Growth of Sobolev norms for the defocusing analytic NLS on $\mathbb{T}^2$

M. Guardia[*], E. Haus[†], M. Procesi[‡]

July 5, 2016

Abstract

We consider the completely resonant defocusing non–linear Schrödinger equation on the two dimensional torus with any analytic gauge invariant nonlinearity. Fix $s > 1$. We show the existence of solutions of this equation which achieve arbitrarily large growth of H^s Sobolev norms. We also give estimates for the time required to attain this growth.

Contents

[*]Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain, marcel.guardia@upc.edu

[†]Dipartimento di Matematica e Applicazioni ''R. Caccioppoli'', Università degli Studi di Napoli ''Federico II'', Via Cintia, Monte S. Angelo, 80126 Napoli, Italy, emanuele.haus@unina.it

[‡]Dipartimento di Matematica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy, mprocesi@mat.uniroma1.it

1 Introduction

Consider the completely resonant defocusing non–linear Schrödinger equation on the torus $\mathbb{T}^2 = (\mathbb{R}/2\pi\mathbb{Z})^2$ (NLS for brevity),

$$-\mathrm{i}u_t + \Delta u = 2d|u|^{2(d-1)}u + 2G'(|u|^2)u, \quad d \in \mathbb{N},\ d \geq 2 \tag{1.1}$$

where $G(y)$ is an analytic function (in the unit ball) with a zero of degree at least $d+1$ (the coefficients $2d$ and 2 are just to have simpler formulas later on).

It is well known [Bou93, BGT04] that equation (1.1) is globally well posed in time in H^s for $s \geq 1$, and defines an infinite dimensional Hamiltonian dynamical system with respect to the energy functional

$$H(u) = \int_{\mathbb{T}^2} \left(\frac{1}{2}|\nabla u|^2 + |u|^{2d} + G(|u|^2) \right) \frac{dx}{(2\pi)^2}.$$

It has also the following first integrals: the *momentum*

$$M(u) = \int_{\mathbb{T}^2} \bar{u} \nabla u \frac{dx}{(2\pi)^2}$$

and the *mass*

$$L(u) = \int_{\mathbb{T}^2} |u|^2 \frac{dx}{(2\pi)^2},$$

which is just the square of the L^2 norm.

The purpose of this paper is to study the problem of growth of Sobolev norms for the equation (1.1). That is, to obtain orbits whose s-Sobolev norm, $s > 1$, defined as usual as

$$\|u\|_{H^s} = \sum_{k \in \mathbb{Z}^k} \langle k \rangle^{2s} |u_k|^2, \quad \text{where } u(x) = \sum_{k \in \mathbb{Z}^2} u_k e^{\mathrm{i}kx} \quad \text{and } \langle k \rangle = \sqrt{1 + |k|^2},$$

grows by an arbitrarily large factor. Note that the H^1 norm is almost constant due to energy conservation.

The importance of growth of Sobolev norms stems from the fact that it implies transfer of energy from low to high modes as time grows, a phenomenon related to the so called weak turbulence.

In [Bou00], Bourgain posed the following question: are there solutions of the cubic nonlinear Schrödinger equation

$$-\mathrm{i}u_t + \Delta u = |u|^2 u$$

in $\mathbb{T}^2$ such that $\|u(t)\|_{H^s} \to +\infty$ as $t \to +\infty$?

This question has been recently positively answered for the cubic NLS on $\mathbb{R} \times \mathbb{T}^2$ in [HPTV13]. It is believed to be also true in the original setting $\mathbb{T}^2$ but the question remains open on any compact manifold.

In the past years there have been a set of results proving the existence of solutions of the *cubic NLS* with arbitrarily large *finite* growth. The first result, proven in [Kuk97], was for large data. Namely given a large constant $\mathcal{K} > 0$ there exists a solution whose initial Sobolev norm is large with respect to $\mathcal{K}$ which after certain time T attains a Sobolev norm satisfying $\|u(T)\|_{H^s} \geq \mathcal{K}\|u(0)\|_{H^s}$. In the context of small initial data, the breakthrough result was proved in [CKS+10] for the cubic NLS. The authors prove that given two constants $\mu \ll 1$ and $\mathcal{C} \gg 1$, there are orbits whose Sobolev norms grow from μ to $\mathcal{C}$ after certain time $T > 0$. Estimates for the time needed to attain such growth are given in [GK15]. Note that small initial Sobolev norm implies that the mass and the energy remain small for all times.

Growth of Sobolev norms has drawn considerable attention since the 90's not only for NLS on the two torus but also in more general settings and for other dispersive PDEs. Let us briefly review the literature on the subject. In [Bou96, Sta97, CDKS01, Bou04, Zho08, CW10, Soh11a, Soh12, Soh11b, CKO12], the authors obtain polynomial upper bounds for the growth of Sobolev norms.

Arbitrarily large finite growth was first proven in [Bou96], for the wave equation with a cubic nonlinearity but with a spectrally defined Laplacian. As we have already mentioned the same result has been obtained for the cubic NLS in [Kuk97, CKS+10, GK15]. The results in [CKS+10, GK15] have been generalized to the cubic NLS with a convolution potential in [Gua14] and the result in [CKS+10] has been generalized to the quintic NLS in [HP14]. Large finite growth of Sobolev norms has also been obtained in [GG12, Poc13] for certain nonlinear half-wave equations. In [CF12], the authors obtain orbits of the cubic NLS which undergo spreading of energy among the modes. Nevertheless, this spreading does not lead to growth of Sobolev norms. Similar phenomena were discussed in [GPT13, GT12, HT13].

Finally, the unbounded growth of Sobolev norms has been recently obtained for the Szegő equation by Gérard and collaborators following the work initiated in [GG10, Poc11, GG15]. Unbounded Sobolev growth, as it has been mentioned before, has been also proven for the cubic NLS in $\mathbb{R} \times \mathbb{T}^2$ in [HPTV13]. In [Han11, Han14] unbounded growth is shown in a pseudo partial differential equation which is a simplification of cubic NLS.

A dual point of view to instability is to construct quasi-periodic orbits. These are solutions which are global in time and whose Sobolev norms are approximately constant. Among the relevant literature we mention [Way90, Pös96, KP96, Bou98, BB13, EK10, GXY11, BB11, Wan10, PX13, BCP14, PP12]. Of particular interest are the recent results obtained through KAM theory which gives information on linear stability close to the quasi-periodic solutions. In particular the paper [PP14] proves the existence of both stable and unstable tori (of arbitrary finite dimension) for the cubic NLS. In principle such unstable tori could be used to construct orbits whose Sobolev norm grows, indeed in finite dimensional systems diffusive orbits are usually constructed by proving that the stable and unstable manifolds of a *chain of unstable tori* intersect. Usually however the intersection of *stable/unstable* manifolds is deduced by dimensional arguments, by constructing chains of co-dimension one tori. In the infinite dimensional case this would mean constructing almost-periodic orbits, which is an open problem except for very special cases such as integrable equations or equations with infinitely many external parameters (see for instance [CP95, Pös02, Bou05]).

In [CKS$^+$10], [GK15], [HP14] (and the present paper) this problem is avoided by taking advantage of the specific form of the equation. First one reduces to an approximate equation, i.e. the Hamitonian flow of the *first order Birkhoff normal form* $H_{\rm Res}$, see (2.5). Then for this dynamical system one proves directly the existence of chains of one dimensional unstable tori (periodic orbits) together with their heteroclinic connections. Next, one proves the existence of a *slider solution* which shadows the heteroclinic chain in a finite time. Finally, one proves the persistence of the slider solution for the full NLS by scaling arguments.

The fact that one may construct a heteroclinic chain for the Birkhoff normal form Hamiltonian (2.5) relies on the property that this Hamiltonian is *non-integrable* but has nonetheless many invariant subspaces on which the dynamics simplifies significantly. More precisely given a set $\mathcal{S} \subset \mathbb{Z}^2$ we define the subspace

$$U_{\mathcal{S}} := \{u \in L^2(\mathbb{T}^2) : \quad u(x) = \sum_{j \in \mathcal{S}} u_j e^{\mathrm{i} j \cdot x}\},$$

and consider the following definitions.

Definition 1.1 (Completeness)**.** *We say that a set $\mathcal{S} \subset \mathbb{Z}^2$ is* complete *if $U_{\mathcal{S}}$ is invariant for the dynamics of $H_{\rm Res}$.*

Definition 1.2 (Action preserving)**.** *A complete set $\mathcal{S} \subset \mathbb{Z}^2$ is said to be* action preserving *if all the actions $|u_j|^2$ with $j \in \mathcal{S}$ are constants of motion for the dynamics of $H_{\rm Res}$ restricted to $U_{\mathcal{S}}$.*

The conditions under which a given set $\mathcal{S}$ is complete or action preserving can be rephrased more explicitly by using the structure of $H_{\rm Res}$.

Definition 1.3 (Resonance)**.** *Given a $2d$-tuple $(j_1, \dots, j_{2d}) \in (\mathbb{Z}^2)^{2d}$ we say that it is a resonance of order d if*

$$\sum_{i=1}^{2d} (-1)^i j_i = 0, \qquad \sum_{i=1}^{2d} (-1)^i |j_i|^2 = 0.$$

Now $\mathcal{S}$ is complete if and only if for any $(2d-1)$-tuple $(j_1, \dots, j_{2d-1}) \in \mathcal{S}^{2d-1}$ there does not exist any $k \in \mathbb{Z}^2 \setminus \mathcal{S}$ such that $(j_1, \dots, j_{2d-1}, k)$ is a resonance. Similarly $\mathcal{S}$ is action preserving if all resonances $(j_1, \dots, j_{2d}) \in \mathcal{S}^{2d}$ are *trivial*, namely there exists a permutation such that $(j_1, \dots, j_d) = (j_{d+1}, \dots, j_{2d})$.

Now a good strategy is to look for a finite dimensional set $\mathcal{S}$ which is complete but not action preserving, where we can prove existence of diffusive orbits. A difficulty stems from the fact that *generic* choices of $\mathcal{S}$ are action preserving (see [PP12]). As a preliminary step one may study simple sets $\mathcal{S}$ where the dynamics is integrable and one can exhibit some growth of Sobolev norms. In particular one would like to produce a set which has two periodic orbits linked by a heteroclinic connection, since this is a natural building block for a heteroclinic chain. A natural choice is to fix a *simple* resonance $\mathcal{S} = \{j_1, \dots, j_{2k}\}$ of order k, namely a resonance which does not factorize as sum of two resonances of lower order. Clearly any set of this form produces non-trivial resonances (of order d) in $\mathcal{S}^{2d}$ for all $d \geq k$. Sets of this type have been studied for the quintic NLS, see [GT12, HP14]. The Hamiltonian H_{Res} restricted to such sets can be explicitly written (with a relatively heavy combinatorics) and one easily sees that there are in fact two periodic orbits, however we are not able to give a general statement about the existence of a heteroclinic connection. Some experiments seem to indicate that *no resonance of order* $k > 2$ produces a heteroclinic connection, while a single resonance of order two (i.e. one rectangle) produces a heteroclinic connection only for $d \leq 5$.

A crucial fact is the following: consider a *large* set $\mathcal{S}$ which is the union of $q \gg d$ rectangles and such that $\mathcal{S}$ does not contain any *simple* resonance apart from these rectangles. Then this system *always* has two periodic orbits linked by a heteroclinic connection. Indeed, after some symplectic reductions, it turns out that H_{Res} is a *small* perturbation of the one obtained for the cubic NLS restricted to a single rectangle. Note that this procedure works only for rectangles: if $\mathcal{S}$ is the union of $q \gg d$ resonances of order $k > 2$ and if these are the only simple resonances in $\mathcal{S}$, then after the same symplectic reductions one is left with a small perturbation of an action preserving system, having again two periodic orbits but no heteroclinic connection between them. While clearly this does not in any way constitute a proof, it gives some interesting negative evidence about the possibility of extending these results to the NLS on the circle.

1.1 Main results

The purpose of this paper is to generalize the results of [CKS+10] and [GK15] to the nonlinear Schrödinger equation (1.1) with any $d \geq 2$. The case $d = 3$ was treated in [HP14] where it is proven a result analogous to the one in [CKS+10].

This is the main result of our paper.

Theorem 1.4. *Let $d \geq 2$ and $s > 1$. There exists $c > 0$ with the following property: for any large $\mathcal{C} \gg 1$ and small $\mu \ll 1$, there exists a global solution $u(t) = u(t, \cdot)$ of (1.1) and a time T satisfying*

$$T \leq e^{\left(\frac{\mathcal{C}}{\mu}\right)^c}$$

such that

$$\|u(0)\|_{H^s} \leq \mu \quad \text{and} \quad \|u(T)\|_{H^s} \geq \mathcal{C}.$$

This result generalizes the results in [GK15] for the cubic NLS. In [GK15] the authors give two results. In the first result (Theorem 1 in [GK15]), they only measure growth of Sobolev norms and do not assume that the initial Soblev norm is small. Then, they obtain *polynomial* time estimates with respect to the growth. In the second result (Theorem 7 in [GK15]) they impose small initial Sobolev norm and large final Sobolev norm and obtain slower time estimates.

In the present setting we cannot get improved estimates as in Theorem 1 of [GK15] by assuming that only the L^2 norm of the initial datum is small. The reason is that the higher the degree of the nonlinearity the more interactions between modes exist. Certainly, more interactions should imply more paths to obtain growth of Sobolev norm and therefore similar or faster time estimates. Nevertheless, they also make the problem harder to handle. The proof of Theorem 1.4 follows the approach developed by [CKS+10] and analyzes very particular orbits which are essentially supported on a finite number of modes (see Section 2). Thus, one needs to keep track of the large number of interactions between modes so that the energy is not spread to a larger and larger number of modes as time evolves. This is more difficult for equation (1.1) with $d \geq 3$ than for the cubic NLS. To avoid this spreading, we have to choose slower orbits.

It is reasonable to expect that polynomial time estimates are still true for equation (1.1) with $d \geq 3$ but one needs some new ideas in the analysis of the finite set of modes on which the orbit attaining growth of Sobolev norms is supported. This is explained in Remark 2.7.

Theorem 1.4 is proven in Section 2. Then, Sections 3-9 contain the proofs of the partial results needed in Section 2.

Acknowledgements: The first author is partially supported by the Spanish MINECO-FEDER Grant MTM2012-31714 and the Catalan Grant 2014SGR504; the second author is supported by ERC under FP7-ERC Project 306414, HamPDEs and Programme STAR, funded by UniNA and Compagnia di San Paolo; the third author is supported by ERC under FP7-ERC Project 306414, HamPDEs.

2 Structure of the proof

2.1 Basic notations

We write the differential equation for the Fourier modes

$$u(t,x) := \sum_{k\in\mathbb{Z}^2} u_k(t) e^{\mathrm{i}(k,x)}, \tag{2.1}$$

associated to (1.1). It is of the form

$$\dot{u}_k = 2\mathrm{i}\partial_{\overline{u_k}} \mathcal{H}(u,\overline{u}). \tag{2.2}$$

Thus, is Hamiltonian with respect to the symplectic form $\Omega = \frac{\mathrm{i}}{2}\sum_{k\in\mathbb{Z}^2} du_k \wedge d\bar{u}_k$ and the Hamiltonian

$$\mathcal{H}(u,\overline{u}) = \mathcal{D}(u,\overline{u}) + \mathcal{G}(u,\overline{u}) \tag{2.3}$$

with

$$\begin{aligned}
\mathcal{D}(u,\overline{u}) &= \frac{1}{2}\sum_{k\in\mathbb{Z}^2} |k|^2 |u_k|^2 \\
\mathcal{G}(u,\overline{u}) &= \sum_{k_i\in\mathbb{Z}^2 : \sum_{i=1}^{2d}(-1)^i k_i = 0} u_{k_1}\bar{u}_{k_2}u_{k_3}\bar{u}_{k_4}\dots u_{k_{2d-1}}\bar{u}_{k_{2d}} \\
&\quad + \int_{\mathbb{T}^2} G(|u|^2)\frac{dx}{(2\pi)^2}.
\end{aligned}$$

We may write, for any $r\in\mathbb{N}$,

$$\begin{aligned}
[u]^{2r} &:= \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2} |u|^{2r} dx \\
&= \sum_{k_i\in\mathbb{Z}^2,\ \sum_i (-1)^i k_i = 0} u_{k_1}\bar{u}_{k_2}u_{k_3}\bar{u}_{k_4}\dots u_{k_{2r-1}}\bar{u}_{k_{2r}} \\
&= \sum_{\substack{\alpha,\beta\in(\mathbb{N})^{\mathbb{Z}^2}: \\ |\alpha|=|\beta|=r, \sum(\alpha_k-\beta_k)k=0}} \binom{r}{\alpha}\binom{r}{\beta} u^\alpha \bar{u}^\beta,
\end{aligned} \tag{2.4}$$

where $\alpha : k \mapsto \alpha_k \in \mathbb{N}$ and $u^\alpha = \prod_k u_k^{\alpha_k}$, same for β. With this notation one clearly has

$$\mathcal{G}(u,\overline{u}) = \sum_{r\geq d} c_r [u]^{2r}, \quad c_d = 1, \quad \sum_{r\geq d} |c_r| < \infty.$$

Remark 2.1. *Since by hypothesis the mass is preserved, we may perform the trivial phase shifts* $u_j \to e^{-2\mathrm{i}f(L)t}u_j$. *In this way the Hamiltonian becomes* $\mathcal{D}+\mathcal{G}_1$ *with* $\mathcal{G}_1 = \mathcal{G} - F(L)$ *where* F *is a primitive of* f.

In the course of the paper we will need the following definition

Definition 2.2. *Given a set of complex symplectic variables* $(z_k, \bar{z}_k)$ *with the symplectic form* $\frac{\mathrm{i}}{2}dz\wedge d\bar{z}$, *we say that a monomial is* action preserving *if it depends only on the* actions $|z_k|^2$. *This naturally defines a projection on the subspace of action preserving polynomials which we denote by* Π_I.

2.2 Birkhoff Normal Form

We perform one step of Birkhoff normal form to reduce the size of the non-resonant terms. We perform it in the ℓ^1 space, which is defined, as usual, by

$$\ell^1 = \left\{ u : \mathbb{Z}^2 \to \mathbb{C} : \|u\|_{\ell^1} = \sum_{k\in\mathbb{Z}^2} |u_k| < \infty \right\}.$$

Recall that ℓ^1 is a Banach algebra with respect to the convolution product. We consider a small ball centered at the origin,

$$B(\eta) = \left\{ u \in \ell^1 : \|u\|_{\ell^1} \leq \eta \right\}.$$

Theorem 2.3. *There exists $\eta > 0$ small enough such that there exists a symplectic change of coordinates $\Gamma : B(\eta) \to B(2\eta) \subset \ell^1$, $u = \Gamma(a)$, which takes the Hamiltonian $\mathcal{H}$ in (2.3) into its Birkhoff normal form up to order $2d$, that is,*

$$\mathcal{H} \circ \Gamma = \mathcal{D} + H_{\text{Res}} + \mathcal{R}, \tag{2.5}$$

where H_{Res} only contains resonant terms, namely

$$H_{\text{Res}} = \sum_{\substack{k_i \in \mathbb{Z}^2, \\ \sum_i (-1)^i k_i = 0 \\ \sum_i (-1)^i |k_i|^2 = 0}} a_{k_1} \bar{a}_{k_2} a_{k_3} \bar{a}_{k_4} \dots a_{k_{2d-1}} \bar{a}_{k_{2d}} = \sum_{\substack{\alpha, \beta \in (\mathbb{N})^{\mathbb{Z}^2} : |\alpha| = |\beta| = d \\ \sum_k (\alpha_k - \beta_k) k = 0, \ \sum_k (\alpha_k - \beta_k) |k|^2 = 0}} \binom{d}{\alpha} \binom{d}{\beta} a^\alpha \bar{a}^\beta . \tag{2.6}$$

The vector field $X_{\mathcal{R}}$, associated to the Hamiltonian $\mathcal{R}$, satisfies

$$\|X_{\mathcal{R}}\|_{\ell^1} \le \mathcal{O}\left(\|a\|_{\ell^1}^{2d+1} \right).$$

Moreover, the change of variables Γ satisfies

$$\|\Gamma - \text{Id}\|_{\ell^1} \le \mathcal{O}\left(\|a\|_{\ell^1}^{2d-1} \right).$$

The proof of this theorem follows the same lines as the proof of Theorem 2 in [GK15].

To study the Hamiltonian $\mathcal{H} \circ \Gamma$, we change to rotating coordinates to remove the quadratic part of the Hamiltonian. We take

$$a_k = r_k e^{\mathrm{i}|k|^2 t}. \tag{2.7}$$

Then, r satisfies the equation associated to the Hamiltonian

$$\mathcal{H}' = H_{\text{Res}} + \mathcal{R}', \tag{2.8}$$

where

$$\mathcal{R}' \left(\{r_k\}_{k \in \mathbb{Z}^2}, t \right) = \mathcal{R} \left(\{ r_k e^{\mathrm{i}|k|^2 t} \}_{k \in \mathbb{Z}^2} \right). \tag{2.9}$$

As a first step we study the dynamics of the truncated Hamiltonian H_{Res}. The associated equation is given by

$$-\mathrm{i}\dot{r} = \mathcal{E}(r) \tag{2.10}$$

where

$$\mathcal{E}_k(r) = 2d \sum_{\substack{k_i \in \mathbb{Z}^2, \\ \sum_{i=1}^{2d-1} (-1)^i k_i = k \\ \sum_{i=1}^{2d-1} (-1)^i |k_i|^2 = |k|^2}} r_{k_1} \bar{r}_{k_2} r_{k_3} \bar{r}_{k_4} \dots r_{k_{2d-1}}. \tag{2.11}$$

The Hamiltonian H_{Res} and the associated equation are scaling invariant with respect to

$$r^\varrho(t) = \varrho^{-1} r(\varrho^{-(2d-2)} t), \qquad \varrho \in \mathbb{R} \setminus \{0\}. \tag{2.12}$$

2.3 The reduction to the Toy Model

Following [CKS+10], we look for a finite set of modes which interact in a very particular and symmetric way. This set was constructed for the cubic case in [CKS+10] and in the quintic case in [HP14]. The higher the degree of the nonlinearity, the more complicated the interaction between the modes is. Here we follow the approach developed in [HP14]. We start by defining an *acceptable* frequency set as follows.

Definition 2.4. *Fix $N \gg 1$, $s > 1$. Then $\mathcal{S} \equiv \mathcal{S}(N) \subset \mathbb{Z}^2$ is acceptable if the following holds:*

1. *$\mathcal{S}$ is the disjoint union of N generations $\mathcal{S} = \cup_{i=1}^N \mathcal{S}_i$, each of them having cardinality $n := 2^{N-1}$*
2. *$\mathcal{S}$ satisfies the* norm explosion *property:*

$$\frac{\sum_{k \in \mathcal{S}_{N-2}} |k|^{2s}}{\sum_{k \in \mathcal{S}_3} |k|^{2s}} > 2^{(N-6)(s-1)}. \tag{2.13}$$

3. The N dimensional subspace

$$U_{\mathcal{S}} := \{r \in \mathbb{C}^{\mathbb{Z}^2} \, : \, r_k = 0 \; \forall k \notin \mathcal{S} \, , \quad r_l = r_j := b_i \;\; \forall i = 1, \ldots, N \, , \; \forall l, j \in \mathcal{S}_i\}. \tag{2.14}$$

is invariant under the flow of the Hamiltonian H_{Res} defined in (2.6).

4. The flow of H_{Res} restricted to $U_{\mathcal{S}}$ is Hamiltonian with respect to the symplectic form $\frac{i}{2}\sum_j db_j \wedge d\bar{b}_j$ with:

$$\begin{aligned} h_{\mathcal{S}}(b) &= d! n^{d-1} \left(\sum_{i=1}^{N} |b_i|^2 \right)^d + \\ &+ n^{d-2} d! d(d-1) \left\{ \left(\sum_{i=1}^{N} |b_i|^2 \right)^{d-2} \left[-\frac{1}{4} \sum_{i=1}^{N} |b_i|^4 + \sum_{i=1}^{N-1} \mathrm{Re}\,(b_i^2 \bar{b}_{i+1}^2) \right] + \frac{1}{n} \mathcal{P}\left(b, \bar{b}, \frac{1}{n} \right) \right\}. \end{aligned} \tag{2.15}$$

where $\mathcal{P}$ satisfies the following properties:

(a) $\mathcal{P}$ is a real coefficients polynomial in all its variables and it is homogeneous of degree $2d$ in $(b, \bar{b})$.

(b) $\mathcal{P}$ is real, namely $\mathcal{P}(b, \bar{b}, \frac{1}{n}) = \mathcal{P}(\bar{b}, b, \frac{1}{n})$.

(c) $\mathcal{P}$ is Gauge preserving, i.e. it Poisson commutes with $J = \sum_{i=1}^{N} |b_i|^2$.

(d) All the monomials in $\mathcal{P}$ are of even degree in each $(b_i, \bar{b}_i)$. This implies that for all $i = 1, \ldots, N$ the subspace $\{b_i = 0\}$ is invariant for the flow of $h_{\mathcal{S}}$.

(e) For $j = 1, \ldots, N-1$, the subspace

$$U_{\mathcal{S}}^j := \{b \in \mathbb{C}^N \, : \, b_i = 0 \, , \quad i \neq j, j+1\}$$

is invariant with respect to the flow of H_{Res}. Moreover, the pullback of this Hamiltonian into $U_{\mathcal{S}}^j$ is j-independent (up to an index translation) and, as a function of $(b_j, \bar{b}_j), (b_{j+1}, \bar{b}_{j+1})$, is symmetric with respect to the exchange $j \longleftrightarrow j+1$.

(f) Given $i \neq j$, consider the monomials in $\mathcal{P}$ which depend only on $(b_i, \bar{b}_i), (b_j, \bar{b}_j)$ and are exactly of degree two in $(b_i, \bar{b}_i)$. Then if $|i-j| \neq 1$ such monomials are action preserving *namely of the form $\chi_{ij} |b_j|^{2d-2} |b_i|^2$ (for some suitable coefficient χ_{ij}). Otherwise, if $|i-j| = 1$ then they are either of the form $\chi_{ij} |b_j|^{2d-2} |b_i|^2$ or of the form $\rho_{ij} |b_j|^{2d-4} \mathrm{Re}\,(b_i^2 \bar{b}_j^2)$. Moreover, $\chi_{ij} \equiv \chi$ is independent of i and j and $\rho_{i,i+1} \equiv \rho$ is independent of i.*

Theorem 2.5. *For each N sufficiently large there exist infinitely many acceptable sets $\mathcal{S}(N)$.*

This theorem is combinatoric in nature and proved in Section 3. The set of modes $\mathcal{S} \subset \mathbb{Z}^2$ is a generalization of the set of modes constructed in [CKS+10]. In [CKS+10] there is only one possible resonant interaction given by the conditions $|k_1|^2 + |k_3|^2 = |k_2|^2 + |k_4|^2$ and $k_1 + k_3 = k_2 + k_4$. Geometrically corresponds to four modes forming a rectangle in $\mathbb{Z}^2$. Now, since $d \geq 3$, there are more possibilities of resonant interactions. Nevertheless, as it is explained in [HP14], the interactions more suitable to achieve growth of Sobolev norms are still the ones which form rectangles. The interaction through more modes not built upon rectangles seem to be more stable. Therefore we consider analogs of the resonant interactions constructed in [CKS+10]. Nevertheless, in the case $d \geq 3$, the rectangles construction presents the following obvious difficulties. On the one hand, linear combinations of rectangular resonance conditions generate new unavoidable resonant relations which make the toy model more difficult to analyze. On the other hand, one needs to construct the set $\mathcal{S}$ such that all the resonant relations which are not constructed upon rectangles are avoided. The higher the degree d the larger amount of such new resonant relations.

Remark 2.6. *In order to obtain the time estimates we also need a quantitative version of Theorem 2.5, i.e. a bound on the size of the modes in $\mathcal{S}$. This is done in Lemma 3.19 and Corollary 3.21.*

Remark 2.7. *In [GK15] an extra condition to the set $\mathcal{S}$ is added. This condition, called by the authors* no spreading condition *says the following. Take $k \in \mathbb{Z}^2 \setminus \mathcal{S}$, then there exist at most four rectangles which have k as a vertex, two vertices in $\mathcal{S}$ and the fourth does not belong to $\mathcal{S}$. This implies that in the Hamiltonian H_{Res}, among the monomials which depend on a_k there are only four which depend also on two modes in $\mathcal{S}$. This implies that when one considers the full Hamiltonian* (2.5) *one has slow spreading of energy from the modes of $\mathcal{S}$ to the modes not belonging to $\mathcal{S}$ since essentially a_k only "receives energy" through these four monomials.*

This condition is not true in the present setting. Nevertheless we expect that a slightly weaker condition holds, replacing four rectangles by a fixed number of rectangles which depends on the degree d but not on the number of generations N. Unfortunately, such no spreading condition is considerably more involved since resonant interactions occur for all choices of $\mathcal{S}$ and classifying them seems a complicated task requiring some new ideas. Note that if this weak no spreading condition *were proved to be true we would obtain polynomial time estimates as in Theorem 1 of [GK15].*

The toy model (2.15) is gauge invariant by condition 4(c) of Definition 2.4. Thus, as explained in Remark 2.1, the first term $d!n^{d-1}\left(\sum_{i=1}^N |b_i|^2\right)^d$, which is a function of the mass $J = \sum_{i=1}^N |b_i|^2$, can be eliminated by a change of coordinates which does not modify the modulus of the components b_i's. Thus, we can consider the toy model with this term subtracted. Now, we rescale time to have a first order independent of n (n has been introduced in Definition 2.4). We consider the new time τ defined by

$$t = \frac{\tau}{n^{d-2} d! d(d-1)}. \tag{2.16}$$

We obtain then, the Hamiltonian

$$h(b) = \left(\sum_{i=1}^N |b_i|^2\right)^{d-2}\left[-\frac{1}{4}\sum_{i=1}^N |b_i|^4 + \sum_{i=1}^{N-1} \mathrm{Re}\,(b_i^2 \bar{b}_{i+1}^2)\right] + \frac{1}{n}\mathcal{P}\left(b, \bar{b}, \frac{1}{n}\right). \tag{2.17}$$

This toy model is a perturbation from the one obtained in [CKS$^+$10] and studied also in [GK15]. Nevertheless, note that it is a perturbation in terms of n^{-1}. Since we want to study the dynamics of this model for rather long time, classical perturbative methods do not apply. This implies that we need to redo and adapt the study done in [GK15] for the toy model for the cubic NLS.

The key point is that the properties of the toy model obtained in Theorem 2.5 (see Definition (2.4)) imply that the toy model (2.17) presents the same dynamical features as the toy model in [CKS$^+$10]. Even if (2.17) may have very complicated dynamics, it has certain invariant subspaces where the dynamics is easy to analyze. Fix mass $J = \sum_{i=1}^N |b_i|^2 = 1$. Then, by property 4(d) of Definition 2.4, the toy model (2.17) has the periodic orbits $\mathbb{T}_j = \{|b_j| = 1 \text{ and } b_i = 0 \text{ for all } i \neq j\}$. One can also consider the invariant subspaces $U_{\mathcal{S}}^j$ where two modes are non zero (see property 4(e) in Definition 2.4). This subspace contains $\mathbb{T}_j$ and $\mathbb{T}_{j+1}$. Furthermore, as it is explained in [GK15], in this subspace the Hamiltonian $h(b)$ becomes a two degrees of freedom Hamiltonian which is integrable since h itself and the mass J are first integrals in involution. Then, one can see that in $U_{\mathcal{S}}^j$ the unstable manifold of $\mathbb{T}_j$ coincides with the stable manifold of $\mathbb{T}_{j+1}$. Thus, we have a sequence of periodic orbits $\{\mathbb{T}_j\}_{j=1}^N$ which are connected by heteroclinic orbits. Orbits shadowing such structure provide growth of Sobolev norms.

Next theorem shows the existence of orbits with such dynamics.

Theorem 2.8. *Fix large $\gamma > 1$. Then, for any N large enough and $\delta = e^{-\gamma N}$ and for any acceptable set $\mathcal{S}$, there exists an orbit $b(\tau)$ of equations (2.17), constants $\mathbb{K} > 0$ and $\nu > 0$, independent of N and δ, and $T_0 > 0$ satisfying*

$$T_0 \leq \mathbb{K} N \ln(1/\delta),$$

such that

$$\begin{aligned} |b_3(0)| &> 1 - \delta^\nu \\ |b_j(0)| &< \delta^\nu \quad \text{for } j \neq 3 \end{aligned} \qquad \text{and} \qquad \begin{aligned} |b_{N-2}(T_0)| &> 1 - \delta^\nu \\ |b_j(T_0)| &< \delta^\nu \quad \text{for } j \neq N-2 \end{aligned}$$

Moreover, there exist times $\tau_j \in [0, T_0]$, $j = 3, \ldots, N-2$, satisfying $\tau_{j+1} - \tau_j \leq \mathbb{K} \ln(1/\delta)$, such that for any $\tau \in [\tau_j, \tau_{j+1}]$ and $k \neq j-1, j, j+1$,

$$|b_k(\tau)| \leq \delta^\nu.$$

This theorem is proved in Section 5.

