$ and T are self-consistent bounds for (26).*

Let $X_0 \subset W_2$ and $T(0) \subset T$ be self-consistent bounds for F, such that $m(T) = m(T(0))$ and $M(T(0)) = M(T)$. Assume that

$$\varphi([0,h], X_0, T) \subset W_2. \tag{130}$$

Let $N_k^\pm$ be such that

$$N_k^- < N_k(x+q) < N_k^+, \qquad \text{for all } k > m,\ x \in W_2 \text{ and } q \in T \tag{131}$$

Assume that for $k > m$ holds

$$\lambda_k \quad < \quad 0. \tag{132}$$

For $k > m$ we define $b_k^{\pm}$, $g_k^{\pm}$, $T(h)_k^{\pm}$ and $T([0,h])_k^{\pm}$ as follows

$$
\begin{aligned}
b_k^{\pm} &= \frac{N_k^{\pm}}{-\lambda_k} && (133)\\
g_k^{\pm} &= \left(T(0)_k^{\pm} - b_k^{\pm}\right) e^{\lambda_k h} + b_k^{\pm} && (134)\\
T(h)_k^{\pm} &= g_k^{\pm}, && (135)\\
T([0,h])_k^{+} &= \begin{cases} T(0)_k^{+}, & \text{if } T(0)_k^{+} \geq b_k^{+},\\ g_k^{+}, & \text{if } T(0)_k^{+} < b_k^{+}, \end{cases} && (136)\\
T([0,h])_k^{-} &= \begin{cases} T(0)_k^{-} & \text{if } T(0)_k^{-} \leq b_k^{-},\\ g_k^{-} & \text{if } T(0)_k^{-} > b_k^{-}. \end{cases} && (137)
\end{aligned}
$$

If

$$T([0,h]) \subset T, \tag{138}$$

then for any $n > M$ holds

$$
\begin{aligned}
\varphi^n([0,h], X_0 \oplus P_n T(0)) &\subset W_2 \oplus P_n T([0,h]), && (139)\\
\varphi^n(h, X_0 \oplus P_n T(0)) &\subset W_2 \oplus P_n T(h). && (140)
\end{aligned}
$$

Moreover, for any $u_0 \in X_0 \oplus T(0)$ there exists $u : [0,h] \to W_2 \oplus T([0,h])$ a solution of (26), such that $u(0) = u_0$ and $u(h) \in W_2 \oplus T(h)$.

Proof: It is enough to prove (139), because (140) follows then immediately from Lemma 15 and the last assertion is a consequence of Lemma 3 applied to self-consistent bounds $W_2 \oplus T$ and conditions (139) and (140).

To prove (139) let us fix $n > M$, $p \in X_0$ and $y \in P_n T(0)$. For sufficiently small $\epsilon > 0$ let $W(\epsilon) \subset \mathbb{R}^m$ and $V(\epsilon) \subset \mathbb{R}^{n-m}$ be such that

$$
\begin{aligned}
W_2 &\subset \operatorname{int} W(\epsilon),\\
W(\epsilon) &\subset B(W_2, \epsilon)\\
P_n T([0,h]) &\subset \operatorname{int} V(\epsilon)\\
V(\epsilon) &\subset B(P_n(T([0,h])), \epsilon)
\end{aligned}
$$

and for $x + q \in W(\epsilon) \oplus V(\epsilon)$ and $k = m, \ldots, n$ holds

$$N_k^- < N_k(x+q) < N_k^+,$$

where $N_k^{\pm}$ are the constants from condition (131).

We define

$$t_1 = \sup\{t \in [0,h] \mid \varphi^n([0,t], p+y) \subset W_2 \oplus P_n T([0,h])\}. \tag{141}$$

To finish the proof it is enough to show that $t_1 = h$. If this is not the case, then there exists $\delta > 0$, $t_1 + \delta \leq h$, such that we have $\varphi^n([0, t+\delta], p+y) \subset W(\epsilon) \oplus V(\epsilon)$. Hence we can use the constants $N_k^{\pm}$ in Lemma 15 for $t \in [0, t_1 + \delta]$ to obtain

$$Q_m \varphi^n([0, t_1 + \delta], p+y) \subset \operatorname{int} P_n T([0,h])).$$

From conditions (130) and (138) it follows that

$$P_m\varphi^n([0,t_1+\delta],p+y) \subset W_2.$$

Hence

$$\varphi^n([0,t_1+\delta],p+y) \subset W_2 \oplus P_n T([0,h]. \tag{142}$$

But this is in the contradiction with the definition of t_1 and our assumption that $t_1+\delta < h$. Hence $t_1 = h$. □

Lemma 23 *Same assumptions and definitions as in Lemma 22. If additionally for $k > M(T)$ we have*

$$N_k^- = -N_k^+, \qquad T(0)_k^- = -T(0)_k^+, \tag{143}$$

then $W_2 \oplus T(h)$ are self-consistent bounds for F.

Proof: Observe that $T(h)_k \subset T([0,h])_k \subset T_k$ for all $k > m$, hence $W_2 \oplus T(h) \subset W_2 \oplus T$. From this it follows that conditions C2 and C3 are satisfied on $W_2 \oplus T(h)$. To finish the proof is enough to notice that $T(h)_k^- = -T(h)_k^+$ for $k > M$. □

6.1 Uniform treatment of the tail, polynomial bounds

In a computer program we can not work directly with an infinite sequence of intervals $[a_k^-, a_k^+]$. We need to have a finite number of formulas describing $[a_k^-, a_k^+]$.

Definition 6 *Let $m \leq M$ be positive integers. The structure, T, consisting of the sequence of pairs $\{a_k^-, a_k^+\}_{k\in I, k>m}$, such that*

- *$a_k^- \leq a_k^+$ for all $k \in I$,*
- *there exists $C \geq 0$ and $s \geq 0$, such that*

$$a_k^+ = -a_k^- = \frac{C}{|k|^s}, \quad \textit{for } k > M \tag{144}$$

will be called the polynomial bound.

For polynomial bound T, by $m(T)$, $M(T)$, $s(T)$ and $C(T)$ we will denote the numbers m, M, s and C, respectively.

We define $T_k^\pm$ by

$$T_k^\pm = a_k^\pm.$$

When discussing algorithms we will also use the expression $T \in PolyBd$ to say that T is a polynomial bound.

We define the near tail of T *by $nearTail(T) = \Pi_{m\leq k\leq M}[a_k^-, a_k^+]$ and* the far tail of T *by $farTail(T) = \Pi_{k>M}[a_k^-, a_k^+]$.*

In my implementation for the KS-equation we consider polynomial bounds with the same m and M. For such class of tails it is easy to define and implement arithmetic and set theoretic operations.

For example the question of the verification of inclusion $T_k \subset b_k$ for $|k| > M$, where $T, b \in PolyBd$ and $M(b) = M(T) = M$ can be handled as follows:

- if $C(b) = 0$, then $T_k \subset b_k$ for $|k| > M$, iff $C(T) = 0$.
- if $C(b) \neq 0$ and $C(T) = 0$, then $T_k \subset b_k$ for $|k| > M$.
- if $C(b) \neq 0$. Let $K = \min\{|k| \quad k \in I, |k| > M\}$. Then $T_k \subset b_k$ for $|k| > M$, iff the following two conditions are satisfied

$$s(T) \geq s(b) \tag{145}$$

$$\frac{C(T)}{K^{s(T)}} \leq \frac{C(b)}{K^{s(b)}} \tag{146}$$

6.2 Uniform computation of b_k.

In the context of a computer assisted proof using the enclosure function based on Lemma 22 we have to explain how the expressions for $b_k^\pm$ and $g_k^\pm$ can be handled using polynomial bounds. In the remainder of this section we will use the notations from Lemma 22. Moreover, we assume that T and N are polynomial bounds such that $m(T) = m(N) = m$ and $M(T) = M(N) = M$ and for all other polynomial bounds introduced below we have these values of m and M.

For the further discussion we assume that λ_k satisfies conditions (3) and (4). We define an auxiliary function $V : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$V(x) = \inf\{v(|k|) \mid k \in I,\ |k| > x\}. \tag{147}$$

Observe that the assumption $\lambda_k < 0$ for $|k| > m$ implies that

$$0 < v(m) \leq v_1. \tag{148}$$

Now we are ready to explain how the formula

$$b_k^\pm = \frac{N_k^\pm}{-\lambda_k} \tag{149}$$

can be treated in a finite programmable way in terms of polynomial bounds.

We define $b \in PolyBd$ as follows

near tail to calculate the near tail of b we evaluate (149) for $k \in I$, $m < k \leq M$

far tail for $k > M$ we set

$$b_k^+ = -b_k^- = \frac{C(b)}{|k|^{s(b)}} = \frac{C(N)}{V(M)|k|^{s(N)+p}}. \tag{150}$$

Observe that with such definition we have $b_k^+ \geq \frac{N_k^+}{-\lambda_k}$ for all $|k| > M$ (with a reversed inequality for b_k^-) and this change corresponds to taking bigger value for N_k^+ in an application of Lemma 15. Hence the formulas for $T(h)$ and $T([0,h])$ give valid enclosures, when we use the polynomial bound $b_k^\pm$ defined above.

6.3 Uniform computation of $T(h)$

In Lemma 22 the following expression was obtained

$$T(h)_k^{\pm} \quad = \quad \left(T(0)_k^{\pm} - b_k^{\pm}\right) e^{\lambda_k h} + b_k^{\pm}. \tag{151}$$

We want to represent $T(h)$ as the polynomial bound. This is achieved by finding a larger set which is a polynomial bound and contains the product of intervals defined by equation (151).

The near tail of $T(h)$ is defined by a direct evaluation of (151). The far tail requires some analaitical work. We have

Lemma 24 *Let I, m, M, λ_k be as above. For any $r \in \mathbb{R}$ and $h > 0$, there exists $E = E(r, h, M) > 0$, such that*

$$e^{\lambda_k h} \leq \frac{E}{|k|^r}, \qquad \textit{for } |k| > M. \tag{152}$$

Proof: It is enough to observe that the function $|k|^r e^{-a|k|^p h}$, where $a = \inf\{v(|k|), |k| > M\} > 0$, is bounded. □

Now we are ready to give a formula for $T(h)_k^{\pm}$ for $k > M$

$$T(h)_k^+ \leq T(0)_k^+ e^{\lambda_k} + b_k^+ \leq \frac{C(T(0))}{|k|^{s(T(0))}} \cdot \frac{E(s(b) - s(T(0)), h, M)}{|k|^{s(b)-s(T(0))}} + \frac{C(b)}{|k|^{s(b)}} = \frac{C(T(0)) \cdot E(s(b) - s(T(0)), h, M) + C(b)}{|k|^{s(b)}}.$$

Hence we set

$$\begin{aligned} C(T(h)) &= C(T(0)) \cdot E(s(b) - s(T(0)), h, M) + C(b), \\ s(T(h)) &= s(b) \end{aligned}$$

7 The enclosure procedure for the tail

The goal of this section is to describe the function, which constructs the rough enclosure (Part 1 of the algorithm) and computes the tail after the time step (Part 5) for dissipative PDEs. The proposed function is based on Lemma 22 and uses the notion of the polynomial bound introduced in Section 6.1.

As in Section 6 throughout this section we assume that the range of k is $I = \mathbb{Z}_+$ and $\dim H_k = 1$. The modification required for other dissipative equations with periodic boundary conditions is obvious and will not be discussed. We have $m, M \in \mathbb{Z}_+$ fixed in advance and all polynomial bounds will use these values.

We assume that we have the enclosure function for the differential inclusion (see Section 5.3)

$$x' \in P_m F(x + T) \tag{153}$$

where $x \in \mathbb{R}^m$ and T is the tail.

We assume that this function has the following declaration
function incl_enclosure$(h \in \mathbb{R}, [x] \subset \mathbb{R}^m, T \in PolyBd, [W_2] \subset \mathbb{R}^m) \in bool$.

This function constructs the set $[W_2] \subset \mathbb{R}^m$, such that

$$\varphi([0,h],[x],T) \subset [W_2]. \tag{154}$$

If it succeeds then *true* is returned and $[W_2]$ is updated, otherwise it returns *false*. In both cases the parameter T is unchanged.

7.1 Case of an a-priori given tail

We have the set $A \subset \mathbb{R}^m$ representing the a-priori bounds used to compute the global tail, $T_G = \Pi_{k>m}[T^-_{G,k}, T^+_{G,k}]$.

We generate $[W_2]$ by calling **function incl_enclosure**$(h,[x],T_G,[W_2]) : bool$ and we check if $[W_2] \subset A$. If this is the case then the pair $([W_2], T_G)$ is validated.

This is the approach used in [Z2]. It turned out to be ineffective when compared to the one with the evolving tail described below.