Note that this theorem can be also stated in terms of the original time t, then the time needed to have such evolution is given by

$$T_0' \leq \mathbb{K}' n^{-(d-2)} N \ln(1/\delta)\,, \qquad \mathbb{K}' = d! d(d-1)\mathbb{K} \tag{2.18}$$

2.4 Proof of Theorem 1.4

Once we have analyzed certain orbits of the toy model, we show that they are a good first order for certain orbits of the original partial differential equation. We use the invariance rescaling (2.12). Consider

$$b^{\varrho}(t) = \varrho^{-1} b\left(\varrho^{-2(d-1)} n^{d-2} d! d(d-1) t\right)$$

where $b(\tau) = b(n^{d-2} d! d(d-1)t)$ is the trajectory given in Theorem 2.8. Then, the trajectory

$$\begin{aligned} \mathtt{r}_k^{\varrho}(t) &= b_j^{\varrho}(t) \text{ for any } k \in \mathcal{S}_j \\ \mathtt{r}_k^{\varrho}(t) &= 0 \text{ for } k \notin \mathcal{S} \end{aligned} \tag{2.19}$$

is a solution of the Hamiltonian H_{Res} given in (2.6). Due to the rescaling, now we study such trajectory in the time range $[0, T]$ with

$$T = \varrho^{2(d-1)} T_0', \tag{2.20}$$

where T_0' is the time introduced in (2.18).

We show that for large enough ϱ, (2.19) is the first order of a true solution of the nonlinear Schrödinger equation (1.1).

Theorem 2.9. *Fix $N \gg 1$ and $\varrho_0 = e^{C 2^{dN} N^2}$ for some large C. Let $\mathtt{r}_k^{\varrho}$ be (2.19), T be (2.20). Then, for all $\varrho \geq \varrho_0$ and for any solution $r(t)$ of (2.5) with initial condition $r(0) \in \ell^1$ satisfying $\|r(0) - \mathtt{r}^{\varrho}(0)\|_{\ell^1} \leq \varrho^{-5/2}$, one has that*

$$\|r(t) - \mathtt{r}^{\varrho}(t)\|_{\ell^1} \leq \varrho^{-3/2}$$

for $0 \leq t \leq T$.

This theorem is proven in Section 9.

To prove Theorem 1.4, it only remains to show that a well chosen trajectory $r(t)$ among those obtained in Theorem 2.9 undergoes growth of Sobolev norms. The proof of this fact is done analogously as in [GK15]. We reproduce here the reasoning for completeness.

Proof of Theorem 1.4. We start by choosing the trajectory which undergoes the growth of Sobolev norms. We consider a solution $u(t)$ of (2.2) satisfying $u(0) = \mathtt{r}^{\varrho}(0)$, where $\mathtt{r}^{\varrho}(t)$ has been defined in (2.19).

We define

$$\mathfrak{S}_j = \sum_{k \in \mathcal{S}_j} |k|^{2s} \quad \text{for } j = 1, \ldots, N.$$

We obtain a bound of the final Sobolev norm $\|u(T)\|_{H^s}$ in terms of $\mathfrak{S}_{N-2}$ as

$$\|u(T)\|_{H^s}^2 \geq \sum_{k \in \mathcal{S}_{N-2}} |k|^{2s} |u_k(T)|^2 \geq \mathfrak{S}_{N-2} \inf_{k \in \mathcal{S}_{N-2}} |u_k(T)|^2 .$$

Now we obtain a lower bound for $|u_k(T)|$, $k \in \mathcal{S}_{N-2}$. To this end, we need to show that we can apply Theorem 2.9 to the solution u. Using the change Γ obtained in Theorem 2.3 and the change of variables (2.7), we can write $u(t)$ as

$$u(t) = \Gamma\left(\left\{r_k(t) e^{\mathrm{i}|k|^2 t}\right\}\right),$$

where $r(t)$ is a solution of system (2.10). Note that, since $u(0) = \mathtt{r}^{\varrho}(0)$, by Theorem 2.3,

$$\begin{aligned} \|r(0) - \mathtt{r}^{\varrho}(0)\|_{\ell^1} &= \|r(0) - u(0)\|_{\ell^1} \\ &= \|r(0) - \Gamma(r)(0)\|_{\ell^1} \\ &\lesssim \|r(0)\|_{\ell^1}^3 . \end{aligned}$$

We compute the ℓ^1 norm of $u(0) = \mathtt{r}^{\varrho}(0)$. From the definition of $\mathtt{r}^{\varrho}(0)$ in (2.19) and Theorem 2.8, we know that $\|\mathtt{r}^{\varrho}(0)\|_{\ell^\infty} \leq \varrho^{-1}$. Moreover, $|\mathrm{supp}\, \mathtt{r}^{\varrho}(0)| = |\mathcal{S}| = N 2^{N-1}$. Thus,

$$\|u(0)\|_{\ell^1} = \|\mathtt{r}^{\varrho}(0)\|_{\ell^1} \leq \varrho^{-1} N 2^{N-1}.$$

Theorem 2.3 implies that Γ is invertible and that Γ^{-1} satisfies $\left\|\Gamma^{-1}(u) - u\right\|_{\ell^1} \leq \mathcal{O}\left(\|u\|_{\ell^1}^3\right)$. Therefore,

$$\|r(0)\|_{\ell^1} \leq \left\|\Gamma^{-1}(u(0))\right\|_{\ell^1} \lesssim \|u(0)\|_{\ell^1} \lesssim \varrho^{-1} N 2^{N-1},$$

which implies, using the definition of ϱ_0 and taking N large enough,

$$\|r(0)-\mathtt{r}^{\varrho}(0)\|_{\ell^1}\lesssim \varrho^{-3}N^3 2^{3(N-1)}\leq \varrho^{-5/2}.$$

This estimate implies that $r(0)$ satisfies the hypothesis of Theorem 2.9. We use this fact to estimate the Sobolev norm of $r(T)$. Using also Theorem 2.3, we split $|u_k(T)|$ as

$$\begin{aligned} |u_k(T)| &\geq |r_k(T)| - \left|\Gamma_k\left(\left\{r_k(T)e^{\mathrm{i}|k|^2T}\right\}\right)(T) - r_k(T)e^{\mathrm{i}|k|^2T}\right| \\ &\geq |\mathtt{r}_k^{\varrho}(T)| - |r_k(T)-\mathtt{r}_k^{\varrho}(T)| \\ &\quad - \left|\Gamma_k\left(\left\{r_k(T)e^{\mathrm{i}|k|^2T}\right\}\right)(T) - r_k(T)e^{\mathrm{i}|k|^2T}\right|. \end{aligned} \tag{2.21}$$

We need a lower bound for the first term of the right hand side and upper bounds for the second and third ones. Using the definition of $\mathtt{r}^{\varrho}$ in (2.19), the relation between T and T_0 established in (2.20) and the results in Theorem 2.8, we have that for $k\in\mathcal{S}_{N-2}$,

$$|\mathtt{r}_k^{\varrho}(T)| = \varrho^{-1}\left|b_{N-1}(T_0)\right| \geq \frac{3}{4}\varrho^{-1}.$$

For the second term in the right hand side of (2.21), it is enough to use Theorem 2.9 to obtain,

$$|r_k(T)-\mathtt{r}_k^{\varrho}(T)| \leq \left(\sum_{k\in\mathbb{Z}^2}|r_k(T)-\mathtt{r}_k^{\varrho}(T)|\right) \leq \frac{\varrho^{-1}}{8}.$$

For the lower bound of the third term, we use the bound for $\Gamma - \mathrm{Id}$ given in Theorem 2.3. Then,

$$\begin{aligned} &\left|\Gamma_k\left(\left\{r_k(T)e^{\mathrm{i}|k|^2T}\right\}\right)(T) - r_k(T)e^{\mathrm{i}|k|^2T}\right| \\ &\quad\leq \left\|\Gamma_k\left(\left\{r_k(T)e^{\mathrm{i}|k|^2T}\right\}\right)(T) - r_k(T)e^{\mathrm{i}|k|^2T}\right\|_{\ell^1} \leq \frac{\varrho^{-1}}{8}. \end{aligned}$$

Thus, we can conclude that

$$\|u(T)\|_{H^s}^2 \geq \frac{\varrho^{-2}}{4}\mathfrak{S}_{N-2}. \tag{2.22}$$

Now we prove that

$$\|u(0)\|_{H^s}^2 \lesssim \varrho^{-2}\mathfrak{S}_3. \tag{2.23}$$

Let us recall that $u(0)=\mathtt{r}^{\varrho}(0)$ and then $\operatorname{supp} u(0)=\mathcal{S}$. Therefore,

$$\|u(0)\|_{H^s}^2 = \sum_{k\in\mathcal{S}}|k|^{2s}\left|u_k(0)\right|^2 = \sum_{k\in\mathcal{S}}|k|^{2s}\left|\mathtt{r}_k^{\varrho}(0)\right|^2.$$

Then, recalling the definition of $\mathtt{r}^{\varrho}$ in (2.19) and the results in Theorem 2.8,

$$\begin{aligned} \sum_{k\in\mathcal{S}}|k|^{2s}\left|\mathtt{r}_k^{\varrho}(0)\right|^2 &\leq \varrho^{-2}\mathfrak{S}_3 + \varrho^{-2}\delta^{2\nu}\sum_{j\neq 3}\mathfrak{S}_j \\ &\leq \varrho^{-2}\mathfrak{S}_3\left(1+\delta^{2\nu}\sum_{j\neq 3}\frac{\mathfrak{S}_j}{\mathfrak{S}_3}\right). \end{aligned}$$

From Theorem 2.5 (see Lemma 3.19) we know that for $j\neq 3$, $\mathfrak{S}_j/\mathfrak{S}_3 \lesssim e^{sN}$. Therefore, to bound these terms we use the definition of δ from Theorem 2.8 taking $\gamma=\widetilde{\gamma}(s-1)$. Since $s-1>0$ is fixed, we can choose such $\widetilde{\gamma}\gg 1$. Then, we have that

$$\|u(0)\|_{H^s} = \sum_{k\in\mathcal{S}}|k|^{2s}\left|\mathtt{r}_k^{\varrho}(0)\right|^2 \sim \varrho^{-2}\mathfrak{S}_3.$$

Using inequalities (2.22) and (2.23), we have that

$$\frac{\|u(T)\|_{H^s}^2}{\|u(0)\|_{H^s}^2} \gtrsim \frac{\mathfrak{S}_{N-2}}{\mathfrak{S}_3},$$

and then, applying Theorem 2.5, we obtain

$$\frac{\|u(T)\|_{H^s}^2}{\|u(0)\|_{H^s}^2} \gtrsim 2^{(s-1)(N-6)} \geq \left(\frac{\mathcal{C}}{\mu}\right)^2 .$$

The last bound is obtained by taking N appropriately large.

Now we have to ensure that $\|u(0)\|_{H^s} \sim \varrho^{-2}\mathfrak{S}_3 \sim \mu$ so that the final norm satisfies $\|u(T)\|_{H^s} \gtrsim \mathcal{C}$. By Corollary 3.21 we know that the modes $k \in \mathcal{S}_3$ satisfy $|k| \sim e^{\eta 2^{8dN} N^{8d+1}}$ for some $\eta > 0$. Let us also note that if Definition 2.4 is satisfied by the set $\mathcal{S} \subset \mathbb{Z}^2$, it is also satisfied by the set

$$\mathcal{S}' = \{qk : k \in \mathcal{S}\}$$

for any given $q \in \mathbb{N}$. Call $u^{\mathcal{S}}$ and $u^{\mathcal{S}'}$ the orbits obtained by reducing into the toy model in the sets $\mathcal{S}$ and $\mathcal{S}'$ respectively. Then, $\|u^{\mathcal{S}'}(0)\|_{H^s} \sim q^s \|u^{\mathcal{S}}(0)\|_{H^s}$. Taking

$$\varrho = e^{\kappa 2^{8dN} N^{8d+1}}, \ \ \kappa \gg 1 \tag{2.24}$$

and adjusting the parameters q and κ, one can impose that $\mu/2 \leq \|u^{\mathcal{S}'}(0)\|_{H^s} \leq \mu$.

Finally, it only remains to estimate the diffusion time T. We have chosen N such that $2^{(s-1)(N-6)} \sim (\mathcal{C}/\mu)^2$. Then, using the definition of T in (2.20) and ϱ in (2.24) and choosing properly c, we obtain

$$|T| \lesssim \varrho^{2(d-1)} N^2 \leq e^{(\mathcal{C}/\mu)^c}$$

for some $c > 0$. This completes the proof of Theorem 1.4. □

3 Generation sets and combinatorics

We now discuss the combinatorial part of the paper, namely we prove Theorem 2.5. As explained in the introduction, we need to choose some frequency set $\mathcal{S} \subset \mathbb{Z}^2$ which is complete (see Definition 1.1) under the flow of (2.6) and not action preserving (namely a certain number of resonances occur). Moreover, we need enough resonances to be able to attain the desired energy transfer. Remember that a resonance is a relation of the form

$$\sum_{\ell=1}^{2d} (-1)^\ell k_\ell = 0 \qquad \sum_{\ell=1}^{2d} (-1)^\ell |k_\ell|^2 = 0 . \tag{3.1}$$

In the case of the cubic NLS (i.e. $d = 2$) the only non-trivial resonances are given by non-degenerate rectangles. For $d > 2$ we have many more options in the choice of the resonant sets. However, as discussed in the introduction, we are still going to use rectangles as building blocks for the construction of the set $\mathcal{S}$ with the same generation structure as in the cubic case, so that every point of $\mathcal{S}$ (except for the first and the last generation) belongs to exactly two rectangles. In the cubic case this means that each mode contributes only to two resonant monomials. Clearly this is false already in the quintic case, as one can see as follows. Assume that

$$k_1 - k_2 + k_3 - k_4 = 0 \qquad |k_1|^2 - |k_2|^2 + |k_3|^2 - |k_4|^2 = 0$$
$$k_4 - k_5 + k_6 - k_7 = 0 \qquad |k_4|^2 - |k_5|^2 + |k_6|^2 - |k_7|^2 = 0$$

are two rectangles with a common vertex. Then, these two relations give the resonant sextuple

$$k_1 - k_2 + k_3 - k_5 + k_6 - k_7 = 0 \qquad |k_1|^2 - |k_2|^2 + |k_3|^2 - |k_5|^2 + |k_6|^2 - |k_7|^2 = 0 .$$

As the degree of the NLS increases, the combinatorics of the resonances that appear as a consequence of the rectangle relations becomes more and more complicated, so we need some formal bookkeeping in order to handle this complex structure.

It will be convenient to work in the space $\mathbb{Z}^m$ with $m = N2^{N-1} = |\mathcal{S}|$. We denote by $\{\mathbf{e}_j\}_{j=1}^m$ the canonical basis of $\mathbb{Z}^m$ and divide the basis elements in N disjoint *abstract generations* $\mathcal{A}_i$ (each containing 2^{N-1} elements) using the convention that $\mathbf{e}_j \in \mathcal{A}_i$ if and only if $(i-1)2^{N-1} + 1 \leq j \leq i2^{N-1}$. Following [PP12] given $\mathcal{S} = \{\mathtt{v}_1, \ldots, \mathtt{v}_m\} \in (\mathbb{R}^2)^m$ we define the linear maps

$$\pi_{\mathcal{S}} : \mathbb{Z}^m \to \mathbb{R}^2 , \quad \pi_{\mathcal{S}}(\mathbf{e}_i) = \mathtt{v}_i , \quad \pi_{\mathcal{S}}^{(2)} : \mathbb{Z}^m \to \mathbb{R}, \quad \pi_{\mathcal{S}}^{(2)}(\mathbf{e}_i) = |\mathtt{v}_i|^2 \tag{3.2}$$

so that $\pi(\mathcal{A}_i) = \mathcal{S}_i$. By convention we denote $\cup_i \mathcal{A}_i = \mathcal{A}$.

Definition 3.1 (Abstract Family)**.** *An abstract family (of generation number $i \in \{1, \dots, N-1\}$) is a vector*

$$f = \mathbf{e}_{j_1} + \mathbf{e}_{j_2} - \mathbf{e}_{j_3} - \mathbf{e}_{j_4}\,, \quad \textit{with} \quad \mathbf{e}_{j_1}, \mathbf{e}_{j_2} \in \mathcal{A}_i\,, \quad \mathbf{e}_{j_3}, \mathbf{e}_{j_4} \in \mathcal{A}_{i+1},$$

and $j_1 \neq j_2$, $j_3 \neq j_4$

We say that $\mathbf{e}_{j_1}, \mathbf{e}_{j_2}$ are the parents of $\mathbf{e}_{j_3}, \mathbf{e}_{j_4}$ and that $\mathbf{e}_{j_3}, \mathbf{e}_{j_4}$ are the children of $\mathbf{e}_{j_1}, \mathbf{e}_{j_2}$. Moreover, we say that $\mathbf{e}_{j_1}$ is the spouse of $\mathbf{e}_{j_2}$ (and vice versa) and that $\mathbf{e}_{j_3}$ is the sibling of $\mathbf{e}_{j_4}$ (and vice versa).

Definition 3.2 (Genealogical tree)**.** *A set $\mathcal{F}$ of abstract families is called a* genealogical tree *provided that:*

1. *For all $i \in \{1, \dots, N-1\}$, every $\mathbf{e}_j \in \mathcal{A}_i$ is a member of one and only one abstract family of generation number i.*
2. *For all $i \in \{2, \dots, N\}$, every $\mathbf{e}_j \in \mathcal{A}_i$ is a member of one and only one abstract family of generation number $i-1$.*
3. *For all $i \in \{2, \dots, N-1\}$ and for all $\mathbf{e}_j \in \mathcal{A}_i$ we have that the* sibling *of $\mathbf{e}_j$ and the* spouse *of $\mathbf{e}_j$ do not coincide.*
4. *For all $i \in \{1, \dots, N\}$ and for all $\mathbf{e}_{j_1}, \mathbf{e}_{j_2} \in \mathcal{A}_i$, there exists a linear isomorphism*

$$g_{j_1 j_2} = g : \mathbb{Z}^m \to \mathbb{Z}^m$$

with the following properties:

   (a) basis elements are mapped to basis elements, namely for all $\mathbf{e}_{k_1} \in \mathcal{A}$ there exists $\mathbf{e}_{k_2} \in \mathcal{A}$ s.t. $g(\mathbf{e}_{k_1}) = \mathbf{e}_{k_2}$;

   (b) for all $\ell \in \{1, \dots, N\}$, one has $g(\mathcal{A}_\ell) = \mathcal{A}_\ell$;

   (c) $g(\mathbf{e}_{j_1}) = \mathbf{e}_{j_2}$;

   (d) $g(\mathcal{F}) = \mathcal{F}$.

Definition 3.3. *Given $\lambda = \sum_j \lambda_j \mathbf{e}_j \in \mathbb{R}^m$ we denote by* $\mathrm{Supp}(\lambda) := \{j = 1, \dots, m : \lambda_j \neq 0\}$ *its support.*

Remark 3.4. *If f_1, f_2 are abstract families of the same generation number, then their supports have empty intersection.*

Remark 3.5. *Item 4 in Definition 3.2 is a symmetry property of the genealogical tree that will be used in order to prove that the* intra-generational equality

$$r_l = r_j := b_i \ \ \forall i = 1, \dots, N\,,\ \forall l, j \in \mathcal{S}_i$$

(see the definition of $U_{\mathcal{S}}$ in (2.14)*) is preserved by the flow of the truncated resonant Hamiltonian* (2.6)*. In the case of cubic and quintic NLS, items 1, 2, 3 of Definition 3.2 are enough to ensure this; however, starting from degree 7, the structure of resonances gets more complicated and some additional symmetry property is needed.*

Definition 3.6 (Generation set)**.** *Consider a set $\mathcal{S} = \{\mathtt{v}_1, \dots, \mathtt{v}_m\} \in (\mathbb{R}^2)^m$ and the linear maps $\pi_{\mathcal{S}}$ and $\pi^{(2)}_{\mathcal{S}}$ defined in* (3.2)*. We say that the set $\mathcal{S}$ is an N generation set if*

$$\pi_{\mathcal{S}}(f) = 0\,, \quad \pi^{(2)}_{\mathcal{S}}(f) = 0\,, \quad \forall f \in \mathcal{F}\,. \tag{3.3}$$

We want to use the same genealogical tree as in [CKS+10]. Namely, we identify our abstract generations $\mathcal{A}_i$ with the Σ_i's defined in Section 4 of [CKS+10] and consider the genealogical tree $\mathcal{F}$ corresponding to the set of *combinatorial nuclear families connecting generations* Σ_i, Σ_{i+1} defined in [CKS+10].

More precisely, let

$$S_1 = \{1, \mathrm{i}\}\,, \qquad S_2 = \{0, \mathrm{i}+1\}\,,$$

then the 2^{N-1} elements of the k-th generation are identified with

$$(z_1, \dots, z_{k-1}, z_k, \dots, z_{N-1}) \in S_2^{k-1} \times S_1^{N-k} := \Sigma_k \tag{3.4}$$

The union of the Σ_k is denoted by Σ. We order Σ by identifying it with the ordered set $\mathcal{A}$ in such a way that each Σ_i is identified with $\mathcal{A}_i$. Now, for all $1 \leq k \leq N-1$, a combinatorial nuclear family of generation number k is a quadruple

$$(z_1, \dots, z_{k-2}, w, z_k, \dots, z_{N-1}) \quad w \in S_1 \cup S_2 \tag{3.5}$$

where all the z_j with $j \neq k-1$ are fixed with $z_j \in S_2$ if $1 \leq j \leq k-2$ and $z_j \in S_1$ if $k \leq j \leq N-1$. Each combinatorial nuclear family identifies a quadruple in $\mathcal{A}$ and hence an abstract family according to Definition 3.1. This fixes a set $\mathcal{F}$.

Lemma 3.7. *The set $\mathcal{F}$ defined by the combinatorial nuclear families is a genealogical tree according to Definition 3.2.*

Proof. Properties 1, 2, 3 follow directly from the definition (see also [CKS+10]). As for property 4 we proceed as follows. Let σ be the permutation of the elements of S_1, S_2 defined by

$$\sigma(0) = \mathrm{i}+1\,,\quad \sigma(\mathrm{i}+1) = 0\,,\quad \sigma(1) = \mathrm{i}\,,\quad \sigma(\mathrm{i}) = 1\,.$$

For all $\ell = 1, \ldots, N-1$ we define the map $\mathfrak{f}_\ell : \Sigma \to \Sigma$ as

$$(z_1, \ldots, z_{\ell-1}, z_\ell, z_{\ell+1}, \ldots, z_{N-1}) \mapsto (z_1, \ldots, z_{\ell-1}, \sigma(z_\ell), z_{\ell+1}, \ldots, z_{N-1}).$$

All the $\mathfrak{f}_\ell$ preserve the sets Σ_i and the combinatorial nuclear families. Moreover they commute with each other. Given any two elements $\mathbf{e}_{j_1}, \mathbf{e}_{j_2} \in \mathcal{A}_i$ we consider the corresponding two elements $\mathtt{e}_{j_1}, \mathtt{e}_{j_2}$ in Σ_i. Then there exists a map $\mathfrak{g}_{j_1,j_2}$, composition of a finite number of $\mathfrak{f}_\ell$, which maps $\mathtt{e}_{j_1}$ to $\mathtt{e}_{j_2}$. By construction these maps preserves the Σ_i and the combinatorial nuclear families. We pull back $\mathfrak{g}_{j_1 j_2}$ to $\mathcal{A}$ and then extend it to $\mathbb{Z}^m$ by linearity. This is the required map $g_{j_1 j_2}$. ☐

3.1 Some geometry

Now we want to prove the existence of sets $\mathcal{S} = \{\mathtt{v}_1, \ldots \mathtt{v}_m\}$ which satisfy all the properties of Definition 2.4. We take advantage of the abstract combinatorial setting which we have defined in the previous section and we use the maps $\pi_{\mathcal{S}}$ and $\pi^{(2)}_{\mathcal{S}}$ given in (3.2).

It is helpful to think of $\mathcal{S}$ as a vector in $\mathbb{R}^{2m}$. Then we claim that the resonance relations (3.3) define a real algebraic manifold $\mathcal{M}$ as

$$\mathcal{M} := \left\{ \mathcal{S} \in \mathbb{R}^{2m} \;:\; \forall f \in \mathcal{F} \quad \pi_{\mathcal{S}}(f) = 0,\; \pi^{(2)}_{\mathcal{S}}(f) = 0 \right\}\,. \tag{3.6}$$

Indeed by imposing the linear equations we reduce to a $(N+1)2^{N-1}$ subspace which we denote by

$$\mathcal{L} := \left\{ \mathcal{S} \in \mathbb{R}^{2m} \;:\; \forall f \in \mathcal{F} \quad \pi_{\mathcal{S}}(f) = 0 \right\}\,. \tag{3.7}$$

Then by imposing the quadratic constraints we further reduce the dimension. We can proceed by induction. Let us suppose that we have enforced all the linear and quadratic constraints for the first i generations (i.e for all abstract families f of generation number $\leq i-1$) and for the first $h < 2^{N-2}$ families of generation number i. Then given a *parental couple*, (which for simplicity of notation we denote) $\mathtt{v}_1, \mathtt{v}_2$ in the i-th generation we have to fix the corresponding children which we denote by $\mathtt{w}_1, \mathtt{w}_2$ in the generation $i+1$. We have the two equations

$$\mathtt{w}_2 = -\mathtt{w}_1 + \mathtt{v}_1 + \mathtt{v}_2\,,\quad (\mathtt{v}_1 - \mathtt{w}_1, \mathtt{v}_2 - \mathtt{w}_1) = 0.$$

so that $\mathtt{w}_2$ is fixed in terms of $\mathtt{w}_1$ which in turn lies on the circle with diameter the segment joining $\mathtt{v}_1, \mathtt{v}_2$. Hence provided that $\mathtt{v}_1 \neq \mathtt{v}_2$ both children $\mathtt{w}_2 \neq \mathtt{w}_1$ are fixed by one angle. Finally (by excluding at most a finite number of points) we can ensure that $\mathtt{w}_1, \mathtt{w}_2$ do not coincide with any of the previously fixed tangential sites.

In conclusion we have $2 \cdot 2^{N-1}$ degrees of freedom from the first generation and then 2^{N-2} angles for each subsequent generation, hence a manifold of dimension $(N+3)2^{N-2}$ with singularities all contained in the proper submanifold $\mathcal{B} := \cup_{i \neq j}\{\mathtt{v}_i - \mathtt{v}_j = 0\} \cap \mathcal{M}$. Moreover $\mathbb{Q}^{2m} \cap \mathcal{M}$ is dense on $\mathcal{M}$. Now a resonance as in formula (3.1) defines a codimension 3 algebraic variety in $\mathbb{R}^{2m}$ as follows.

Definition 3.8. *Given $k \in \mathbb{N}$, we denote by $\mathcal{R}_k$ the set of vectors $\lambda \in \mathbb{Z}^m$ with*

$$\sum_i \lambda_i = 0\,,\quad \sum_i |\lambda_i| \leq 2k\,.$$

We say that $\lambda \in \mathcal{R}_d$ is resonant within $\mathcal{S}$ if

$$\pi_{\mathcal{S}}(\lambda) = 0\,,\quad \pi^{(2)}_{\mathcal{S}}(\lambda) = 0\,.$$

Note that any resonance within $\mathcal{S}$ given by equation (3.1) can be written in this form. Some resonances cannot be avoided: they are the ones whose associated algebraic variety contains $\mathcal{M}$.

Definition 3.9. *We denote by*

$$\langle \mathcal{F} \rangle = \mathrm{Span}(f \in \mathcal{F}; \mathbb{Q}) \cap \mathbb{Z}^m \ .$$

Remark 3.10. *Since both $\pi_{\mathcal{S}}$ and $\pi_{\mathcal{S}}^{(2)}$ are linear maps then*

$$\pi_{\mathcal{S}}(\lambda) = 0\,, \quad \pi_{\mathcal{S}}^{(2)}(\lambda) = 0\,, \quad \forall \lambda \in \langle \mathcal{F} \rangle.$$

Note that all the abstract families are in $\mathcal{R}_d$ for all $d \geq 2$. Moreover, all the elements of $\langle \mathcal{F} \rangle \cap \mathcal{R}_d$ correspond to resonances that cannot be avoided, since they are obtained as linear combination of the relations defining family rectangles. The next lemma gives properties of the unavoidable resonances.

Lemma 3.11. *The following statements hold:*

(i) a genealogical tree $\mathcal{F}$ is a set of linearly independent abstract families;

(ii) all nonzero vectors $\lambda \in \langle \mathcal{F} \rangle$ have support $|\mathrm{Supp}(\lambda)| \geq 4$ and $|\mathrm{Supp}(\lambda)| = 4$ if and only if λ is a multiple of an abstract family.

(iii) we have that $\langle \mathcal{F} \rangle = \mathrm{Span}(f \in \mathcal{F}; \mathbb{Z})$;

(iv) assume that $\lambda \in \langle \mathcal{F} \rangle$ is such that all the elements of $\mathrm{Supp}(\lambda)$ except at most two belong to the same generation: then λ is a multiple of an abstract family;

(v) let $\lambda \in \langle \mathcal{F} \rangle$ and let $v = \sum_{j \in \mathcal{A}} \lambda_j \mathbf{e}_j$ be its decomposition on the basis $\{\mathbf{e}_j\}_j$. Then for all $1 \leq i \leq N$ one has

$$\sum_{j \in \mathcal{A}_i} |\lambda_j| \in 2\mathbb{N}\,,$$

namely the ℓ^1-norm of the projection of λ on the i-th generation is an even number.

The proof of this lemma is delayed to Section 3.3.

Now we prove that all the other resonances can be avoided.

Definition 3.12 (Non-degeneracy)**.** *We say that a generation set $\mathcal{S}$ is* non-degenerate *if*

(i) For all $\lambda \in \mathcal{R}_{2d} \setminus \langle \mathcal{F} \rangle$ one has $\pi_{\mathcal{S}}(\lambda) \neq 0$.

(ii) For all $\mu \in \mathbb{Z}^m$ such that $\sum_i \mu_i = 1$ and $\sum_i |\mu_i| \leq 2d - 1$ one has that either

$$K_{\mathcal{S}}(\mu) := |\pi_{\mathcal{S}}(\mu)|^2 - \pi_{\mathcal{S}}^{(2)}(\mu) \neq 0$$

or there exists $1 \leq j \leq m$ such that $\mu - \mathbf{e}_j \in \langle \mathcal{F} \rangle$.

Remark 3.13. *Note that the non-degeneracy condition (ii) implies the completeness of $\mathcal{S}$; if $d = 2$ (i.e. the cubic NLS) actually this is all that is needed (and indeed only this condition is imposed in [CKS+10]). Condition (i) is a faithfulness condition (namely it ensures that all $\lambda \in \mathcal{R}_{2d} \setminus \langle \mathcal{F} \rangle$ are not resonant within $\mathcal{S}$) and could probably be weakened.*

The fact that there exist generation sets $\mathcal{S}$ has been proved in [CKS+10] together with a weaker non-degeneracy condition in the case $d = 2$. In this section we prove that one can construct non-degenerate generation sets for the NLS of any degree.

Theorem 3.14. *Consider the manifold $\mathcal{M}$ introduced in (3.6). Then, there exists a proper algebraic manifold $\mathcal{D} \subset \mathcal{M}$ (of codimension one in $\mathcal{M}$) such that all $\mathcal{S} \in \mathcal{M} \setminus \mathcal{D}$ are non-degenerate generation sets.*

The proof of this Theorem is delayed to Section 3.4. Now we are ready to prove Theorem 2.5.

3.2 Proof of Theorem 2.5

We need to prove the existence of a set $\mathcal{S} \subset \mathbb{Z}^2$, $|\mathcal{S}| = m = N2^{N-1}$, satisfying all the properties in Definition 2.4. It is convenient to consider $\mathcal{S}$ as a point in $\mathbb{Z}^{2m}$.

Lemma 3.15. *Consider $\mathcal{S}$ belonging to $(\mathcal{M} \setminus \mathcal{D}) \cap \mathbb{Z}^{2m}$, where $\mathcal{M}$ is the variety defined in (3.6) and $\mathcal{D}$ is the subvariety given by Theorem 3.14. Then, $\mathcal{S}$ satisfies condition 1 in Definition 2.4.*

Proof. Condition 1 is equivalent to saying $\mathtt{v}_i - \mathtt{v}_j \neq 0$ for all $i \neq j$. This can be also written as $\pi_{\mathcal{S}}(\mathbf{e}_i - \mathbf{e}_j) \neq 0$. Item 2 in Lemma 3.11 implies that $\mathbf{e}_i - \mathbf{e}_j \notin \langle \mathcal{F} \rangle$. Moreover, $\mathbf{e}_i - \mathbf{e}_j \in \mathcal{R}_{2d}$. Then, condition 1 of Definition 2.4 follows from item (i) in Definition 3.12. □

Lemma 3.16. *Consider $\mathcal{S}$ belonging to $(\mathcal{M}\setminus\mathcal{D})\cap\mathbb{Z}^{2m}$, where $\mathcal{M}$ is the variety defined in* (3.6) *and $\mathcal{D}$ is the subvariety given by Theorem 3.14. Then, $\mathcal{S}$ satisfies condition 3 in Definition 2.4.*

Proof. We prove the fulfillment of condition 3 in two steps. First, we show that the larger subspace

$$V_{\mathcal{S}} = \left\{ r \in \mathbb{C}^{\mathbb{Z}^2} : r_k = 0, \forall k \notin \mathcal{S} \right\}$$

is invariant. This fact follows from item (ii) of Definition 3.12. Indeed, consider a resonance as in (3.1) where $k = k_1 \notin \mathcal{S}$ and $k_2, \ldots, k_{2d} \in \mathcal{S}$. Then, by construction there exists $\mu \in \mathbb{Z}^m$, $|\mu| \leq 2d-1$ and $\sum \mu_i = 1$ such that

$$k = \sum \mu_i \mathtt{v}_i, \qquad |k|^2 = \sum \mu_i |\mathtt{v}_i|^2.$$

Substituting the linear equation in the quadratic one, we obtain $K_{\mathcal{S}}(\mu) = 0$. This contradicts item (ii) of Definition 3.12. Thanks to item (i) of Definition 3.12, the Hamiltonian restricted to $V_{\mathcal{S}}$ is

$$H_{\mathcal{S}} := \sum_{\substack{\alpha-\beta\in\langle\mathcal{F}\rangle\cap\mathcal{R}_d \\ \alpha_j,\beta_j\geq 0\,,\ |\alpha|_1=|\beta|_1=d}} \binom{d}{\alpha}\binom{d}{\beta} r^\alpha \bar{r}^\beta\,, \tag{3.8}$$

where $r^\alpha = \Pi r_{\mathtt{v}_i}^{\alpha_i}$. Indeed the relations

$$\pi_{\mathcal{S}}(\alpha-\beta) = \sum_{j=1}^m (\alpha_j - \beta_j)\mathtt{v}_j = 0\,, \quad \pi_{\mathcal{S}}^{(2)}(\alpha-\beta) = \sum_{j=1}^m (\alpha_j-\beta_j)|\mathtt{v}_j|^2 = 0$$

hold if and only if $\alpha - \beta \in \langle\mathcal{F}\rangle$.

It still remains to show that $U_{\mathcal{S}} \subset V_{\mathcal{S}}$ defined in (2.14) is also invariant. Fix any $j_1, j_2 \in \mathcal{A}_i$, we need to prove that

$$\partial_{\bar{r}_{j_1}} H_{\mathcal{S}}\Big|_{U_{\mathcal{S}}} = \partial_{\bar{r}_{j_2}} H_{\mathcal{S}}\Big|_{U_{\mathcal{S}}}. \tag{3.9}$$

Now consider the map $g := g_{j_1 j_2}$ of Definition 3.2 item 4. We can extend this map to monomials (and, by linearity, to polynomials) by setting

$$\forall \alpha, \beta \in \mathbb{N}^m\,, \quad g(r^\alpha \bar{r}^\beta) = r^{g(\alpha)}\bar{r}^{g(\beta)}.$$

where we recall that

$$g(\alpha) := \sum_j \alpha_j g(\mathbf{e}_j) = \sum_j \alpha_{g^{-1}(j)} \mathbf{e}_j$$

(same for β). In particular, we have $(g(\alpha))_{j_2} = \alpha_{j_1}$ and $(g(\beta))_{j_2} = \beta_{j_1}$. It is also easy to see that for all ℓ, setting

$$\mathfrak{a}_\ell := \sum_{k\in\mathcal{A}_\ell} \alpha_k\,, \qquad \mathfrak{b}_\ell := \sum_{k\in\mathcal{A}_\ell} \beta_k\,,$$

one has $g(\mathfrak{a}_\ell) = \mathfrak{a}_\ell$, same for $\mathfrak{b}_\ell$.

For each monomial $\mathfrak{m} := \mathfrak{m}_{\alpha,\beta} = r^\alpha \bar{r}^\beta$ one has

$$\partial_{\bar{r}_{j_1}} \mathfrak{m}\big|_{U_{\mathcal{S}}} = \frac{\beta_{j_1}}{\bar{b}_i} \prod_{\ell=1}^N b_\ell^{\mathfrak{a}_\ell} \bar{b}_\ell^{\mathfrak{b}_\ell} = \frac{(g(\beta))_{j_2}}{\bar{b}_i} \prod_{\ell=1}^N b_\ell^{g(\mathfrak{a}_\ell)} \bar{b}_\ell^{g(\mathfrak{b}_\ell)} = \partial_{\bar{r}_{j_2}} g(\mathfrak{m})\big|_{U_{\mathcal{S}}}\,. \tag{3.10}$$

Note that $\beta_{j_1} \neq 0$ implies $\mathfrak{b}_\ell > 0$ so that all the expressions in (3.10) are monomials.

Moreover, g preserves the Hamiltonian $H_{\mathcal{S}}$ i.e.

$$g(H_{\mathcal{S}}) = \sum_{\substack{\alpha-\beta\in\langle\mathcal{F}\rangle\cap\mathcal{R}_d \\ \alpha_j,\beta_j\geq 0\,,\ |\alpha|_1=|\beta|_1=d}} \binom{d}{\alpha}\binom{d}{\beta} r^{g(\alpha)}\bar{r}^{g(\beta)} = \sum_{\substack{\alpha-\beta\in\langle\mathcal{F}\rangle\cap\mathcal{R}_d \\ \alpha_j,\beta_j\geq 0\,,\ |\alpha|_1=|\beta|_1=d}} \binom{d}{\alpha}\binom{d}{\beta} r^\alpha \bar{r}^\beta = H_{\mathcal{S}} \tag{3.11}$$

since

$$g(\alpha) - g(\beta) \in \langle\mathcal{F}\rangle\cap\mathcal{R}_d \iff \alpha - \beta \in \langle\mathcal{F}\rangle\cap\mathcal{R}_d\,,$$
$$\binom{d}{\alpha} = \binom{d}{g(\alpha)}\,, \qquad \binom{d}{\beta} = \binom{d}{g(\beta)}\,.$$

Finally, we use (3.10) and (3.11) in order to prove (3.9):

$$\partial_{\bar{r}_{j_1}} H_{\mathcal{S}}\Big|_{U_{\mathcal{S}}} = \partial_{\bar{r}_{j_2}} g(H_{\mathcal{S}})\Big|_{U_{\mathcal{S}}} = \partial_{\bar{r}_{j_2}} H_{\mathcal{S}}\Big|_{U_{\mathcal{S}}} .$$

□

In order to prove condition 4 we first analyze the Hamiltonian H_{Res} we have the following lemma.

Lemma 3.17. *The Hamiltonian H_{Res} has the following form*

$$H_{\text{Res}} - d!(L^{(2)})^d = d!d(d-1)(L^{(2)})^{d-2}\Big[-\frac{1}{4}\sum_{k\in\mathbb{Z}^2}|r_k|^4 + \sum_{\substack{k_1,k_2,k_3,k_4\in\mathbb{Z}^2\\ k_1\neq k_3,k_4\ ,\ k_1+k_2=k_3+k_4\\ |k_1|^2+|k_2|^2=|k_3|^2+|k_4|^2}} r_{k_1}r_{k_2}\bar{r}_{k_r}\bar{r}_{k_4}\Big] + R \quad (3.12)$$

where the term in square brackets is the cubic *NLS while R contains only terms of the following types:*

- *Action preserving terms (see Definition 2.2) of the form $|r|^{2\alpha}$ with $\alpha! := \prod_{k\in\mathbb{Z}^2}\alpha_k! > 2$.*
- *Non-action preserving terms whose degree in the actions is less than $d-2$.*
- *Non-action preserving terms $r^\alpha\bar{r}^\beta$ of degree $d-2$ in the actions and such that $\alpha!\beta! > 1$.*

Proof. First we note that a resonance is action preserving if (up to permutations of the even and the odd indexes between themselves) one has

$$\{k_1, k_1, k_2, k_2, \ldots, k_d, k_d\}.$$

We can evidence an *integrable part* of the Hamiltonian H_{Res} as

$$H_{\text{Int}} := H_{\text{Int}}(|r_k|^2) = \sum_{\substack{\alpha\in(\mathbb{N})^{\mathbb{Z}^2}:\\ |\alpha|=d}} \binom{d}{\alpha}^2 |r|^{2\alpha},$$

which contains all the terms in (2.6) with $\alpha = \beta$. Note that H_{Int} is a symmetric function of the actions $\{|r_k|^2\}_{k\in\mathbb{Z}^2}$. It will be convenient to denote

$$\|r\|_{\ell^{2j}}^{2j} = L^{(2j)} = \sum_{k\in\mathbb{Z}^2}|r_k|^{2j}$$

so that, for instance, $L^{(2)}$ equals the conserved quantity $\|r\|_{L^2}^2$. It is well known that the functions $L^{(2j)}$ generate the symmetric polynomials in the actions. Hence we can express the integrable Hamiltonian as a polynomial in the $L^{(2j)}$. Since $L^{(2)}$ is a constant of motion (and we will perform a symmetry reduction with respect to it) it will be convenient to evidence the two terms of highest degree in $L^{(2)}$.

$$(L^{(2)})^m = \sum_{|\alpha|=m}\binom{m}{\alpha}|r|^{2\alpha} = m!\sum_{\substack{|\alpha|=m\\ \alpha!=1}}|r|^{2\alpha} + \sum_{\substack{|\alpha|=m\\ \alpha!>1}}\binom{m}{\alpha}|r|^{2\alpha}.$$

By direct computation,

$$\begin{aligned} H_{\text{Int}} - d!(L^{(2)})^d &= \sum_{\alpha!>1}\binom{d}{\alpha}^2|r|^{2\alpha}(1-\alpha!) \\ &= -\frac{(d!)^2}{4}\sum_{\substack{|\alpha|=d,\\ \alpha!=2}}|r|^{2\alpha} + \sum_{\alpha!>2}\binom{d}{\alpha}^2|r|^{2\alpha}(1-\alpha!). \end{aligned}$$

Note that

$$\sum_{\substack{|\alpha|=d,\\ \alpha!=2}}|r|^{2\alpha} = \sum_{k\in\mathbb{Z}^2}|r_k|^4\sum_{\substack{|\beta|=d-2\\ (\beta+2\mathbf{e}_k)!=2}}|r|^{2\beta},$$

where $\mathbf{e}_k \in (\mathbb{N})^{\mathbb{Z}^2}$ is the k'th basis vector. We compare the above expression with

$$(L^{(2)})^{d-2}L^{(4)} = \sum_k \sum_{|\beta|=d-2} \binom{d-2}{\beta} |r|^{2\beta+4\mathbf{e}_k}$$
$$= (d-2)! \sum_{\substack{|\beta|=d-2 \\ (\beta+2\mathbf{e}_k)!=2}} |r|^{2\beta+4\mathbf{e}_k} + \sum_{\substack{|\beta|=d-2 \\ (\beta+2\mathbf{e}_k)!>2}} \binom{d-2}{\beta} |r|^{2\beta+4\mathbf{e}_k}.$$

Thus,

$$H_{\mathrm{Int}} - d!(L^{(2)})^d = -\frac{d!d(d-1)}{4} L^{(4)} \left(L^{(2)}\right)^{d-2} + \sum_{\alpha!>2} c_\alpha |r|^{2\alpha} \tag{3.13}$$

We could continue our computation. However we only need that $\sum_{\alpha!>2} c_\alpha |r|^{2\alpha}$, as polynomial in the $L^{(i)}$, is of degree at most $d-3$ in $L^{(2)}$.

We can perform a similar procedure for the non-integrable part of the Hamiltonian

$$H_{\mathrm{Res}} - H_{\mathrm{Int}} = \sum_{\substack{\alpha\neq\beta\in(\mathbb{N})^{\mathbb{Z}^2}:|\alpha|=|\beta|=d \\ \sum_k(\alpha_k-\beta_k)k=0\,,\ \sum_k(\alpha_k-\beta_k)|k|^2=0}} \binom{d}{\alpha}\binom{d}{\beta} r^\alpha \bar{r}^\beta.$$

We first evidence the terms of higher degree in the action variables which are clearly:

$$\sum_{\substack{k_1,k_2,k_3,k_4\in\mathbb{Z}^2 \\ k_1\neq k_3,k_4\,,\ k_1+k_2=k_3+k_4 \\ |k_1|^2+|k_2|^2=|k_3|^2+|k_4|^2}} \sum_{|\alpha|=d-2} \binom{d}{\alpha+\mathbf{e}_{k_1}+\mathbf{e}_{k_2}} \binom{d}{\alpha+\mathbf{e}_{k_3}+\mathbf{e}_{k_4}} |r|^{2\alpha} r_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} =$$

$$= \sum_{\substack{k_1,k_2,k_3,k_4\in\mathbb{Z}^2 \\ k_1\neq k_3,k_4\,,\ k_1+k_2=k_3+k_4 \\ |k_1|^2+|k_2|^2=|k_3|^2+|k_4|^2}} \Bigg[(d!)^2 \sum_{\substack{|\alpha|=d-2 \\ (\alpha+\mathbf{e}_{k_1}+\mathbf{e}_{k_2})!=1 \\ (\alpha+\mathbf{e}_{k_3}+\mathbf{e}_{k_4})!=1}} |r|^{2\alpha} r_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} +$$
$$+ \sum_{\substack{|\alpha|=d-2 \\ (\alpha+\mathbf{e}_{k_1}+\mathbf{e}_{k_2})!(\alpha+\mathbf{e}_{k_3}+\mathbf{e}_{k_4})!>1}} \binom{d}{\alpha+\mathbf{e}_{k_1}+\mathbf{e}_{k_2}} \binom{d}{\alpha+\mathbf{e}_{k_3}+\mathbf{e}_{k_4}} |r|^{2\alpha} r_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} \Bigg]$$

We proceed as for the integrable terms evidencing the highest order term in $L^{(2)}$, we have

$$(L^{(2)})^{d-2} \sum_{\substack{k_1,k_2,k_3,k_4\in\mathbb{Z}^2 \\ k_1\neq k_3,k_4\,,\ k_1+k_2=k_3+k_4 \\ |k_1|^2+|k_2|^2=|k_3|^2+|k_4|^2}} r_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} =$$

$$\sum_{\substack{k_1,k_2,k_3,k_4\in\mathbb{Z}^2 \\ k_1\neq k_3,k_4\,,\ k_1+k_2=k_3+k_4 \\ |k_1|^2+|k_2|^2=|k_3|^2+|k_4|^2}} \Bigg[(d-2)! \sum_{\substack{|\alpha|=d-2 \\ \alpha!=1}} |r|^{2\alpha} u_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} + \sum_{\substack{|\alpha|=d-2 \\ \alpha!>1}} \binom{d-2}{\alpha} |r|^{2\alpha} r_{k_1} r_{k_2} \bar{r}_{k_3} \bar{r}_{k_4} \Bigg]$$

We have proved our thesis, in formulæ the remainder R is given by

$$R = \sum_{\substack{|\alpha|=|\beta|=d\,,\ |\alpha-\beta|>4 \\ \sum_k(\alpha_k-\beta_k)k=0\,,\ \sum_k(\alpha_k-\beta_k)|k|^2=0}} R_{\alpha,\beta} r^\alpha \bar{r}^\beta + \sum_{\substack{|\alpha|=|\beta|=d\,,\ |\alpha-\beta|=4\,,\alpha!\beta!>1 \\ \sum_k(\alpha_k-\beta_k)k=0\,,\ \sum_k(\alpha_k-\beta_k)|k|^2=0}} R_{\alpha,\beta} r^\alpha \bar{r}^\beta + \sum_{|\alpha|=d\,,\ \alpha!>2} R_\alpha |r|^{2\alpha}$$

□

Lemma 3.18. *Consider $\mathcal{S}$ belonging to $(\mathcal{M} \setminus \mathcal{D}) \cap \mathbb{Z}^{2m}$, where $\mathcal{M}$ is the variety defined in* (3.6) *and $\mathcal{D}$ is the subvariety given by Theorem 3.14. Then, $\mathcal{S}$ satisfies condition 4 in Definition 2.4.*

Proof. We consider the Hamiltonian $H_{\rm Res}$ of formula (3.12) restricted to the subspace $U_{\mathcal{S}}$. The new Hamiltonian $h_{\mathcal{S}}$ is defined as $h_{\mathcal{S}} = n^{-1} H_{\rm Res}|_{U_{\mathcal{S}}}$ where $n = 2^{N-1}$. Note that the factor n^{-1} is not a time rescaling. It needs to be added in order to obtain the symplectic form $\frac{\mathrm{i}}{2}\sum_j db_j \wedge d\bar{b}_j$. Note that $h_{\mathcal{S}}$ is homogeneous of degree $2d$ in $(b, \bar{b})$.

One can analyze explicitly the toy-model Hamiltonian $h_{\mathcal{S}}$:

$$h_{\mathcal{S}}(b) = \frac{1}{n} \sum_{\substack{\mathfrak{a},\mathfrak{b} \in \mathbb{N}_0^N \\ \sum_i \mathfrak{a}_i = \sum_i \mathfrak{b}_i = d}} C_{\mathfrak{a},\mathfrak{b}} b^{\mathfrak{a}} \bar{b}^{\mathfrak{b}}, \quad C_{\mathfrak{a},\mathfrak{b}} := \sum_{\substack{\alpha,\beta \in \mathbb{N}_0^m :\ \alpha-\beta \in \langle \mathcal{F} \rangle \\ \sum_{j\in\mathcal{A}_i} \alpha_j = \mathfrak{a}_i \\ \sum_{j\in\mathcal{A}_i} \beta_j = \mathfrak{b}_i}} \binom{d}{\alpha}\binom{d}{\beta}, \tag{3.14}$$

where by an abuse of notation with $j \in \mathcal{A}_i$ we mean $\mathbf{e}_j \in \mathcal{A}_i$ and hence $(i-1)n + 1 \le j \le in$. The important fact is that $h_{\mathcal{S}}$ is a polynomial in $n = 2^{N-1}$ (the number of elements in each generation) and, since n is very large, we only need to compute the leading orders. The degree of $h_{\mathcal{S}}$ in n is at most $d-1$ (the coefficients $C_{\mathfrak{a},\mathfrak{b}}$'s have degree at most d). The terms that we will need to compute explicitly are the coefficients of n^{d-1} and of n^{d-2} in $h_{\mathcal{S}}$ (which amounts to computing the coefficients of n^d and of n^{d-1} in $C_{\mathfrak{a},\mathfrak{b}}$). The crucial remark, informally stated, is that the degree in n is lower for terms whose combinatorics imposes more constraints; this happens by two mechanisms:

- by decreasing $\mathrm{Supp}(\alpha)$ i.e. the cardinality (which is at most d) of the $\{\alpha_i \neq 0\}$ (or, symmetrically, $\mathrm{Supp}(\beta)$);
- by increasing $\alpha - \beta \in \langle \mathcal{F} \rangle$, indeed if we know that the indexes k_i satisfy some family relations then by fixing the family we fix four of the indexes.

We know that $H_{\rm Res}$ Hamiltonian has the expression (3.12) and moreover it is easy to see that all the terms in R contribute at most n^{d-3}. Then we have

$$\begin{aligned} h_{\mathcal{S}}(b) = \quad & d! n^{d-1} \left(\sum_{i=1}^N |b_i|^2 \right)^d + \\ + \quad & n^{d-2} d! d(d-1) \left(\sum_{i=1}^N |b_i|^2 \right)^{d-2} \left[-\frac{1}{4} \sum_{i=1}^N |b_i|^4 + \sum_{i=1}^{N-1} \mathrm{Re}\,(b_i^2 \bar{b}_{i+1}^2) \right] + \mathcal{O}(n^{d-3}). \end{aligned}$$

We still have to analyze the terms contained in $\mathcal{O}(n^{d-3})$ in order to check that the polynomial $\mathcal{P}$ of Definition 2.4 satisfies the properties 4(a)-4(f). Properties 4(a), 4(b), 4(c) are completely straightforward, while 4(d) follows directly from Lemma 3.11, item (v).

As for property 4(e), the fact that $U_{\mathcal{S}}^j$ is invariant follows from 4(d). Note that the only elements of $\langle \mathcal{F} \rangle$ entirely supported on the generations $\mathcal{A}_j, \mathcal{A}_{j+1}$ are of the form $\sum_k \lambda_k f_k$ with $\lambda_k \in \mathbb{Z}$, where the f_k's are the vectors representing the 2^{N-2} families of generation number j (see Definition 3.1). Note that the f_k's have disjoint support. Then, using (3.8), it is immediate to see that the expression of the Hamiltonian as a function of $(b_j, \bar{b}_j)$ and $(b_{j+1}, \bar{b}_{j+1})$ relies on a purely combinatorial computation, independent of j (up to an index translation). This combinatorial structure is left invariant if one exchanges parents with children in all the families f_k: this gives the symmetry with respect to the exchange $j \longleftrightarrow j+1$.

To conclude, we prove 4(f). Given $i \neq j$, we consider monomials in $h_{\mathcal{S}}$ which depend only on $(b_i, \bar{b}_i), (b_j, \bar{b}_j)$ and are exactly of degree two in $(b_i, \bar{b}_i)$. Monomials of this form can only come from monomials in (3.8) such that $\alpha - \beta \in \langle \mathcal{F} \rangle$ is supported entirely on the j-th generation except for at most two elements. Therefore we can apply Lemma 3.11, (iv) and deduce that $\alpha - \beta$ is either zero or (up to the sign) an abstract family. The case $\alpha - \beta = 0$ corresponds to action preserving monomials of (3.8) and produces monomials of the form $\chi_{ij} |b_j|^{2d-2} |b_i|^2$ (for some suitable coefficient χ_{ij}), while the case $\alpha - \beta = \pm f$ with $f \in \mathcal{F}$ is possible only if $|i - j| = 1$ and produces terms of the form $\rho_{ij} |b_j|^{2d-4} \mathrm{Re}\,(b_i^2 \bar{b}_j^2)$ (for some suitable coefficient ρ_{ij}). The fact that $H_{\rm Int}$ is a symmetric polynomial in the variables $|r_k|^2$ implies that $\rho_{ij} \equiv \rho$ is independent of i and j. Finally, the fact that $\rho_{i,i+1} \equiv \rho$ is independent of i follows from 4(e). □

In conclusion, any $\mathcal{S} \in (\mathcal{M} \setminus \mathcal{D}) \cap \mathbb{Z}^{2m}$ satisfies conditions 1,3,4 of Definition 2.4. The fact that $(\mathcal{M} \setminus \mathcal{D}) \cap \mathbb{Z}^{2m}$ is non-empty follows from the density of $\mathcal{M} \cap \mathbb{Q}^{2m}$ on $\mathcal{M}$ and from the fact that $\mathcal{M}$ and $\mathcal{D}$ are homogeneous (if v belongs to the manifold, tv also belongs to the manifold for all $t \in \mathcal{R}$). The existence of sets $\mathcal{S}$ satisfying also item 2 follows the same reasoning as in [CKS$^+$10]. In order to give quantitative estimates for the norm of the points in $\mathcal{S}$, we denote by $\{\mathtt{j}_1, \dots, \mathtt{j}_m\}$ the prototype embedding obtained by mapping each $\mathtt{e}_i = (z_1, \dots, z_{k-1}, z_k, \dots, z_{N-1}) \in \Sigma$ (notation as in (3.4)) to $\mathtt{j}_i \in \mathbb{Z}^2$ via

$$\mathtt{e}_i \mapsto \mathtt{j}_i = (\mathrm{Re} \prod_i z_i, \mathrm{Im} \prod_i z_i) \in \mathbb{Z}^2 \qquad i = 1, \dots, m$$

(note that this in this list the vectors $\mathtt{j}_i$ are NOT distinct but have high multiplicity).

Lemma 3.19. *There exists $R < (N2^N)^{16dN(N2^N)^{8d}}$, such that one may choose a non-degenerate generation set*

$$\mathcal{S} = \{\mathtt{v}_1, \dots, \mathtt{v}_m\} \in (\mathcal{M} \setminus \mathcal{D}) \cap \mathbb{Z}^{2m}$$

satisfying

$$|\mathtt{v}_i - R\mathtt{j}_i| \le 3^{-N} R \qquad \forall i = 1, \dots, m \ . \tag{3.15}$$

Remark 3.20. *For $N \gg 1$, the condition* (3.15) *implies that the norm explosion property* (2.13) *is satisfied.*

Proof. We preliminarily notice that the resonance relations are a set of at most quadratic equations in the variables $\mathtt{v}_i$'s. Thus, if we assume to have fixed (with the inductive procedure described in Section 3.1) the first k variables so that the non-degeneracy conditions given by Definition 3.12, then in adding the $k+1$-th variable, in order to enforce the non-degeneracy conditions, we must verify that it does not satisfy $K \le k^{7d}$ at most quadratic relations. Moreover by definition $\mathcal{M}$ is a homogeneous manifold, namely if $v \in \mathcal{M}$ then $tv \in \mathcal{M}$ for all $t \in \mathbb{R}$. Finally we notice that also the resonance relations are homogeneous, hence if $v \in \mathcal{M} \setminus \mathcal{D}$ then also $tv \in \mathcal{M} \setminus \mathcal{D}$.

We start by considering a neighborhood of radius 10^{-N} of each $\mathtt{j}_i$ with $i \in \mathcal{A}_1$. Then rescaling by $R_1 = 10^{2dN}$ we can ensure that in each neighborhood there are more than 2^{8dN} integer points so that we can surely choose in these neighborhoods integer points $\mathtt{w}_i^{(1)}$ such that

$$|\mathtt{w}_i^{(1)} - R_1 \mathtt{j}_i| \le 10^{-N} R_1 \quad \forall i \in \mathcal{A}_1$$

and the $\mathtt{w}_i^{(1)}$ satisfy the non-degeneracy conditions.

We proceed by induction. At each generation $j \ge 2$ we have $\mathtt{w}_i^{(j-1)} \in \mathbb{Z}^2$ with $i \in \cup_{h=1}^{j-1} \mathcal{A}_h$ so that

1_{j-1} $|\mathtt{w}_i^{(j-1)} - R_{j-1}\mathtt{j}_i| \le 3^{j-2} \cdot 10^{-N} R_{j-1}$,

2_{j-1} $\{\mathtt{w}_i^{(j-1)}\}_{i \in \cup_{h=1}^{j-1} \mathcal{A}_h}$ is a non-degenerate generation set with $j-1$ generations.

Then we claim that we can choose $\mathtt{w}_i^{(j-1)} \in \mathbb{Q}^2$ for $i \in \mathcal{A}_j$ so that

(i) $|\mathtt{w}_i^{(j-1)} - R_{j-1}\mathtt{j}_i| \le 3^{j-1} \cdot 10^{-N} R_{j-1}$,

(ii) setting $K = (N2^N)^{16d(N2^N)^{8d}}$, we have that $K\mathtt{w}_i^{(j-1)} \in \mathbb{Z}^2$.

(iii) $\{\mathtt{w}_i^{(j-1)}\}_{i \in \cup_{h=1}^{j} \mathcal{A}_h}$ is a non-degenerate generation set with j generations.

If our claim holds true, we set $R_j = KR_{j-1}$ and $\mathtt{w}_i^{(j)} = K\mathtt{w}_i^{(j-1)}$ for $i \in \cup_{h=1}^{j} \mathcal{A}_h$. By construction items 1_j and 2_j hold . We conclude our proof by fixing $R = R_N$ and $\mathtt{v}_i = \mathtt{w}_i^{(N)}$ for all $i \in \mathcal{A}$.

It remains to prove our claim. To pass from a $j-1$ generation set to one with j generations we have to use the family relations and for each couple of parents produce the corresponding two children. Let us fix two parents $\mathtt{w}_{i_1}^{(j-1)} \rightsquigarrow p_1$ and $\mathtt{w}_{i_2}^{(j-1)} \rightsquigarrow p_2$. This means fixing two opposite points on the circle $(v - p_1, v - p_2) = 0$. By construction, if we choose as children c_1, c_2 the two opposite points such that $c_1 - c_2$ is orthogonal to $p_1 - p_2$ then

$$|c_k - R_{j-1}\mathtt{j}_{\ell_k}| \le 2 \cdot 3^{j-1} \cdot 10^{-N} R_{j-1} \qquad k = 1, 2$$

(where $\mathtt{j}_{\ell_1}, \mathtt{j}_{\ell_2}$ are the two corresponding children in the prototype embedding), however we cannot guarantee that these two points satisfy item (iii). We can write the rational points on the circle as

$$P_t := p_1 - \frac{(p_1 - p_2, t)t}{|t|^2}, \quad t = (m_1, m_2) \in \mathbb{Z}^2 \ .$$

Noting that

$$c_k = \frac{p_1 + p_2}{2} \pm O\frac{p_1 - p_2}{2}, \quad O = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

we can compute the τ corresponding to one of the two children, say

$$P_\tau := c_1 = \frac{p_1 + p_2}{2} + O\frac{p_1 - p_2}{2}$$

and we get

$$(p_1 - p_2, \tau) = (p_1 - p_2, O\tau) \longrightarrow \tau := (O - \mathbb{I})(p_1 - p_2)$$

(note that in this way the x, y coordinates of τ are NOT coprime!). We now consider the points P_{τ_k} for $\pm k = D, \ldots, 2D$ defined by

$$\tau_k = (k\mathbb{I} + O)\,\tau$$

so that as k varies, τ_k identifies different points on the circle. By direct computation one has that

$$P_{\tau_k} = p_1 + \frac{k+1}{2(k^2+1)}\tau_k$$

and

$$\mathrm{dist}(P_{\tau_k}, P_\tau) = \frac{1}{\sqrt{k^2+1}}|p_1 - p_2|$$

If we fix $D > 20^N$ we are sure that each point P_{τ_k} satisfies item (i). Now each non-degeneracy condition removes at most two points on the circle and we have at most $(N2^N)^{7d}$ conditions, hence we can ensure the existence of non-degenerate P_{τ_k} by fixing $D := (N2^N)^{8d} > 20^N$, finally since $K \geq (D!)^2 \gtrsim \mathrm{mcm}(k^2+1)_{k=D}^{2D}$ also item (ii) is verified. Therefore, the thesis follows by noting that $R_N = R_1 K^{N-1} < K^N$ and that $3^{N-1} \cdot 10^{-N} < 3^{-N}$.

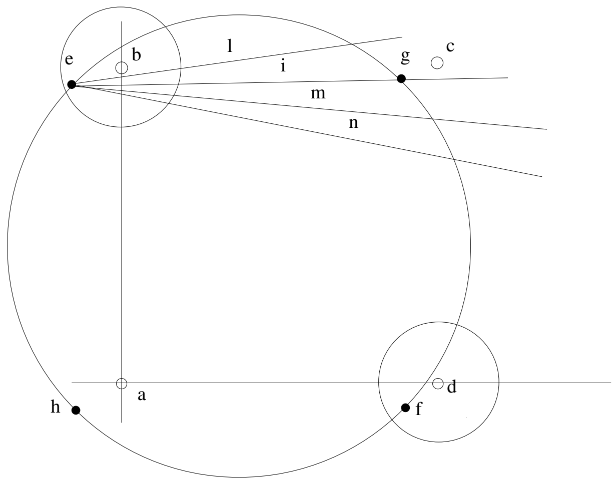

Figure 1: Our procedure for finding rational points with bounded denominators, here we wish to place a point in the second generation *close* to $\mathtt{j}_3$.