7.2 Basic functions

We assume that we have a function computing the nonlinear term in (26) with the following declaration
function $N([z] \subset \mathbb{R}^m, T \in PolyBd) \in PolyBd$,
where $[z]$ and T are such that such that for all $k > m$ holds

$$\inf_{(x,y)\in[z]\oplus T} N_k(x,y) > N^-([z],T), \qquad \sup_{(x,y)\in[z]\oplus T} N_k(x,y) < N^+([z],T), \tag{155}$$

and for $k > M$ we have

$$N^+_k([z],T) = -N^-_k([z],T) = \frac{C(N)}{k^{s(N)}}. \tag{156}$$

For the KS equation in our implementation we have $s(N) = s(T) - 2$, but it is possible to obtain $s(N) = s(T) - 1$.

There is an unpleasant feature of our implementation of $N([z],T)$ (but it appears to be a rather inherent for such approach): it happens that (see formula (233) in Section 8): we have two tails $T_2 \subset T_1$, such that $T_{1,k} = T_{2,k}$ for $m < k \leq M+1$ (the near-tails of T_1 and T_2 are the same), but $s(T_1) < s(T_2)$ (the far tail in T_2 is decaying faster than that in T_1), but nevertheless

$$N^+_{M+1}([z],T_1) < N^+_{M+1}([z],T_2), \tag{157}$$

which later produces worse isolation intervals (for $k \approx M+1$) for T_2 than for T_1. This phenomenon results from the following fact, when we try to bound N_k by $\frac{C}{k^s}$, then taking larger s forces larger C, which may result in larger value of N_k for $k \approx M+1$.

To handle the above issue we introduce the function (the method) **decpower**, which for the polynomial bound T will produce a new polynomial bound T' with a slower decay rate for the far tail. Namely if $T' = T.decpower(d)$, then

$$
\begin{aligned}
T &\subset T' && (158)\\
T'_k &= T_k, \qquad \text{for } m < k \leq M+1 && (159)\\
s(T') &= s(T) - d. && (160)
\end{aligned}
$$

The following obvious lemma tells how to check condition $T([0,h]) \subset T$ from Lemma 22.

Lemma 25 *The same assumptions and definitions as in Lemma 22. Assume that*

$$
\begin{aligned}
&T(0) \subset T\\
&\textit{if } T(0)^+_k < b^+_k, \textit{ then } T^+_k \geq g^+_k \textit{ for } k > m\\
&\textit{if } T(0)^-_k > b^-_k, \textit{ then } T^-_k \leq g^-_k \textit{ for } k > m.
\end{aligned}
$$

Then

$$T([0,h]) \subset T.$$

Now we are ready to describe the algorithm for the tail validation.

function validate_tail($h \in \mathbb{R}, [z] \subset \mathbb{R}^m, T(0) \in PolyBd, T \in PolyBd,$ $gen\_new \in bool) \in bool$
Input parameters:

- $h > 0$ is the time step
- $[z] \subset \mathbb{R}^m$ represents the a-priori bounds for $x_i(0,h])$ for $i = 1,\ldots,m$
- $T(0)$ is the initial condition for the tail,
- T is the candidate for $T([0,h])$,
- *gen_new* tells whether to generate (update) T.

Output: *true* is returned if T is validated, otherwise *false* is returned. Additionally if *gen_new* is equal to *true*, then T is updated as follows : in case it is validated, then we find better (smaller) T, otherwise we produce the new guess for T. If *gen_new* is equal to *false*, then T is left unchanged. For the precise meaning of validation see Theorem 27

the body of the function:

- we set
$validated = false$, $farTailValidated = false$, $kvalidated[k] = false$ or $m < k \leq M$.

- computation of $N \in PolyBd$, $b \in PolyBd$ and g_k for $m < k < M$

$$\begin{aligned} N_k^{\pm} &= N_k^{\pm}([z], T) && (161)\\ b_k^{\pm} &= \frac{N_k^{\pm}}{-\lambda_k}, \qquad \text{for } m < k \leq M && (162)\\ g_k^{\pm} &= \left(T(0)_k^{\pm} - b_k^{\pm}\right) e^{\lambda_k h} + b_k^{\pm}, \qquad \text{for } m < k \leq M && (163)\end{aligned}$$

  To define b_k for $k > M$ we proceed along the lines described in Section 6.2. We set

$$b_k^+ - b_k^- = \frac{C(N)}{V(M) k^{s(N)+p}} = \frac{C(b)}{k^{s(b)}}, \tag{164}$$

  where for the KS equation from (28) we have $V(M) = \nu - \frac{1}{(M+1)^2}$

  Observe that with such b_k we have for $k > M$

$$\frac{N_k^+}{-\lambda_k} \leq b_k^+ \tag{165}$$

  and the equality holds for $k = M + 1$ only.

- *validation*, we set $validated = true$ if the assumptions in Lemma 25 are satisfied, because if it is the case then from Lemma 22 we obtain the desired enclosure.

  Below we discuss this verification in some detail.

  The first check $T(0) \subset T$ is discussed in Section 6.1. In it does not hold then we exit the function returning *false*.

  Next we have to check the following conditions for all $k > m$

$$\begin{aligned} &\text{if} \quad T(0)_k^+ < b_k^+, \quad \text{then} \quad T_k^+ \geq g_k^+ && (166)\\ &\text{if} \quad T(0)_k^- > b_k^-, \quad \text{then} \quad T_k^- \leq g_k^-. && (167)\end{aligned}$$

  For the near tail ($m < k \leq M$) we verify the above conditions one by one, setting $kvalidated[k] = true$ when (166) and (167) are satisfied for k and $kvalidated[k] = false$, otherwise.

  For the far tail ($k > M$) we proceed as follows. First of all observe that due to symmetry of all polynomial bounds involved it is enough to verify condition (166), only.

  We have three cases:

  I. $s(b) > s(T(0))$ and $C(T(0)) \neq 0$.
  We check that

$$T_k^+ \geq g_k^+, \qquad \text{for } M + 1 \leq k \leq L, \tag{168}$$

where

$$L = \left(\frac{C(b)}{C(T(0))} \right)^{\frac{1}{s(b)-s(T(0))}} \tag{169}$$

If (168) is satisfied we set $farTailValidated = true$.

To justify the above condition let us notice that if $k \geq M + 1$ and $k \geq L$, then $T(0)^+_k \geq b^+_k$ for $k \geq M + 1$. Observe also that if $L < M+1$, then condition (168) is satisfied, because there are no k's in this range.

II. $s(b) = s(T(0))$ or $C(T(0)) = 0$

If

$$C(T(0)) \geq C(b), \tag{170}$$

then we set $farTailValidated = true$, because in this situation we have $T(0)^+_k \geq b^+_k$ for $k \geq M + 1$, hence there is nothing more to check.

If (170) does not hold, then we should check whether $T^+_k \geq g^+_k$. Since in this case we have $T(0)^+_k < g^+_k < b^+_k$, we will instead check the stronger condition

$$T^+_k \geq b^+_k, \quad \text{for } k > M, \tag{171}$$

which is equivalent to the following two conditions

$$s(T) \leq s(b), \qquad T^+_{M+1} \geq b^+_{M+1}. \tag{172}$$

If the above conditions are satisfied then we set $farTailValidated = true$.

III. $s(b) < s(T(0))$ and $C(T(0)) \neq 0$.

Let us define

$$L = \left(\frac{C(T(0))}{C(b)} \right)^{\frac{1}{s(T(0))-s(b)}}. \tag{173}$$

It is easy to see that

$$T(0)^+_k \geq b^+_k, \qquad \text{for } M < k \leq L \tag{174}$$

$$T(0)^+_k < b^+_k, \qquad \text{for } k > L \text{ and } k > M. \tag{175}$$

Hence for $k > M$ and $k > L$ we have to check that $T^+_k \geq g^+_k$. Like in the previous case we replace g^+_k by b^+_k and we obtain the following two conditions

$$s(T) \leq s(b), \qquad T^+_{p+1} \geq b^+_{p+1}, \tag{176}$$

where $p = max(M, \text{int}(L))$ and $\text{int}(L)$ is the largest integer less than or equal to L.

If the above conditions are satisfied then we set $farTailValidated = true$.

- *update of* T. There are two update modes depending on the current value of the boolean variable *validated*.

  If $validated = true$, then we proceed as follows (compare formulas (136) and (137) in Lemma 22):
  For $i = m+1$ to M we update $T_k^{\pm}$ as follows

$$\begin{aligned} &\text{if } b_k^+ \leq T(0)_k^+ \text{ then } , && T_k^+ = T(0)_k^+, && (177)\\ &\text{if } b_k^+ > T(0)_k^+ \text{ then } , && T_k^+ = g_k^+, && (178)\\ &\text{if } b_k^- \geq T(0)_k^- \text{ then } , && T_k^- = T(0)_k^-, && (179)\\ &\text{if } b_k^- < T(0)_k^- \text{ then } , && T_k^- = g_k^-. && (180)\end{aligned}$$

  For the far tail we perform the modification only if $b_k \subset T(0)_k$, for $k > M$ and $C(T(0)) \neq 0$. If this is the case then we leave $s(T)$ unchanged and we set

$$C(T) = C(T(0))(M+1)^{s(T)-s(T(0))}. \tag{181}$$

  With this modification we obtain $farTail(T(new)) \subset fartail(T(old))$ and $T_{M+1} = T(0)_{M+1}$.

  Now if $validated = false$, then we modify only these coordinates in the near tail, for which the validation failed ($kvalidated[k] = false$). The far tail in T is modified only when it is validated ($farTailValidated = true$) and $b_k \subset T(0)_k$ for $k > M$. Below we present the details.

  We have two parameters $0 < d_g < 1$ and $d_2 > 1$ (in my program $d_g = 0.1$, $d_2 = 1.01$). For $k = m+1$ to M, such that $kvalidated[k] = false$ we do the following

$$\begin{aligned} &\text{if } b_k^+ > T_k^+ \text{ then } , && T_k^+ = (1-d_g)g_k^+ + d_g b_k^+, && (182)\\ &\text{if } b_k^- < T_k^- \text{ then } , && T_k^- = (1-d_g)g_k^- + d_g b_k^-, && (183)\\ &T_k = \mathit{inflate}(T_k, d_2). && && (184)\end{aligned}$$

  If $farTailValidated = true$ and $b_k \subset T(0)_k$ for $k > M$ holds and $C(T(0)) \neq 0$, then we modify T as follows: we leave $s(T)$ unchanged and we set

$$C(T) = C(T(0))(M+1)^{s(T)-s(T(0))}. \tag{185}$$

  If $farTailValidated = false$, then we define the new far tail so that $b_k \cup T(0)_k \subset T_k$. For this end we leave $s(T)$ unchanged and we set

$$C(T) = \max\left(d_2 C(b)(M+1)^{s(T)-s(b)}, C(T(0))(M+1)^{s(T)-s(T(0))}\right).$$

Remark 26 *Let us remark that it is essential for our function to work that we keep $s(T)$ unchanged instead of setting $s(T) = s(T_0)$, because increasing s may result in worse estimates for N_k for $k \approx M+1$, see comments at begin of this subsection.*

- return *validated*.

End of the function validate_tail

Theorem 27 *Assume that* ***validate_tail***$(h, [z], T(0), T, gen\_new)$ *returns true. Let $n > M$, let $(x(t), y(t)) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$ for $t \in [0, h]$ be a solution of*

$$x' = P_m F(x, y) \tag{186}$$

$$y' = P_n Q_m F(x, y), \tag{187}$$

such that $x(t) \in [z]$ for $t \in [0, h]$ and $y(0) \in T(0)$, then

$$y(t) \in T, \qquad \text{for } t \in [0, h]. \tag{188}$$

Proof: For the proof it is enough to compare the checks performed in the validation part with Lemmas 25 and 22. In particular it is easy to see, that if $validated = true$, then $T([0, h]) \subset T$, where $T([0, h])$ is defined as in Lemma 22. In the update part the substitution $neartail(T) = neartail(T([0, h]))$ is performed for the near tail and for the far tail we substitute it with some enclosure of $T([0, k])_k$ for $k > M$. □

7.3 The enclosure algorithm

We assume that we have the function $guessfarTail$, which produces a reasonable initial guess for the far tail. For the KS equation on the line with odd and periodic boundary conditions such a function is given in Section 8.2.

function enclosure_with_tail$(h : real, [x] \subset \mathbb{R}^m, T(0) : PolyBd, [W_2] \subset \mathbb{R}^m, T : PolyBd, Tisgoodinitialguess : bool) : bool$;
begin
    $max\_iter = T(0).M/2$;
    $maxdcount = 3$;
    if !Tisgoodinitialguess
        $Tinitial = guessfarTail([x], T(0))$;
    else Tinitial=T;
    $validated = false$;
    $dcount = 0$;
    while(!validated **and** $(dcount \leq maxdcount)$) **do**
        $[W_2] = [x] + [0, h] \cdot P_m F([x])$;
        $T = Tinitial$;
        $T.decpower(dcount)$;
        // the initial guess for the tail
        if validate_far_tail$(h, [W_2], T(0), T)$ **then**
            validate_tail$(h, [W_2], T(0), T, true)$;
        // we have now the initial guess for the tail in variable T
            $i = 1$;

```
      while ((!validated) and (i ≤ max_iter)) do
        if  incl_enclosure(h, [x], T, [W_2])  then
          validated=validate_tail([W_2], T(0), T, true);
        i = i + 1;
      end while;
    end if;
    dcount = dcount + 1
  end while

  i = 1; /* the refinement loop */
  max_iter = 1;
  while (i ≤ max_iter) do
    incl_enclosure(h, [x], T, [W_2]);
    validate_tail([W_2], T(0), T, true);
    i = i + 1;
  end while;
  return validated;
end
```

From Theorem 27 and the above algorithm we obtain immediately the following

Theorem 28 *Let $h > 0$, assume that $[x] \oplus T(0)$ are self-consistent bounds, $m = m(T(0))$ and $M = M(T(0))$. Assume that* ***enclosure_with_tail***$(h, [x], T(0), [W_2], T, Tisgoodinitialguess)$ *returns* true.