□

Corollary 3.21. *For $N \gg 1$ there exists an acceptable (see Definition 2.4) generation set $\mathcal{S} = \mathcal{S}(N)$ such that*

$$|\mathtt{v}| < (N2^N)^{16dN(N2^N)^{8d}+1} \qquad \forall \mathtt{v} \in \mathcal{S}\ . \tag{3.16}$$

Proof. It follows directly from Lemma 3.19 and Remark 3.20. □

3.3 Proof of Lemma 3.11

We first prove (i), (ii). Assume by contradiction that the abstract families of $\mathcal{F}$ are linearly dependent. This means that there exist some $\alpha_k \in \mathbb{Q}, f_k \in \mathcal{F}$ such that

$$\mathscr{L} := \sum_k \alpha_k f_k = 0 \ , \quad \exists k_0 \ : \alpha_{k_0} \neq 0 \, .$$

Now, let $i_<$ and $i_>$ be respectively the minimal and the maximal generation numbers of the families appearing in the linear combination $\mathscr{L}$ (with nonzero coefficient). It follows from Remark 3.4 that the support of $\mathscr{L}$ contains at least two elements of the generation $i_<$ and two elements of the generation $i_> + 1$, which implies $\mathscr{L} \neq 0$, which is absurd. Now, since the abstract families of $\mathcal{F}$ are linearly independent, they form a basis of $\langle \mathcal{F} \rangle$. Therefore each $\lambda \in \langle \mathcal{F} \rangle$ can be written in a unique way as a linear combination

$$\lambda = \sum_k \alpha_k f_k \ . \tag{3.17}$$

Then, $|\mathrm{Supp}(\lambda)| \geq 4$ since as above it contains at least two elements in $\mathcal{A}_{i_<}$ and two elements in $\mathcal{A}_{i_>+1}$.

Now, suppose $|\mathrm{Supp}(\lambda)| = 4$. This means that we have exactly two elements in the generation $i_<$ and two in the generation $i_> + 1$, and no elements in the possible intermediate generations. We claim that $i_< = i_>$. In order to prove our claim, we first notice that, if for some i the linear combination (3.17) contains h different families (namely, there are h different families with nonzero coefficient α_k) of generation number i and $k \neq h$ different families of generation number $i+1$, then $\mathrm{Supp}(\lambda)$ contains at least 2 elements of the generation $i+1$. Then we notice that, if for some i the expression (3.17) contains exactly one family of generation number i and exactly one family of generation number $i+1$, then $\mathrm{Supp}(\lambda)$ contains at least 2 elements of the generation $i+1$ (since sibling and spouse cannot coincide). It follows that, since $\mathrm{Supp}(\lambda)$ does not contain elements from the intermediate generations, there cannot be intermediate generations, i.e. $i_< = i_> = \bar{\imath}$. Finally, thanks to Remark 3.4, in order to have only two elements of generation number $\bar{\imath}$ and two elements of generation number $\bar{\imath}+1$, there must be exactly one family.

We now prove (iii). Consider $\lambda \in \langle \mathcal{F} \rangle$. Setting $\mathfrak{g}(f_k)$ to be the generation number of the family f_k, we can write in a unique way

$$\lambda = \sum_k \alpha_k f_k = \sum_{\mathfrak{g}(f_k) = i_>} \alpha_k f_k + \sum_{\mathfrak{g}(f_k) \neq i_>} \alpha_k f_k.$$

Now, by definition of $\langle \mathcal{F} \rangle$, we have $\lambda = \sum_j \lambda_j \mathbf{e}_j$ with $\lambda_j \in \mathbb{Z}$. One easily sees, by Remark 3.4, that for all $j \in \mathcal{A}_{i_>+1}$ one has $\lambda_j = -\alpha_k$ for one (and only one) f_k of generation number $i_>$. Then

$$\lambda - \sum_{\mathfrak{g}(f_k) = i_>} \alpha_k f_k \in \langle \mathcal{F} \rangle$$

and the claim follows by recursion on the maximal age.

Then, we prove (iv). Assume $\lambda \neq 0$, otherwise the thesis is obvious. We have observed that $\mathrm{Supp}(\lambda)$ must contain at least two elements of the generation $i_<$ and at least two elements of the generation $i_> + 1$. The assumption in (iv) implies that one of these two generations contains exactly two elements of $\mathrm{Supp}(\lambda)$ and that moreover, for all $j \neq i_<, i_> + 1$, $\mathrm{Supp}(\lambda)$ contains no elements of the generation j. We assume that the generation $i_<$ contains exactly two elements of $\mathrm{Supp}(\lambda)$ which means that the linear combination defining λ contains one and only one family f_{k_0} of generation number $i_<$ (the case with generation $i_> + 1$ is symmetric), appearing with the coefficient $\alpha_{k_0} \neq 0$. Then we distinguish two cases: either $i_< = i_>$ or $i_< \neq i_>$. If $i_< = i_>$, then the thesis follows easily by Remark 3.4. If $i_< \neq i_>$, then the generation $i_< + 1$ contains no element of $\mathrm{Supp}(\lambda)$. But this means that the two children in f_{k_0} (call them $\mathbf{c}_1, \mathbf{c}_2$) must be canceled out, which implies that the two families f_{k_1}, f_{k_2} in which $\mathbf{c}_1, \mathbf{c}_2$ appear as parents ($k_1 \neq k_2$ since siblings do not marry each other) have the same (non-zero) coefficient $\alpha_{k_1} = \alpha_{k_2} = \alpha_{k_0}$. But then also the two spouses of $\mathbf{c}_1, \mathbf{c}_2$ must cancel out: consider for instance the spouse of $\mathbf{c}_1$ and call it $\mathbf{s}_1$. We have that $\mathbf{s}_1$ appears as a child in a family of generation number $i_<$: we call this family f_{k_3}. Note that $k_3 \neq k_0$ since $\mathbf{s}_1 \notin \{\mathbf{c}_1, \mathbf{c}_2\}$. The fact that $\mathbf{s}_1$ is canceled out implies that $\alpha_{k_3} = \alpha_{k_1} \neq 0$, but this is absurd since f_{k_0} is the only family of generation number $i_<$ to appear in the linear combination. This completes the proof of (iv).

Finally, the property (v) is a simple remark when λ is a single family vector and trivially generalizes to the case $\lambda \in \langle \mathcal{F} \rangle$. This completes the proof of the lemma.

3.4 Proof of Theorem 3.14

Now we prove the existence of non degenerate generation sets, according to Definition 3.12. A preliminary but very important step is to show that the linear and quadratic relations defining $\mathcal{M}$, see (3.6), do not imply any linear relation except those given by $\pi_{\mathcal{S}}(\lambda) = 0$ for all $\lambda \in \langle \mathcal{F} \rangle$. It is clear that if this were not true one could not impose condition (i) of Definition 3.12. Specifically we prove

Lemma 3.22. *Consider a codimension one subspace $\Sigma \subset \mathbb{R}^{2m}$ and the sets $\mathcal{M}$ and $\mathcal{L}$ defined respectively in* (3.6) *and* (3.7). *Then, $\mathcal{M} \subset \Sigma$ implies $\mathcal{L} \subset \Sigma$.*

The strategy of this proof relies on the choice of a good set of variables for $\mathcal{L}$ (and consequently $\mathcal{M}$) as explained in Section 3.1. Namely, up to a reordering of the vectors $\mathbf{e}_j$, the matrix whose rows are the abstract families (see Definition 3.1) is in *row echelon form*, see (3.19). This gives a recursive rule to fix the dependent variables (the *pivots*) as well as the circles for the remainig variables. Then we write the relation defining Σ in the independent variables of $\mathcal{L}$ and the condition $\mathcal{L} \not\subset \Sigma$ means that the coefficients are not all zero. Then with respect to the *youngest* variable Σ defines a line, while the quadratic relations a *non-degenerate* circle, which obviously cannot be contained in a line.

Proof of Lemma 3.22. We will prove that if $\mathcal{L} \not\subset \Sigma$, then $\mathcal{M} \not\subset \Sigma$. Namely, we will prove that, for any given codimension one subspace Σ which does not contain $\mathcal{L}$, we can choose $\mathcal{S} \in \mathcal{M} \setminus \Sigma$.

If we denote by $\mathtt{v}_j^{(1)}, \mathtt{v}_j^{(2)}$ the two components of $\mathtt{v}_j \in \mathbb{R}^2$, a codimension one subspace $\Sigma \subset \mathbb{R}^{2m}$ is defined by an equation of the form

$$\sum_{j=1}^{m} \sum_{k=1}^{2} \lambda_{jk} \mathtt{v}_j^{(k)} = 0 \ . \tag{3.18}$$

Now, for simplicity of notation and without loss of generality, we reorder the basis $\{\mathbf{e}_j\}_{j=1}^m$ of $\mathbb{Z}^m$ so that two siblings belonging to the same abstract family always have consecutive subindices. In matrix notation, the condition of $\mathcal{S}$ being a generation set can be denoted

$$\pi_{\mathcal{S}}(F^T) = 0\,, \quad \pi_{\mathcal{S}}^{(2)}(F^T) = 0\ ,$$

where F is a matrix whose rows are given by the abstract families and F^T denotes its transpose. We choose to order the rows of F so that the matrix is in *lower row echelon* form (see figure).

$\mathtt{w}_1$	$\mathtt{w}_2$	$\mathtt{w}_3$	$\mathtt{w}_4$	$\mathtt{w}_5$	$\mathtt{p}_5$	$\mathtt{w}_6$	$\mathtt{p}_6$	$\mathtt{w}_7$	$\mathtt{p}_7$	$\mathtt{w}_8$	$\mathtt{p}_8$
1	1	0	0	−1	−1	0	0	0	0	0	0
0	0	1	1	0	0	−1	−1	0	0	0	0
0	0	0	0	1	0	1	0	−1	−1	0	0
0	0	0	0	0	1	0	1	0	0	−1	−1

(3.19)

Each row of a matrix in lower row echelon form has a *pivot*, i.e. the first nonzero coefficient of the row starting from the right. Being in lower row echelon form means that the pivot of a row is always strictly to the right of the pivot of the row above it. In the matrix F defined by the abstract families, the pivots are all equal to -1 and they correspond to one and only one child from each family. In order to exploit this fact, we accordingly rename the elements of the generation set by writing $\mathcal{S} = (\mathtt{p}, \mathtt{w}) \in \mathbb{R}^{2a} \times \mathbb{R}^{2b}$, with $a = (N-1)2^{N-2}$, $b = m - a = (N+1)2^{N-2}$, where the $\mathtt{p}_j \in \mathbb{R}^2$ are the elements of the generation set corresponding to the pivots and the $\mathtt{w}_\ell \in \mathbb{R}^2$ are all the others, i.e. all the elements of the first generation and one and only one child (the non-pivot) from each family. Here, the index ℓ ranges from 1 to b, while the index j ranges from $2^{N-1}+1$ to b (note that $a + 2^{N-1} = b$), so that a couple $(\mathtt{p}_j, \mathtt{w}_\ell)$ corresponds to a couple of siblings if and only if $j = \ell$. Then, the linear relations $\pi_{\mathcal{S}}(F^T) = 0$ can be used to write each $\mathtt{p}_j$ as a linear combination containing only the $\mathtt{w}_\ell$'s with $\ell \le j$:

$$\mathtt{p}_j = \sum_{\ell \le j} \mu_\ell \mathtt{w}_\ell, \quad \mu_\ell \in \mathbb{Q}\ . \tag{3.20}$$

Finally, the quadratic relations $\pi_{\mathcal{S}}^{(2)}(F^T) = 0$ constrain each $\mathtt{w}_\ell$ with $\ell > 2^{N-1}$ (i.e. not in the first generation) to a circle depending on the $\mathtt{w}_j$ with $j < \ell$; note that this circle has positive radius provided that the parents of $\mathtt{w}_\ell$ are distinct. Then, equation (3.20) together with Lemma 3.11 (i) implies that the left hand side of equation (3.18) can be rewritten in a unique way as a linear combination of the $\mathtt{w}_\ell$'s only. Thus, we have

$$\sum_{\ell=1}^{b}\sum_{k=1}^{2}\eta_{\ell,k}\mathtt{w}_\ell^{(k)} = 0 \ . \tag{3.21}$$

Hence, the assumption that $\mathcal{L} \not\subset \Sigma$ is equivalent to the fact that $\eta \in \mathbb{R}^{2b}$ does not vanish. Let

$$\bar{\ell} := \max\{\ell \mid (\eta_{\ell,1}, \eta_{\ell,2}) \neq (0,0)\} \ .$$

If $\bar{\ell} \leq 2^{N-1}$, then $\mathtt{w}_{\bar{\ell}}$ is in the first generation. Since there are no restrictions (either linear or quadratic) on the first generation, the statement is trivial. Hence assume $\bar{\ell} > 2^{N-1}$. As we have discussed previously, we can assume (by removing from $\mathcal{M}$ a proper submanifold of codimension one) that $\mathtt{v}_h \neq \mathtt{v}_k$ for all $h \neq k$. Then the quadratic constraint on $\mathtt{w}_{\bar{\ell}} \in \mathbb{R}^2$ gives a circle of positive radius. Since (3.21) defines a line in the variable $\mathtt{w}_{\bar{\ell}}$ we can ensure that the relation (3.21) is not fulfilled by excluding at most two points on this circle. Thus we are able to construct $\mathcal{S} \in \mathcal{M} \setminus \Sigma$.

□

Remark 3.23. *By Lemma 3.22, if a linear equation $\sum_j \lambda_j \mathtt{v}_j \equiv 0$ identically on $\mathcal{M}$ then it must be $\lambda \in \langle \mathcal{F} \rangle$.*

Now we are ready to prove Theorem 3.14.

Proof of Theorem 3.14. Set

$$\mathcal{D}_0 = \cup_{\lambda \in \mathcal{R}_{2d} \setminus \langle \mathcal{F} \rangle}\{\mathcal{S} \in \mathcal{M} : \ \ \pi_{\mathcal{S}}(\lambda) = 0\}.$$

By Remark 3.23, this is an algebraic manifold of codimension one in $\mathcal{M}$. Moreover, by definition, in $\mathcal{M} \setminus \mathcal{D}_0$ the condition (i) of Definition 3.12 is satisfied.

To impose the second condition we proceed by induction. As in our geometric construction of $\mathcal{M}$ (see Section 3.1), we suppose to have fixed i generations and $0 \leq h < 2^{N-2}$ families with children in the $i+1$-th generation. This means that we have fixed $\pi_{\mathcal{S}}(\mathbf{e}_j)$ for all $j \leq 2^{N-1}i$ and for some subset of cardinality $2h$ of $\mathbf{e}_j$ in the $i+1$-th generation. Let us denote by A the set of indexes j such that $\pi_{\mathcal{S}}(\mathbf{e}_j)$ has been fixed. Our inductive hypothesis is that all the non-degeneracy conditions with support contained in A are satisfied. In particular, this implies that all the $\mathtt{v}_j$ with $j \in A$ are distinct. Let us denote by $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{A}_{i+1}$ the next children we wish to generate and by $\mathbf{e}_{j_1}, \mathbf{e}_{j_2} \in \mathcal{A}_i$ their *parental couple*. We wish to fix $\mathtt{w}_1 = \pi_{\mathcal{S}}(\mathbf{c}_1)$ and $\mathtt{w}_2 = \pi_{\mathcal{S}}(\mathbf{c}_2)$ so that the non-degeneracy conditions hold. Due to the linear relations $\mathtt{w}_2 = -\mathtt{w}_1 + \mathtt{v}_{j_1} + \mathtt{v}_{j_2}$ while the quadratic relations read $(\mathtt{v}_{j_1} - \mathtt{w}_1, \mathtt{v}_{j_2} - \mathtt{w}_1) = 0$. Let us consider $\mu \in \mathbb{Z}^m$ of the form

$$\mu = \sum_{j \in A}\xi_j \mathbf{e}_j + a\mathbf{c}_1 + b\mathbf{c}_2 \ , \quad \sum_{j \in A}\xi_j + a + b = 1 \ , \quad \sum_{j \in A}|\xi_j| + |a| + |b| \leq 2d - 1 \tag{3.22}$$

and study $K_{\mathcal{S}}(\mu)$.

Recall that $\mathtt{w}_2 = -\mathtt{w}_1 + \mathtt{v}_{j_1} + \mathtt{v}_{j_2}$ and $|\mathtt{w}_2|^2 = -|\mathtt{w}_1|^2 + |\mathtt{v}_{j_1}|^2 + |\mathtt{v}_{j_2}|^2$. We have

$$\begin{cases} \pi_{\mathcal{S}}(\mu) = \sum_{j \in A}\xi_j \mathtt{v}_j + (a-b)\mathtt{w}_1 + b(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}) \\ \pi_{\mathcal{S}}^{(2)}(\mu) = \sum_{j \in A}\xi_j |\mathtt{v}_j|^2 + (a-b)|\mathtt{w}_1|^2 + b(|\mathtt{v}_{j_1}|^2 + |\mathtt{v}_{j_2}|^2). \end{cases}$$

We set $\alpha := a - b$ and

$$\lambda := \sum_{j \in A}\lambda_j \mathbf{e}_j = \sum_{j \in A}\xi_j \mathbf{e}_j + b(\mathbf{e}_{j_1} + \mathbf{e}_{j_2})$$

so that

$$K_{\mathcal{S}}(\mu) = |\pi_{\mathcal{S}}(\lambda) + \alpha\mathtt{w}_1|^2 - \pi_{\mathcal{S}}^{(2)}(\lambda) - \alpha|\mathtt{w}_1|^2.$$

If $\alpha = 0$, then $a = b$ and therefore $\lambda - \mu = a(\mathbf{e}_{j_1} + \mathbf{e}_{j_2} - \mathbf{c}_1 - \mathbf{c}_2) \in \langle \mathcal{F} \rangle$ and $K_{\mathcal{S}}(\mu) = K_{\mathcal{S}}(\lambda)$. Moreover, we have that $\sum_j \lambda_j = 1$ and $\sum_j |\lambda_j| \leq 2d - 1$. Since the support of λ is contained in A, the non-degeneracy condition (ii) of Definition 3.12 for the vector μ follows from the inductive hypothesis.

Otherwise, assume $\alpha \neq 0$. Then, since $|\mathtt{w}_1|^2 = (\mathtt{v}_{j_1} + \mathtt{v}_{j_2}, \mathtt{w}_1) - (\mathtt{v}_{j_1}, \mathtt{v}_{j_2})$, we get

$$\begin{aligned} K_{\mathcal{S}}(\mu) &= |\pi_{\mathcal{S}}(\lambda)|^2 + \alpha^2 |\mathtt{w}_1|^2 + 2\alpha(\pi_{\mathcal{S}}(\lambda), \mathtt{w}_1) - \pi_{\mathcal{S}}^{(2)}(\lambda) - \alpha|\mathtt{w}_1|^2 \\ &= K_{\mathcal{S}}(\lambda) + 2\alpha(\pi_{\mathcal{S}}(\lambda), \mathtt{w}_1) + \alpha(\alpha-1)[(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}, \mathtt{w}_1) - (\mathtt{v}_{j_1}, \mathtt{v}_{j_2})] \\ &= K_{\mathcal{S}}(\lambda) + \alpha\Big(2\pi_{\mathcal{S}}(\lambda) + (\alpha-1)(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}), \mathtt{w}_1\Big) - \alpha(\alpha-1)(\mathtt{v}_{j_1}, \mathtt{v}_{j_2}). \end{aligned}$$

If $2\pi_{\mathcal{S}}(\lambda) + (\alpha-1)(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}) \neq 0$ then $K_{\mathcal{S}}(\mu) = 0$ defines a line in the plane $\mathtt{w}_1 \in \mathbb{R}^2$. Then the non-degeneracy condition amounts to fixing $\mathtt{w}_1$ so that $K_{\mathcal{S}}(\mu) \neq 0$ i.e. by excluding at most two points on the circle $(\mathtt{w}_1 - \mathtt{v}_{j_1}, \mathtt{w}_1 - \mathtt{v}_{j_2}) = 0$.

Suppose now that $2\pi_{\mathcal{S}}(\lambda) + (\alpha-1)(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}) = 0$, then $K_{\mathcal{S}}(\mu)$ does not depend on the choice of $\mathtt{w}_1$. We have to show that either $K_{\mathcal{S}}(\mu) \neq 0$ for all $\mathcal{S} \in \mathcal{M} \setminus \mathcal{D}_0$ or we get the special case allowed in Definition 3.12 (ii). We claim that[1]

$$\eta := 2\lambda + (\alpha-1)(\mathbf{e}_{j_1} + \mathbf{e}_{j_2}) \in \mathcal{R}_{2d}.$$

Indeed, since $\sum_j \lambda_j + \alpha = 1$ then $\sum_j \eta_j = 0$. Moreover $\eta = 2\sum_{j\in A} \xi_j \mathbf{e}_j + (a+b-1)(\mathbf{e}_{j_1} + \mathbf{e}_{j_2})$ which, by (3.22), implies

$$\sum_j |\eta_j| \leq 2(\sum_j |\xi_j| + |a| + |b| + 1) \leq 4d.$$

Now by definition for all $\mathcal{S} \in \mathcal{M} \setminus \mathcal{D}_0$, we have $\pi_{\mathcal{S}}(\eta) = 0$ if and only if $\eta \in \langle \mathcal{F} \rangle$. This in turn implies that not only

$$\pi_{\mathcal{S}}(\eta) = 2\pi_{\mathcal{S}}(\lambda) + (\alpha-1)(\mathtt{v}_{j_1} + \mathtt{v}_{j_2}) = 0$$

but also (see Remark 3.10)

$$\pi_{\mathcal{S}}^{(2)}(\eta) = 2\pi_{\mathcal{S}}^{(2)}(\lambda) + (\alpha-1)(|\mathtt{v}_{j_1}|^2 + |\mathtt{v}_{j_2}|^2) = 0.$$

Hence

$$K_{\mathcal{S}}(\lambda) = \frac{(\alpha-1)^2}{4}|\mathtt{v}_{j_1} + \mathtt{v}_{j_2}|^2 + \frac{\alpha-1}{2}(|\mathtt{v}_{j_1}|^2 + |\mathtt{v}_{j_2}|^2)$$

and in conclusion

$$K_{\mathcal{S}}(\mu) = \frac{(\alpha-1)(\alpha+1)}{4}|\mathtt{v}_{j_1} - \mathtt{v}_{j_2}|^2.$$

We have that $\mathbf{e}_{j_1} - \mathbf{e}_{j_2} \in \mathcal{R}_{2d}$ and Lemma 3.11 (ii) implies $\mathbf{e}_{j_1} - \mathbf{e}_{j_2} \notin \langle \mathcal{F} \rangle$. Therefore, for $\mathcal{S} \in \mathcal{M} \setminus \mathcal{D}_0$, we have $\mathtt{v}_{j_1} \neq \mathtt{v}_{j_2}$ (see Remark 3.23). Then $K_{\mathcal{S}}(\mu)$ vanishes on $\mathcal{M} \setminus \mathcal{D}_0$ only if $\alpha = \pm 1$.

If $\alpha = 1$ then $\lambda \in \mathcal{R}_{2d}$ and $\pi_{\mathcal{S}}(\lambda) = 0$, which holds true in $\mathcal{M} \setminus \mathcal{D}_0$ if only if $\lambda \in \langle \mathcal{F} \rangle$; then

$$\mu - \mathbf{c}_1 = \lambda - b(\mathbf{e}_{j_1} + \mathbf{e}_{j_2} - \mathbf{c}_1 - \mathbf{c}_2) \in \langle \mathcal{F} \rangle$$

and the non-degeneracy condition (ii) in Definition 3.12 holds.

If $\alpha = -1$, then one symmetrically defines

$$\tilde{\lambda} := \sum_{j\in A} \tilde{\lambda}_j \mathbf{e}_j = \sum_{j \in A} \xi_j \mathbf{e}_j + a(\mathbf{e}_{j_1} + \mathbf{e}_{j_2})$$

and proceeding as above one obtains $\tilde{\lambda} \in \mathcal{R}_{2d}$ and $\pi_{\mathcal{S}}(\tilde{\lambda}) = 0$, which implies $\tilde{\lambda} \in \langle \mathcal{F} \rangle$; finally

$$\mu - \mathbf{c}_2 = \tilde{\lambda} - a(\mathbf{e}_{j_1} + \mathbf{e}_{j_2} - \mathbf{c}_1 - \mathbf{c}_2) \in \langle \mathcal{F} \rangle$$

which again ensures that the non-degeneracy condition (ii) holds. □

4 Dynamics of the toy model

4.1 Invariant subspaces for one and two generations

We now study the invariant subspaces of $H_{\mathcal{S}}$ where $\mathcal{S}$ is an acceptable set, see Definition 2.4. The simplest non trivial orbits are those where we fix $j = 1, \dots, N$ and set $b_i = 0$ for all $i \neq j$. This is an invariant subspace by Definition 2.4 item 4(d). By gauge invariance and reality (resp. items 4(c) and 4(b) of Definition 2.4)

[1]This motivates our choice of $\mathcal{R}_{2d}$ in Definition 3.12.

we have that the Hamiltonian restricted to this subspace is a single monomial $|b_j|^{2d}$ with a real non-zero coefficient. Moreover, this coefficient does not depend on j (it follows, for instance, by Definition 2.4 item 4(f)). We have proved, for any fixed surface level of the mass, the existence of N periodic orbits all with the same frequency. We denote by $\mathbb{T}_j$ the corresponding periodic orbit with $|b_j| = 1$ and $b_i = 0$ for $i \neq j$.

We can now suppose that all the b_j's are zero except two consecutive ones. By Definition 2.4 item 4(e) we can restrict ourselves to the case when these two generations are the first and the second. Thus, we get the Hamiltonian

$$h(b_1, \bar{b}_1, b_2, \bar{b}_2) = (|b_1|^2 + |b_2|^2)^{d-2}\Big(-\frac{1}{4}(|b_1|^4 + |b_2|^4) + \mathrm{Re}\,(b_1^2\bar{b}_2^2)\Big) + \frac{1}{n}\mathcal{P}(b_1, \bar{b}_1, b_2, \bar{b}_2, \frac{1}{n}).$$

with the constant of motion $J = |b_1|^2 + |b_2|^2$. We know by Definition 2.4 item 4 that h is symmetric with respect to the exchange of b_1 and b_2, real and gauge invariant. Moreover it has even degree in its variables. Thus we can have only a finite number of possible fundamental *building blocks* which appear through sums and products:

1. Integrable terms $L^{(j)} := |b_1|^{2j} + |b_2|^{2j}$;
2. Non-integrable terms which are multiples of a family relation $\mathrm{Re}\,[(b_1\bar{b}_2)^{2k}]$.

We remark that all the integrable terms $L^{(j)}$ can be written in terms of $|b_1|^2|b_2|^2$ and J in the same way the non integrable terms are polynomials in $|b_1|^2|b_2|^2$ and $\mathrm{Re}\,[(b_1\bar{b}_2)^2]$. We now reduce the degrees of freedom passing to one complex variable c, one (cyclic) angle ϑ and the conserved quantity J. Explicitly we have

$$J = |b_1|^2 + |b_2|^2,\ b_1 = \sqrt{J - |c|^2}e^{\mathrm{i}\vartheta}\ ,\ b_2 = ce^{\mathrm{i}\vartheta}. \tag{4.1}$$

This change of variables is symplectic and the new symplectic form is $\frac{1}{2}(dJ \wedge d\vartheta + \mathrm{i}dc \wedge d\bar{c})$. Note that, writing both $|b_1|^2|b_2|^2$ and $\mathrm{Re}\,[(b_1\bar{b}_2)^2]$ in terms of the new variables, the term $J - |c|^2$ always factors out. This means that, subtracting the constant terms depending only on J, we get the Hamiltonian

$$h(J, c, \bar{c}) = (J - |c|^2)\Big(J^{d-2}\big(\frac{1}{2}|c|^2 + \mathrm{Re}\,(c^2)\big) + \frac{1}{n}Q(J, |c|^2, \mathrm{Re}\,(c^2))\Big),$$

where the dependence of Q on its arguments is polynomial and homogeneous of degree $d-1$ and we have $Q(J, 0, 0) = 0$. Now we may extract the linear terms in $|c|^2, \mathrm{Re}\,(c^2)$ from Q (note that for the quintic NLS $d = 3$ these are the only possible terms) and restricting to the surface level $J = 1$ we get an expression of the form

$$h(c, \bar{c}) = \kappa_n(1 - |c|^2)\Big(a_n|c|^2 + \mathrm{Re}\,(c^2) + \frac{1}{n}\mathcal{Q}(|c|^2, \mathrm{Re}\,(c^2))\Big)\ . \tag{4.2}$$

where

$$\mathcal{Q}(|c|^2, \mathrm{Re}\,(c^2)) = Q(1, |c|^2, \mathrm{Re}\,(c^2)) - \partial_2 Q(1, 0, 0)|c|^2 - \partial_3 Q(1, 0, 0)\mathrm{Re}\,(c^2).$$

Note that $\mathcal{Q}$ has a zero of order two in its variables $\kappa_n = 1 + \mathcal{O}(1/n)$ and $a_n = (\frac{1}{2} + \mathcal{O}(1/n))$. It is natural to pass the quadratic part of the Hamiltonian in hyperbolic normal form by defining

$$\mathrm{Re}\,(\omega^2) := -a_n\ , \quad \text{namely} \quad \omega = e^{\mathrm{i}\theta} \ \text{ with } \ \theta = \frac{1}{2}\arccos(-a_n) = \frac{\pi}{3} + \mathcal{O}(n^{-1}) \tag{4.3}$$

and setting

$$\begin{aligned} c &= \frac{1}{\sqrt{\mathrm{Im}\,(\omega^2)}}(\omega q + \bar{\omega}p) \\ \bar{c} &= \frac{1}{\sqrt{\mathrm{Im}\,(\omega^2)}}(\bar{\omega}q + \omega p). \end{aligned} \tag{4.4}$$

Lemma 4.1. *The change of variables given by* (4.4) *is symplectic i.e.* $\frac{\mathrm{i}}{2}dc\wedge d\bar{c} = dp\wedge dq$ *and the Hamiltonian in the new variables is given by*

$$h(p, q) = \kappa_n\Big(1 - \frac{1}{\mathrm{Im}\,(\omega^2)}(p^2 + q^2 + 2\mathrm{Re}\,(\omega^2)pq)\Big)\Big(2\mathrm{Im}\,(\omega^2)pq + \frac{1}{n}P(pq, p^2 + q^2)\Big) \tag{4.5}$$

with P having a zero of degree at least two in its arguments.

Proof. We have:

$$|c|^2 = \frac{1}{\mathrm{Im}\,(\omega^2)}(p^2+q^2+2\mathrm{Re}\,(\omega^2)pq)$$
$$\mathrm{Re}\,(c^2) = \frac{1}{\mathrm{Im}\,(\omega^2)}(\mathrm{Re}\,(\omega^2)(p^2+q^2)+2pq)$$

which imply

$$-\mathrm{Re}\,(\omega^2)|c|^2+\mathrm{Re}\,(c^2) = \frac{2}{\mathrm{Im}\,(\omega^2)}(1-\mathrm{Re}\,(\omega^2)^2)pq = 2\mathrm{Im}\,(\omega^2)pq$$

and hence substituting into (4.2) we get the thesis. □

The flow generated by the Hamiltonian (4.5) leaves invariant the ellipse $\mathcal{E}$ with equation $p^2+q^2+2\mathrm{Re}\,(\omega^2)pq = \mathrm{Im}\,(\omega^2)$, which corresponds to the periodic orbit $\mathbb{T}_2$, while the hyperbolic fixed point $(0,0)$ corresponds to the periodic orbit $\mathbb{T}_1$. We are going to prove the existence of heteroclinic connections linking $(0,0)$ to a point in $\mathcal{E}$, i.e. sliding from the periodic orbit $\mathbb{T}_1$ to the periodic orbit $\mathbb{T}_2$.

Neglecting the term $(1/n)P(pq,p^2+q^2)$ in (4.5), one easily sees that there is a heteroclinic connection lying on $q=0$, flowing from the point $(0,0)$ as $t\to-\infty$ to the hyperbolic critical point $(p,q)=(\mathrm{Im}\,(\omega^2),0)$ on $\mathcal{E}$ as $t\to+\infty$. Now, we can deal with the full system using perturbative methods.

Lemma 4.2. *The Hamiltonian system given by (4.5) has a hyperbolic critical point $(p^*,q^*)=(\sqrt{3}/2,0)+\mathcal{O}(n^{-1})$, which belongs to $\mathcal{E}$, and a heteroclinic connection which tends to this point in forward time and to the point $(p,q)=(0,0)$ in backward time. Moreover, this connection can be written as a graph*

$$q=\xi(p),\ \ p\in[0,p^*]$$

and it satisfies $\sup_{p\in[0,p^*]}|\xi(p)|=\mathcal{O}(n^{-1})$.

The proof of this lemma is straightforward.

Remark 4.3. *Since the Hamiltonian (4.5) is symmetric in (p,q), then there is also the hyperbolic critical point $(p,q)=(q^*,p^*)=(0,\sqrt{3}/2)+\mathcal{O}(n^{-1})$, which belongs to $\mathcal{E}$, and a heteroclinic connection which tends to this point in backward time and to the point $(p,q)=(0,0)$ in forward time. Such heteroclinic connection can be written as a graph*

$$p=\xi(q),\ \ q\in[0,p^*].$$

4.2 Adapted coordinates for the j-th periodic orbit

In this section we study the dynamics of the toy models

$$h(b)=\left(\sum_{i=1}^N|b_i|^2\right)^{d-2}\left[-\frac{1}{4}\sum_{i=1}^N|b_i|^4+\sum_{i=1}^{N-1}\mathrm{Re}\,(b_i^2\bar{b}_{i+1}^2)\right]+\frac{1}{n}\mathcal{P}\left(b,\bar{b},\frac{1}{n}\right).$$

Following [CKS+10], we will take advantatge of the mass conservation to make a symplectic reduction. This will allow us to obtain certain good systems of coordinates.

To make the symplectic reduction we fix the mass $\mathcal{M}(b)=1$. Note that the toy model is invariant by certain rescaling and time reparameterization. So, from orbits in $\mathcal{M}(b)=1$ we can obtain orbits for any mass. Now, we perform the change of coordinates close to the j periodic orbit

$$\left(b_1,\bar{b}_1,\dots,b_N,\bar{b}_N\right)\mapsto\left(c_1^{(j)},\bar{c}_1^{(j)},\dots,J,\theta^{(j)},\dots,c_N^{(j)},\bar{c}_N^{(j)}\right)$$

defined by

$$b_j=\sqrt{J-\sum_{k\neq j}\left|c_k^{(j)}\right|^2}\,e^{\mathrm{i}\theta^{(j)}},\ \ b_k=c_k^{(j)}e^{\mathrm{i}\theta^{(j)}}\ \ \text{for all}\ \ k\neq j, \tag{4.6}$$

where $\theta_j^{(j)}$ is the angular variable over the periodic orbit and $J=\sum_{k=1}^N|b_k|^2$ is the mass. It can be checked that this change of coordinates is symplectic. From now we omit the superscript (j) when it is clear in the neighborhood of which saddle we are dealing with. The new Hamiltonian is independent of θ since the mass

J is a first integral. Fixing $J=1$, the system for the variables $c=(c_1,\dots,c_{j-1},c_{j+1},\dots c_N)$ is Hamiltonian with respect to

$$
\begin{aligned}
H^{(j)}(c) = &-\frac{1}{4}\sum_{k\neq j}|c_k|^4 - \frac{1}{4}\left(1-\sum_{k\neq j}|c_k|^2\right)^2 + \sum_{k\neq j,j+1}\mathrm{Re}\,(c_k^2\bar{c}_{k-1}^2)\\
&+\left(1-\sum_{k\neq j}|c_k|^2\right)\mathrm{Re}\,\left(c_{j-1}^2+c_{j+1}^2\right)+\frac{1}{n}\widetilde{\mathcal{P}}\left(c,\bar{c},\frac{1}{n}\right)
\end{aligned}
\tag{4.7}
$$

and the symplectic form $\Omega=\sum_{k\neq j}\frac{\mathrm{i}}{2}dc_k\wedge d\bar{c}_k$, where $\widetilde{\mathcal{P}}$ is the polynomial $\mathcal{P}$ introduced in Theorem 3.14 expressed in the new variables c. The Hamiltonian system can be split as

$$H^{(j)}(c)=H_2^{(j)}(c)+H_4^{(j)}(c)$$

where $H_2^{(j)}(c)$ contains the quadratic monomials and $H_4^{(j)}(c)$ contains the higher order terms, that is monomials of even degree from 4 to $2d$. Statements 4(e) and 4(f) of Definition 2.4 imply the following lemma.

Lemma 4.4. *The Hamiltonian $H_2^{(j)}(c)$ is of the form*

$$H_2^{(j)}(c)=a_n\sum_{k\neq j}|c_k|^2+a_n\kappa_n\mathrm{Re}\,\left(c_{j-1}^2+c_{j+1}^2\right)$$

where $a_n=1/2+\mathcal{O}(n^{-1})$, $\kappa_n=1+\mathcal{O}(n^{-1})$.

Being close to $\mathbb{T}_j$ corresponds to $c\sim 0$. To analyze this local behavior, we diagonalize the linear part at the critical point. Note that for c_k, $k\neq j-1,j+1$ it is already diagonalized so we apply the change of variables (4.4) to the adjacent modes $c_{j\pm1}$. We obtain the new quadratic part of the Hamiltonian

$$H_2^{(j)}(p,q,c)=a_n\sum_{k\in\mathcal{P}_j}|c_k|^2+\lambda_n\left(p_1q_1+p_2q_2\right),\quad \lambda_n=2\;\mathrm{Im}(\omega^2)=2\sqrt{1-a_n^2}=\sqrt{3}+\mathcal{O}(n^{-1}). \tag{4.8}$$

Note that this change of coordinates transform Ω into the symplectic form

$$\Omega=\sum_{k\neq j-1,j,j+1}\frac{\mathrm{i}}{2}dc_k\wedge d\bar{c}_k+dp_1\wedge dq_1+dp_2\wedge dq_2.$$

To study the Hamiltonian expressed in the new variables we introduce

$$\mathcal{P}_j=\{1\leq k\leq N; k\neq j-1,j,j+1\},$$

which is the set of subindexes of the elliptic modes. From now on we will denote by q and p all the stable and unstable coordinates $q=(q_1,q_2)$ and $p=(p_1,p_2)$ respectively and by c all the elliptic modes, namely c_k with $k\in\mathcal{P}_j$.

Lemma 4.5. *The change* (4.4) *transforms the Hamiltonian* (4.7) *into the Hamiltonian*

$$H^{(j)}(p,q,c)=H_2^{(j)}(p,q,c)+H_4^j(p,q,c) \tag{4.9}$$

with homogeneous polynomials $H_2^{(j)}(p,q,c)$ given by (4.8) *and*

$$H_4^{(j)}(p,q,c)=H_{\mathrm{hyp}}^{(j)}(p,q)+H_{\mathrm{ell}}^{(j)}(c)+H_{\mathrm{mix}}^{(j)}(p,q,c)$$

where

$$\begin{aligned}
H_{\text{hyp}}^{(j)}(p,q) =& -2p_1q_1(p_1^2+q_1^2-p_1q_1)-2p_2q_2(p_2^2+q_2^2-p_2q_2)\\
&+\sum_{k,\ell=0}^{2}\nu_{k\ell}p_1^kq_1^{2-k}p_2^\ell q_2^{2-\ell}+\mathcal{O}\left(\frac{1}{n}(p_1+q_1+p_2+q_2)^4\right)\\
H_{\text{ell}}^{(j)}(c) =& -\frac{1}{4}\sum_{k\in\mathcal{P}_j}|c_k|^4-\frac{1}{4}\left(\sum_{k\in\mathcal{P}_j}|c_k|^2\right)^2\\
&+\sum_{k\in\mathcal{P}_j\setminus\{j+2\}}\operatorname{Re}\left(c_k^2\overline{c_{k-1}}^2\right)+\mathcal{O}\left(\frac{1}{n}\sum_{k,k'\in\mathcal{P}_j}|c_k|^2|c_{k'}|^2\right)\\
H_{\text{mix}}^{(j)}(p,q,c) =& -\sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2(q_1p_1+q_2p_2)\\
&+\frac{2\sqrt{3}}{3}\operatorname{Re}\left((\omega_0p_1+\overline{\omega}_0q_1)^2c_{j-2}^2\right)+\frac{2\sqrt{3}}{3}\operatorname{Re}\left((\omega_0p_2+\overline{\omega}_0q_2)^2c_{j+2}^2\right)\\
&+\mathcal{O}\left(\frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2(p_1+q_1+p_2+q_2)^2\right)
\end{aligned}$$

for some constants $\nu_{k\ell}\in\mathbb{R}$ *and* $\omega_0=e^{\mathrm{i}\frac{\pi}{3}}$. *It can be easily checked that all* $\nu_{k\ell}$ *satisfy* $\nu_{k\ell}\neq 0$.

Now for the Hamiltonian (4.9), the periodic $\mathbb{T}_j$ has become the critical point $(p,q,c)=(0,0,0)$ which is of mixed type (four hyperbolic eigenvalues and $2N-6$ elliptic eigenvalues). Thanks to Lemma 4.2 and the particular form of the Hamiltonian (4.9) the hyperbolic directions give connections to the neighboring periodic orbits $\mathbb{T}_{j\pm1}$. In the full phase space the heteroclinic connection between $(0,0,0)$ and $\mathbb{T}_{j+1}$ can be parameterized as a graph by

$$(p_1,q_1,p_2,q_2,c)=(0,0,p_2,\xi(p_2),0).$$

Recall that in the cubic case, this connection is just given by $c_k=q_1=p_1=q_2=0$ (see [CKS+10]).

Following [CKS+10, GK15] we look for orbits which shadow this concatenation of heteroclinic orbits.

4.3 The iterative argument: almost product structure

To prove Theorem 2.8 and shadow the concatenation of heteroclinic connections, we follow the approach in [GK15]. That is, we consider several co-dimension one sections $\{\Sigma_j^{\text{in}}\}_{j=1}^N$ and transition maps $\mathcal{B}^j$ from one section Σ_j^{in} to the next one Σ_{j+1}^{in}. The maps are given by the flow associated to the Hamiltonian (4.9). We consider sets $\{\mathcal{V}_j\}_j$, $\mathcal{V}_j\subset\Sigma_j^{\text{in}}$, $j=1,\dots,N-1$, of a very particular form, which in [GK15] were called sets with *almost product structure* (see Definition 4.7 below). Moreover, we impose that these sets satisfy that $\mathcal{V}_{j+1}\subset\mathcal{B}^j(\mathcal{V}_j)$ and that none of them is empty. Each set $\mathcal{V}_j$ is located close to the stable manifold of the periodic orbit $\mathbb{T}_j$. Composing all these maps we will be able to find orbits claimed to exist in Theorem 2.8. Note that these sets will be slightly different from the ones in [GK15] due to the deviation of the heteroclinic connections given in Lemma 4.2.

In this section we keep the superindexes (j) in the variables since it involves two consecutive adapted system of coordinates. We start by defining transversal sections to the flow. We use the coordinates adapted to the saddle j, $(p^{(j)},q^{(j)},c^{(j)})$ to define these sections. In Lemma 4.2 we have seen that the heteroclinic connections which connect $(p^{(j)},q^{(j)},c^{(j)})=(0,0,0)$ with the previous and next saddles are $\mathcal{O}(n^{-1})$ close to $(p_1^{(j)},p_2^{(j)},q_2^{(j)},c^{(j)})=(0,0,0,0)$ and $(p_1^{(j)},q_1^{(j)},q_2^{(j)},c^{(j)})=(0,0,0,0)$ respectively. Thus, we define the map $\mathcal{B}^j$ from the section

$$\Sigma_j^{\text{in}}=\left\{q_1^{(j)}=\sigma\right\} \tag{4.10}$$

to the section

$$\Sigma_{j+1}^{\text{in}}=\left\{q_1^{(j+1)}=\sigma\right\}.$$

Here $\sigma>0$ is a small parameter that will be determined later on. We do not define the map $\mathcal{B}^j$ in the whole section but in a set $\mathcal{V}_j\subset\Sigma_j^{\text{in}}$, which lies close to the heteroclinic that connects the saddle $j-1$ to the saddle j. Then,

$$\mathcal{B}^j:\mathcal{V}_j\subset\Sigma_j^{\text{in}}\to\Sigma_{j+1}^{\text{in}}$$

and we choose the sets $\mathcal{V}_j$ recursively in such a way that

$$\mathcal{V}_{j+1} \subset \mathcal{B}^j(\mathcal{V}_j). \tag{4.11}$$

This condition allows us to compose all the maps $\mathcal{B}^j$.

The sets $\mathcal{V}_j$ will have a product-like structure as is stated in the next definition, introduced in [GK15]. Before stating it, we fix j and introduce notations adapted to j.

Definition 4.6. *We call b_j the* primary mode, *$b_{j\pm1}$* secondary modes, *$b_{j\pm2}$* adjacent modes *and all the others* peripheral modes. *If $k < j$ we say that b_k is a* trailing mode *while if $k > j$ we say that b_k is a* leading mode. *Finally we set*

$$\mathcal{P}_j^- = \{k = 1, \dots, j-3\} \qquad \mathcal{P}_j^+ = \{k = j+3, \dots, N\}, \quad \mathcal{P}_j = \mathcal{P}_j^- \cup \{j \pm 2\} \cup \mathcal{P}_j^+.$$

For a point $(p^{(j)}, q^{(j)}, c^{(j)}) \in \Sigma_j^{\text{in}}$, we define $c_-^{(j)} = (c_1^{(j)}, \dots, c_{j-2}^{(j)})$ and $c_+^{(j)} = (c_{j+2}^{(j)}, \dots, c_N^{(j)})$. We define also the projections $\pi_\pm(p^{(j)}, q^{(j)}, c^{(j)}) = c_\pm^{(j)}$ and $\pi_{\text{hyp},+} = (p^{(j)}, q^{(j)}, c_+^{(j)})$.

Definition 4.7. *Fix positive constants $r \in (0,1)$, δ and σ and consider a multi-parameter set of positive constants*

$$\mathcal{I}_j = \left\{C^{(j)}, m_{\text{ell}}^{(j)}, M_{\text{ell},\pm}^{(j)}, m_{\text{adj}}^{(j)}, M_{\text{adj},\pm}^{(j)}, m_{\text{hyp}}^{(j)}, M_{\text{hyp}}^{(j)}\right\}. \tag{4.12}$$

We associate to the set $\mathcal{I}_j$ a smooth function $g_j(p_2, q_2) = g_{\mathcal{I}_j}(p_2, q_2)$, which is defined in (7.5).

Then, we say that a (non-empty) set $\mathcal{U} \subset \Sigma_j^{\text{in}}$ has an $\mathcal{I}_j$-product-like structure if it satisfies the following two conditions:

C1

$$\mathcal{U} \subset \mathbb{D}_j^1 \times \dots \times \mathbb{D}_j^{j-2} \times \mathcal{N}_j^+ \times \mathbb{D}_j^{j+2} \times \dots \times \mathbb{D}_j^N,$$

where

$$\mathbb{D}_j^k = \left\{\left|c_k^{(j)}\right| \le M_{\text{ell},\pm}^{(j)} \delta^{(1-r)/2}\right\} \quad \textit{for } k \in \mathcal{P}_j^\pm$$
$$\mathbb{D}_j^{j\pm2} \subset \left\{\left|c_{j\pm2}^{(j)}\right| \le M_{\text{adj},\pm}^{(j)} \left(C^{(j)}\delta\right)^{1/2}\right\}$$

and

$$\begin{aligned}\mathcal{N}_j^+ = \Big\{ &\left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}\right) \in \mathbb{R}^4 : \\ \xi(\sigma) - C^{(j)}\delta\left(\ln(1/\delta) + M_{\text{hyp}}^{(j)}\right) &\le p_1^{(j)} \le \xi(\sigma) - C^{(j)}\delta\left(\ln(1/\delta) - M_{\text{hyp}}^{(j)}\right), \\ q_1^{(j)} = \sigma,\ &g_j(p_2^{(j)}, q_2^{(j)}) = 0,\ |p_2^{(j)}|, |q_2^{(j)}| \le M_{\text{hyp}}^{(j)}\left(C^{(j)}\delta\right)^{1/2}\Big\}.\end{aligned} \tag{4.13}$$

C2

$$\mathcal{N}_j^- \times \mathbb{D}_{j,-}^{j+2} \times \dots \times \mathbb{D}_{j,-}^N \subset \pi_{\text{hyp},+}\mathcal{U},$$

where

$$\mathbb{D}_{j,-}^k = \left\{\left|c_k^{(j)}\right| \le m_{\text{ell}}^{(j)} \delta^{(1-r)/2}\right\} \quad \textit{for } k \in \mathcal{P}_j^+$$
$$\mathbb{D}_{j,-}^{j+2} = \left\{\left|c_{j+2}^{(j)}\right| \le m_{\text{adj}}^{(j)} \left(C^{(j)}\delta\right)^{1/2}\right\}$$

and

$$\begin{aligned}\mathcal{N}_j^- = \Big\{ &\left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}\right) \in \mathbb{R}^4 : \\ \xi(\sigma) - C^{(j)}\delta\left(\ln(1/\delta) + m_{\text{hyp}}^{(j)}\right) &\le p_1^{(j)} \le \xi(\sigma) - C^{(j)}\delta\left(\ln(1/\delta) - m_{\text{hyp}}^{(j)}\right), \\ q_1^{(j)} = \sigma,\ &g_j(p_2^{(j)}, q_2^{(j)}) = 0,\ |p_2^{(j)}|, |q_2^{(j)}| \le m_{\text{hyp}}^{(j)}\left(C^{(j)}\delta\right)^{1/2}\Big\}.\end{aligned} \tag{4.14}$$

where ξ is the function introduced in Lemma 4.2.

If one compares this definition to the one in [GK15] the only difference appears in the $p_1^{(j)}$ variable. The reason is the deviation of the separatrix connection. The domains $\mathcal{V}_j$ of the maps $\mathcal{B}^j$ will have $\mathcal{I}_j$-product-like structure as defined in Definition 4.7. Thus, we need to obtain the multi-parameter sets $\mathcal{I}_j$. They will be defined recursively. In Theorem 2.8 we are looking for an orbit which starts close to the periodic orbit $\mathbb{T}_3$, thus the recursively defined multi-parameter sets $\mathcal{I}_j$ will start with a set $\mathcal{I}_3$.

Definition 4.8. *Fix $\gamma > 0$, any constants $r, r' \in (0,1)$ satisfying $r < \ln 2/(2\gamma)$ and $0 < r' < \ln 2/\gamma - 2r$, $K > 0$ and small $\delta, \sigma > 0$. We say that a collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\ldots,N-2}$ defined in (4.12) is (σ, δ, K)-recursive if for $j = 3, \ldots, N-2$ the constants $C^{(j)}$ satisfy*

$$\begin{aligned} C^{(j)}/K \leq C^{(j+1)} &\leq KC^{(j)} \\ 0 < m_{\rm hyp}^{(j+1)} &\leq m_{\rm hyp}^{(j)} \end{aligned}$$

and all the other parameters should be strictly positive and are defined recursively as

$$\begin{aligned} M_{\rm ell,\pm}^{(j+1)} &= M_{\rm ell,\pm}^{(j)} + K\delta^{r'} \\ m_{\rm ell}^{(j+1)} &= m_{\rm ell}^{(j)} - K\delta^{r'} \\ M_{\rm adj,+}^{(j+1)} &= 2M_{\rm ell,+}^{(j)} + K\delta^{r'} \\ M_{\rm adj,-}^{(j+1)} &= KM_{\rm hyp}^{(j)} \\ m_{\rm adj}^{(j+1)} &= \frac{1}{2}m_{\rm ell}^{(j)} - K\delta^{r'} \\ M_{\rm hyp}^{(j+1)} &= KM_{\rm adj,+}^{(j)}. \end{aligned}$$

This definition coincides with the definition of [GK15]. The only difference appears in the conditions on the parameters r an r'. The reason is that one needs to take into account the $\mathcal{O}(n^{-1})$ terms in the Hamiltonian (4.9).

The next theorem defines recursively the product-like sets $\mathcal{V}_j$, so that condition (4.11) is satisfied.

Theorem 4.9 (Iterative Theorem)**.** *Fix a large $\gamma > 0$, a small $\sigma > 0$, two constants $r, r' \in (0,1)$ satisfying $r < \ln 2/(2\gamma)$, $0 < r' < \ln 2/\gamma - 2r$ and set $\delta = e^{-\gamma N}$. There exist strictly positive constants K and $C^{(3)}$ independent of N satisfying*

$$\delta^r K^{N-3} \leq C^{(3)} \leq \delta^{-r} K^{-(N-3)}, \tag{4.15}$$

and a multi-parameter set $\mathcal{I}_3$ (as defined in (4.12)) with the following property: there exists a (σ, δ, K)-recursive collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\ldots,N-2}$ and $\mathcal{I}_j$-product-like sets $\mathcal{V}_j \subset \Sigma_j^{\rm in}$ such that for each $j = 3, \ldots, N-3$ we have

$$\mathcal{V}_{j+1} \subset \mathcal{B}^j(\mathcal{V}_j).$$

Moreover, the time spent to reach the section $\Sigma_{j+1}^{\rm in}$ can be bounded by

$$|T_{\mathcal{B}^j}| \leq K \ln(1/\delta)$$

for any $(p, q, c) \in \mathcal{V}_j$ and any $j = 3, \ldots, N-3$.

The condition

$$C^{(j)}/K < C^{(j+1)} < KC^{(j)}$$

implies

$$K^{-(j-2)}C^{(3)} \leq C^{(j+1)} \leq K^{j-2}C^{(3)}.$$

Namely, at each saddle, the orbits we are studying may lie further from the heteroclinic orbit. Nevertheless, since $\delta = e^{-\gamma N}$ and (4.15), these constant does not grow too much. Indeed,

$$\delta^r \leq C^{(j)} \leq \delta^{-r}, \tag{4.16}$$

where $r > 0$ is taken small. We use the bound (4.16) throughout the proof of Theorem 4.9.

Theorem 2.8 is a straightforward consequence of Theorem 4.9.

Proof of 2.8. It is enough to take as a initial condition b^0 a point in the set $\mathcal{V}_3 \subset \Sigma_3^{\rm in}$ obtained in Theorem 4.9. Then, thanks to this theorem we know that there exists a time T_0 satisfying

$$T_0 \sim N \ln(1/\delta),$$

such that the corresponding orbit satisfies that $b(T_0) \in \mathcal{V}_{N-2} \subset \Sigma_{N-2}^{\rm in}$. Note that in this section there are two components of b with size independent of δ. Nevertheless, from the proof of Theorem 4.9 in Section 7 it can be easily seen that if we shift the time interval $[0, T_0]$ to $[\rho\ln(1/\delta), \rho\ln(1/\delta) + T_0]$, for any $\rho < \lambda$ independent of n, there exists $\nu > 0$ such that the orbit $b(t)$ satisfies the statements given in Theorem 2.8 $\square$

5 Proof of Theorem 4.9: local and global maps

To prove Theorem 4.9 we proceed as in [GK15] and split it into two inductive lemmas. The first part analyzes the evolution of the trajectories close to the saddle j and the second one the travel along the heteroclinic orbit. Thus, we study $\mathcal{B}^j$ as a composition of two maps, which we call local and global map.

We consider an intermediate section transversal to the flow

$$\Sigma_j^{\rm out} = \left\{ p_2^{(j)} = \sigma \right\}. \tag{5.1}$$

Then, we consider the local map

$$\mathcal{B}_{\rm loc}^j : \mathcal{V}_j \subset \Sigma_j^{\rm in} \longrightarrow \Sigma_j^{\rm out}, \tag{5.2}$$

and the global map

$$\mathcal{B}_{\rm glob}^j : \mathcal{U}^j \subset \Sigma_j^{\rm out} \longrightarrow \Sigma_{j+1}^{\rm in}. \tag{5.3}$$

Then, the map $\mathcal{B}^j$ considered in Theorem 4.9 is just $\mathcal{B}^j = \mathcal{B}_{\rm glob}^j \circ \mathcal{B}_{\rm loc}^j$. To compose the two maps we need that the set $\mathcal{U}^j$, introduced in (5.3), has a modified product-like structure. To define its properties, we consider the projection

$$\widetilde{\pi}\left(c_-^{(j)}, p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c_+^{(j)}\right) = \left(p_2^{(j)}, q_2^{(j)}, c_+^{(j)}\right).$$

Definition 5.1. *Fix constants $r \in (0,1)$, $\delta > 0$ and $\sigma > 0$ and consider a multi-parameter set of positive constants*

$$\widetilde{\mathcal{I}}_j = \left\{ \widetilde{C}^{(j)}, \widetilde{m}_{\rm ell}^{(j)}, \widetilde{M}_{{\rm ell},\pm}^{(j)},\ \widetilde{m}_{\rm adj}^{(j)},\ \widetilde{M}_{{\rm adj},\pm}^{(j)},\ \widetilde{m}_{\rm hyp}^{(j)},\ \widetilde{M}_{\rm hyp}^{(j)} \right\}.$$

Then, we say that a (non-empty) set $\mathcal{U} \subset \Sigma_j^{\rm out}$ has a $\widetilde{\mathcal{I}}_j$-product-like structure provided it satisfies the following two conditions:

C1

$$\mathcal{U} \subset \widetilde{\mathbb{D}}_j^1 \times \ldots \times \widetilde{\mathbb{D}}_j^{j-2} \times \widetilde{\mathcal{N}}_j^+ \times \widetilde{\mathbb{D}}_j^{j+2} \times \ldots \times \widetilde{\mathbb{D}}_j^N$$

where

$$\widetilde{\mathbb{D}}_j^k = \left\{ \left| c_k^{(j)} \right| \leq \widetilde{M}_{{\rm ell},\pm}^{(j)} \delta^{(1-r)/2} \right\} \quad \textit{for } k \in \mathcal{P}_j^\pm$$
$$\widetilde{\mathbb{D}}_j^{j\pm2} \subset \left\{ \left| c_{j\pm2}^{(j)} \right| \leq \widetilde{M}_{{\rm adj},\pm}^{(j)} \left(\widetilde{C}^{(j)} \delta \right)^{1/2} \right\},$$

and

$$\widetilde{\mathcal{N}}_j^+ = \Big\{ (p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}) \in \mathbb{R}^4 : \left| p_1^{(j)} \right|, \left| q_1^{(j)} \right| \leq \widetilde{M}_{\rm hyp}^{(j)} \left(\widetilde{C}^{(j)} \delta \right)^{1/2},$$
$$p_2^{(j)} = \sigma, \xi(\sigma) - \widetilde{C}^{(j)}\, \delta \left(\ln(1/\delta) + \widetilde{M}_{\rm hyp}^{(j)} \right) \leq q_2^{(j)} \leq \xi(\sigma) - \widetilde{C}^{(j)}\, \delta \left(\ln(1/\delta) - \widetilde{M}_{\rm hyp}^{(j)} \right) \Big\},$$

C2

$$\{\sigma\} \times \left[\xi(\sigma) - \widetilde{C}^{(j)}\, \delta \left(\ln(1/\delta) + \widetilde{m}_{\rm hyp}^{(j)} \right), \xi(\sigma) - \widetilde{C}^{(j)}\, \delta \left(\ln(1/\delta) - \widetilde{m}_{\rm hyp}^{(j)} \right) \right] \times \widetilde{\mathbb{D}}_{j,-}^{j+2} \times \ldots \times \widetilde{\mathbb{D}}_{j,-}^N \subset \widetilde{\pi}(\mathcal{U})$$

where

$$\widetilde{\mathbb{D}}_{j,-}^k = \left\{ \left| c_k^{(j)} \right| \leq \widetilde{m}_{\rm ell}^{(j)} \delta^{(1-r)/2} \right\} \quad \textit{for } k \in \mathcal{P}_j^+$$
$$\widetilde{\mathbb{D}}_{j,-}^{j+2} = \left\{ \left| c_{j+2}^{(j)} \right| \leq \widetilde{m}_{\rm adj}^{(j)} \left(\widetilde{C}^{(j)} \delta \right)^{1/2} \right\}.$$

With this definition, we can state the following two lemmas. Combining these two lemmas we deduce Theorem 4.9.

Lemma 5.2. *Let $\gamma, \sigma, r, r', \delta$ be as in Theorem 4.9. Fix any natural j with $3 \le j \le N-3$ and consider any parameter set $\mathcal{I}_j$ with $M_{\mathrm{hyp}}^{(j)} \ge 1$ and a $\mathcal{I}_j$-product-like set $\mathcal{V}_j \subset \Sigma_j^{\mathrm{in}}$. Then, for N big enough, there exists:*

- *A constant $K > 0$ independent of N and j but which might depend on σ.*
- *A parameter set $\widetilde{\mathcal{I}}_j$ whose constants satisfy*

$$\begin{aligned} C^{(j)}/2 \le \widetilde{C}^{(j)} &\le 2C^{(j)} \\ 0 < \widetilde{m}_{\mathrm{hyp}}^{(j)} &\le m_{\mathrm{hyp}}^{(j)} \end{aligned}$$

and

$$\begin{aligned} \widetilde{M}_{\mathrm{hyp}}^{(j)} &= K \\ \widetilde{M}_{\mathrm{ell},\pm}^{(j)} &= M_{\mathrm{ell},\pm}^{(j)} + K\delta^{r'} \\ \widetilde{m}_{\mathrm{ell}}^{(j)} &= m_{\mathrm{ell}}^{(j)} - K\delta^{r'} \\ \widetilde{M}_{\mathrm{adj},\pm}^{(j)} &= M_{\mathrm{adj},\pm}^{(j)}(1+4\sigma) \\ \widetilde{m}_{\mathrm{adj}}^{(j)} &= m_{\mathrm{adj}}^{(j)}\ (1-4\sigma), \end{aligned}$$

- *A $\widetilde{\mathcal{I}}_j$-product-like set $\mathcal{U}_j$ for which the map $\mathcal{B}_{\mathrm{loc}}^j$ satisfies*

$$\mathcal{U}_j \subset \mathcal{B}_{\mathrm{loc}}^j\left(\mathcal{V}_j\right). \tag{5.4}$$

Moreover, the time to reach the section Σ_j^{out} can be bounded as

$$\left| T_{\mathcal{B}_{\mathrm{loc}}^j} \right| \le K \ln(1/\delta).$$

The proof of this lemma follows the same approach than the proof of Lemma 4.7 in [GK15]. First, in Section 6, we set the elliptic modes c to zero, and we study the saddle map associated to the corresponding system. We call this system *Hyperbolic Toy Model*. It has two degrees of freedom. As happens in [GK15], the saddle is resonant since both stable eigenvalues coincide. We used the ideas developed in [GK15] to overcome this problem. They are based on techniques developed by Shilnikov [Šil67]. Then, in Section 7 we use the results obtained for the Hyperbolic Toy Model to deal with the full system and prove Lemma 5.2.

Now we state the iterative lemma for the global maps $\mathcal{B}_{\mathrm{glob}}^j$.

Lemma 5.3. *Let $\gamma, \sigma, r, r', \delta$ be as in Theorem 4.9. Fix any natural j with $3 \le j \le N-3$ and consider any parameter set $\widetilde{\mathcal{I}}_j$ and a $\widetilde{\mathcal{I}}_j$-product-like set $\mathcal{U}_j \subset \Sigma_j^{\mathrm{out}}$. Then, for N large enough, there exists:*

- *A constant $\widetilde{K}$ depending on σ, but independent of N and j.*
- *A parameter set $\mathcal{I}_{j+1}$ whose constants satisfy*

$$\begin{aligned} \widetilde{C}^{(j)}/\widetilde{K} \le C^{(j+1)} &\le \widetilde{K}\widetilde{C}^{(j)} \\ 0 < m_{\mathrm{hyp}}^{(j+1)} &\le \widetilde{m}_{\mathrm{hyp}}^{(j)} \end{aligned}$$

and

$$\begin{aligned} M_{\mathrm{ell},-}^{(j+1)} &= \max\left\{\widetilde{M}_{\mathrm{ell},-}^{(j)} + \widetilde{K}\delta^{r'}, \widetilde{K}\widetilde{M}_{\mathrm{adj},-}^{(j)}\right\} \\ M_{\mathrm{ell},+}^{(j+1)} &= \widetilde{M}_{\mathrm{ell},+}^{(j)} + \widetilde{K}\delta^{r'} \\ m_{\mathrm{ell}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} - \widetilde{K}\delta^{r'} \\ M_{\mathrm{adj},+}^{(j+1)} &= \widetilde{M}_{\mathrm{ell},+}^{(j)} + \widetilde{K}\delta^{r'} \\ M_{\mathrm{adj},-}^{(j+1)} &= \widetilde{K}\widetilde{M}_{\mathrm{hyp}}^{(j)} \\ m_{\mathrm{adj}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} + \widetilde{K}\delta^{r'} \\ M_{\mathrm{hyp}}^{(j+1)} &= \max\left\{\widetilde{K}\widetilde{M}_{\mathrm{adj},+}^{(j)}, \widetilde{K}\right\} \end{aligned}$$

- *A $\mathcal{I}_{j+1}$-product-like set $\mathcal{V}_{j+1} \subset \Sigma_{j+1}^{\mathrm{in}}$ for which the map $\mathcal{B}_{\mathrm{glob}}^{j}$ satisfies*

$$\mathcal{V}_{j+1} \subset \mathcal{B}_{\mathrm{glob}}^{j}\left(\mathcal{U}_j\right). \tag{5.5}$$

Moreover, the time spent to reach the section $\Sigma_{j+1}^{\mathrm{in}}$ can be bounded as

$$\left|T_{\mathcal{B}_{\mathrm{glob}}^{j}}\right| \leq \widetilde{K}.$$

The proofs of this lemma is postponed to Section 8.

Now it only remains to deduce from Lemmas 5.2 and 5.3 the Iterative Theorem 4.9.