Then for any $n > M$, $x(0) \in [x]$ and $y(0) \in P_n T(0)$ holds

$$\varphi^n([0,h], x(0) \oplus P_n y(0)) \quad \subset \quad [W_2] \oplus P_n T. \tag{189}$$

Moreover, $[W_2] \oplus T$ are self-consistent bounds.

7.4 Computation of $T(h)$

Assume that $[x_0] \oplus T(0)$ and $[W_2] \oplus T([0,h])$ are self-consistent bounds, such that for $n > M$ holds

$$\varphi^n([0,h], x(0) \oplus P_n T(0)) \subset [W_2] \oplus P_n T([0,h]). \tag{190}$$

From Lemma 22 it follows that

$$T(h)^{\pm}_k = \left(T(0)^{\pm}_k - b^{\pm}_k\right) e^{\lambda_k h} + b^{\pm}_k, \qquad \text{for } k > m, \tag{191}$$

where $b \in PolyBd$ satisfies

$$N \;=\; N([W_2], T([0,h])) \in PolyBd \tag{192}$$

$$\frac{N_k}{-\lambda_k} \;\subset\; b_k, \quad \text{for } m < k. \tag{193}$$

To enclosure $T(h)$ we proceed along the lines outlined in Section 6.3. We need to find $E = E(r, h, M)$ defined in Lemma 24 for the KS equation given by (28).

As the first step in this direction we prove the following lemma.

Lemma 29 *Assume $\lambda_k = -\nu k^4 + k^2$, where $\nu > 0$. Let $r, E, h \in \mathbb{R}$, $E > 0$, $h > 0$. Assume that for some $K > 0$ holds*

$$e^{h\lambda_K} \leq \frac{E}{K^r} \tag{194}$$

$$-4h\nu K^4 + 2hK^2 + r \leq 0, \tag{195}$$

$$4\nu K^2 \geq 1. \tag{196}$$

Then for any $k > K$ holds

$$e^{h\lambda_k} \leq \frac{E}{k^r}. \tag{197}$$

Proof: It is enough to show that the function $f(k) = e^{h\lambda_k} k^r$ is nonincreasing for $k \geq K$.

We have

$$f'(k) = \lambda_k' h e^{h\lambda_k} k^r + r e^{h\lambda_k} k^{r-1} = (\lambda_k' k h + r) k^{r-1} e^{h\lambda_k}$$

Hence $f'(k) \leq 0$ if

$$g(k) = -4h\nu k^4 + 2hk^2 + r \leq 0. \tag{198}$$

We want the above condition to hold for $k \geq K$. We will show it by proving that $g'(k) \leq 0$ for $k > K$, because in view of (195) we know that $g(K) \leq 0$.

Observe that $g'(k) < 0$ iff the following condition holds

$$4hk(-4\nu k^2 + 1) \leq 0 \tag{199}$$

since we are interested in $k > 0$, hence we obtain

$$4\nu k^2 \geq 1, \qquad \text{for } k \geq K. \tag{200}$$

□

We look for $C(T(h))$, such that for $k > M$ we have

$$T(h)_k^+ < T(0)_k^+ e^{\lambda_k h} + b_k^+ \leq \frac{C(T(h))}{k^{s(b)}}. \tag{201}$$

To compute $C(T(h))$ we use Lemma 29 with $r = s(b) - s(T(0))$. We check whether $K \leq M + 1$ (if this is not the case we return the *failure* message). Hence we have to verify that

$$-4h\nu(M+1)^4 \quad + \quad 2h(M+1)^2 + s(b) - s(T(0)) \leq 0, \tag{202}$$

$$4\nu(M+1)^2 \quad \geq \quad 1. \tag{203}$$

If above conditions are satisfied then we set

$$E = e^{h\lambda_{M+1}} (M+1)^{s(b) - s(T(0))}. \tag{204}$$

Now from (151) it follows that we can set

$$T_k^{\pm}(h) = \pm \frac{C(T(0))E + C(b)}{k^{s(b)}}. \tag{205}$$

Observe that with the above definition of $T(h)$ there is no guarantee that $T(h)_{k>M} \subset T_{k>M}$, hence in the final step we set

$$T(h) := T[0,h] \cap T(h).$$

Remark 30 *To obtain the feeling about conditions (202) and (203) let us consider typical numbers for the KS equation (28). For example for the possible chaotic case for the KS we have $r = 2$, $v \approx \nu \approx 0.03$, $M = 3m = 36$, $h \approx \frac{1}{2\lambda_m}$. We obtain*

$$4\nu M^2 \approx 155.5$$

$$-4h\nu M^4 + 2hM^2 + r = -4\frac{M^4\nu}{2\nu(\frac{M}{3})^4} + \frac{2M^2}{2\nu(\frac{M}{3})^4} + 2 = -162 + \frac{81}{\nu M^2} + 2 \approx -158$$

So conditions (202) and (203) are satisfied with large margin.

7.5 Estimates during the time step

From the results of Section 6 (the monotonicity of the bounds) it follows that we have the following refinement of the enclosure $T = T([0,h])$

$$T([0,h]) \subset T(0) \cup T(h). \tag{206}$$

We will use it in the section region in the computation of the bounds for the Poincaré map (see [ZLo, Section 5]).

8 Finding a good guess for the far tail for the KS equation

We will discuss here the question: *How to obtain a good initial guess for the far tail?*

By a good guess we understand $T \in PolyBd$, such that condition (138) in Lemma 22 is likely to be satisfied. In this section we consider the KS equation (28) and we derive heuristic conditions, which will guarantee that

$$\frac{N_k([z],T)}{-\lambda_k} \subset T_k, \qquad \text{for } k > M. \tag{207}$$

Observe that (207) together with condition $T(0)_k \subset T_k$ for $k > M$ (this is a minimal requirement for T being the tail enclosing evolution of $T(0)$) implies that

$$T([0,h])_k \subset T_k, \qquad \text{for } k > M. \tag{208}$$

As the result of the analysis of (207) we will obtain:

- the relation between possible values of M and s for T, (see condition (238))
- the function realizing the guess of the far tail (see Section 8.2)

The KS equation with odd and periodic boundary conditions in the Fourier domain can be written as (compare (28))

$$a'_k = k^2(1-\nu k^2)a_k - k(FS(k) - 2\cdot IS(k)), \qquad k = 1,2,\ldots \tag{209}$$

where

$$\begin{aligned} FS(k) &= \sum_{n=1}^{k-1} a_n a_{k-n} & (210)\\ IS(k) &= \sum_{n=1}^{\infty} a_n a_{n+k} & (211)\\ B_k &= -FS(k) + 2IS(k) & (212)\\ N_k &= kB_k & (213) \end{aligned}$$

We fix $T \in PolyBd$. Let $N = N(W,T) \in PolyBd$, where $W \subset \mathbb{R}^m$ is a compact set, which will not be important in the following discussion. In the sequel we assume that $C = C(T)$ and $s = s(T)$. First we need to find $D = C(N)$ (see [ZM, Corollary 3.7]), such that

$$|B_k| \leq \frac{D}{k^{s-1}}. \tag{214}$$

Here we will organize the computation of D slightly differently than in [ZM] to get a better feeling about the dependence of D on C and s. There are also some mistakes in the printed version of [ZM] on page 279 in formulas for D_1 and D_2, which where derived from (correct) Lemma 3.6 in [ZM] (but fortunately these errors were not present in the actual programm, which was based on the correct Lemma 3.6).

First, we seek the bounds for $FS(k)$ and $IS(k)$

$$|FS(k)| \leq \frac{D_1}{k^{s-1}}, \qquad |IS(k)| \leq \frac{D_2}{k^{s-1}} \tag{215}$$

and then we obtain

$$|B_k| \leq \frac{D_1 + 2D_2}{k^{s-1}}. \tag{216}$$

The following lemmas has been proven in [ZM].

Lemma 31 *[ZM, Lemma 3.4] Let $M < k \leq 2M$. Then*

$$FS(k) \subset 2 \sum_{k-M\leq n<k/2} a_n a_{k-n} + e(k)a_{k/2}^2 + 2C \sum_{n=1}^{k-M-1} \frac{|a_n|}{(k-n)^s}[-1,1],$$

where $e(k) = 1$ if k is even and $e(k) = 0$ when k is odd.

Lemma 32 *[ZM, Lemma 3.5] Let $k > 2M$. Then*

$$FS(k) \subset \frac{C}{k^{s-1}} \left(\frac{2^{s+1}}{2M+1} \sum_{n=1}^{M} |a_n| + \frac{C4^s}{(2M+1)^{s+1}} + \frac{C2^s}{(s-1)M^s} \right) [-1, 1].$$

Lemma 33 *[ZM, Lemma 3.6] Let $k > M$. Then*

$$IS(k) \subset \frac{C}{k^{s-1}(M+1)} \left(\frac{C}{(M+1)^{s-1}(s-1)} + \sum_{n=1}^{M} |a_n| \right) [-1, 1].$$

Let us set

$$D_1(k \leq 2M) = \max\{k^{s-1}|FS(k)|, \quad M < k \leq 2M\}, \tag{217}$$

$$D_1(k > 2M) = C \left(\frac{2^{s+1}}{2M+1} \sum_{n=1}^{M} |a_n| + \frac{C4^s}{(2M+1)^{s+1}} + \frac{C2^s}{(s-1)M^s} \right) \tag{218}$$

From the above lemmas it follows immediately that

$$D_1 = \max(D_1(k > 2M), D_1(k \leq 2M)) \tag{219}$$

$$D_2 = \frac{C}{(M+1)} \left(\frac{C}{(M+1)^{s-1}(s-1)} + \sum_{n=1}^{M} |a_n| \right) \tag{220}$$

8.1 Dependence of D_i on C and s

We will make several assumptions regarding the candidate tail (these numbers are typical for attracting periodic orbits for $\nu = 0.1 - 0.127$)

$$\sum_{n=1}^{M} \sup_{W \oplus T} |a_n| \approx A \approx 2 \tag{221}$$

$$C < 10^{15} \tag{222}$$

$$s \approx 12 \tag{223}$$

$$M \approx 40 \tag{224}$$

$$\frac{C}{M^s} < 10^{-8} \tag{225}$$

Consider first D_2. It is easy to see that the term linear in C is dominant, namely

$$\left(\sum_{n=1}^{M} |a_n| \right) / \left(\frac{C}{(M+1)^{s-1}(s-1)} \right) \approx \frac{A(M+1)^{s-1}(s-1)}{C} <$$

$$\frac{2 \cdot 40^{10} \cdot 10}{10^{15}} = 2^{21} 10^{-4} \approx 2 \cdot 10^6 \cdot 10^{-4} = 200$$

Hence we have

$$D_2 \approx \frac{A}{M+1} \cdot C < \frac{C}{10} \tag{226}$$

Remark 34 *D_2 appears to depend linearly on C. The dependence on s appears to be insignificant.*

Now we take a closer look at $D_1(k > 2M)$ observe first that the third term is considerably larger than the second one. Namely we have

$$\frac{C4^s}{(2M+1)^{s+1}} < \frac{C2^s}{M^s(2M+1)} = \frac{s-1}{2M+1} \cdot \frac{C2^s}{(s-1)M^s} \tag{227}$$

The first term is dominating the other two. Namely we have

$$\frac{\frac{2^{s+1}}{2M+1}\sum_{n=1}^{M}|a_n|}{\frac{C2^s}{(s-1)M^s}} \approx \frac{2A(s-1)}{(2M+1)\frac{C}{M^s}} \approx 10^8 \tag{228}$$

Hence we obtain the following

Remark 35 *It appears that*

$$D_1(k > 2M) \approx \frac{2^{s+1}A}{2M+1}C \approx 2^{12}C$$

Moreover, it is also clear that $D_1(k > 2M)$ is several orders of magnitude larger than D_2. There is also the significant dependence on s.

The expression for $D_1(k \leq 2M)$ appears to be more difficult to analyze. For further discussion let us define

$$FS_1(k) = \left| 2 \sum_{k-M \leq n < k/2} a_n a_{k-n} + e(k)a_{k/2}^2 \right| \tag{229}$$

$$FS_C(k) = 2C \sum_{n=1}^{k-M-1} \frac{|a_n|}{(k-n)^s}. \tag{230}$$

We have

$$|FS(k)| \leq FS_1(k) + FS_C(k).$$

Let us make a few observations :

- $FS_1(k)$ contains the largest number of terms for $k = M+1$, hence we expect that it achieves is the maximum value for $k = M+1$. It is much less obvious where the maximum for $k^{s-1}FS_1(k)$ will be, but in my experiments it turns out that it is achieved also for $k = M+1$,
- in the term $FS_C(k)$ the number of terms increases with k, but since there are only a few dominating modes (say d) we can approximate $FS_C(k)$ and $k^{s-1}FS_C(k)$ for $k > M-1+d$ by

$$FS_C(k) \approx 2C\left(\sum |a_n|\right)/k^s \tag{231}$$

$$k^{s-1}FS_C(k) \approx 2C\left(\sum |a_n|\right)/k \leq \frac{2A}{M+1}C \tag{232}$$

From above considerations it follows that

$$D_1(k \leq 2M) \approx (M+1)^{s-1} F S_1(M+1) + \frac{2A}{M+1} C.$$

Summarizing it appears that

$$D(C,s) \approx \max\left(D_1(k \leq 2M), \frac{2^{s+1}A}{2M+1} C \right), \tag{233}$$

hence for large C

$$D(C,s) \approx \frac{2^{s+1}A}{2M+1} C. \tag{234}$$

Consider now isolation equation (207) for $k > M$

$$\frac{C}{k^s} \geq \frac{D}{k^{s-2} k^4 \left(\nu - \frac{1}{k^2}\right)}. \tag{235}$$

An easy computation shows that it is equivalent to

$$k^2 \left(\nu - \frac{1}{k^2}\right) C \geq D, \qquad k > M. \tag{236}$$

It turns out that it is enough to check the above inequality for $k = M+1$, only. Hence we obtain

$$(M+1)^2 \left(\nu - \frac{1}{(M+1)^2}\right) C \geq D \tag{237}$$

After using (234) we obtain

$$\frac{(M+1)^2 (2M+1) \left(\nu - \frac{1}{(M+1)^2}\right)}{A} \geq 2^{s+1}. \tag{238}$$

8.2 The function for generation of the initial guess.

It appears that equation (238) should serve as the basic test, whether the enclosure is possible with the given values of s and M, because it guarantees that a 'large' enclosure for the far tail always exists. In this situation it is enough to increase C in some geometric fashion until we enter into the linear regime for $D(C)$.

It is also easy to compute C_L, where the linear regime approximately begin. From (233) we obtain

$$C_L = \frac{D_1(k \leq 2M) \cdot (2M+1)}{2^{s+1}A}. \tag{239}$$

The above considerations lead to the following procedure for guessing the enclosure for the far tail.

Guess of the enclosure for the tail
Input: $[x] \subset \mathbb{R}^m$, $T(0) \in PolyBd$
Output: $T \in PolyBd$, this is a candidate for $T([0,h])$.

0. $m(T) = m(T(0))$, $M(T) = M(T(0))$

1. Computation of A, $D_1(k \leq 2M)$

2. Computation of $s(T)$. We seek the largest integer, s_{max}, such that condition (238) holds and $s_{max} \leq s(T(0))$. We set $s(T) = s_{max}$.

3. Computation of $C(T)$. We take the maximum of $C(T(0))$ and $3C_L$, where C_L is given by (239).

Warning: If we start with an empty near tail and we just evaluate A and D_1 on the initial condition, then we end up with $D_1(k \leq 2M) = 0$, which leads $C_L = c(T) = 0$. But even in this case a correct value of $s(T)$ is generated, and while this is really not a good guess since the tail is empty, the enclosure function works, because the update step in function **validate_tail** produces a new candidate, T, such that $C(T) \neq 0$.

8.3 Other equations.

The analysis presented in this section was restricted to the KS equation (9) and used heavily the fact that N was quadratic. But it is quite obvious that this approach could be generalized to a general polynomial function N. Observe that in this case we should obtain the following expression

$$|N_k| \leq \frac{CA_1 + C^2A_2 + \cdots + C^pA_p}{k^{s-r}} \tag{240}$$

where A_i are functions of M, s. Just as in the case of the KS equation the functions A_i for $i > 1$ will contain positive powers M in the denominator, hence for bounded set of possible C the terms A_iC^i for $i > 1$ could be made as small as need by taking sufficiently large M and then we are in the situation already considered for the KS equation earlier in this section.

9 Computation of the Poincaré map

The goal of this section is to discuss the question of the computation of the Poincaré map for (26).

We fix parameters m and M, and we assume that $m = m(S)$ and $M = M(S)$ for all self-consistent bounds, S, appearing in the sequel.

To compute the Poincaré map we need the estimates for the trajectory during the time step. As such estimates one can use the rough enclosure obtained in Part 1 of the algorithm, but these estimates are usually too crude and moreover can be quite easily improved. For the tail this was discussed in Section 7.5 and for the main variables ($x \in X_m$) the procedure is described in Section 6.7 in [Z2].

Consider a sequence $0 = t_0 < t_1 < \cdots < t_N$ and let $h_i = t_i - t_{i-1}$ be the corresponding time steps. Assume that we apply our algorithm for rigorous integration of (26) to some initial condition $X_0 \oplus T_0$ using the time steps h_i. To facilitate the further discussion we introduce the following notation

- by $\widehat{\varphi}(t_i, X_0 \oplus T_0)$ we will denote the result of i-th iteration of our algorithm for the sequence of time steps $h_1, \dots, h_i$
- for any $h > 0$ and self-consistent polynomial bounds V by $\widehat{\varphi}([0,h], V)$ we will denote the enclosure for $\varphi([0,h], V)$ obtained by our algorithm.

Using the above conventions we have:
For any $n > M$ and $x_0 \in X_0 \oplus T_0$

$$\varphi^n(t_i, P_n(x_0)) \in \widehat{\varphi}(t_i, X_0 \oplus T_0), \quad i = 1, \dots, N \tag{241}$$

$$\varphi^n([t_{i-1}, t_i], P_n(x_0)) \subset \widehat{\varphi}([0, h_i], \widehat{\varphi}(t_{i-1}, X_0 \oplus T_0)), \quad i = 1, \dots, N. \tag{242}$$

Extending in a natural way the above notation we set

$$\widehat{\varphi}([t_i, t_{i+k}], X_0 \oplus T_0) = \bigcup_{l=1}^{k} \widehat{\varphi}([0, h_{i+l}], \widehat{\varphi}(t_{i+l-1}, X_0 \oplus T_0)). \tag{243}$$

We have for any $n > M$

$$\varphi^n([t_i, t_{i+k}], P_n(X_0 \oplus T_0)) \subset \widehat{\varphi}([t_i, t_{i+k}], X_0 \oplus T_0). \tag{244}$$

In the computer assisted proofs of the existence of periodic orbits we consider the Poincaré maps for all Galerkin projections P_n of (26) for $n > M$. Moreover, we want to obtain such bounds for all such Galerkin projections in a single application of the algorithm as in Theorem 28. For this purpose we will always define the section of (26) in terms of X_m.

We define the section $\theta \subset H$ as follows:
Let $\alpha : X_m \to \mathbb{R}$ be a C^1-function.

- $\theta = \{x \in H \mid \alpha(P_m x) = 0\}$
- $P_m(\theta)$ is a submanifold in X_m of the codimension one.

In our computation for the KS equation we always use linear (affine) α. For the purpose of computation of the Poincaré map we need to add some transversality condition with respect to (26). But since the vector field defined by (26) might be not defined on θ, we rather formulate an easy theorem containing the transversality condition as an assumption, which have to be verified during the execution of the algorithm.

In paper [Z2] since the tail was fixed we considered the Poincaré map on section θ, denoted there by $\mathcal{G}_\theta$, as a multivalued map defined on $P_m(\theta)$ with values in $P_m(\theta)$. In this paper we will rather treat $\mathcal{G}_\theta$ be as a multivalued map with both the domain and the range being infinite dimensional.

Definition 7 *Consider (26) and the section θ. For $n > M$ let us denote the Poincaré map for φ^n by $G_{n,\theta}$. Then we define the Poincaré map $\mathcal{G}_\theta$ as follows:*

$$x \in \operatorname{dom} \mathcal{G}_\theta, \quad \textit{iff} \quad P_n x \in \operatorname{dom} G_{n,\theta} \quad \textit{for all } n > M$$
$$\mathcal{G}_\theta(x) = \textit{convex hull}\left(\{G_{n,\theta}(P_n(x)) \mid n > M\}\right), \quad x \in \operatorname{dom} G_\theta$$

For two sections θ_1 and θ_2 analogously define $G_{n,\theta_1 \to \theta_2}$ and $\mathcal{G}_{\theta_1 \to \theta_2}$. In this notation we have $\mathcal{G}_\theta = \mathcal{G}_{\theta \to \theta}$.

Theorem 36 *Consider (26). Let $X_0 \oplus T_0$ be self-consistent bounds, such that there exists N, a sequence of real numbers t_i for $i = 1, \ldots, N$ and two sequences of self-consistent bounds $X_i \oplus T_i$ for $i = 1, \ldots, N$ and $W_i \oplus V_i$ for $i = 1, \ldots, N-1$ such that*

$$0 < t_1 < \cdots < t_N, \tag{245}$$
$$\widehat{\varphi}(t_1, X_0 \oplus T_0) \quad \subset \quad X_1 \oplus T_1, \tag{246}$$
$$\widehat{\varphi}([0, t_{i+1} - t_i], X_i \oplus T_i) \quad \subset \quad W_i \oplus V_i, \qquad i = 1, \ldots, N-1, \tag{247}$$
$$\widehat{\varphi}(t_{i+1} - t_i, X_i \oplus T_i) \quad \subset \quad X_{i+1} \oplus T_{i+1}, \qquad i = 1, \ldots, N-1. \tag{248}$$

Assume that

$$\alpha(X_1) < 0, \qquad \alpha(X_N) > 0 \tag{249}$$
$$\nabla\alpha(P_m(x)) \cdot P_m F(x) > 0, \qquad \textit{for all } x \in \bigcup_{i=1}^{N-1} W_i \oplus V_i, \tag{250}$$

then for any $n > M$ and any $x \in P_n(X_0 \oplus V_0)$ there exists a uniquely defined $t^n_\theta(x)$, such that $t_1 < t^n_\theta(x) < t_N$ and

$$\varphi^n(t^n_\theta(x), x) \in \theta. \tag{251}$$

Moreover, the map $t_{n,\theta} : P_n(X_0 \oplus T_0) \to \mathbb{R}$ is continuous. Consequently the map $G_{n,\theta} : P_n(X_0 \oplus T_0) \to P_n\theta$, given by $G_{n,\theta}(x) = \varphi^n(t_{n,\theta}(x), x)$ is well defined and continuous and

$$G_{n,\theta}(P_n(X_0 \oplus T_0)) \subset P_n\left(\theta \cap \bigcup_{i=1}^{N-1} W_i \oplus V_i\right) \tag{252}$$

Proof: From conditions (246–248) it follows that

$$\varphi^n(t_1, P_n(X_0 \oplus T_0)) \quad \subset \quad X_1 \oplus T_1, \tag{253}$$
$$\varphi^n([0, t_{i+1} - t_i], P_n(X_i \oplus T_i)) \quad \subset \quad W_i \oplus V_i, \qquad i = 1, \ldots, N-1, \tag{254}$$
$$\varphi^n(t_{i+1} - t_i, P_n(X_i \oplus T_i)) \quad \subset \quad X_{i+1} \oplus T_{i+1}, \qquad i = 1, \ldots, N-1 \tag{255}$$

To finish the proof observe that

$$\frac{d\alpha \circ \varphi^n(t, x)}{dt} = \nabla\alpha(P_m(\varphi^n(t, x))) \cdot P_m F(\varphi^n(t, x)) > 0$$

□

In the context of Theorem 36 the algorithm computing $\mathcal{G}_\theta$ will give

$$\mathcal{G}_\theta(X_0 \oplus T_0)) \subset \widehat{\mathcal{G}_\theta}(X_0 \oplus T_0) = \theta \cap \bigcup_{i=1}^{N-1} W_i \oplus V_i, \tag{256}$$

where by $\widehat{\mathcal{G}_\theta}$ we denote the bounds for $\mathcal{G}_\theta$ computed by our algorithm.

We also introduce the following

Definition 8 *Same assumptions as in Theorem 36. We define the* transition time $t_{\mathcal{G}_\theta}$ *by*

$$t_{\mathcal{G}_\theta} = (t_1, t_N)$$

The extension of Theorem 36 and Definition 8 to maps obtained as the transition between sections θ_1 and θ_2 is straightforward and is left to the reader.

An important issue in this context is the realization of the intersection appearing in formula (256). In our implementation θ is always linear. We have found the following approach to be the most efficient: we introduce a new coordinate system (an affine transformation) in X_m, such that if $(z_1, \dots, z_t)$ denote the new coordinates, then $\theta = \{z_1 = 0\}$. We will refer to these coordinates as *the section coordinates*. Moreover, if we are close to the section (in the section region), then we express all enclosures $W_i \oplus V_i$ in these coordinates. This means that formulas for rigorous estimates during the time steps from [Z2, Sec. 6.7] have to be evaluated directly in the section coordinates. In this situation the intersection (256) is just a projection onto $(z_2, \dots, z_t)$ of all sets $V_i \oplus V_i$.