Proof of Theorem 4.9. We choose the multi-index $\mathcal{I}_3$ so that we can apply iteratively the Lemmas 5.2 and 5.3. Indeed, from the recursive formulas in Lemma 5.2 and 5.3 it is clear that it is enough to choose a parameter set $\mathcal{I}_3$ satisfying

$$1 < M_{\mathrm{ell},+}^{(3)} \ll M_{\mathrm{adj},+}^{(3)} \ll M_{\mathrm{hyp}}^{(3)} \ll M_{\mathrm{adj},-}^{(3)} \ll M_{\mathrm{ell},-}^{(3)}$$

and

$$0 < m_{\mathrm{ell}}^{(3)} < 3m_{\mathrm{adj}}^{(3)}.$$

From the choice of the constants in $\mathcal{I}_3$ and the recursion formulas in Lemmas 5.2 and 5.3, we have that $M_{\mathrm{hyp}}^{(j)} \geq 1$ for any $j = 3, \dots N-2$. This fact along with conditions (5.4) and (5.5), allow us to apply Lemmas 5.2 and 5.3 iteratively so that we obtain the (δ, σ, K)-recursive collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\dots,N-2}$ and the $\mathcal{I}_j$-product-like sets $\mathcal{V}_j \subset \Sigma_j^{\mathrm{in}}$. In particular, note that the recursion formulas stated in Theorem 4.9 can be easily deduced from the recursion formulas given in Lemmas 5.2 and 5.3 and the choice of $\mathcal{I}_3$.

Finally, we bound the time

$$|T_{\mathcal{B}^j}| \leq \left|T_{\mathcal{B}_{\mathrm{loc}}^{j}}\right| + \left|T_{\mathcal{B}_{\mathrm{glob}}^{j}}\right| \leq K\ln(1/\delta) + \widetilde{K}.$$

This completes the proof of Theorem 4.9. □

5.1 Straightening the heteroclinic connections

To prove Lemmas 5.2 and 5.3 the first step is to straighten the heteroclinic connections which connect with the future and past saddles. They have been analyzed in Lemma 4.2.

We perform the change of coordinates $(P_1, Q_1, P_2, Q_2) = \Xi(p_1, q_1, p_2, q_2)$ defined as

$$\begin{aligned} P_1 &= p_1 - \xi(q_1) \\ Q_1 &= q_1 \\ P_2 &= p_2 \\ Q_2 &= q_2 - \xi(p_2), \end{aligned} \tag{5.6}$$

which straightens the heteroclinic connections. This change is symplectic.

Lemma 5.4. *If one performs the change of coordinates* (5.6)*, one obtains a new Hamiltonian system of the form*

$$H^{(j)}(P,Q,c) = H_2^{(j)}(P,Q,c) + H_4^{(j)}(P,Q,c) \tag{5.7}$$

with

$$H_2^{(j)}(P,Q,c) = a_n \sum_{k\in\mathcal{P}_j} |c_k|^2 + \lambda_n \left(P_1Q_1 + P_2Q_2\right)$$

and

$$H_4^{(j)}(P,Q,c) = H_{\mathrm{hyp}}^{(j)}\left(P,Q\right) + H_{\mathrm{ell}}^{(j)}(c) + H_{\mathrm{mix}}^{(j)}\left(P,Q,c\right)$$

where

$$\begin{aligned}
H^{(j)}_{\rm hyp}(P,Q) =& -2P_1Q_1(P_1^2+Q_1^2-P_1Q_1)-2P_2Q_2(P_2^2+Q_2^2-P_2Q_2)\\
&+\sum_{k,\ell=0}^{2}\nu_{k\ell}P_1^kQ_1^{2-k}P_2^\ell Q_2^{2-\ell}+\mathcal{O}\left(\frac{1}{n}(P_1+Q_1)^2(P_2+Q_2)^2\right)\\
&+\mathcal{O}\left(\frac{1}{n}(P_1Q_1(P_1+Q_1)^2+P_2Q_2(P_2+Q_2)^2)\right)\\
H^{(j)}_{\rm ell}(c) =& -\frac{1}{4}\sum_{k\in\mathcal{P}_j}|c_k|^4-\frac{1}{4}\left(\sum_{k\in\mathcal{P}_j}|c_k|^2\right)^2\\
&+\sum_{k\in\mathcal{P}_j\setminus\{j+2\}}\mathrm{Re}\left(c_k^2\overline{c_{k-1}}^2\right)+\mathcal{O}\left(\frac{1}{n}\sum_{k,k'\in\mathcal{P}_j}|c_k|^2|c_{k'}|^2\right)\\
H^{(j)}_{\rm mix}(p,q,c) =& -\sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2\left(Q_1P_1+Q_2P_2\right)\\
&+\frac{2\sqrt{3}}{3}\mathrm{Re}\left((\omega_0P_1+\overline{\omega}_0Q_1)^2c_{j-2}^2\right)+\frac{2\sqrt{3}}{3}\mathrm{Re}\left((\omega_0P_2+\overline{\omega}_0Q_2)^2c_{j+2}^2\right)\\
&+\mathcal{O}\left(\frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2(P_1+Q_1+P_2+Q_2)^2\right)
\end{aligned}$$

To fix notation, we define the vector field associated to Hamiltonian (5.7),

$$\begin{aligned}
\dot P_1 =&\ \lambda_nP_1+\mathcal{Z}_{{\rm hyp},P_1}+\mathcal{Z}_{{\rm mix},P_1} = \lambda_nP_1+\partial_{Q_1}H^{(j)}_{\rm hyp}+\partial_{Q_1}H^{(j)}_{\rm mix}\\
\dot Q_1 =&\ -\lambda_nQ_1+\mathcal{Z}_{{\rm hyp},Q_1}+\mathcal{Z}_{{\rm mix},Q_1} = -\lambda_nQ_1-\partial_{P_1}H^{(j)}_{\rm hyp}-\partial_{P_1}H^{(j)}_{\rm mix}\\
\dot P_2 =&\ \lambda_nP_2+\mathcal{Z}_{{\rm hyp},P_2}+\mathcal{Z}_{{\rm mix},P_2} = \lambda_nP_2+\partial_{Q_2}H^{(j)}_{\rm hyp}+\partial_{Q_2}H^{(j)}_{\rm mix}\\
\dot Q_2 =&\ -\lambda_nQ_2+\mathcal{Z}_{{\rm hyp},Q_2}+\mathcal{Z}_{{\rm mix},Q_2} = -\lambda_nQ_2-\partial_{P_2}H^{(j)}_{\rm hyp}-\partial_{P_2}H^{(j)}_{\rm mix}\\
\dot c_k =&\ 2{\rm i}a_nc_k+\mathcal{Z}_{{\rm ell},c_k}+\mathcal{Z}_{{\rm mix},c_k} = 2{\rm i}a_nc_k-2{\rm i}\partial_{c_k}H^{(j)}_{\rm ell}-2{\rm i}\partial_{c_k}H^{(j)}_{\rm mix}.
\end{aligned}\tag{5.8}$$

6 The local dynamics of the hyperbolic toy model

If we set to zero the elliptic modes in the Hamiltonian obtained in Lemma 4.2, we obtain the Hamiltonian

$$H(P,Q)=\lambda_n(P_1Q_1+P_2Q_2)+H^{(j)}_{\rm hyp}(P,Q)\tag{6.1}$$

Therefore, the associated vector field is

$$\begin{aligned}
\dot P_1 &= \lambda_nP_1+\mathcal{Z}_{{\rm hyp},P_1}\\
\dot Q_1 &= -\lambda_nQ_1+\mathcal{Z}_{{\rm hyp},Q_1}\\
\dot P_2 &= \lambda_nP_2+\mathcal{Z}_{{\rm hyp},P_2}\\
\dot Q_2 &= -\lambda_nQ_2+\mathcal{Z}_{{\rm hyp},Q_2},
\end{aligned}\tag{6.2}$$

where $\mathcal{Z}_{{\rm hyp},P_i}=\partial_{Q_i}H^{(j)}_{\rm hyp}(P,Q)$ and $\mathcal{Z}_{{\rm hyp},Q_i}=-\partial_{P_i}H^{(j)}_{\rm hyp}(P,Q)$.

As we have explained, $(P,Q)=(0,0)$ is a hyperbolic critical point for the hyperbolic toy model. We want to study the local dynamics. The first step is to perform a $\mathcal{C}^k$ resonant normal form to remove the nonresonant terms, as is done in [GK15]. Note that here we encounter the same type of resonace. We use a result by Bronstein and Kopanskii [BK92], see Theorem 6 of [GK15], which implies the following.

Lemma 6.1. *There exists a $\mathcal{C}^2$ change of coordinates*

$$(P_1,Q_1,P_2,Q_2)=\Psi_{\rm hyp}(x_1,y_1,x_2,y_2)=(x_1,y_1,x_2,y_2)+\widetilde\Psi_{\rm hyp}(x_1,y_1,x_2,y_2)$$

which transforms the vector field (6.2) *into the vector field*

$$\mathcal{X}_{\rm hyp}(z) = Dz + R_{\rm hyp}, \tag{6.3}$$

where z denotes $z = (x_1, y_1, x_2, y_2)$, D is the diagonal matrix $D = \mathrm{diag}(\lambda_n, -\lambda_n, \lambda_n, -\lambda_n)$ and $R_{\rm hyp}$ is a polynomial, which only contains resonant monomials. It can be split as

$$R_{\rm hyp} = R^0_{\rm hyp} + R^1_{\rm hyp}, \tag{6.4}$$

where $R^0_{\rm hyp}$ is the first order, which is given by

$$R^0_{\rm hyp}(z) = \begin{pmatrix} R^0_{{\rm hyp},x_1}(z) \\ R^0_{{\rm hyp},y_1}(z) \\ R^0_{{\rm hyp},x_2}(z) \\ R^0_{{\rm hyp},y_2}(z) \end{pmatrix} = \begin{pmatrix} 4x_1^2y_1 + 2\nu_{02}y_1x_2^2 + \nu_{11}x_1x_2y_2 \\ -4x_1y_1^2 - 2\nu_{20}x_1y_2^2 - \nu_{11}y_1x_2y_2 \\ 4y_2x_2^2 + 2\nu_{20}x_1^2y_2 + \nu_{11}x_1y_1x_2 \\ -4x_2y_2^2 - \nu_{02}y_1^2x_2 - \nu_{11}x_1y_1y_2 \end{pmatrix},$$

and $R^1_{\rm hyp}$ is the remainder and satisfies $R^1_{{\rm hyp},x_i} = \mathcal{O}\left(x^3y^2\right)$ and $R^1_{{\rm hyp},y_i} = \mathcal{O}\left(x^2y^3\right)$.

Moreover, the function $\widetilde{\Psi}_{\rm hyp} = (\widetilde{\Psi}_{{\rm hyp},x_1}, \widetilde{\Psi}_{{\rm hyp},y_1}, \widetilde{\Psi}_{{\rm hyp},x_2}, \widetilde{\Psi}_{{\rm hyp},y_2})$ satisfies

$$\begin{aligned}
\widetilde{\Psi}_{{\rm hyp},x_1}(z) &= \mathcal{O}\left(x_1^3, x_1y_1, x_1(x_2^2+y_2^2), y_1y_2(x_2+y_2)\right) \\
\widetilde{\Psi}_{{\rm hyp},y_1}(z) &= \mathcal{O}\left(y_1^3, x_1y_1, y_1(x_2^2+y_2^2), x_1x_2(x_2+y_2)\right) \\
\widetilde{\Psi}_{{\rm hyp},x_2}(z) &= \mathcal{O}\left(x_2^3, x_2y_2, x_2(x_1^2+y_1^2), y_1y_2(x_1+y_1)\right) \\
\widetilde{\Psi}_{{\rm hyp},y_2}(z) &= \mathcal{O}\left(y_2^3, x_2y_2, y_2(x_1^2+y_1^2), x_1x_2(x_1+y_1)\right).
\end{aligned}$$

Remark 6.2. *All functions involved in this lemma, and also all functions involved in the forthcoming sections depend on the parameter n. We omit this dependence to simplify the notation. Note that when we use the notation $f = \mathcal{O}(g)$ we mean that there exists a constant $C > 0$ independent of n, δ and σ such that $|f| \le C|g|$.*

We analyze the local dynamics for the vector field (6.3) and then we will deduce the dynamics in the original variables. Note that after normal form, the vector field (6.3) is of the same form as the corresponding vector field in [GK15]. Therefore, we can use the results from that paper. As we have said in Section 4.3, we follow the notation of multiparameter sets from that paper.

In Section 4.3, we have considered the sets $\mathcal{N}_j^- \subset \mathcal{N}_j^+$ to define the almost product structure. Since in this section we have set the elliptic modes to zero, that is, $c = 0$, we consider a set $\mathcal{N}_j'$ satisfying

$$\mathcal{N}_j^- \cap \{c = 0\} \subset \mathcal{N}_j' \subset \mathcal{N}_j^+ \cap \{c = 0\}$$

Now, we need to express it in the new coordinates (x, y). We denote the inverse of the change $\Psi_{\rm hyp}$, obtained in Lemma 6.1, by $\Upsilon = \mathrm{Id} + \widetilde{\Upsilon} = \mathrm{Id} + \left(\widetilde{\Upsilon}_{x_1}, \widetilde{\Upsilon}_{y_1}, \widetilde{\Upsilon}_{x_2}, \widetilde{\Upsilon}_{y_2}\right)$ and we define

$$\widehat{C}^{(j)} = \widetilde{C}^{(j)}\left(1 + \widetilde{\Upsilon}_{x_1}(0, \sigma, 0, 0)\right).$$

Note that $\widehat{C}^{(j)} = \widetilde{C}^{(j)}(1 + \mathcal{O}(\sigma))$. We also define $f_1(\sigma) = \Upsilon_{y_1}(0, \sigma, 0, 0)$. This correspond to the first order of shift in the Poincaré section due to the normal form. That is, the section $y_1 = f_1(\sigma)$ approximates the section $\Upsilon(\Sigma_j^{\rm in})$ (recall that the change (5.6) has not moved the transversal section). We define the set of points in the normal form variables (x, y) whose dynamics we want to analyze by

$$\begin{aligned}
\widehat{\mathcal{N}}_j = \Big\{ &|x_1 + \widehat{C}^{(j)}\,\delta\,(\ln(1/\delta)| \le \widehat{C}^{(j)}\,\delta\,K_\sigma, \quad |x_2 - x_2^*| \le 2\,M^{(j)}_{\rm hyp}\frac{\left(\widehat{C}^{(j)}\delta\right)^{1/2}}{\ln(1/\delta)}, \\
&|y_1 - f_1(\sigma)| \le K_\sigma \widehat{C}^{(j)}\delta\ln(1/\delta), \qquad |y_2| \le 2\,M^{(j)}_{\rm hyp}\left(\widehat{C}^{(j)}\delta\right)^{1/2}\Big\},
\end{aligned} \tag{6.5}$$

The constant x_2^* will be choosen later analogously to [GK15]. The choice will allow us to obtain a cancellation which avoids deviation from the invariant manifolds.

The outcoming section gets also slightly modified by the normal form. To this end, we need to define the function $f_2(\sigma)$ as $f_2(\sigma) = \Upsilon_{x_2}(0, 0, \sigma, 0)$. As will be seen, the coordinate x_2 behaves almost linearly as

$x_2 \sim x_2^0 e^{\lambda\tau}$ (recall that we have rescaled time by (2.16) so now the time variable is τ). Therefore, the time needed to reach the section $x_2 = f_2(\sigma)$ is given approximately by

$$T_j(x_2^0) = \frac{1}{\lambda} \ln\left(\frac{f_2(\sigma)}{x_2^0}\right). \tag{6.6}$$

In order to analyze the action of Φ_τ^{hyp}, i.e. the flow associated to (6.3), on points $\widehat{\mathcal{N}}_j$ we proceed as in [GK15]. We choose x_2^* as the unique positive solution of

$$\left(x_2^*\right)^2 T_j(x_2^*) = \frac{\widehat{C}^{(j)}\, \delta\, \ln(1/\delta)}{2\,\nu_{02}\, f_1(\sigma)}. \tag{6.7}$$

We perform the change of coordinates $x_i = e^{\lambda\tau} u_i,\ y_i = e^{-\lambda\tau} v_i$ and thus we obtain the integral equations

$$\begin{aligned} u_i &= x_i^0 + \int_0^T e^{-\lambda\tau} R_{\mathrm{hyp},x_i}\left(ue^{\lambda\tau}, ve^{-\lambda\tau}\right) d\tau \\ v_i &= y_i^0 + \int_0^T e^{\lambda\tau} R_{\mathrm{hyp},y_i}\left(ue^{\lambda\tau}, ve^{-\lambda\tau}\right) d\tau. \end{aligned} \tag{6.8}$$

In the linear case u_i's and v_i's are constant. We use these variables to find a fixed point argument. We define the contractive operator in two steps. This approach is inspired by Shilnikov [Šil67]. First we define an auxiliary (non-contractive) operator

$$\mathcal{F}_{\mathrm{hyp}} = \left(\mathcal{F}_{\mathrm{hyp},u_1}, \mathcal{F}_{\mathrm{hyp},v_1}, \mathcal{F}_{\mathrm{hyp},u_2}, \mathcal{F}_{\mathrm{hyp},v_2}\right)$$

as

$$\begin{aligned} \mathcal{F}_{\mathrm{hyp},u_i}(u,v) &= x_i^0 + \int_0^T e^{-\lambda\tau} R_{\mathrm{hyp},x_i}\left(ue^{\lambda\tau}, ve^{-\lambda\tau}\right) d\tau \\ \mathcal{F}_{\mathrm{hyp},v_i}(u,v) &= y_i^0 + \int_0^T e^{\lambda\tau} R_{\mathrm{hyp},y_i}\left(ue^{\lambda\tau}, ve^{-\lambda\tau}\right) d\tau. \end{aligned} \tag{6.9}$$

As happens in [GK15], for the u_1 and v_2 components the main terms are not given by the initial condition but by the integral terms. In other words, the dynamics near the saddle is not well approximated by the linearized dynamics and the operator is not contractive. Following ideas from Shilnikov [Šil67], to have a contractive operator, we modify slightly two of the components of $\mathcal{F}_{\mathrm{hyp}}$ by considering

$$\widetilde{\mathcal{F}}_{\mathrm{hyp}} = \left(\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}, \widetilde{\mathcal{F}}_{\mathrm{hyp},v_1}, \widetilde{\mathcal{F}}_{\mathrm{hyp},u_2}, \widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}\right)$$

as

$$\begin{aligned} \widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(u,v) &= \mathcal{F}_{\mathrm{hyp},u_1}(u_1, \mathcal{F}_{\mathrm{hyp},v_1}(u,v), \mathcal{F}_{\mathrm{hyp},u_2}(u,v), v_2) \\ \widetilde{\mathcal{F}}_{\mathrm{hyp},v_1}(u,v) &= \mathcal{F}_{\mathrm{hyp},v_1}(u,v) \\ \widetilde{\mathcal{F}}_{\mathrm{hyp},u_2}(u,v) &= \mathcal{F}_{\mathrm{hyp},u_2}(u,v) \\ \widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(u,v) &= \mathcal{F}_{\mathrm{hyp},v_2}(u_1, \mathcal{F}_{\mathrm{hyp},v_1}(u,v), \mathcal{F}_{\mathrm{hyp},u_2}(u,v), v_2) \end{aligned}$$

The fixed points of these operators are exactly the same as the fixed points of $\mathcal{F}_{\mathrm{hyp}}$ and, then, are solutions of equation (6.8).

The operator $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ is contractive in a suitable Banach space. We define the following weighted norms. To fix notation, we denote by $\|\cdot\|_\infty$ the standard supremum norm. Then define

$$\begin{aligned} \|h\|_{\mathrm{hyp},u_1} &= \sup_{\tau\in[0,T_j]} \left|\left(-\widehat{C}^{(j)}\delta\ln(1/\delta) + 2\nu_{02} f_1(\sigma)\left(x_2^*\right)^2\tau + \widehat{C}^{(j)}\delta\right)^{-1} h(\tau)\right| \\ \|h\|_{\mathrm{hyp},v_1} &= f_1(\sigma)^{-1}\|h\|_\infty \\ \|h\|_{\mathrm{hyp},u_2} &= \left(x_2^*\right)^{-1}\|h\|_\infty \\ \|h\|_{\mathrm{hyp},v_2} &= \left(\left(y_1^0\right)^2 x_2^0 T_j\right)^{-1}\|h\|_\infty \end{aligned} \tag{6.10}$$

and the norm

$$\|(u,v)\|_* = \sup_{i=1,2} \left\{ \|u_i\|_{\mathrm{hyp},u_i}, \|v_i\|_{\mathrm{hyp},v_i} \right\}. \tag{6.11}$$

This gives rise to the following Banach space

$$\mathcal{Y}_{\mathrm{hyp}} = \left\{ (u,v) : [0,T] \to \mathbb{R}^4; \|(u,v)\|_* < \infty \right\}.$$

The contractivity of $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ is a consequence of the following two auxiliary propositions, whose proofs are given in [GK15].

Proposition 6.3. *Assume* (6.7)*, then there exists a constant $\kappa_0 > 0$ independent of σ, δ and j such that for δ and σ small enough, the operator $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ satisfies*

$$\|\widetilde{\mathcal{F}}(0)\|_* \leq \kappa_0.$$

Proposition 6.4. *Consider $w, w' \in B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ and let us assume* (6.7)*, then taking $\delta \ll \sigma$, the operator $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ satisfies*

$$\|\widetilde{\mathcal{F}}_{\mathrm{hyp}}(w) - \widetilde{\mathcal{F}}_{\mathrm{hyp}}(w')\|_* \leq K_\sigma \left(\widehat{C}^{(j)} \delta \right)^{1/2} \ln^2(1/\delta) \|w - w'\|_*.$$

These two propositions show that $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ is contractive from $B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ to itself. Therefore, it has a unique fixed point in $B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ which we denote by w^*. This fixed point argument gives precise estimates for the local dynamics of the Hyperbolic Toy Model (6.1). We use these estimates in order to study the behaviour of the full Toy Model (5.7) in Section 7.

7 The local dynamics for the toy model

We study the dynamics of the local map and we prove Lemma 5.2. We rely on the previous analysis of the hyperbolic toy model (6.1) done in Section 4.3. In this section, we consider the Hamiltonian (5.7), that is, we incorporate the elliptic modes.

Our goal is to study the map $\mathcal{B}^j_{\mathrm{loc}}$. We adapt its study from [GK15]. As in [GK15], the key point of this study is that the elliptic modes remain almost constant through the saddle map, which implies that they do not make much influence on the hyperbolic ones. In comparison to [GK15], the vector field (5.8) has some extra terms. Even if they are small, one needs to treat them carefully since the local map involves a rather long time.

As a first step we perform the change obtained in Lemma 6.1 to the vector field (5.8).

Lemma 7.1. *Let Ψ_{hyp} be the map defined in Lemma 6.1. Then, if one performs the change of coordinates*

$$(P_1, Q_1, P_2, Q_2, c) = (\Psi_{\mathrm{hyp}}(x_1, y_1, x_2, y_2), c), \tag{7.1}$$

to the vector field (5.8)*, obtains a vector field of the form*

$$\begin{aligned} \dot{z} &= Dz + R_{\mathrm{hyp}}(z) + R_{\mathrm{mix},z}(z,c) \\ \dot{c}_k &= 2\mathrm{i}a_n c_k + \mathcal{Z}_{\mathrm{ell},c_k}(c) + R_{\mathrm{mix},c}(z,c), \end{aligned}$$

where z denotes $z = (z_1, z_2) = (x_1, y_1, x_2, y_2)$, $D = \mathrm{diag}(\lambda_n, -\lambda_n, \lambda_n, -\lambda_n)$, R_{hyp} has been given in Lemma

6.1, $\mathcal{Z}_{\mathrm{ell},c_k}$ is defined in (5.8)*, and $R_{\mathrm{mix},z}$ and R_{mix,c_k} are defined as*

$$
\begin{aligned}
R_{\mathrm{mix},x_1} &= A_{x_1}(z)\overline{c_{j-2}}^2 + \overline{A_{x_1}(z)}c_{j-2}{}^2 - \sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2\Psi_{x_1}(z) + \frac{1}{n}D_{x_1}(z,c)\\
R_{\mathrm{mix},y_1} &= A_{y_1}(z)\overline{c_{j-2}}^2 + \overline{A_{y_1}(z)}c_{j-2}{}^2 + \sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2\Psi_{y_1}(z) + \frac{1}{n}D_{y_1}(z,c)\\
R_{\mathrm{mix},x_2} &= A_{x_2}(z)\overline{c_{j+2}}^2 + \overline{A_{x_2}(z)}c_{j+2}{}^2 - \sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2\Psi_{x_2}(z) + \frac{1}{n}D_{x_2}(z,c)\\
R_{\mathrm{mix},y_2} &= A_{y_2}(z)\overline{c_{j+2}}^2 + \overline{A_{y_2}(z)}c_{j+2}{}^2 + \sqrt{3}\sum_{k\in\mathcal{P}_j}|c_k|^2\Psi_{y_2}(z) + \frac{1}{n}D_{y_2}(z,c)\\
R_{\mathrm{mix},c_k} &= \mathrm{i}c_k P(z) + \frac{1}{n}D_{c_k}(z,c) \quad \textit{for } k\neq j\pm 2\\
R_{\mathrm{mix},c_{j\pm 2}} &= \mathrm{i}c_{j\pm 2}P(z) - \mathrm{i}\bar{c}_{j\pm 2}Q_\pm(z) + \frac{1}{n}D_{c_{j\pm 2}}(z,c)
\end{aligned}
$$

where $\Psi_{\mathrm{hyp},z}$ are the functions defined in Lemma 6.1, A_z satisfy

$$A_{x_i} = \mathcal{O}(x_i,y_i) \quad \textit{and} \quad A_{y_i} = \mathcal{O}(x_i,y_i),$$

the functions D_z satisfy

$$D_z = \mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|c_k|^2(x_i+y_i)\right),\ D_{c_k}(z,c) = \mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|c_k|(x_i+y_i)^2\right)$$

and P and $Q_\pm$ satisfy

$$P(z) = \mathcal{O}(xy)\,,\ \ Q_-(z) = \mathcal{O}(x_1,y_1) \quad \textit{and} \quad Q_+(z) = \mathcal{O}(x_2,y_2)\,.$$

The proof of this lemma is straightforward taking into account the form of the vector field (5.8) and the properties of Ψ_{hyp} given in Lemma 6.1.

As happens in [GK15], there is a rather strong interaction between the hyperbolic and the elliptic modes due to the terms R_{mix,x_i} and R_{mix,y_i}. As explained in [GK15], the importance of these terms can be seen as follows. The manifold $\{x=0, y=0\}$ is normally hyperbolic [Fen74, Fen77, HPS77] for the linear truncation of the vector field obtained in Lemma 7.1 and its stable and unstable manifolds are defined as $\{x=0\}$ and $\{y=0\}$. For the full vector field, the manifold $\{x=0, y=0\}$ is persistent. Moreover it is still normally hyperbolic thanks to [Fen74, Fen77, HPS77]. Nevertheless, the associated invariant manifolds deviate from $\{x=0\}$ and $\{y=0\}$ due to the terms R_{mix,x_i} and R_{mix,y_i}. To overcome this problem, we slightly modify the change (7.1) to straighten these invariant manifolds completely.

Lemma 7.2. *There exist a change of coordinates of the form*

$$(P_1,Q_1,P_2,Q_2,c) = (\Psi(x_1,y_1,x_2,y_2,c),c) = (x_1,y_1,x_2,y_2,c) + \left(\widetilde{\Psi}(x_1,y_1,x_2,y_2,c),0\right) \tag{7.2}$$

which transforms the vector field (5.8) *into a vector field of the form*

$$
\begin{aligned}
\dot{z} &= Dz + R_{\mathrm{hyp}}(z) + \widetilde{R}_{\mathrm{mix},z}(z,c)\\
\dot{c}_k &= 2\mathrm{i}a_n c_k + \mathcal{Z}_{\mathrm{ell},c_k}(c) + \widetilde{R}_{\mathrm{mix},c_k}(z,c),
\end{aligned}
\tag{7.3}
$$

where R_{hyp} and $\mathcal{Z}_{\mathrm{ell}}$ are the functions defined in (6.4) *and* (5.8) *respectively, and*

$$
\begin{aligned}
\widetilde{R}_{\mathrm{mix},x_1} &= B_{x_1}(z,c)\overline{c_{j-2}}^2 + \overline{B_{x_1}(z,c)}c_{j-2}{}^2 + \sum_{k\in\mathcal{P}_j}|c_k|^2 C_{x_1}(z,c) + \frac{1}{n}F_{x_1}(z,c)\\
\widetilde{R}_{\mathrm{mix},y_1} &= B_{y_1}(z,c)\overline{c_{j-2}}^2 + \overline{B_{y_1}(z,c)}c_{j-2}{}^2 + \sum_{k\in\mathcal{P}_j}|c_k|^2 C_{y_1}(z,c) + \frac{1}{n}F_{y_1}(z,c)\\
\widetilde{R}_{\mathrm{mix},x_2} &= B_{x_2}(z,c)\overline{c_{j+2}}^2 + \overline{B_{x_2}(z,c)}c_{j+2}{}^2 + \sum_{k\in\mathcal{P}_j}|c_k|^2 C_{x_2}(z,c) + \frac{1}{n}F_{x_2}(z,c)\\
\widetilde{R}_{\mathrm{mix},y_2} &= B_{y_2}(z,c)\overline{c_{j+2}}^2 + \overline{B_{y_2}(z,c)}c_{j+2}{}^2 + \sum_{k\in\mathcal{P}_j}|c_k|^2 C_{y_2}(z,c) + \frac{1}{n}F_{y_2}(z,c)\\
\widetilde{R}_{\mathrm{mix},c_k} &= \mathrm{i}c_k\widetilde{P}(z,c) + \frac{1}{n}F_{c_k}(z,c) \quad \textit{for } k\neq j\pm 2\\
\widetilde{R}_{\mathrm{mix},c_{j\pm2}} &= \mathrm{i}c_{j\pm2}\widetilde{P}(z,c) - \mathrm{i}\overline{c}_{j\pm2}\widetilde{Q}_{\pm}(z,c) + \frac{1}{n}F_{j\pm2}(z,c),
\end{aligned}
$$

where the functions B_z and C_z satisfy

$$
\begin{aligned}
B_{x_1}(z,c) &= \mathcal{O}(x_1 + y_1x_2z_2) & B_{x_2}(z,c) &= \mathcal{O}(x_2 + y_2x_1z_1)\\
B_{y_1}(z,c) &= \mathcal{O}(y_1 + x_1y_2z_2) & B_{y_2}(z,c) &= \mathcal{O}(y_2 + x_2y_1z_1)\\
C_{x_1}(z,c) &= \mathcal{O}(x_1 + y_1x_2z_2) & C_{x_2}(z,c) &= \mathcal{O}(x_2 + y_2x_1z_1)\\
C_{y_1}(z,c) &= \mathcal{O}(y_1 + x_1y_2z_2) & C_{y_2}(z,c) &= \mathcal{O}(y_2 + x_2y_1z_1)
\end{aligned}
$$

the functions D_z and D_c satisfy

$$
F_{x_i} = \mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|c_k|^2(x_1+x_2)\right),\ F_{y_i} = \mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|c_k|^2(y_1+y_2)\right),\ F_{c_k}(z,c) = \mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|c_k|(x_i+y_i)^2\right)
$$

and $\widetilde{P}$ and $\widetilde{Q}_{\pm}$ satisfy

$$
\widetilde{P}(z,c) = \mathcal{O}(xy)\,,\ \ \widetilde{Q}_{-}(z,c) = \mathcal{O}(x_1,y_1) \quad \textit{and} \quad \widetilde{Q}_{+}(z) = \mathcal{O}(x_2,y_2)\,.
$$

Moreover, the function $\widetilde{\Psi}$ satisfies

$$
\begin{aligned}
\widetilde{\Psi}_{x_1} &= \mathcal{O}\left(x_1^3, x_1y_1, x_1(x_2^2+y_2^2), y_1y_2(x_2+y_2), c_{j-2}^2y_1, \sum_{k\in\mathcal{P}}|c_k|^2y_1y_2^2, \frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2y_i\right)\\
\widetilde{\Psi}_{y_1} &= \mathcal{O}\left(y_1^3, x_1y_1, y_1(x_2^2+y_2^2), x_1x_2(x_2+y_2), c_{j-2}^2x_1, \sum_{k\in\mathcal{P}}|c_k|^2x_1x_2^2, \frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2x_i\right)\\
\widetilde{\Psi}_{x_2} &= \mathcal{O}\left(x_2^3, x_2y_2, x_2(x_1^2+y_1^2), y_1y_2(x_1+y_1), c_{j+2}^2y_1, \sum_{k\in\mathcal{P}}|c_k|^2y_2y_1^2, \frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2y_i\right)\\
\widetilde{\Psi}_{y_2} &= \mathcal{O}\left(y_2^3, x_2y_2, y_2(x_1^2+y_1^2), x_1x_2(x_1+y_1), c_{j+2}^2x_1, \sum_{k\in\mathcal{P}}|c_k|^2x_2x_1^2, \frac{1}{n}\sum_{k\in\mathcal{P}_j}|c_k|^2x_i\right).
\end{aligned}
$$

Proof. It is enough to compose two change of coordinates. The first change is the change (7.2) considered in Lemma 7.1. The second one is the one which straightens the invariant manifolds of a normally hyperbolic invariant manifold [Fen74, Fen77, HPS77]. Then, to obtain the required estimates, it suffices to combine Lemmas 6.1 and 7.1 with the standard results about normally hyperbolic invariant manifolds. □

The change obtained in Lemma 7.2 straightens the stable and unstable invariant manifolds of $\{x = 0, y = 0\}$. This allows us to perform the detailed study of the transition map close to the saddle that we need. As in [GK15], we define a set $\widehat{\mathcal{V}}_j$ such that

$$
\Upsilon\circ\Xi(\mathcal{V}_j)\subset\widehat{\mathcal{V}}_j, \tag{7.4}
$$

where $\mathcal{V}_j$ is the set defined in Lemma 5.2, Υ is the inverse of the coordinate change Ψ obtained in Lemma 7.2 and Ξ is the change of coordinates defined in (5.6). Then, we apply the flow $\widehat{\Phi}^\tau$ associated to the vector field (7.3) to points in $\widehat{\mathcal{V}}_j$. To obtain the inclusion (7.4) we define the function $g_j(p_2, q_2)$ involved in the definition of $\mathcal{V}_j$.