This approach has also an another advantage. When one uses the Brouwer Theorem to prove the existence of a fixed point for the smooth map P (Poincaré map) one needs, B, a set homeomorphic to a ball such that $P(B) \subset B$. Usually the shape of B has to be carefully chosen. Assume that x_0 is a good approximation of this fixed point and $v_1, \dots, v_n$ are approximate eigenvectors of $dP(x_0)$. Then good candidate set for B is given by

$$B = \{x_0 + \sum_{i=1}^{n} a_i v_i \mid a_i \in [-\delta_i, \delta_i]\}, \tag{257}$$

for some $\delta_i > 0$ for $i = 1, \dots, n$.

It is then desirable to express the computed value of $P(B)$ directly using the linear coordinate frame induced by $v_1, \dots, v_n$.

10 Periodic orbits for the KS equation - topological theorems

In this section and the following ones we report on the computer assisted proofs of the existence of multiple periodic orbits for the KS equation (9) with periodic and odd boundary conditions (10). These orbits are obtained using the algorithm for rigorous integration of dissipative PDEs described in earlier sections.

In the Fourier domain the system (9–10) is given by (28) and has the reflectional symmetry R, which acts as follows

$$a_{2k} \to a_{2k}, \qquad a_{2k+1} \to -a_{2k+1}, \qquad k \in \mathbb{Z}_+. \tag{258}$$

We consider $\nu \in [0.02991, 0.128]$. For description of various periodic orbits and Silnikov connections, which indicates the existence of chaotic dynamics, for $\nu \in (0.111, 0.133)$ the reader is referred to [JJK] and the literature cited there. One should be aware that in [JJK] the KS equation is written in a different form and the parameter $\alpha = 4/\nu$ is used. In this parameter range we focus on the periodic orbits branch denoted in [JJK] by γ_{Hopf}. This branch, consisting of R-symmetric attracting periodic orbits, bifurcates off the positive bimodal fixed point branch for $\nu \approx 0.13254$. As ν decreases the branch γ_{Hopf} undergoes the period doubling bifurcation at $\nu \approx 0.1223$ losing its stability, which is inherited by an asymmetric periodic orbit. Along the branch γ_{Hopf} we were able to prove the existence of periodic orbits, both stable and unstable, respectively before and past the period doubling bifurcation. We were also able to prove the existence of an orbit on the non-symmetric branch bifurcating from γ_{Hopf}.

Other periodic orbits, whose existence is proved in this paper are unrelated to γ_{Hopf} and were chosen, with the objective to be in the chaotic region [CCP] or close to it.

To prove the existence of periodic orbits, which in numerical simulations appears to be attracting, a Brouwer-type theorem was used - see section 10.1. For the apparently unstable orbits we use the covering relations [ZGi] and the Miranda Theorem [Mi] - see section 10.2.

Throughout this section we use the notation for Poincaré maps introduced in Section 9.

10.1 Brouwer-type existence theorem

We fix parameters m and M, and we assume that these parameters are used for all self-consistent bounds appearing in the computations.

Theorem 37 *Consider (26), assume that conditions (3), (4) and (5) hold. Let $s_0 = d + p + 1$. Let θ be a section. Assume that there exists set $B \oplus T_0 \subset P_m(\theta) \oplus Y_m$, such that*

- *$B \oplus T_0$ are self-consistent bounds,*
- *B is homeomorphic to $(m-1)$-dimensional closed ball,*
- *$\mathcal{G}_{\theta\to\theta}(B \oplus T_0) \subset B \oplus T_0$*
- *there exists $a > 0$ and $b > 0$, such that $a < t < b$ for all $t \in t_{\mathcal{G}_{\theta\to\theta}}(B \oplus T_0)$*
- *for all $0 < t < t_{\mathcal{G}_{\theta\to\theta}}(B \oplus T_0)$ holds $\varphi(t, B \oplus T_0) \cap \theta = \emptyset$.*

Then there exists $t^* \in t_{\mathcal{G}_{\theta\to\theta}(B\oplus T_0)}$, $u : \mathbb{R} \to H$ *a solution of (26), such that* $u(0) \in B \oplus T_0$ *and* $u(t^*) = u(0)$ *(hence* $u(t)$ *is periodic) .*

Moreover, if all self-consistent bounds used in the computation of $\mathcal{G}_{\theta\to\theta}(B \oplus T_0)$ *were polynomial bounds with* $s \geq s_0$*, then* u *defines a classical solution of (1).*

Proof: Let $t_{n,\theta\to\theta}(x)$ be the Poincaré return time to section $P_n(\theta)$ for $x \in X_n$ for φ^n. From Theorem 36 we obtain for $n > M$

$$t_{n,\theta\to\theta}(P_n(B \oplus T_0)) \subset t_{\mathcal{G}_{\theta\to\theta}},$$
$$G_{n,\theta\to\theta}(P_n(B \oplus T_0)) \subset P_n(B \oplus T_0).$$

From the Brouwer Theorem applied to $G_{n,\theta\to\theta}$ on $P_n(B \oplus T_0)$ for each $n > M$ we obtain a periodic orbit for n-th Galerkin projection. Let us denote this orbit by u^n. We have $u^n : \mathbb{R} \to X_n$ and t^n, such that $u(0) = u(t^n) \in P_n(B \oplus T_0)$, $t^n \in t_{\mathcal{G}_{\theta\to\theta}}$. By picking up a subsequence we can assume that $t^n \to t^*$.

Let $t_{max} = right(t_{\mathcal{G}_{\theta\to\theta}})$. Observe that the set $\varphi([0, t_{max}], B \oplus T_0)$ is a finite sum of self-consistent bounds (one for each time step), hence from Lemma 4, it follows that we can pick up in (u^n) a convergent subsequence, which is converging to u^* a solution of (26). It is easy to see (see the proof Thm. 8 in [Z2]) that u^* is periodic of period t^*.

The assertion regarding the classical solution is an immediate consequence of Theorem 11. □

To obtain orbits with the reflectional symmetry, R, we will use the following modification of Theorem 37.

Theorem 38 *Consider (26), assume that conditions (3), (4) and (5) hold. Let* $s_0 = d + p + 1$.

Assume that there exists a symmetry $R : H \to H$*, such that* $R(\mathrm{dom}\,(F)) = \mathrm{dom}\,(F)$ *and* $F \circ R = R \circ F$ *on* $\mathrm{dom}\,(F)$.

Let θ *be a section. Assume that there exists set* $B \oplus T_0 \subset P_m(\theta) \oplus Y_m$*, such that*

- $B \oplus T_0$ *are self-consistent bounds,*
- B *is homeomorphic to* $(m-1)$*-dimensional closed ball,*
- $R \circ \mathcal{G}_{\theta\to R\theta}(B \oplus T_0) \subset B \oplus T_0$
- *there exists* $a > 0$ *and* $b > 0$*, such that* $a < t < b$ *for all* $t \in t_{\mathcal{G}_{\theta\to R\theta}}(B \oplus T_0)$
- *for all* $0 < t < t_{\mathcal{G}_{\theta\to R\theta}}(B \oplus T_0)$ *holds* $\varphi(t, B \oplus T_0) \cap R\theta = \emptyset$.

Then there exists $t^* \in t_{\mathcal{G}_{\theta\to R\theta}(B\oplus T_0)}$, $u : \mathbb{R} \to H$ *a solution of (26), such that* $u(0) \in B \oplus T_0$ *and* $u(t^*) = Ru(0)$*, hence* u *is* R*-symmetric periodic orbit.*

Moreover, if all self-consistent bounds used in the computation of $\mathcal{G}_{\theta\to\theta}(B \oplus T_0)$ *were polynomial bounds with* $s \geq s_0$*, then* u *defines a classical solution of (1).*

Proof: The same as the proof of Theorem 37 with obvious modifications. □

10.2 Covering relations and the Miranda Theorem

The notion of the covering relation was introduced in papers [Z0, Z1]. Here we follow the most recent and the most general version introduced in [ZGi] and the reader is referred there for proofs.

Definition 9 *A h-set, N, is the object consisting of the following data*

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

$$c_N(|N|) = \overline{B_{u(N)}}(0,1) \times \overline{B_{s(N)}}(0,1).$$

We set

$$\begin{aligned} N_c &= \overline{B_{u(N)}}(0,1) \times \overline{B_{s(N)}}(0,1), \\ N_c^- &= \partial \overline{B_{u(N)}}(0,1) \times \overline{B_{s(N)}}(0,1) \\ N_c^+ &= \overline{B_{u(N)}}(0,1) \times \partial \overline{B_{s(N)}}(0,1) \\ N^- &= c_N^{-1}(N_c^-), \quad N^+ = c_N^{-1}(N_c^+) \end{aligned}$$

Hence a h-set, N, is a product of two closed balls in some coordinate system. The numbers, $u(N)$ and $s(N)$, stand for the dimensions of nominally unstable and stable directions, respectively. The subscript c refers to the new coordinates given by homeomorphism c_N. Usually we will identify the h-set with its support. According to this convention $|N| = N$.

For the unstable periodic orbits for the KS-equation considered in this paper it is enough to consider h-sets with $u = 1$, so we have only one nominally expanding direction. This restriction enables us to give sufficient conditions for the existence of covering relations, which are easy to verify.

Definition 10 *Let N be a h-set, such that $u(N) = 1$. We set*

$$\begin{aligned} N_c^{le} &= \{-1\} \times \overline{B_{s(N)}(0,1)} \\ N_c^{re} &= \{1\} \times \overline{B_{s(N)}(0,1)} \\ S(N)_c^l &= (-\infty, -1) \times \mathbb{R}^{s(N)} \\ S(N)_c^r &= (1, \infty) \times \mathbb{R}^{s(N)}. \end{aligned}$$

We define

$$\begin{gathered} N^{le} = c_N^{-1}(N_c^{le}), \quad N^{re} = c_N^{-1}(N_c^{re}), \\ S(N)^l = c_N^{-1}(S(N)^l), \quad S(N)^r = c_N^{-1}(S(N)^r). \end{gathered}$$

We will call N^{le}, N^{re}, $S(N)^l$ and $S(N)^r$ the left edge, the right edge, the left side and right side of N, respectively.

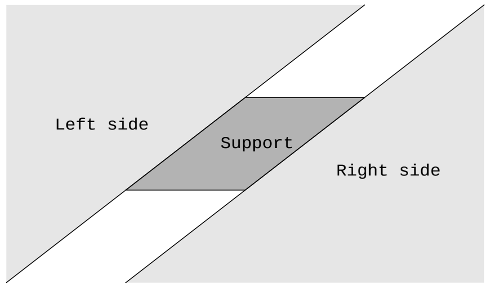

Figure 1: An example of an h-set on the plane.

It is easy to see that $N^- = N^{le} \cup N^{re}$.

We will not recall here the definition of covering relation in full generality, as we restrict ourselves to the case of $u = 1$, and will reformulate Theorem 16 from [ZGi] as the definition.

Definition 11 *Let N, M be two h-sets in $\mathbb{R}^n$, such that $u(N) = u(M) = 1$ and $s(N) = s(M) = s = n - 1$. Let $f : |N| \to \mathbb{R}^n$ be continuous.*

We say that N f-covers M with degree $w = \pm 1$, denoted by

$$N \overset{f,w}{\Longrightarrow} M,$$

if there exists $q_0 \in \overline{B}_s(0, 1)$, such that the following conditions are satisfied

$$f(c_N([-1, 1] \times \{q_0\})) \subset \operatorname{int}(S(M)^l \cup |M| \cup S(M)^r) \tag{259}$$

$$f(|N|) \cap M^+ = \emptyset, \tag{260}$$

and one of the following two conditions holds

$$f(N^{le}) \subset S(M)^l \quad \text{and} \quad f(N^{re}) \subset S(M)^r \tag{261}$$

$$f(N^{le}) \subset S(M)^r \quad \text{and} \quad f(N^{re}) \subset S(M)^l. \tag{262}$$

$w = 1$ if condition (261) is satisfied and $w = -1$ if condition (262) holds.

Quite often we will drop w in the symbol of the covering relation.

Remark 39 *A usual picture of a h-set on the plane with $u(N) = s(N) = 1$ is given in Figure 1. A typical picture illustrating covering relation on the plane with one 'unstable' direction is given on Figure 2.*

From Theorem 9 in [ZGi] we immediately obtain the following Miranda Theorem [Mi].