Define the set $\widehat{\mathcal{V}}_j = \mathbb{D}_1^1 \times \ldots \times \mathbb{D}_j^{j-2} \times \widehat{\mathcal{N}}_j \times \mathbb{D}_j^{j+2} \times \ldots \times \mathbb{D}_j^N$, where $\widehat{\mathcal{N}}_j$ is the set defined in (6.5) and $\mathbb{D}_j^k$ are defined as

$$\mathbb{D}_j^k = \left\{|c_k| \le M_{\mathrm{ell},\pm}\delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^\pm$$
$$\mathbb{D}_j^{j\pm 2} = \left\{|c_{j\pm 2}| \le M_{\mathrm{adj},\pm}\left(\widehat{C}^{(j)}\delta\right)^{1/2}\right\}.$$

Define the function $g_j(p_2, q_2)$ involved in the definition of the set $\mathcal{V}_j$ as

$$g_j(p_2, q_2) = p_2 + a_p(\sigma)p_2 + a_q(\sigma)q_2 - x_2^* \tag{7.5}$$

where x_2^* is the constant defined in (6.7) and

$$a_p(\sigma) = \partial_{p_2}\widetilde{\Upsilon}_{p_2}(0, \sigma, 0, 0, 0)$$
$$a_q(\sigma) = \partial_{q_2}\widetilde{\Upsilon}_{p_2}(0, \sigma, 0, 0, 0),$$

where $\Upsilon = \mathrm{Id} + \widetilde{\Upsilon}$.

Lemma 7.3. *With the above notations for δ small enough condition (7.4) is satisfied.*

Proof. It is a straightforward consequence of Lemmas 6.1 and 7.2. □

After straightening the invariant manifold, next lemma studies the saddle map in the transformed variables for points belonging to $\mathcal{V}_j$.

Lemma 7.4. *Let us consider the flow $\widehat{\Phi}_\tau$ associated to (7.3) and a point $(z^0, c^0) \in \widehat{\mathcal{V}}_j$. Then for δ and σ small enough, the point $(z^f, c^f) = \widehat{\Phi}_{T_j}(z^0, c^0)$, where $T_j = T_j(x_2^0)$ is the time defined in (6.6), satisfies*

$$|x_1^f|, |y_1^f| \le K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}$$
$$|x_2^f - f_2(\sigma)| \le K_\sigma\delta^{r'}$$
$$\left|y_2^f + \frac{f_1(\sigma)}{f_2(\sigma)}\widehat{C}^{(j)}\delta\ln(1/\delta)\right| \le \frac{f_1(\sigma)}{f_2(\sigma)}\delta.$$

and

$$\left|c_k^f - c_k^0 e^{2\mathrm{i}a_n T_j}\right| \le K_\sigma\delta^{(1-r)/2+r'} \quad \text{for } k \in \mathcal{P}_j^\pm$$
$$\left|c_{j\pm 2}^f - c_{j\pm 2}^0 e^{2\mathrm{i}a_n T_j}\right| \le 2M_{\mathrm{adj},\pm}\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}.$$

The proof of this lemma follows the same lines as the analogous result in [GK15]. It is explained in Section 7.1

Now, to complete the proof of Lemma 5.2 we need two final steps. First we undo the change of coordinates performed in Lemma 7.2 to express the estimates of the saddle map in the original variables. The second step is to adjust the time so that the image belongs to the section Σ_j^{out}. These two final steps are done in the next two following lemmas.

They proofs follow the same lines as the proofs of Lemmas 6.5 and 6.6 in [GK15]. One only needs to take into account the extra terms appearing in the change of coordinates Ψ, given in Lemma 7.2.

Lemma 7.5. *Let us consider the flow Φ_τ associated to (5.8) and a point $(P^0, Q^0, c^0) \in \Xi(\widehat{\mathcal{V}}_j)$, where Ξ the change in (5.6) and $\mathcal{V}_j$ is the set considered in Theorem 4.9. Then for δ and σ small enough, the point $(P^f, Q^f, c^f) = \Phi_{T_j}(P^0, Q^0, c^0)$, where T_j is the time defined in (6.6), satisfies*

$$|P_1^f|, |Q_1^f| \le K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}$$
$$|P_2^f - \sigma| \le K_\sigma\delta^{r'}$$
$$|Q_2^f + \widetilde{C}^{(j)}\delta\ln(1/\delta)| \le \widetilde{C}^{(j)}\ \delta\ K_\sigma.$$

for certain constant $\widetilde{C}^{(j)}$ *satisfying* $C^{(j)}/2 \le \widetilde{C}^{(j)} \le 2C^{(j)}$ *and*

$$\begin{aligned} \left| c_k^f - c_k^0 e^{2\mathrm{i}a_n T_j} \right| \le & K_\sigma \delta^{(1-r)/2+r'} \quad \text{for } k \in \mathcal{P}_j^{\pm} \\ \left| c_{j\pm 2}^f - c_{j\pm 2}^0 e^{2\mathrm{i}a_n T_j} \right| \le & 2M_{\mathrm{adj},\pm}\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}. \end{aligned}$$

Once we have obtained good estimates for the approximate time map in the original variables, we adjust it to obtain image points belonging to the section Σ_j^{out}.

Lemma 7.6. *Let us consider a point* $\left(P^f, Q^f, c^f\right) \in \Phi^{T_j} \circ \Xi(\mathcal{V}_j)$*, where* Φ^τ *is the flow of* (5.8)*,* T_j *is the time defined in* (6.6)*,* Ξ *the change in* (5.6) *and* $\mathcal{V}_j$ *is the set considered in Theorem 4.9.*

Then, there exists a time T'*, which depends on the point* $\left(P^f, Q^f, c^f\right)$*, such that*

$$(P^*, Q^*, c^*) = \Phi^{T'}\left(P^f, Q^f, c^f\right) \in \Sigma_j^{\mathrm{out}}.$$

Moreover, there exists a constant K_σ *such that*

$$|T'| \le K_\sigma \delta^r$$

and

$$\begin{aligned} \left| c_k^* - c_k^f \right| &\le K_\sigma \delta^{1-r} \quad \text{for } k \in \mathcal{P}_j \\ \left| P_1^* - P_1^f \right| &\le K_\sigma \left(C^{(j)}\delta\right)^{1/2} \delta^{1-r} \\ \left| Q_1^* - Q_1^f \right| &\le K_\sigma \left(C^{(j)}\delta\right)^{1/2} \delta^{1-r} \\ P_2 &= \sigma \\ \left| Q_2^* - Q_2^f \right| &\le K_\sigma C^{(j)} \delta^{2-r} \ln(1/\delta). \end{aligned}$$

To finish the proof of Lemma 5.2, it is enough to undo the change (5.6) and to proceed as in [GK15]. Recall that the change (5.6) only alters two coordinates.

7.1 Proof of Lemma 7.4

It follows the same lines as the proof of Lemma 6.4 in [GK15]. We only need to check that the additional terms are small enough so that the fixed point argument goes through. We make the variation of constants change of coordinates

$$x_i = e^{\lambda_n \tau} u_i, \; y_i = e^{-\lambda_n \tau} v_i, \; c_k = e^{2\mathrm{i}a_n \tau} s_k \tag{7.6}$$

to obtain the integral equation

$$\begin{aligned} u_i &= x_i^0 + \int_0^{T_j} e^{-\lambda_n \tau} \left(R_{\mathrm{hyp},x_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\lambda \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right)\right) d\tau \\ v_i &= y_i^0 + \int_0^{T_j} e^{\lambda_n \tau} \left(R_{\mathrm{hyp},y_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}\right) + \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right)\right) d\tau \\ s_k &= c_k^0 + \int_0^{T_j} e^{-2\mathrm{i}a_n \tau} \left(\mathcal{Z}_{\mathrm{ell},c_k}\left(se^{2\mathrm{i}a_n \tau}\right) + \widetilde{R}_{\mathrm{mix},c_k}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right)\right) d\tau. \end{aligned} \tag{7.7}$$

The terms $R_{\mathrm{hyp},z}$ are the ones considered in Section 6. So we use the properties of these functions obtained in that section. We use the integration time T_j introduced in (6.6).

We use (7.7) to set up a fixed point argument in two steps. First we define $\mathcal{G} = (\mathcal{G}_{\mathrm{hyp}}, \mathcal{G}_{\mathrm{ell}})$ as

$$\begin{aligned} \mathcal{G}_{\mathrm{hyp},u_i}(u,v,s) &= x_i^0 + \int_0^{T_j} e^{-\lambda_n \tau} \left(R_{\mathrm{hyp},x_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right)\right) d\tau \\ &= \mathcal{F}_{\mathrm{hyp},u_i}(u,v) + \int_0^{T_j} e^{-\lambda_n \tau} \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right) d\tau \\ \mathcal{G}_{\mathrm{hyp},v_i}(u,v,s) &= y_i^0 + \int_0^{T_j} e^{\lambda_n \tau} \left(R_{\mathrm{hyp},y_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\lambda_n \tau}, e^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right)\right) d\tau \\ &= \mathcal{F}_{\mathrm{hyp},v_i}(u,v) + \int_0^{T_j} e^{\lambda_n \tau} \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\lambda_n \tau}, ve^{-\lambda_n \tau}, se^{2\mathrm{i}a_n \tau}\right) d\tau, \end{aligned}$$

where $\mathcal{F}_{\rm hyp}$ is the operator defined in (6.9), and

$$\mathcal{G}_{{\rm ell},c_k}(u,v,s) = c_k^0 + \int_0^{T_j} e^{-2{\rm i}a_n\tau}\left(\mathcal{Z}_{{\rm ell},c_k}\left(se^{2{\rm i}a_n\tau}\right) + \widetilde{R}_{{\rm mix},c_k}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2{\rm i}a_n\tau}\right)\right)d\tau.$$

We proceed as in Section 6, and we modify it by defining

$$\begin{aligned}
\widetilde{\mathcal{G}}_{{\rm hyp},u_1}(u,v,s) &= \mathcal{G}_{{\rm hyp},u_1}(u_1, \mathcal{G}_{{\rm hyp},v_1}(u,v,s), \mathcal{G}_{{\rm hyp},u_2}(u,v,s), v_2, s)\\
\widetilde{\mathcal{G}}_{{\rm hyp},v_1}(u,v,s) &= \mathcal{G}_{{\rm hyp},v_1}(u,v,s)\\
\widetilde{\mathcal{G}}_{{\rm hyp},u_2}(u,v,s) &= \mathcal{G}_{{\rm hyp},u_2}(u,v,s)\\
\widetilde{\mathcal{G}}_{{\rm hyp},v_2}(u,v,s) &= \mathcal{G}_{{\rm hyp},v_2}(u_1, \mathcal{G}_{{\rm hyp},v_1}(u,v,s), \mathcal{G}_{{\rm hyp},u_2}(u,v,s), v_2, s)\\
\widetilde{\mathcal{G}}_{\rm ell}(u,v,s) &= \mathcal{G}_{\rm ell}(u,v,s)
\end{aligned}$$

which will be contractive. We denote the new operator by

$$\widetilde{\mathcal{G}} = \left(\widetilde{\mathcal{G}}_{{\rm hyp},u_1}, \widetilde{\mathcal{G}}_{{\rm hyp},u_2}, \widetilde{\mathcal{G}}_{{\rm hyp},v_1}, \widetilde{\mathcal{G}}_{{\rm hyp},v_2}, \widetilde{\mathcal{G}}_{\rm ell}\right), \tag{7.8}$$

whose fixed points coincide with those of $\mathcal{G}$.

We extend the norm defined in (6.10), as in [GK15], by defining

$$\begin{aligned}
\|h\|_{{\rm ell},\pm} &= \left(M_{{\rm ell},\pm}\delta^{(1-r)/2}\right)^{-1}\|h\|_\infty\\
\|h\|_{{\rm adj},\pm} &= M_{{\rm adj},\pm}^{-1}\left(\widehat{C}^{(j)}\delta\right)^{-1/2}\|h\|_\infty
\end{aligned}$$

and

$$\|(u,v,s)\|_* = \sup_{\substack{k\in\mathcal{P}_j^\pm\\ i=1,2}}\left\{\|u_i\|_{{\rm hyp},u_i}, \|v_i\|_{{\rm hyp},v_i}, \|s_k\|_{{\rm ell},\pm}, \|s_{j\pm 2}\|_{{\rm adj},\pm}\right\}$$

which, abusing notation, is denoted as the norm in (6.11). We also define the Banach space

$$\mathcal{Y} = \left\{(u,v,s) : [0,T] \to \mathbb{C}^{N-3}\times\mathbb{R}^4; \|(u,v,s)\|_* < \infty\right\}.$$

We state the two following propositions, which imply the contractivity of $\widetilde{\mathcal{G}}$. The proof of the first one is straightforward taking into account the definition of $\widetilde{\mathcal{G}}$ and Lemma 6.3. The proof of the second one is deferred to end of the section.

Proposition 7.7. *Let us consider the operator $\widetilde{\mathcal{G}}$ defined in* (7.8)*. Then, the components of $\widetilde{\mathcal{G}}(0)$ are given by*

$$\begin{aligned}
\widetilde{\mathcal{G}}_{{\rm hyp},u_1}(0) &= \widetilde{\mathcal{F}}_{{\rm hyp},u_i}(0)\\
\widetilde{\mathcal{G}}_{{\rm hyp},v_1}(0) &= y_1^0\\
\widetilde{\mathcal{G}}_{{\rm hyp},u_2}(0) &= x_2^0\\
\widetilde{\mathcal{G}}_{{\rm hyp},v_2}(0) &= \widetilde{\mathcal{F}}_{{\rm hyp},v_2}(0)\\
\widetilde{\mathcal{G}}_{{\rm ell},c_k}(0) &= c_k^0.
\end{aligned}$$

Thus, there exists a constant $\kappa_1 > 0$ independent of σ, δ and j such that the operator $\widetilde{\mathcal{G}}$ satisfies

$$\left\|\widetilde{\mathcal{G}}(0)\right\|_* \leq \kappa_1.$$

Proposition 7.8. *Let us consider $w_1, w_2 \in B(2\kappa_1) \subset \mathcal{Y}$, a constant r' satisfying $0 < r' < \ln 2/\gamma - 2r$ and δ as defined in Theorem 2.8. Then taking σ small enough and N big enough such that $0 < \delta = e^{-\gamma N} \ll 1$,* *there exist a constant $K_\sigma > 0$ which is independent of j and N, but might depend on σ, and a constant K independent of j, N and σ, such that the operator $\widetilde{\mathcal{G}}$ satisfies*

$$\begin{aligned}
&\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},u_i}(u,v,s)-\widetilde{\mathcal{G}}_{\mathrm{hyp},u_i}(u',v',s')\right\|_{\mathrm{hyp},u_i,v_i} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s)-(u',v',s')\right\|_* \\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},v_i}(u,v,s)-\widetilde{\mathcal{G}}_{\mathrm{hyp},v_i}(u',v',s')\right\|_{\mathrm{hyp},u_i,v_i} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s)-(u',v',s')\right\|_* \\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{ell},c_k}(u,v,s)-\widetilde{\mathcal{G}}_{\mathrm{ell},c_k}(u',v',s')\right\|_{\mathrm{ell},\pm} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s)-(u',v',s')\right\|_*, \quad \text{for } k \in \mathcal{P}_j^{\pm} \\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{adj},\pm}(u,v,s)-\widetilde{\mathcal{G}}_{\mathrm{adj},\pm}(u',v',s')\right\|_{\mathrm{adj},\pm} \leq \\
&\qquad \leq K\sigma \left\|(u,v,s)-(u',v',s')\right\|_* .
\end{aligned}$$

Thus, since $0 < \delta \ll \sigma$,

$$\left\|\widetilde{\mathcal{G}}(w_2)-\widetilde{\mathcal{G}}(w_1)\right\|_* \leq 2K\sigma\|w_2-w_1\|_*$$

and therefore, for σ small enough, it is contractive.

The previous two propositions show that the operator $\widetilde{\mathcal{G}}$ is contractive. Let us denote by (u^*, v^*, s^*) its unique fixed point in the ball $B(2\kappa_1) \subset \mathcal{Y}$. Now, it only remains to obtain the estimates stated in Lemma 7.4. The estimates for the hyperbolic variables are obtained as in [GK15]: it is enough to undo the change of coordinates (7.6) and to recall the definition of the norm 6.10. For the elliptic ones it is enough to take into account that

$$\begin{aligned}
c_k^f &= c_k(T_j) = s_k(T_j)e^{2\mathrm{i}a_nT_j} \\
&= \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j)e^{2\mathrm{i}a_nT_j} + \left(\mathcal{G}_{\mathrm{ell},c_k}(u^*,v^*,s^*)(T_j) - \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j)\right)e^{2\mathrm{i}a_nT_j} \\
&= c_k^0 e^{2\mathrm{i}a_nT_j} + \left(\mathcal{G}_{\mathrm{ell},c_k}(u^*,v^*,s^*)(T_j) - \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j)\right)e^{2\mathrm{i}a_nT_j}
\end{aligned}$$

and bound the second term using the Lipschitz constant obtained in Proposition 7.8.

We finish the section by proving Proposition 7.8, which completes the proof of Lemma 7.4.

Proof of Proposition 7.8. As we have done in the proof of Proposition 6.4, first, we stablish bounds for any $(u,v,s) \in B(2\kappa_1) \subset \mathcal{Y}$ in the supremmum norm, which will be used to bound the Lipschitz constant of each component of $\widetilde{\mathcal{G}}$. Indeed, if $(u,v,s) \in B(2\kappa_1) \subset \mathcal{Y}$, it satisfies

$$\begin{aligned}
|u_1| &\leq K_\sigma \widehat{C}^{(j)} \delta \ln(1/\delta) \\
|v_1| &\leq K\sigma \\
|u_2| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \\
|v_2| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln(1/\delta),
\end{aligned}$$

where $K > 0$ is a constant independent of σ, and

$$\begin{aligned}
|s_k| &\leq K_\sigma \delta^{(1-r)/2} \quad \text{for} \quad k \in \mathcal{P}_j^{\pm} \\
|s_{j\pm 2}| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \leq K_\sigma \delta^{(1-r)/2}.
\end{aligned}$$

We bound the Lipschitz constant for each component of $\widetilde{\mathcal{G}}_{\mathrm{ell}}$. We split each component of the operator between the elliptic, hyperbolic and mixed part. For the elliptic part the additional terms with respect to the toy model in [GK15] are of the same type as the terms in [GK15] (plus an extra n^{-1}). Therefore, they can be bounded as done in [GK15] to obtain

$$\left\|\int_0^{T_j} e^{-2\mathrm{i}a_n\tau}\left(\mathcal{Z}_{\mathrm{ell},c_k}\left(s_k e^{2\mathrm{i}a_n\tau}\right) - \mathcal{Z}_{\mathrm{ell},c_k}\left(s' e^{2\mathrm{i}a_n\tau}\right)\right)dt\right\|_{\mathrm{ell},\pm} \leq K_\sigma \delta^{1-r} N T_j \|(u,v,s)-(u',v',s')\|_*.$$

and

$$\left\| \int_0^{T_j} e^{-2\mathrm{i}a_n\tau} \left(\mathcal{Z}_{\mathrm{ell},c_{j\pm 2}}\left(se^{2\mathrm{i}a_n\tau}\right) - \mathcal{Z}_{\mathrm{ell},c_{j\pm 2}}\left(s'e^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{adj},\pm} \leq K_\sigma \delta^{1-r} N T_j \|(u,v,s)-(u',v',s')\|_*.$$

Now we bound the mixed terms. We can write them as $\widetilde{R}_{\mathrm{mix},c_k} = \widetilde{R}^0_{\mathrm{mix},c_k} + \widetilde{R}^1_{\mathrm{mix},c_k}$, where $\widetilde{R}^0_{\mathrm{mix},c_k}$ is the order in first in n^{-1}, that is, it is the term considered in [GK15], and $\widetilde{R}^1_{\mathrm{mix},c_k}$ contains the rest. In [GK15] it is seen that

$$\begin{aligned}\left\| \int_0^{T_j} e^{-2\mathrm{i}a_n\tau} \left(\widetilde{R}^0_{\mathrm{mix},c_k}(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}) - \widetilde{R}^0_{\mathrm{mix},c_k}(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau})\right) d\tau \right\|_{\mathrm{ell},\pm} \\ \leq K_\sigma \widehat{C}^{(j)} \delta \ln^3(1/\delta) \|(u,v,s)-(u',v',s')\|_*.\end{aligned}$$

for non adjacent modes and

$$\begin{aligned}\left\| \int_0^{T_j} e^{-2\mathrm{i}a_n\tau} \left(\widetilde{R}^0_{\mathrm{mix},c_{j\pm 2}}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}^0_{\mathrm{mix},c_{j\pm 2}}\left(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{adj},-} \\ \leq K\sigma \|(u,v,s)-(u',v',s')\|_*,\end{aligned}$$

where $K > 0$ is a constant independent of σ.

Now we bound the term $\widetilde{R}^1_{\mathrm{mix},c_k}$ stated in Lemma 7.2. We can see that for either for non-adjacent elliptic modes,

$$\begin{aligned}&\left\|\widetilde{R}^1_{\mathrm{mix},c_k}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}^1_{\mathrm{mix},c_k}\left(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau}\right)\right\|_{\mathrm{ell},\pm} \\ &\leq K_\sigma n^{-1} \sum_{i=1,2} \left(\|u_i - u'_i\|_{\mathrm{hyp},u_i} + \|v_i - v'_i\|_{\mathrm{hyp},v_i}\right) \\ &\quad + K_\sigma n^{-1} \sum_{\ell\in\mathcal{P}_j^\pm} \left\|s_\ell - s'_\ell\right\|_{\mathrm{ell},\pm} \\ &\leq K_\sigma N n^{-1} \left\|(u,v,s)-(u',v',s')\right\|_*.\end{aligned}$$

For the adjacent modes, recalling the bounds for $C^{(j)}$ in (4.15), we have that

$$\begin{aligned}&\left\|\widetilde{R}^1_{\mathrm{mix},c_{j\pm 2}}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}^1_{\mathrm{mix},c_{j\pm 2}}\left(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau}\right)\right\|_{\mathrm{ell},\pm} \\ &\leq K_\sigma N \delta^{-2r} n^{-1} \left\|(u,v,s)-(u',v',s')\right\|_*.\end{aligned}$$

Therefore, using that $\delta = e^{-\gamma N}$ and (6.6), we have that for $k \in \mathcal{P}_j^\pm$,

$$\begin{aligned}\left\| \int_0^{T_j} e^{-2\mathrm{i}a_n\tau} \left(\widetilde{R}_{\mathrm{mix},c_k}(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}) - \widetilde{R}_{\mathrm{mix},c_k}(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau})\right) d\tau \right\|_{\mathrm{ell},\pm} \\ \leq K_\sigma n^{-1} \ln^2(1/\delta) \|(u,v,s)-(u',v',s')\|_*.\end{aligned}$$

and for the adjacent modes

$$\begin{aligned}\left\| \int_0^{T_j} e^{-2\mathrm{i}a_n\tau} \left(\widetilde{R}_{\mathrm{mix},c_k}(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}) - \widetilde{R}_{\mathrm{mix},c_k}(u'e^{\lambda_n\tau}, v'e^{-\lambda_n\tau}, s'e^{2\mathrm{i}a_n\tau})\right) d\tau \right\|_{\mathrm{ell},\pm} \\ \leq K_\sigma n^{-1} \delta^{-2r} \ln^2(1/\delta) \|(u,v,s)-(u',v',s')\|_*.\end{aligned}$$

So, using the definition the properties of r and r' stated in Lemma 5.2, we can conclude that either for $k \in \mathcal{P}_j^\pm$,

$$\left\|\mathcal{G}_{\mathrm{ell},c_k}(u,v,s) - \mathcal{G}_{\mathrm{ell},c_k}(u',v',s')\right\|_{\mathrm{ell},\pm} \leq K_\sigma \delta^{r'} \|(u,v,s)-(u',v',s')\|_*.$$

and for the adjacent modes

$$\left\|\mathcal{G}_{\mathrm{ell},c_{j-2}}(u,v,s) - \mathcal{G}_{\mathrm{ell},c_{j-2}}(u',v',s')\right\|_{\mathrm{adj},-} \leq K\sigma \|(u,v,s)-(u',v',s')\|_*.$$

Now we bound the Lipschitz constant for the hyperbolic components of the operator. Note that we only need to bound the terms involving $\widetilde{R}_{\mathrm{mix},z}$ since the other terms of the operator have been bounded in Proposition 6.4. As for the elliptic modes we split them as $\widetilde{R}_{\mathrm{mix},z} = \widetilde{R}^0_{\mathrm{mix},z} + \widetilde{R}^1_{\mathrm{mix},z}$, where $\widetilde{R}^0_{\mathrm{mix},z}$ is the term in [GK15] and $\widetilde{R}^1_{\mathrm{mix},z}$ is the remainder which contains the terms of order $\mathcal{O}(n^{-1})$.

We start with the Lipschitz constants of $\mathcal{G}_{\mathrm{hyp},v_i}$. In [GK15] it is shown that

$$\begin{aligned}
&\left\| \int_0^{T_j} e^{\lambda_n \tau} \left(\widetilde{R}^0_{\mathrm{mix},y_1}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}^0_{\mathrm{mix},y_1}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{hyp},v_1} \\
&\qquad \leq K_\sigma \delta^{1-r} \ln^2(1/\delta) \|(u,v,s) - (u',v',s')\|_* \\
&\left\| \int_0^{T_j} e^{\lambda_n \tau} \left(\widetilde{R}^0_{\mathrm{mix},y_2}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}^0_{\mathrm{mix},y_2}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{hyp},v_2} \\
&\qquad \leq K_\sigma \delta^{1/2-2r} \ln(1/\delta) \|(u,v,s) - (u',v',s')\|_*.
\end{aligned}$$

Now, using the bounds on F_{y_i} given in Lemma 7.2, the definition of T_j in (6.6) and the upper and lower bounds for $C^{(j)}$ in (4.15), we bound the $\widetilde{R}^1_{\mathrm{mix},y_i}$ terms as

$$\begin{aligned}
&\left| \int_0^{T_j} e^{\lambda_n \tau} \left(\widetilde{R}^1_{\mathrm{mix},y_i}(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}) - \widetilde{R}^1_{\mathrm{mix},y_i}(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau})\right) d\tau \right| \\
&\qquad \leq K_\sigma N n^{-1} \delta^{1/2-3r/2} \|(u,v,s) - (u',v',s')\|_*.
\end{aligned}$$

Therefore, applying norms and using condition on δ from Theorem 2.8 and the condition on r' in Lemma 5.2, we obtain

$$\begin{aligned}
&\left\| \int_0^{T_j} e^{\lambda_n \tau} \left(\widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{hyp},v_1} \\
&\qquad \leq K_\sigma N n^{-1} \delta^{1/2-3r/2} \|(u,v,s) - (u',v',s')\|_* \\
&\qquad \leq K_\sigma \delta^{r'} \|(u,v,s) - (u',v',s')\|_* \\
&\left\| \int_0^{T_j} e^{\lambda_n \tau} \left(\widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right) - \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\lambda_n\tau}, ve^{-\lambda_n\tau}, se^{2\mathrm{i}a_n\tau}\right)\right) d\tau \right\|_{\mathrm{hyp},v_2} \\
&\qquad \leq K_\sigma \delta^{-2r} n^{-1} \|(u,v,s) - (u',v',s')\|_* \\
&\qquad \leq K_\sigma \delta^{r'} \|(u,v,s) - (u',v',s')\|_*.
\end{aligned}$$

Then, taking into account the results of Lemma 6.4, one can conclude that

$$\begin{aligned}
&\left\| \widetilde{\mathcal{G}}_{\mathrm{hyp},v_1}(u,v,s) - \widetilde{\mathcal{G}}_{\mathrm{hyp},v_1}(u',v',s') \right\|_{\mathrm{hyp},v_1} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s) - (u',v',s')\right\|_* \\
&\left\| \widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(u,v,s) - \widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(u',v',s') \right\|_{\mathrm{hyp},v_2} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s) - (u',v',s')\right\|_* .
\end{aligned}$$

Proceeding in the same way, one can obtain that

$$\begin{aligned}
&\left\| \widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(u,v,s) - \widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(u',v',s') \right\|_{\mathrm{hyp},u_1} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s) - (u',v',s')\right\|_* \\
&\left\| \widetilde{\mathcal{G}}_{\mathrm{hyp},u_2}(u,v,s) - \widetilde{\mathcal{G}}_{\mathrm{hyp},u_2}(u',v',s') \right\|_{\mathrm{hyp},u_2} \leq \\
&\qquad \leq K_\sigma \delta^{r'} \left\|(u,v,s) - (u',v',s')\right\|_* .
\end{aligned}$$

This completes the proof. □

8 Study of the global map: proof of Lemma 5.3

We devote this section to prove Lemma 5.3. It follows the same lines as the proof of Lemma 4.8 in [GK15]. The main difference is that now the heteroclinic connection is not straightened in the original coordinates and therefore we use the coordinates (P, Q), obtained in Lemma 4.2, to prove Lemma 5.3. Recall that the initial section Σ_j^{out}, defined in (5.1), is expressed in the variables adapted to the j^{th} saddle, that is $(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)})$, whereas the final section Σ_{j+1}^{in}, defined in (4.10), is expressed in the variables adapted to the $(j+1)^{st}$ saddle, that is $(p_1^{(j+1)}, q_1^{(j+1)}, p_2^{(j+1)}, q_2^{(j+1)}, c^{(j+1)})$. The change of variables between these two system of coordinates is given in [GK15], and stated in the next lemma. To simplify notation we define

$$(p_1, q_1, p_2, q_2, c) = \left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)}\right)$$

and

$$(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = \left(p_1^{(j+1)}, q_1^{(j+1)}, p_2^{(j+1)}, q_2^{(j+1)}, c^{(j+1)}\right)$$

and we denote by Θ^j the change of coordinates that relates them, namely

$$(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = \Theta^j(p_1, q_1, p_2, q_2, c).$$