Theorem 40 *If $N \overset{f,w}{\Longrightarrow} N$, then there exists $x \in \operatorname{int} N$ such that $f(x) = x$.*

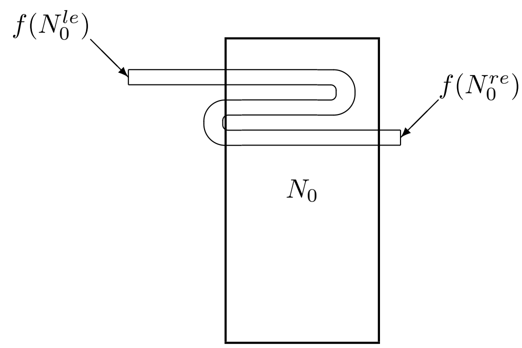

Figure 2: An example of an $f-$covering relation: $N_0 \stackrel{f,1}{\Longrightarrow} N_0$

In the context of computation of the Poincaré map for (15) we have the following easy

Lemma 41 *Assume that $B_1 \oplus T_1$, $B_2 \oplus T_2$ are self-consistent bounds and B_1, B_2 are $(m-1)$-dimensional h-sets. $M = M(T_1) = M(T_2)$ and $u(B_1) = u(B_2) = 1$*
Assume that

$$\mathcal{G}_{\theta_1 \to \theta_2}(B_1 \oplus T_1) \quad \subset \quad \left(\text{int}\, S(B_2)^l \cup B_2 \cup S(B_2)^r\right) \oplus \Pi_{k>m}(T_{2,k}^{-}, T_{2,k}^{+}) \tag{263}$$

and one of the following two conditions holds

$$P_m \mathcal{G}_{\theta_1 \to \theta_2}(B_1^{le} \oplus T_1) \subset S(B_2)^l \quad \text{and} \quad P_m \mathcal{G}_{\theta_1 \to \theta_2}(B_1^{re} \oplus T_1) \subset S(B_2)^r \tag{264}$$

$$P_m \mathcal{G}_{\theta_1 \to \theta_2}(B_1^{le} \oplus T_1) \subset S(B_2)^r \quad \text{and} \quad P_m \mathcal{G}_{\theta_1 \to \theta_2}(B_1^{re} \oplus T_1) \subset S(B_2)^l, \tag{265}$$

then for every $n > M$

$$B_1 \oplus P_n(T_1) \stackrel{G_{n,\theta_1 \to \theta_2}}{\Longrightarrow} B_2 \oplus P_n(T_2) \tag{266}$$

Definition 12 *We say that*

$$B_1 \oplus T_1 \stackrel{\mathcal{G}_{\theta_1 \to \theta_2}}{\Longrightarrow} B_2 \oplus T_2$$

if $B_1 \oplus T_1$, $B_2 \oplus T_2$ satisfy assumptions of Lemma 41 for the map $\mathcal{G}_{\theta_1 \to \theta_2}$.

Now we are ready to state

Theorem 42 *Consider (26), assume that conditions (3), (4) and (5) hold. Let $s_0 = d + p + 1$. Let θ be a section. Assume that there exists set $B \oplus T \subset P_m(\theta) \oplus Y_m$, such that*

- *$B \oplus T$ are self-consistent bounds,*
- *B is an $(m-1)$-dimensional h-set,*
- *$B \oplus T_0 \overset{\mathcal{G}_{\theta\to\theta}}{\Longrightarrow} B \oplus T_0$,*
- *there exists $a > 0$ and $b > 0$, such that $a < t < b$ for all $t \in t_{\mathcal{G}_{\theta\to\theta}}(B \oplus T_0)$,*
- *for all $0 < t < t_{\mathcal{G}_{\theta\to\theta}}(B \oplus T_0)$ holds $\varphi(t, B \oplus T_0) \cap \theta = \emptyset$.*

Then there exists $t^* \in t_{\mathcal{G}_{\theta\to\theta}(B\oplus T_0)}$, $u : \mathbb{R} \to H$ a solution of (26), such that $u(0) \in B \oplus T_0$ and $u(t^*) = u(0)$.

Moreover, if all self-consistent bounds used in the computation of $\mathcal{G}_{\theta\to\theta}(B \oplus T_0)$ were polynomial bounds with $s \geq s_0$, then u defines a classical periodic solution of (1).

Proof: For each n the existence of periodic orbit for n-the Galerkin projection follows directly from Lemma 41 and Theorem 40. Then we continue as in the proof of Theorem 37. □

To obtain orbits with the reflectional symmetry, R, we will use the following modification of Theorem 42.

Theorem 43 *Consider (26), assume that conditions (3), (4) and (5) hold. Let $s_0 = d + p + 1$.*

Assume that there exists a symmetry $R : H \to H$, such that $R(\mathrm{dom}\,(F)) = \mathrm{dom}\,(F)$ and $F \circ R = R \circ F$ on $\mathrm{dom}\,(F)$.

Let θ be a section. Assume that there exists set $B \oplus T \subset P_m(\theta) \oplus Y_m$, such that

- *$B \oplus T$ are self-consistent bounds,*
- *B is an $(m-1)$-dimensional h-set,*
- *$B \oplus T_0 \overset{R\circ\mathcal{G}_{\theta\to R\theta}}{\Longrightarrow} B \oplus T_0$*
- *there exists $a > 0$ and $b > 0$, such that $a < t < b$ for all $t \in t_{\mathcal{G}_{\theta\to R\theta}}(B \oplus T_0)$*
- *for all $0 < t < t_{\mathcal{G}_{\theta\to R\theta}}(B \oplus T_0)$ holds $\varphi(t, B \oplus T_0) \cap R\theta = \emptyset$.*

Then there exists $t^* \in t_{\mathcal{G}_{\theta\to R\theta}(B\oplus T_0)}$, $u : \mathbb{R} \to H$ a solution of (26), such that $u(0) \in B \oplus T_0$ and $u(t^*) = Ru(0)$, hence u is R-symmetric periodic orbit.

Moreover, if all self-consistent bounds used in the computation of $\mathcal{G}_{\theta\to\theta}(B \oplus T_0)$ were polynomial bounds with $s \geq s_0$, then u defines a classical solution of (1).

11 The outline of the computer assisted part of the proofs

In our computations we used the formulas for the Galerkin errors developed in [ZM, ZAKS]. The programm was written in c++, the *gnu* compiler was used. The computations have been performed using the interval arithmetics from package developed at the Jagiellonian University, Krakow, Poland [CAPD]. This interval package was based on the double precision arithmetic. The 3MHz Linux machine was used for the unstable orbits and 1.7GHz Windows machine for the stable ones.

The general scheme of the proof is the same as in [Z2]. Since we want to discuss both symmetric and non-symmetric orbits at the same time we set $R = \mathrm{id}$ for non-symmetric orbits. The proof consists of the following steps:

1. the initialization: setting up the parameters: dimensions m and M, finding an approximate periodic orbit, choosing the section θ_1, finding suitable coordinates on θ_1

2. the construction of initial tail T

3. the construction of a set $N_0 \oplus T_0$, such that for attracting orbits holds:

$$R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0 \oplus T_0) \subset N_0 \oplus T_0. \tag{267}$$

For unstable orbits we require that

$$\begin{aligned} R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0 \oplus T_0) \quad &\subset \\ \mathrm{int}\left(S(N_0)^l \cup N_0 \cup S(N_0)^r\right) \quad &\oplus \quad \Pi_{k>m}(T^-_{0,k}, T^+_{0,k}). \end{aligned} \tag{268}$$

This step includes the rigorous integration of (28).

4. for unstable orbits only, the verification that one of the following conditions is satisfied

$$\begin{aligned} P_m R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{le} \oplus T_0) \subset S(N_0)^l \quad \text{and} \\ P_m R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{re} \oplus T_0)) \subset S(N_0)^r \end{aligned} \tag{269}$$

$$\begin{aligned} P_m R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{le} \oplus T_0) \subset S(N_0)^r \quad \text{and} \\ P_m R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{re} \oplus T_0) \subset S(N_0)^l. \end{aligned} \tag{270}$$

This step includes rigorous integration of (28).

5. the conclusion of the proof, an application of one of Theorem 37, 38, 42 or 43

11.1 Part 1 - the initialization

We set the values of m, the time step h, the order of numerical method r and d - the number of coordinates in the diagonalization of DG as in Tables 3 and 9. We set $M = 3m$.

Starting with x_0, a good candidate for periodic orbit for m-th Galerkin projection of the KS equation, we construct the section θ_1 and *section coordinates* as in [Z2, Sec. 5.1].

θ_1 is a linear section through x_0 orthogonal to $P_m F(x_0)$. In our proofs we have found it most efficient to choose the point x_0 on the section $\sigma = \{a_1 - a_3 = 0, \quad (a_1 - a_3)' > 0\}$. We define the section θ_1 as a section perpendicular to $P_m F(x_0)$ at x_0, namely we set

$$\alpha(x) = (P_m F(x_0)|x) - (P_m F(x_0)|x_0), \quad \alpha' > 0. \tag{271}$$

The main difference in this part of the proof, when compared to [Z2], is in the choice of *section coordinates*, previously we had used *orthogonalized eigenvectors*, now we use *normalized eigenvectors*.

The *normalized eigenvectors coordinates* are constructed as follows: x_0 is an approximate fixed point for the map $g = RG_{m,\theta_1 \to R\theta_1} : \theta_1 \supset U \to \theta_1$. We introduce a new orthogonal coordinate frame such that x_0 is at the origin. The first coordinate direction is $P_m F(x_0)$. To obtain the other directions we remove from the canonical basis $\{e_i\}_{i=1,\ldots,m}$ the vector e_{i_0}, such that

$$|(P_m(F(x_0)|e_{i_0})| = \max_{i=1,\ldots,m} |(P_m(F(x_0)|e_i)|.$$

Next we apply the Gramm-Schmidt orthogonalization procedure to the system $P_m F(x_0), e_1, \ldots, e_{i_0-1}, e_{i_0+1}, \ldots, e_m$. The resulting vectors define new coordinate directions. Observe that in these coordinates the section is given by condition $x_1 = 0$. On section θ we use $(y_1, \ldots, y_{m-1}) = (x_2, \ldots, x_m)$ as *the temporary coordinates*.

Next we compute nonrigorously an approximate Jacobian matrix $Dg(x_0)$ using r-th order Taylor method and the time step h. The matrix $Dg(x_0) \in \mathbb{R}^{(m-1)\times(m-1)}$ is expressed using the temporary coordinates. From the matrix $Dg(x_0)$ we extract $\widetilde{D} \in \mathbb{R}^{d \times d}$ in an upper left corner, hence

$$\widetilde{D}_{ij} = Dg(x_0)_{ij}, \qquad \text{for } i, j = 1, \ldots, d. \tag{272}$$

Next we apply to $\widetilde{D}$ a diagonalization procedure based on the QR-decomposition algorithm [R] to obtain the approximate eigenvectors $v_1, v_2, \ldots, v_d$ corresponding to approximate eigenvalues $\lambda_1, \ldots, \lambda_d$, which are ordered as follows

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_d|.$$

The vectors v_i are normalized as follows. Let $|v|$ be the euclidian norm. If $\lambda_i \in \mathbb{R}$, then we require $|v_i| = 1$. If we have a pair of complex eigenvalues $\lambda_{j+1} = \overline{\lambda_j}$, then eigenvectors v_j and v_{j+1} are such that

$$Dg(x_0) \cdot (v_j + i(v_{j+1})) = \lambda_j \cdot (v_j + iv_{j+1})$$
$$\max(|v_j|, |v_{j+1}|) = 1.$$

Some of the diagonalization data for a symmetric periodic orbit for $\nu = 0.127$ can be found in Table 3 and 4 in [Z2].

Vectors $\{v_1, \ldots, v_d\}$ define a new coordinate system on $\mathbb{R}^d$ and together with coordinates $y_{d+1}, \ldots, y_{m-1}$ define the new coordinates on θ_1, such that our candidate for fixed point at the origin, i.e. $x_0 = 0$. We will denote these coordinates by c_i and we will call them *the section coordinates*.

11.2 Part 2 - the construction of initial tail T.

The initial tail was constructed using the routine described in [Z2, Section 5.2]. In this routine for all periodic orbits we used the following settings: the partition parameter $p = 50$, the stretching parameter $e = 1.25$ and $n_{iso} = 0$. It produced $W \oplus \Pi_{k>m}[a_k^-, a_k^+]$. In the present proof we used only $T_0 = \Pi_{k>m}[a_k^-, a_k^+]$.

In the proofs it will turns out that this initial tail will shrink by a huge factor, see Tables 5, 6, 11 and 12. Essentially the role of this step was to verify that for a given value of M finding the topologically self-consistent bounds is possible, i.e. there exists $s > s_0$ such that (238) holds.

11.3 Part 3 - the construction of $N_0 \oplus T_0$

Our goal is to construct a set $N \subset \theta_1$, such that

$$N_0 \oplus T_0 \quad \subset \quad \mathrm{dom}\,(\mathcal{G}_{\theta_1 \to R\theta_1}) \tag{273}$$

and either (267) or (268) holds, respectively for attracting and unstable orbits.

We constructed N_0 as a result of the following simple algorithm (the section coordinates are used to represent sets N_0 and N_1).

Algorithm

1. *Initialization.* We assign the values for δ, h the time step, the order of numerical method for the computation of $\mathcal{G}_{\theta_1 \to R\theta_1}$ and *iter*, the number of times the loop consisting of steps 2 and 3 described below should be executed, as in Tables 3 and 9.

We initialize $N_0 \oplus T_0$ as follows

$$N_0 \oplus T_0 = [-\delta, \delta]^{m-1} \oplus T, \tag{274}$$

where T is the initial tail obtained in Section 11.2.

2. *Computation of the Poincaré map.* We compute $N_1 \oplus T_1 = R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0 \oplus T_0)$. If the computation was terminated successfully then we go to step **3**, otherwise the execution of the algorithm is interrupted and *false* is returned.

3. If

$$N_0 \oplus T_0 \subset N_1 \oplus T_1, \tag{275}$$

then the execution of the algorithm is interrupted and *false* is returned.

In (275) then we continue as follows.

- For an attracting orbit we check whether

$$N_1 \oplus T_1 \subset N_0 \oplus T_0, \tag{276}$$

  then we set $N_0 \oplus T_0 = N_1 \oplus T_1$ we go to step **4**.

- In the case of an unstable orbit we check whether (268) holds, which in terms of $N_1 \oplus T_1$ can be expressed as follows

$$N_1 \oplus T_1 \subset \text{int}\ \left(S(N_0)^l \cup N_0 \cup S(N_0)^r\right) \oplus \Pi_{k>m}(T^-_{0,k}, T^+_{0,k}). \tag{277}$$

  If it holds, then we set $N_0 \oplus T_0 = N_1 \oplus T_1$ we go to step **4**.

If condition (276) or (277) is not satisfied then we set

$$N_0 = N_1 \cap N_0,$$

and if T_1 is not a subset of T_0, then we define a new value for T_0 as follows:

$$T_0 = \text{PolyBd}(T_0 \cup T_1), \tag{278}$$

$$T_0 = \text{inflate}(T_0, 2), \tag{279}$$

where by $\text{PolyBd}(T_0 \cup T_1)$ we denote the smallest polynomial bounds containing $T_0 \cup T_1$ and an inflation of polynomial bounds is understood componentwise.

Next, we skip back to step **2**.

4. *Further improvement.* We compute several times: $N_0 \oplus T_0 = R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0 \oplus T_0)$ and return **true**.

End of algorithm

Let us comment about (278–279). In principle since in each iteration set N_0 is smaller we should always obtain that $T_1 \subset T_0$, which is a natural consequence following fact:

$$\text{if} \quad A \subset B, \qquad \text{then} \quad f(A) \subset f(B). \tag{280}$$

But our algorithm does not fulfill the above condition - see Remark 19. Hence we increase the tail for the next iteration to make sure that $T_1 \subset T_0$, which was always the case.

11.4 Part 4 - the verification of conditions for image of the boundary for unstable orbits

We compute $R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{le} \oplus T_0)$ and $R \circ \mathcal{G}_{\theta_1 \to R\theta_1}(N_0^{re} \oplus T_0)$. In both cases we input the whole left and right edges as the initial conditions.

Table 1: Coordinates of u_0 - the approximation of the initial condition for the periodic orbit in Theorem 44.

a_1=2.012117e-01	a_2=1.289991e+00
a_3=2.012012e-01	a_4= -3.778725e-01
a_5=-4.230743e-02	a_6= 4.316285e-02
a_7=6.939942e-03	a_8=-4.156653e-03
a_9= -7.944539e-04	a_{10}=3.316281e-04

12 Example theorems and data from the proofs

In our program preforming the existence proof we input the value of ν or the interval of ν using the decimal format, but the computer uses the finite binary representation, hence in fact instead of computing with $\nu = 0.127$ we perform the computation with the binary representation of 0.127, which may differ from 0.127 by the number of order 10^{-16} in double precision arithmetic. In order to deal with this issue in the statement of the theorems we introduce the function $bin : \mathbb{R} \to \mathbb{R}$, which for $x \in \mathbb{R}$ assign its binary representation in the double precision arithmetic. We will use this function in the statements of the theorems about periodic orbits and the reader should also be aware that all real parameters of computation in tables, i.e. the variables ν, h, δ, included in this section are to be replaced by its binary representations.

12.1 Exemplary theorems about the attracting orbits

In this section we present two exemplary theorems about the existence of the apparently attracting periodic orbits for the KS equation. We use the phrase *apparently attracting* to highlight the fact that in numerical simulations it is clearly visible that the orbit is attracting, but we are not able to prove that.

First theorem is about the orbit with the reflectional symmetry.

Theorem 44 *Let $u_0(x) = \sum_{k=1}^{10} -2a_k \sin(kx)$, where a_k are given in Table 1.*

Then for any $\nu \in \mathrm{bin}\,(0.127) + [-\mathrm{bin}\,(10^{-7}), \mathrm{bin}\,(10^{-7})]$ there exists function $u^*(t,x)$, a classical solution of (9 - 10), such that

$$\|u_0 - u^*(0,\cdot)\|_{L_2} < 2.87e-03, \qquad \|u_0 - u^*(0,\cdot)\|_{C^0} < 2.03e-03 \tag{281}$$

and u^* is periodic with respect to t with period $T \in 2 \cdot [1.1217, 1.1226]$ and has the reflectional symmetry, R .

The theorem below present an example of the orbit without the reflectional symmetry.

Theorem 45 *Let $u_0(x) = \sum_{k=1}^{13} -2a_k \sin(kx)$, where a_k are given in Table 2.*

Table 2: Coordinates of u_0 - the approximation of the initial condition for the periodic orbit in Theorem 45.

a_1= 2.559309e-01	a_2= 1.096697e+00
a_3= 2.559306e-01	a_4= -3.079614e-01
a_5=-4.780285e-02	a_6= 3.002051e-02
a_7=7.352645e-03	a_8=-2.530194e-03
a_9= -7.561949e-04	a_{10}=1.624857e-04
a_{11}= 6.833016e-05	a_{12}= -8.789143e-06
a_{13}=-5.429530e-06	

Table 3: The parameters used in the proofs of the existence of apparently attracting periodic orbits for the KS equation. The *Sym* column tells whether the periodic orbit has the reflectional symmetry R , *iter* contains the number of iterates in the algorithm required to fulfill assumptions of Theorem 37 or 38, *comp. time* - the computation time for one iterate on 1.7 GHz Windows machine. For the meaning of other columns see text. In the first row the expression 0.127 ± 10^{-7} means that the whole interval $[0.127 - 10^{-7}, 0.127 + 10^{-7}]$ was plugged-in for ν.

ν	Sym	m	h	order	δ	iter	d	comp. time
0.127 ± 10^{-7}	Yes	10	8e-4	5	4e-4	3	9	32 sec
0.127	Yes	10	1e-3	4	2e-4	1	9	23 sec
0.125	Yes	11	1e-3	5	2e-4	3	8	39 sec
0.1215	No	13	4e-4	6	5e-5	3	10	459 sec
0.032	Yes	23	1.5e-4	5	8e-5	3	12	729 sec

Then for $\nu = \text{bin}\,(0.1215)$ *there exists a function* $u^*(t,x)$*, a classical solution of (9 - 10), such that*

$$\|u_0 - u^*(0,\cdot)\|_{L_2} < 1.01e-04, \qquad \|u_0 - u^*(0,\cdot)\|_{C^0} < 6.28e-05 \tag{282}$$

and u^* *is periodic with respect to* t *with period* $T \in [3.0744, 3.0745]$*. Moreover, this orbit does not have the reflectional symmetry* R*.*

Proof: The existence was obtained using Theorem 37. To prove that the orbit does not have the reflectional symmetry R, we checked that $R \circ G_{\theta_1 \to R\theta_1}(N_0 \oplus T_0) \cap (N_0 \oplus T_0) = \emptyset$. This check was performed by computer. □

In fact more is proved than it is stated in the above theorems, whereas for example any information about the tail is missing. Some partial information on it is contained in Tables 5 and 6 where the far tail described by $\frac{C_e}{k^{s_e}}$ and from the near tail we have data for $a^{\pm}_{m+1}$. More comprehensive data are contained in the

Table 4: Some data from the proof of the existence of apparently attracting periodic orbits for the KS equation. The columns L_2, H_1, C^0, C^1 contain the estimate on the distance in the corresponding norm, between the center of $P_m(N \oplus T_0)$ and the periodic orbit. In the first row the expression 0.127 ± 10^{-7} means that the whole interval $[0.127 - 10^{-7}, 0.127 + 10^{-7}]$ was plugged-in for ν.

ν	period	L_2	H_1	C^0	C^1
0.127 ± 10^{-7}	$2 \cdot [1.1217, 1.1226]$	2.87e-03	9.83e-03	2.07e-03	8.09e-03
0.127	$2 \cdot [1.1218, 1.1225]$	1.89e-03	8.16e-03	1.54e-03	7.36e-03
0.125	$2 \cdot [1.2382, 1.2386]$	5.79e-04	2.11e-03	4.26e-04	1.74e-03
0.1215	$[3.0744, 3.0745]$	1.01e-04	2.99e-04	6.28e-05	2.20e-04
0.032	$2 \cdot [0.4092, 0.4094]$	9.58e-04	6.70e-03	1.03e-03	8.01e-03

Table 5: Some data from the proof of the existence of apparently attracting periodic orbits for the KS equation. We compare the parameters describing the far tail for the initial T_0 (subscript i) and after the proof (subscript e). In the first row the expression 0.127 ± 10^{-7} means that the whole interval $[0.127 - 10^{-7}, 0.127 + 10^{-7}]$ was plugged-in for ν.

ν	C_i	s_i	$a^+_{i,M+1}$	C_e	s_e	$a^+_{e,M+1}$
0.127 ± 10^{-7}	1.3512e+11	12	1.715456e-07	7.542058e+02	12	9.575236e-16
0.127	1.351196e+11	12	1.715449e-07	7.349756e+02	12	9.331094e-16
0.125	1.012962e+11	12	4.244692e-08	7.600541e+02	13	9.367389e-18
0.1215	6.550411e+10	12	3.904349e-09	7.051674e-01	13	1.050781e-21
0.032	5.039215e+15	12	3.640713e-07	1.208202e+05	12	8.728969e-18

Table 6: Some data from the proof of the existence of apparently attracting periodic orbits for the KS equation. We compare the $a^{\pm}_{m+1}$ in the near tail for the initial T_0 (subscript i) and after the proof (subscript e). In the first row the expression 0.127 ± 10^{-7} means that the whole interval $[0.127 - 10^{-7}, 0.127 + 10^{-7}]$ was plugged-in for ν.

ν	$[a^-_{i,m+1}, a^+_{i,m+1}]$	$[a^-_{e,m+1}, a^+_{e,m+1}]$
0.127 ± 10^{-7}	-2.766770e-07 +2.347905e-04*[-1,1]	7.952441e-05 +1.331969e-06*[-1,1]
0.127	-2.766767e-07 +2.347903e-04*[-1,1]	7.958763e-05 +1.250929e-06*[-1,1]
0.125	2.662920e-05 +5.520568e-05*[-1,1]	-1.564897e-05 +1.812118e-07*[-1,1]
0.1215	1.537023e-06 +7.212759e-06*[-1,1]	3.478783e-07 +5.907734e-09*[-1,1]
0.032	-8.280459e-05 +1.286577e-04*[-1,1]	-1.141252e-05 +1.864317e-07*[-1,1]

companion files, where also the complete results of each iteration of $R \circ \mathcal{G}_{\theta_1 \to R\theta_1}$ are given.

We also have proved the existence of periodic orbits for $\nu = 0.127$, $\nu = 0.125$ and $\nu = 0.032$, see companion files for more details. There were several objectives behind these choices of the parameter values. For a given value of ν the main objective always was the minimalization of the computation time, having in mind that the eventual computer assisted proof of the existence of the symbolic dynamics or of an unstable periodic orbit, will require the partition of the domain into pieces for the computation of the Poincaré map.

- $\nu = 0.127$. This is a stable symmetric orbit on γ_{Hopf}. This case was done mainly for the comparison with [Z2]. The speed up factor (taking into account different speed of the machines used) is around 6.
- $\nu = 0.127 + [-10^{-7}, 10^{-7}]$. The same orbit as for $\nu = 0.127$. We tried to see how much we can extend the ν-interval of the existence of periodic orbit in single computation. We expected much larger interval of the diameter around 10^{-4}, but we could not do better than 10^{-7}. We do not have a clear idea why this happens, probably it has something to do with the wrapping effect.
- $\nu = 0.125$. This is a stable symmetric orbit on γ_{Hopf}. This case differs from $\nu = 0.127$ as follows: here the pair of leading eigenvalues is complex $\lambda_{1,2} \approx -0.051 \pm i * 0.0725$, there they were real. This case was the test for the choice of the good coordinates in case of complex eigenvalues.
- $\nu = 0.1215$. This is non-symmetric stable periodic orbit on branch bifurcating from γ_{Hopf}. Contrary to all other cases this one required the rigorous integration of the full Poincaré map. Using the previous approach from [Z2] this was impossible. It turns out also that the leading eigenvalue is complex and is approximately equal to $-0.041 \pm i * 0.312$.
- $\nu = 0.032$. This is the parameter value very close to the range $\nu \approx 0.0291$, where the chaotic dynamics was numerically observed in [CCP]. This computation required $m = 23$ and resulted in the longest computation time per iterate (see Table 3).

12.2 Two exemplary theorems about unstable orbits

Below we present two exemplary theorems on the existence of apparently unstable periodic orbits with and without the reflectional symmetry. We use the phrase *apparently unstable* to highlight the fact that in numerical simulations it is clearly visible that the orbit is unstable, but we are not able to prove that.

Theorem 46 *Let* $u_0(x) = \sum_{k=1}^{11} -2a_k \sin(kx)$*, where* a_k *are given in Table 7.*

Then for $\nu = \mathrm{bin}\,(0.1215)$ *there exists a function* $u^*(t,x)$*, a classical solution of (9 - 10), such that*

$$\|u_0 - u^*(0,\cdot)\|_{L_2} < 1.27 \cdot 10^{-3}, \quad \|u_0 - u^*(0,\cdot)\|_{C^0} < 8.26 \cdot 10^{-4} \qquad (283)$$

Table 7: Coordinates of u_0 - the approximation of the initial condition for the periodic orbit in Theorem 46.

$a_1 = 2.450027e-01$	$a_2 = 1.041500e+00$
$a_3 = 2.449985e-01$	$a_4 = -2.760754e-01$
$a_5 = -4.371320e-02$	$a_6 = 2.531380e-02$
$a_7 = 6.345919e-03$	$a_8 = -1.996779e-03$
$a_9 = -6.177148e-04$	$a_{10} = 1.184863e-04$
$a_{11} = 5.269771e-05$	

Table 8: Coordinates of u_0 - the approximation of the initial condition for the periodic orbit in Theorem 47.

$a_1 = 2.608268e-01$	$a_2 = 1.115112$
$a_3 = 2.608267e-01$	$a_4 = -3.208590e-01$
$a_5 = -4.953884e-02$	$a_6 = 3.199156e-02$
$a_7 = 7.802341e-03$	$a_8 = -2.766005e-03$
$a_9 = -8.196012e-04$	$a_{10} = 1.826998e-04$
$a_{11} = 7.575075e-05$	$a_{12} = -1.023722e-05$
$a_{13} = -6.157446e-06$	

and u^* is periodic with respect to t with period $T \in 2 \cdot [1.545879, 1.546770]$ and has the reflectional symmetry, R .

Proof: We check the assumption of Theorem 43. □

Theorem 47 *Let $u_0(x) = \sum_{k=1}^{13} -2a_k \sin(kx)$, where a_k are given in Table 8.*
Then for $\nu = \text{bin}\,(0.1212)$ there exists a function $u^*(t,x)$, a classical solution of (9 - 10), such that

$$\|u_0 - u^*(0,\cdot)\|_{L_2} < 2.6e-4, \qquad \|u_0 - u^*(0,\cdot)\|_{C^0} < 1.6e-4 \tag{284}$$

and u^* is periodic with respect to t with period $T \in [3.12211, 3.12219]$. Moreover, this orbit does not have the reflectional symmetry R.

Proof: We use Theorem 42 to obtain the existence of periodic orbit .

To prove that the orbit does not posses the reflectional symmetry R, we checked that $R \circ \hat{G}_{\theta_1 \to R\theta_1}(N_0 \oplus T_0) \cap (N_0 \oplus T_0) = \emptyset$. This check was performed with computer assistance. □

We have proved the existence of apparently unstable periodic orbits for several parameter values.

- $\nu = 0.1215$. This symmetric orbit belongs to γ_{Hopf}.

Table 9: The parameters used in the proofs of the existence of apparently unstable periodic orbits for the KS equation. The *Sym* column tells whether the periodic orbit has the reflectional symmetry R , *iter* contains the number of iterates in the algorithm required to fulfill assumptions of Theorem 42 or 43, *comp. time* - the computation time for one iterate on 3 GHz Linux machine. For the meaning of other columns see text.

ν	Sym	m	h	order	δ	iter	d	comp. time
0.02991	Yes	25	1e-4	5	5e-5	3	12	760 sec
0.1212 a	No	13	4e-4	5	2e-5	3	12	210 sec
0.1212 s	Yes	14	2e-4	5	5e-5	1	13	300 sec
0.1215	Yes	11	8e-4	5	1e-4	1	10	32 sec

Table 10: Some data from the proof of the existence of apparently unstable periodic orbits for the KS equation. The columns L_2, H_1, C^0, C^1 contain the estimate on the distance in the corresponding norm, between the center of $P_m(N \oplus T_0)$ and the periodic orbit.

ν	period	L_2	H_1	C^0	C^1
0.02991	2x[0.449024,0.449066]	5.7e-04	3.5e-03	5.9e-04	4.1e-03
0.1212 a	[3.122113,3.122183]	2.6e-04	7.6-04	1.6e-04	5.4e-04
0.1212 s	2x[1.581356,1.581916]	5.4e-04	1.7e-03	3.6e-04	1.2e-03
0.1215	2x[1.545879,1.546770]	1.3e-03	3.9e-03	8.3e-04	2.9e-03

Table 11: Some data from the proof of the existence of apparently unstable periodic orbits for the KS equation. We compare the parameters describing the far tail for the initial T_0 (subscript i) and after the proof (subscript e).

ν	C_i	s_i	$a^+_{i,M+1}$	C_e	s_e	$a^+_{e,M+1}$
0.02991	1.109228e+16	12	2.987155e-07	7.780416e+05	13	2.756932e-19
0.1212 a	7.029879e+10	12	4.190135e-09	8.634687e-01	13	1.286669e-21
0.1212 s	4.211406e+10	12	1.053915e-09	2.207509e-01	14	2.987748e-24
0.1215	1.314143e+11	12	5.506752e-08	1.505515e+02	13	1.855492e-18

Table 12: Some data from the proof of the existence of apparently unstable periodic orbits for the KS equation. We compare the $a^{\pm}_{m+1}$ in the near tail for the initial T_0 (subscript i) and after the proof (subscript e).

ν	$[a^-_{i,m+1}, a^+_{i,m+1}]$	$[a^-_{e,m+1}, a^+_{e,m+1}]$
0.02991	-3.747203e-05 +8.910417e-05*[-1,1]	-1.402969e-06 +4.101779e-08*[-1,1]
0.1212 a	1.624263e-06 +7.666259e-06*[-1,1]	4.292065e-07 +6.908816e-09*[-1,1]
0.1212 s	3.261036e-10 +2.360142e-06*[-1,1]	2.533699e-07 +1.544905e-09*[-1,1]
0.1215	4.528402e-06 +7.478529e-05*[-1,1]	-5.757639e-06 +2.451658e-07*[-1,1]

- $\nu = 0.1212$. For this parameter value we have proven the existence of three different periodic orbits. A pair on non-symmetric orbits (one is obtained from another by the application by R) and the symmetric one. The non-symmetric orbits apparently belong to the chaotic attractor, while the symmetric does not.
- $\nu = 0.02991$, this the parameter value considered in [CCP]. The orbit is on the chaotic attractor.

12.3 Final comments

Tables 5, 6, 11 and 12 show how much the tail has been improved during the computation and this is the basic reason why the method proposed here is so much better than the one from [Z2]. When we compare the initial tail with the tail at the end of the proof we see the improvement of several orders of magnitude (2-3 orders for diameter of a_{m+1} and much more for the far tail). This results and is also a consequence of the significant decrease of the Galerkin projection errors, which is due to the fact that we allow the tail to evolve and the Galerkin errors are computed locally, while in [Z2] the tail was fixed and the Galerkin errors were computed globally.

References

[CAPD] CAPD - Computer Assisted Proofs in Dynamics, a package for rigorous numeric, `http://capd.wsb-nlu.edu.pl`.

[CEES] P. Collet, J.-P. Eckmann, H. Epstein, J. Stubbe *Analyticity for the Kuramoto–Sivashinsky equation*, Physica D 67, (1993), 321–326

[CCP] F. Christiansen, P. Cvitanovic, V. Putkaradze, *Spatiotemporal chaos in terms of unstable recurrent patterns*, Nonlinearity 10, (1997), 55–70

[FNST] C. Foias, B. Nicolaenko, G. Sell, R. Temam *Inertial manifolds for the Kuramoto–Sivashinsky equation and an estimate of their lowest dimension.* J. Math. Pures Appl. 67, (1988), 197–226

[FT] C. Foias and R. Temam, Gevrey Class Regularity for the Solutions of the Navier-Stokes Equations, *Journal of Functional Analysis*, Vol. 87, No. 2, 1989, 359–369,

[HN] J. Hyman, B. Nicolaenko, *The Kuramoto–Sivashinsky equation; A bridge between PDEs and dynamical systems* , Physica 18D, (1986), 113–126

[JJK] M.E. Johnsohn, M. Jolly, I. Kevrekidis, *The oseberg transition: visualization of the global bifurcations for the Kuramoto-Sivashinsky equation*, Int. Journal of Bifurcation and Chaos 11, 1–18 (2001)

[JKT] M. Jolly, I. Kevrekidis, E. Titi, *Approximate inertial manifolds for the Kuramoto–Sivashinsky equation: analysis and computations*, Physica 44D, 38-60, (1990)

[K] H.-O. Kreiss, Fourier expansions of the solutions of the Navier-Stokes equations and their exponential decay rate, in *Analyse mathmatique et applications,* 245–262, Gauthier-Villars, Montrouge, 1988.

[KT] Y. Kuramoto, T. Tsuzuki, Persistent propagation of concentration waves in dissipative media far from thermal equilibrium, *Prog. Theor. Phys.*, **55**,(1976), 365

[Lo] R.J. Lohner, *Computation of Guaranteed Enclosures for the Solutions of Ordinary Initial and Boundary Value Problems*, in: Computational Ordinary Differential Equations, J.R. Cash, I. Gladwell Eds., Clarendon Press, Oxford, 1992.

[Lo1] R.J. Lohner, *Einschliessung der Lösung gewonhnlicher Anfangs- and Randwertaufgaben und Anwendungen* , Universität Karlsruhe (TH), these 1988

[Mi] C. Miranda, Un'osservazione su un teorema di Brouwer. (French) *Boll. Un. Mat. Ital.* (2) 3, (1940). 5–7

[Mo] R.E. Moore, *Methods and Applications of Interval Analysis,* SIAM, Philadelphia, 1979

[MZ] M. Mrozek, P. Zgliczyński, Set arithmetic and the enclosing problem in dynamics, *Annales Pol. Math.*, 2000, 237–259

[R] A. Ralston, *A First Course in Numerical Analysis*, 1965 McGraw-Hill, Inc

[Sa] Y. Sannikov, Mathematical Problems Related to Navier-Stokes Equations, *Thesis, Princeton University*, 2000

[Si] L. Sirovich, *Chaotic dynamics of coherent structures*, Physica D 37, (1989), 126–145

[S] G.I. Sivashinsky, Nonlinear analysis of hydrodynamical instability in laminar flames – 1. Derivation of basic equations, *Acta Astron.* 4 (1977), no. 11-12, 1177–1206

[W] W. Walter, *Differential and integral inequalities,* Springer-Verlag Berlin Heidelberg New York, 1970

[Z0] P. Zgliczyński, Fixed point index for iterations of maps, topological horseshoe and chaos, *Topological Methods in Nonlinear Analysis* **8**, (1996), 169–177

[Z1] P. Zgliczyński, Sharkovskii's Theorem for multidimensional perturbations of 1-dim maps, *Ergodic Theory and Dynamical Systems*, (1999), 19, 1655–1684

[ZM] P. Zgliczyński and K. Mischaikow, Rigorous Numerics for Partial Differential Equations: the Kuramoto-Sivashinsky equation. *Foundations of Computational Mathematics,* (2001) 1:255-288

[ZAKS] P. Zgliczyński, Attracting fixed points for the Kuramoto-Sivashinsky equation - a computer assisted proof, *SIAM Journal on Applied Dynamical Systems*, (2002) Volume 1, Number 2 pp. 215-235, *http://epubs.siam.org/sam-bin/dbq/article/40176*

[ZLo] P. Zgliczyński, C^1-Lohner algorithm, *Foundations of Computational Mathematics,* (2002) 2:429–465

[ZNS] P. Zgliczyński, Trapping regions and an ODE-type proof of an existence and uniqueness for Navier-Stokes equations with periodic boundary conditions on the plane, *Univ. Iag. Acta Math.*, 41 (2003) 89-113

[ZGal] P. Zgliczyński, On smooth dependence on initial conditions for dissipative PDEs, an ODE-type approach, *J. Diff. Eq.*, 195/2 (2003), 271–283

[Z2] P. Zgliczyński, Rigorous numerics for dissipative Partial Differential Equations II. Periodic orbit for the Kuramoto-Sivashinsky PDE - a computer assisted proof, *Foundations of Computational Mathematics,* 4 (2004), 157–185

[ZGi] P. Zgliczyński and M. Gidea, Covering relations for multidimensional dynamical systems, *J. Diff. Eq.* 202/1(2004), 33–58

[ZPLo] P. Zgliczyński, Lohner Algorithm for perturbations of ODEs and differential inclusions, *http://www.im.uj.edu.pl/˜zgliczyn*