Lemma 8.1. *The change of coordinates Θ^j is given by*

$$\begin{aligned}
\Theta^j_{\widetilde{c}_k}(p_1, q_1, p_2, q_2, c) &= \frac{\overline{\omega} q_2 + \omega p_2}{\widetilde{r}\sqrt{2\mathrm{Im}\,(\omega^2)}} c_k \qquad\qquad \text{for } k \in \mathcal{P}^{\pm}_{j+1} \cup \{j+3\}\\
\Theta^j_{\widetilde{c}_{j-1}}(p_1, q_1, p_2, q_2, c) &= \frac{\overline{\omega} q_2 + \omega p_2}{\widetilde{r}\mathrm{Im}\,(\omega^2)} (\omega q_1 + \overline{\omega} p_1)\\
\Theta^j_{\widetilde{p}_1}(p_1, q_1, p_2, q_2, c) &= \frac{r}{\widetilde{r}} q_2\\
\Theta^j_{\widetilde{q}_1}(p_1, q_1, p_2, q_2, c) &= \frac{r}{\widetilde{r}} p_2\\
\Theta^j_{\widetilde{p}_2}(p_1, q_1, p_2, q_2, c) &= \frac{1}{2}\left(\frac{\mathrm{Re}\, z}{\mathrm{Re}\,\omega} + \frac{\mathrm{Im}\, z}{\mathrm{Im}\,\omega}\right)\\
\Theta^j_{\widetilde{q}_2}(p_1, q_1, p_2, q_2, c) &= \frac{1}{2}\left(\frac{\mathrm{Re}\, z}{\mathrm{Re}\,\omega} - \frac{\mathrm{Im}\, z}{\mathrm{Im}\,\omega}\right),
\end{aligned}$$

where ω has been defined in (4.3) *and*

$$\begin{aligned}
r^2 =& 1 - \sum_{k \neq j-1, j, j+1} |c_k|^2 - \frac{1}{\mathrm{Im}\,(\omega^2)}(p_1^2 + q_1^2 + 2\mathrm{Re}\,(\omega^2) p_1 q_1)\\
&- \frac{1}{\mathrm{Im}\,(\omega^2)}(p_2^2 + q_2^2 + 2\mathrm{Re}\,(\omega^2) p_2 q_2)\\
\widetilde{r}^2 =& \frac{1}{\mathrm{Im}\,(\omega^2)}\left(p_2^2 + q_2^2 + 2\mathrm{Re}\,(\omega^2) p_2 q_2\right)\\
z =& \frac{c_{j+2}}{\widetilde{r}} \left(\overline{\omega} q_2 + \omega p_2\right).
\end{aligned}$$

To prove Lemma 5.3, we want to use the system of coordinates given in Lemma 4.2 for the old variables. We define

$$(P_1, Q_1, P_2, Q_2) = \left(P_1^{(j)}, Q_1^{(j)}, P_2^{(j)}, Q_2^{(j)}\right)$$

For the new ones, we want to stick with $(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c})$ since those are the ones used to state Lemma 5.3. We denote by $\widetilde{\Theta}^j$ the change of coordinates that relates them, namely

$$(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = \Theta^j(P_1, Q_1, P_2, Q_2, c).$$

Corollary 8.2. *The change of coordinates $\widetilde{\Theta}^j$ is given by*

$$
\begin{aligned}
\widetilde{\Theta}^j_{\widetilde{c}_k}(P_1,Q_1,P_2,Q_2,c) &= \frac{\overline{\omega}(Q_2+\xi(P_2))+\omega P_2}{\widetilde{r}\sqrt{\mathrm{Im}\,(\omega^2)}}c_k \qquad \textit{for } k\in\mathcal{P}^{\pm}_{j+1}\cup\{j+3\}\\
\widetilde{\Theta}^j_{\widetilde{c}_{j-1}}(P_1,Q_1,P_2,q_2,c) &= \frac{\overline{\omega}(Q_2+\xi(P_2))+\omega P_2}{\widetilde{r}\mathrm{Im}\,(\omega^2)}\left(\omega Q_1+\overline{\omega}(P_1+\xi(Q_1))\right)\\
\widetilde{\Theta}^j_{\widetilde{p}_1}(P_1,Q_1,P_2,Q_2,c) &= \frac{r}{\widetilde{r}}(Q_2+\xi(P_2))\\
\widetilde{\Theta}^j_{\widetilde{q}_1}(P_1,Q_1,P_2,Q_2,c) &= \frac{r}{\widetilde{r}}P_2\\
\widetilde{\Theta}^j_{\widetilde{P}_2}(P_1,Q_1,P_2,Q_2,c) &= \frac{1}{2}\left(\frac{\mathrm{Re}\,z}{\mathrm{Re}\,\omega}+\frac{\mathrm{Im}\,z}{\mathrm{Im}\,\omega}\right)\\
\widetilde{\Theta}^j_{\widetilde{Q}_2}(P_1,Q_1,P_2,Q_2,c) &= \frac{1}{2}\left(\frac{\mathrm{Re}\,z}{\mathrm{Re}\,\omega}-\frac{\mathrm{Im}\,z}{\mathrm{Im}\,\omega}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
r^2 =& 1-\sum_{k\neq j-1,j,j+1}|c_k|^2-\frac{1}{\mathrm{Im}\,(\omega^2)}((P_1+\xi(Q_1))^2+Q_1^2+2\mathrm{Re}\,(\omega^2)(P_1+\xi(Q_1))Q_1)\\
&-\frac{1}{\mathrm{Im}\,(\omega^2)}(P_2^2+(Q_2+\xi(P_2))^2+2\mathrm{Re}\,(\omega^2)P_2(Q_2+\xi(P_2)))\\
\widetilde{r}^2 =& \frac{1}{\mathrm{Im}\,(\omega^2)}\left(p_2^2+(Q_2+\xi(P_2))^2+2\mathrm{Re}\,(\omega^2)P_2(Q_2+\xi(P_2))\right)\\
z =& \frac{c_{j+2}}{\widetilde{r}}\left(\overline{\omega}(Q_2+\xi(P_2))+\omega P_2\right).
\end{aligned}
$$

Note that in the new variables, we will need to check that the sets we obtain in the final section are close to the separatrix defined in Lemma 4.2. This will be a consequence of the next lemma.

Lemma 8.3. *The function ξ introduced in Lemma 4.2 satisfies*

$$\xi(\widetilde{q}_1)=\frac{r_0}{\widetilde{r}_0}\xi\left(\frac{\widetilde{r}_0}{r_0}\widetilde{q}_1\right)$$

where r_0 and $\widetilde{r}_0$ are defined by the following equations

$$
\begin{aligned}
\widetilde{r}_0^2 =& \frac{1}{\mathrm{Im}\,(\omega^2)}\left(\left(\frac{\widetilde{r}_0}{r_0}\widetilde{q}_1\right)^2+\xi^2(\frac{\widetilde{r}_0}{r_0})+2\mathrm{Re}\,(\omega^2)\frac{\widetilde{r}_0}{r_0}\widetilde{q}_1\xi\left(\frac{\widetilde{r}_0}{r_0}\right)\right)\\
r_0^2 =& 1-\widetilde{r}_0^2
\end{aligned}
$$

Proof. Note that the separatrix we are traveling close to is defined by $q_2=\xi(p_2)$. Applying the change obtained in Lemma 8.1, we obtain that in the new variables it must satisfy

$$\widetilde{p}_1=\frac{r_0}{\widetilde{r}_0}\xi\left(\frac{\widetilde{r}_0}{r_0}\widetilde{q}_1\right)$$

where r_0 and $\widetilde{r}_0$ are just the functions r and $\widetilde{r}$ introduced in Lemma 8.1 evaluated over the separatrix. Moreover, using that the hyperbolic toy model at each saddle is the same, we know that in the new variables the separatrix can be parameterized as a graph as $\widetilde{p}_1=\xi(\widetilde{q}_1)$. Since the graph parameterization is unique, we obtain the formula stated in the lemma. □

Now, we express the section $\Sigma^{\mathrm{in}}_{j+1}$ in the variables (P_1,Q_1,P_2,Q_2,c) using the change $\widetilde{\Theta}^j$ obtained in Lemma 8.2.

Corollary 8.4. *Fix $\sigma>0$ and define the set*

$$\widetilde{\Sigma}^{\mathrm{in}}_{j+1}=\left(\widetilde{\Theta}^j\right)^{-1}\left(\Sigma^{\mathrm{in}}_{j+1}\cap\mathcal{W}_{j+1}\right),$$

where Σ_{j+1}^{in} *is the section defined in* (4.10) *and*

$$\mathcal{W}_{j+1} = \left\{|P_1| \leq \eta, |Q_1| \leq \eta, |Q_2| \leq \eta, |c_k| \leq \eta \text{ for } k \in \mathcal{P}_j^{\pm} \text{ and } k = j \pm 2\right\}.$$

Then, for $\eta > 0$ *small enough,* $\mathcal{W}_{j+1}$ *can be expressed as a graph as*

$$P_2 = w(P_1, Q_1, Q_2, c).$$

Moreover, there exist constants κ', κ'' *independent of* η *satisfying*

$$0 < \kappa' < \sqrt{\operatorname{Im}(\omega^2) - \sigma^2} < \kappa'' < 1$$

such that, for any $(P_1, Q_1, Q_2, c) \in \mathcal{W}_{j+1}$*, the function* w *satisfies*

$$\kappa' < w(P_1, Q_1, Q_2, c) < \kappa''.$$

Once we have defined the section $\widetilde{\Sigma}_{j+1}^{\text{in}}$, we can define the map

$$\begin{array}{rccc} \widetilde{\mathcal{B}}_{\text{glob}}^j : & \Xi(\mathcal{U}_j) \subset \Sigma_j^{\text{out}} & \longrightarrow & \widetilde{\Sigma}_{j+1}^{\text{in}} \\ & (P_1, Q_1, Q_2, c) & \mapsto & \widetilde{\mathcal{B}}_{\text{glob}}^j(P_1, Q_1, Q_2, c) \end{array}$$

induced by the flow (5.8). Thanks to Corollary 8.4, one can easily deduce that the time $T_{\widetilde{\mathcal{B}}_{\text{glob}}^j} = T_{\widetilde{\mathcal{B}}_{\text{glob}}^j}(Q_1, P_1, P_2, c)$ spent by the map $\widetilde{\mathcal{B}}_{\text{glob}}^j$ for any point $(Q_1, P_1, P_2, c) \in \Xi(\mathcal{U}_j) \subset \Sigma_j^{\text{out}}$ is independent of δ, j and N. Since the difference between $\widetilde{\mathcal{B}}_{\text{glob}}^j$ and $\mathcal{B}_{\text{glob}}^j$ is just a change of coordinates, we have that the time spent by $\mathcal{B}_{\text{glob}}^j$ is the same $T_{\widetilde{\mathcal{B}}_{\text{glob}}^j}$.

Now we study the behavior of the map $\widetilde{\mathcal{B}}_{\text{glob}}^j$.

Proposition 8.5. *Let us consider a parameter set* $\widetilde{\mathcal{I}}_j$ *(as defined in Definition 5.1) and a* $\widetilde{\mathcal{I}}_j$*-product-like set* $\mathcal{U}_j$*. Then, there exists a constant* $\widetilde{K}_\sigma$ *independent of* j*,* N *and* δ *and a constant* $D^{(j)}$ *satisfying*

$$\widetilde{C}^{(j)}/\widetilde{K}_\sigma \leq D^{(j)} \leq \widetilde{K}_\sigma \widetilde{C}^{(j)},$$

such that the set $\widetilde{\mathcal{B}}_{\text{glob}}^j \circ \Xi(\mathcal{U}_j) \subset \widetilde{\Sigma}_{j+1}^{\text{in}}$*, where* Ξ *is the change defined in* (5.6)*, satisfies the following conditions:*

C1

$$\widetilde{\mathcal{B}}_{\text{glob}}^j \circ \Xi(\mathcal{U}_j) \subset \widehat{\mathbb{D}}_j^1 \times \ldots \times \widehat{\mathbb{D}}_j^{j-2} \times \mathcal{S}_j \times \widehat{\mathbb{D}}_j^{j+2} \times \ldots \times \widehat{\mathbb{D}}_j^N$$

where

$$\widehat{\mathbb{D}}_j^k = \left\{|c_k| \leq \left(\widetilde{M}_{\text{ell},\pm}^{(j)} + \widetilde{K}_\sigma \delta^{r'}\right) \delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^{\pm}$$

$$\widehat{\mathbb{D}}_j^{j\pm 2} \subset \left\{|c_{j\pm 2}| \leq \widetilde{K}_\sigma \widetilde{M}_{\text{adj},\pm}^{(j)} \left(\widetilde{C}^{(j)}\delta\right)^{1/2}\right\},$$

and

$$\begin{aligned} \mathcal{S}_j = \Big\{&(P_1, Q_1, P_2, Q_2) \in \mathbb{R}^4 : |P_1|, |Q_1| \leq \widetilde{K}_\sigma \widetilde{M}_{\text{hyp}}^{(j)} \left(\widetilde{C}^{(j)}\delta\right)^{1/2}, \\ &P_2 = w(P_1, Q_1, Q_2, c), -D^{(j)}\,\delta\left(\ln(1/\delta) - \widetilde{K}_\sigma\right) \leq Q_2^{(j)} \leq -D^{(j)}\,\delta\left(\ln(1/\delta) + \widetilde{K}_\sigma\right)\Big\}, \end{aligned}$$

C2 *Let us define the projection* $\widetilde{\pi}(P, Q, c) = (P_2, Q_2, c_{j+2}, \ldots, c_N)$*. Then,*

$$\begin{aligned} \Big[-D^{(j)}\,\delta\left(\ln(1/\delta) - 1/\widetilde{K}_\sigma\right), -D^{(j)}\,\delta\left(\ln(1/\delta) + 1/\widetilde{K}_\sigma\right)\Big] \times \{P_2 = w(P_1, Q_1, Q_2, c)\} \times \mathbb{D}_{j,-}^{j+2} \times \ldots \times \mathbb{D}_{j,-}^N \\ \subset \widetilde{\pi}\left(\widetilde{\mathcal{B}}_{\text{glob}}^j \circ \Xi(\mathcal{U}_j)\right) \end{aligned}$$

where

$$\mathbb{D}_{j,-}^k = \left\{\left|c_k^{(j)}\right| \leq \left(\widetilde{m}_{\text{ell}}^{(j)} - \widetilde{K}_\sigma \delta^{r'}\right) \delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^+$$

$$\mathbb{D}_{j,-}^{j+2} = \left\{\left|c_{j+2}^{(j)}\right| \leq \widetilde{m}_{\text{adj}}^{(j)} \left(C^{(j)}\delta\right)^{1/2} / \widetilde{K}_\sigma\right\}.$$

This proposition is proved for the cubic nonlinear Schrödinger equation toy model in [GK15]. One can easily check that the prove is also valid for the vector field (5.8). Therefore, the prove in [GK15] also applies to our setting.

Now we complete the proof of Lemma 5.3. We need to show that the set $\mathcal{B}^j_{\text{glob}}(\mathcal{U}_j) \subset \Sigma^{\text{in}}_{j+1}$ satisfies similar properties to the ones of the set $\widetilde{\mathcal{B}}^j_{\text{glob}} \circ \Xi(\mathcal{U}_j)$ and also to obtain a parameter set $\mathcal{I}_{j+1}$ and $\mathcal{I}_{j+1}$-product like set $\mathcal{V}_j \subset \Sigma^{\text{in}}_{j+1}$ which satisfies condition (5.5). These two last steps are summarized in the next lemma.

Lemma 8.6. *Let us consider a parameter set $\mathcal{I}_{j+1}$ whose constants satisfy*

$$
\begin{aligned}
D^{(j)}/2 \leq C^{(j+1)} &\leq 2D^{(j)} \\
0 < m^{(j+1)}_{\text{hyp}} &\leq \widetilde{m}^{(j)}_{\text{hyp}}
\end{aligned}
$$

and

$$
\begin{aligned}
M^{(j+1)}_{\text{ell},-} &= \max\left\{\widetilde{M}^{(j)}_{\text{ell},-} + \widetilde{K}_\sigma \delta^{r'}, \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\text{adj},-}\right\} \\
M^{(j+1)}_{\text{ell},+} &= \widetilde{M}^{(j)}_{\text{ell},+} + \widetilde{K}_\sigma \delta^{r'} \\
m^{(j+1)}_{\text{ell}} &= \widetilde{m}^{(j)}_{\text{ell}} - \widetilde{K}_\sigma \delta^{r'} \\
m^{(j+1)}_{\text{adj},+} &= \widetilde{m}^{(j)}_{\text{ell},+} + \widetilde{K}_\sigma \delta^{r'} \\
M^{(j+1)}_{\text{adj},-} &= \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\text{hyp}} \\
m^{(j+1)}_{\text{adj}} &= \widetilde{m}^{(j)}_{\text{ell}} + \widetilde{K}_\sigma \delta^{r'} \\
M^{(j+1)}_{\text{hyp}} &= \max\left\{\widetilde{K}_\sigma \widetilde{M}^{(j)}_{\text{adj},+}, \widetilde{K}_\sigma\right\}.
\end{aligned}
$$

Then, the set

$$
\mathcal{V}_{j+1} = \mathcal{B}^j_{\text{glob}}(\mathcal{U}_j) \cap \left\{g_j(p_2^{(j+1)}, q_2^{(j+1)}) = 0\right\} \cap \left\{\left|c^{(j+1)}_{j+3}\right| \leq M^{(j+1)}_{\text{adj},+}\left(C^{(j+1)}\delta\right)^{1/2}\right\},
$$

where g_j is the function defined in (7.5)*, is a $\mathcal{I}_{j+1}$-product-like set and satisfies condition* (5.5)

Proof. It is enough to apply the change of coordinates Θ^j given in Lemma 8.1 and take into account that ξ satisfies the equality given by Lemma 8.3 and $|\xi| = \mathcal{O}(n^{-1})$. □

9 The approximation argument: proof of Theorem 2.9

Write the equation associated to Hamiltonian (2.8) as

$$
-\mathrm{i}\dot{r}_n = \mathcal{E}_n(r) + \widetilde{\mathcal{R}}_n(r), \tag{9.1}
$$

where $\mathcal{E}$ is the function defined in (2.11) and $\widetilde{\mathcal{R}}$ is the vector field associated to the Hamiltonian $\mathcal{R}'$ defined in (2.9). We want to study the closeness of the orbit $\mathtt{r}^\varrho(t)$ obtained in (2.19), which is a solution of $-\mathrm{i}\dot{\mathtt{r}}^\varrho = \mathcal{E}(\mathtt{r}^\varrho)$, with an orbit $\widetilde{r}(t)$ of equation (9.1) which satisfies $\|\widetilde{r}(0) - \mathtt{r}^\varrho(0)\|_{\ell^1} \leq \varrho^{-5/2}$. Define the function ξ as

$$
\xi = \widetilde{r} - \mathtt{r}^\varrho, \tag{9.2}
$$

which satisfies $\|\xi(0)\|_{\ell^1} \leq \varrho^{-5/2}$. We proceed as in [GK15] and we apply Gronwall-like estimates to bound the ℓ^1 norm of $\xi(t)$.

The equation for ξ can be written as $\dot{\xi} = \mathcal{Z}^0(t) + \mathcal{Z}^1(t)\xi + \mathcal{Z}^2(\xi, t)$ with

$$
\mathcal{Z}^0(t) = \widetilde{\mathcal{R}}\left(\mathtt{r}^\varrho\right) \tag{9.3}
$$

$$
\mathcal{Z}^1(t) = D\mathcal{E}\left(\mathtt{r}^\varrho\right) \tag{9.4}
$$

$$
\mathcal{Z}^2(\xi, t) = \mathcal{E}\left(\mathtt{r}^\varrho + \xi\right) - \mathcal{E}\left(\mathtt{r}^\varrho\right) - D\mathcal{E}\left(r^\varrho\right)\xi + \widetilde{\mathcal{R}}\left(\mathtt{r}^\varrho + \xi\right) - \widetilde{\mathcal{R}}\left(\mathtt{r}^\varrho\right) \tag{9.5}
$$

Applying the ℓ^1 norm to this equation, we obtain

$$
\frac{d}{dt}\|\xi\|_{\ell^1} \leq \left\|\mathcal{Z}^0(t)\right\|_{\ell^1} + \left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} + \left\|\mathcal{Z}^2(\xi, t)\right\|_{\ell^1}. \tag{9.6}
$$

The next three lemmas give estimates for each term in the right hand side of this equation.

Lemma 9.1. *The function* $\mathcal{Z}^0$ *defined in* (9.3) *satisfies* $\left\|\mathcal{Z}^0\right\|_{\ell^1} \leq C\varrho^{-(2d+1)}2^{(2d+1)N}$.

The proof of this lemma is analogous to the proof of Lemma B.1 in [GK15].

Lemma 9.2. *The linear operator* $\mathcal{Z}^1(t)$ *satisfies* $\left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} \leq C\varrho^{-(2d-2)}2^{N(2d-2)}\|\xi\|_{\ell^1}$

Proof. Taking into account the definition of $\mathcal{E}$ in (2.11), we have that

$$\left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} \leq \|\mathtt{r}^\varrho\|_{\ell^1}^{2d-2}\|\xi\|_{\ell^1}$$

For each $t \in [0,T]$, we have that there exists j^* such that, for any $k \in \mathcal{S}_{j^*}$, $|\mathtt{r}_k^\varrho| \leq \varrho$. For any other j and $k \in \mathcal{S}_{j^*}$, $|\mathtt{r}_k^\varrho| \leq \varrho\delta^\nu$. Recall that $\mathtt{r}_k^\varrho = 0$ for all $k \notin \mathcal{S}$. Then, since $|\mathcal{S}_j| \leq 2^{N-1}$, we have that $\|\mathtt{r}^\varrho\|_{\ell^1} \lesssim \varrho^{-1}2^{N-1}$, which implies $\left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} \lesssim C\varrho^{-(2d-2)}2^{N(2d-2)}\|\xi\|_{\ell^1}$. □

To obtain estimates for $\mathcal{Z}^2(\xi,t)$ defined in (9.5), we apply a bootstrap argument as done in [CKS+10]. Assume that for $0 < t < T^*$ we have

$$\|\xi(t)\|_{\ell^1} \leq C\varrho^{-3/2}2^N. \tag{9.7}$$

For $t = 0$ we know that it is already satisfied since $\|\xi(0)\|_{\ell^1} \leq \varrho^{-5/2}$. *A posteriori* we will show that the time T in (2.20) satisfies $0 < T < T^*$ and therefore the bootstrap assumption holds.

Lemma 9.3. *Assume that condition* (9.7) *is satisfied. Then the operator* $\mathcal{Z}^2(\xi,t)$ *satisfies*

$$\left\|\mathcal{Z}^2(\xi,t)\right\|_{\ell^1} \leq C\varrho^{-(2d-2)-1/2}2^{N(2d-2)}\|\xi\|_{\ell^1}.$$

Proof. The proof of this lemma follows the same lines as the proof of Lemma B.3 in [GK15]. We split $\mathcal{Z}^2$ in (9.5) as $\mathcal{Z}^2 = \mathcal{Z}^{21} + \mathcal{Z}^{22}$ with

$$\begin{aligned}\mathcal{Z}^{21}(\xi,t) =&\mathcal{E}\left(\mathtt{r}^\varrho+\xi\right) - \mathcal{E}\left(r^\varrho\right) - D\mathcal{E}\left(\mathtt{r}^\varrho\right)\xi\\ \mathcal{Z}^{22}(\xi,t) =&\widetilde{\mathcal{R}}\left(\mathtt{r}^\varrho+\xi\right) - \widetilde{\mathcal{R}}\left(\mathtt{r}^\varrho\right).\end{aligned}$$

By the definition of $\mathcal{E}$ in (2.11), we have that

$$\|\mathcal{Z}^{21}\|_{\ell^1} \leq C\sum_{j=2}^{2d-1}\|r^\varrho\|_{\ell^1}^{2d-1-j}\|\xi\|_{\ell^1}^j.$$

In the proof of Lemma 9.2, we have seen that $\|r^\varrho\|_{\ell^1} \leq \varrho^{-1}2^{N-1}$. Using this estimate and the bootstrap assumption (9.7) we obtain

$$\|\mathcal{Z}^{21}\|_{\ell^1} \lesssim \varrho^{-(2d-2)-1/2}2^{N(2d-2)}\|\xi\|_{\ell^1}.$$

Proceeding analogously one can see that $\|\mathcal{Z}^{22}\|_{\ell^1} \lesssim \varrho^{-2d}2^{2Nd}\|\xi\|_{\ell^1}$. Since we assume that $\varrho^{-2d}2^{2N} \ll 1$, these two estimates imply the statement of the lemma. □

We apply the estimates obtained in these three lemmas and the bootstrap assumption (9.7) to equation (9.6). We obtain

$$\frac{d}{dt}\|\xi\|_{\ell^1} \leq C\varrho^{-(2d+1)}2^{N(2d+1)} + C\varrho^{-(2d-2)}2^{N(2d-2)}\|\xi\|_{\ell^1}.$$

We apply Gronwall estimates. We take $\|\xi\|_{\ell^1} = \zeta e^{C\varrho^{-(2d-2)}2^{N(2d-2)}t}$ and therefore

$$\dot\zeta \leq \dot\zeta e^{C\varrho^{-(2d-2)}2^{N(2d-2)}t} \leq C\varrho^{-(2d+1)}2^{N(2d+1)}.$$

Integrating and taking into account the estimates for T in (2.20) and that $\|\zeta(0)\|_{\ell^1} = \|\xi(0)\|_{\ell^1} \leq C\varrho^{-5/2}$, we have that for $t \in [0,T]$,

$$\|\zeta(t)\|_{\ell^1} \leq \|\zeta(0)\|_{\ell^1} + C\varrho^{-(2d+1)}2^{N(2d+1)}T \leq C\varrho^{-5/2} + C\varrho^{-3}2^{N(d+3)}N^2 \leq \varrho^{-5/2}.$$

Then, using again the estimate for T in (2.20), for $t \in [0,T]$,

$$\|\xi(t)\|_{\ell^1} \leq \varrho^{-5/2}e^{C\varrho^{-(2d-2)}2^{N(2d-2)}T} \leq \varrho^{-5/2}e^{C2^{dN}N^2}.$$

Since we have assumed that $\varrho \geq \varrho_0 = e^{C2^{dN}N^2}$, we obtain that $\|\xi(t)\|_{\ell^1} \leq \varrho^{-3/2}$ for all $t \in [0,T]$. This completes the proof of Theorem 2.9.

References

[BB11] M. Berti and L. Biasco. Branching of Cantor manifolds of elliptic tori and applications to PDEs. *Comm. Math. Phys.*, 305(3):741–796, 2011.

[BB13] M. Berti and Ph. Bolle. Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential. *J. Eur. Math. Soc. (JEMS)*, 15(1):229–286, 2013.

[BCP14] M. Berti, L. Corsi, and M. Procesi. An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds. *Comm. Math. Phys.*, online first, 2014.

[BGT04] N. Burq, P. Gérard, and N. Tzvetkov. Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds. *Amer. J. Math.*, 126(3):569–605, 2004.

[BK92] I. U. Bronsteĭn and A. Ya. Kopanskiĭ. Finitely smooth normal forms of vector fields in the vicinity of a rest point. In *Global analysis—studies and applications, V*, volume 1520 of *Lecture Notes in Math.*, pages 157–172. Springer, Berlin, 1992.

[Bou93] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and application to nonlinear evolution equations. i. schrödinger equations. *Geom. Funct. Anal.*, 3(2):107–156, 1993.

[Bou96] J. Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices*, 6:277–304, 1996.

[Bou98] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math. (2)*, 148(2):363–439, 1998.

[Bou00] J. Bourgain. Problems in Hamiltonian PDE's. *Geom. Funct. Anal.*, Special Volume, Part I:32–56, 2000. GAFA 2000 (Tel Aviv, 1999).

[Bou04] J. Bourgain. Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations. *Ergodic Theory Dynam. Systems*, 24(5):1331–1357, 2004.

[Bou05] J. Bourgain. On invariant tori of full dimension for 1D periodic NLS. *J. Funct. Anal.*, 229(1):62–94, 2005.

[CDKS01] J. E. Colliander, J.-M. Delort, C. E. Kenig, and G. Staffilani. Bilinear estimates and applications to 2D NLS. *Trans. Amer. Math. Soc.*, 353(8):3307–3325 (electronic), 2001.

[CF12] R. Carles and E. Faou. Energy cascades for NLS on the torus. *Discrete Contin. Dyn. Syst.*, 32(6):2063–2077, 2012.

[CKO12] J. Colliander, S. Kwon, and T. Oh. A remark on normal forms and the "upside-down" I-method for periodic NLS: growth of higher Sobolev norms. *J. Anal. Math.*, 118(1):55–82, 2012.

[CKS$^+$10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.*, 181(1):39–113, 2010.

[CP95] L. Chierchia and P. Perfetti. Second order Hamiltonian equations on $\mathbf{T}^\infty$ and almost-periodic solutions. *J. Differential Equations*, 116(1):172–201, 1995.

[CW10] F. Catoire and W.-M. Wang. Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori. *Commun. Pure Appl. Anal.*, 9(2):483–491, 2010.

[EK10] L. H. Eliasson and S. B. Kuksin. KAM for the nonlinear Schrödinger equation. *Ann. of Math. (2)*, 172(1):371–435, 2010.

[Fen77] N. Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen74] N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[GG10] P. Gérard and S. Grellier. The cubic Szegő equation. *Ann. Sci. Éc. Norm. Supér. (4)*, 43(5):761–810, 2010.

[GG12] P. Gérard and S. Grellier. Effective integrable dynamics for a certain nonlinear wave equation. *Anal. PDE*, 5(5):1139–1155, 2012.

[GG15] P. Gérard and S. Grellier. An explicit formula for the cubic Szegő equation. *Trans. Amer. Math. Soc.*, 367(4):2979–2995, 2015.

[GK15] M. Guardia and V. Kaloshin. Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. *J. Eur. Math. Soc. (JEMS)*, 17(1):71–149, 2015.

[GPT13] B. Grébert, É. Paturel, and L. Thomann. Beating effects in cubic Schrödinger systems and growth of Sobolev norms. *Nonlinearity*, 26(5):1361–1376, 2013.

[GT12] B. Grébert and L. Thomann. Resonant dynamics for the quintic nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(3):455–477, 2012.

[Gua14] M. Guardia. Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential. *Comm. Math. Phys.*, 329(1):405–434, 2014.

[GXY11] J. Geng, X. Xu, and J. You. An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. *Adv. Math.*, 226(6):5361–5402, 2011.

[Han11] Z. Hani. Global and dynamical aspects of nonlinear Schrödinger equations on compact manifolds. *Ph. D. thesis UCLA*, 2011.

[Han14] Z. Hani. Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations. *Arch. Ration. Mech. Anal.*, 211(3):929–964, 2014.

[HP14] E. Haus and M. Procesi. Growth of Sobolev norms for the quintic NLS on $\mathbb{T}^2$. To appear in *Analysis and PDE*, Preprint available at `http://arxiv.org/abs/1405.1538`, 2014.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1977.

[HPTV13] Z. Hani, B. Pausader, N. Tzvetkov, and N. Visciglia. Modified scattering for the cubic Schrödinger equation on product spaces and applications. Preprint available at `http://arxiv.org/abs/1311.2275`, 2013.

[HT13] E. Haus and L. Thomann. Dynamics on resonant clusters for the quintic non linear Schrödinger equation. *Dyn. Partial Differ. Equ.*, 10(2):157–169, 2013.

[KP96] S.B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math. (2)*, 143(1):149–179, 1996.

[Kuk97] S.B. Kuksin. Oscillations in space-periodic nonlinear Schrödinger equations. *Geom. Funct. Anal.*, 7(2):338–363, 1997.

[Poc11] O. Pocovnicu. Explicit formula for the solution of the Szegö equation on the real line and applications. *Discrete Contin. Dyn. Syst.*, 31(3):607–649, 2011.

[Poc13] O. Pocovnicu. First and second order approximations for a nonlinear wave equation. *J. Dynam. Differential Equations*, 25(2):305–333, 2013.

[Pös96] J. Pöschel. A KAM-theorem for some nonlinear partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 23(1):119–148, 1996.

[Pös02] J. Pöschel. On the construction of almost periodic solutions for a nonlinear Schrödinger equation. *Ergodic Theory Dynam. Systems*, 22(5):1537–1549, 2002.

[PP12] M. Procesi and C. Procesi. A normal form for the Schrödinger equation with analytic nonlinearities. *Comm. Math. Phys.*, 312(2):501–557, 2012.

[PP14] M. Procesi and C. Procesi. A KAM algorithm for the resonant non–linear Schrödinger equation. *Advances in Math.*, to appear, 2014.

[PX13] M. Procesi and X. Xu. Quasi-Töplitz Functions in KAM Theorem. *SIAM J. of Math. Anal.*, 45(4):2148 – 2181, 2013.

[Šil67] L. P. Šil′nikov. On a problem of Poincaré-Birkhoff. *Mat. Sb. (N.S.)*, 74 (116):378–397, 1967.

[Soh11a] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on $\mathbb{R}$. *Indiana Univ. Math. J.*, 60(5):1487–1516, 2011.

[Soh11b] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on S^1. *Differential Integral Equations*, 24(7-8):653–718, 2011.

[Soh12] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to 2D Hartree equations. *Discrete Contin. Dyn. Syst.*, 32(10):3733–3771, 2012.

[Sta97] G. Staffilani. Quadratic forms for a 2-D semilinear Schrödinger equation. *Duke Math. J.*, 86(1):79–107, 1997.

[Wan10] W.-M. Wang. Supercritical Nonlinear Schrödinger equations I: Quasi-Periodic Solutions. arXiv:1007.0156, 2010.

[Way90] C.E. Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via kam theory. *Comm. Math. Phys.*, 127(3):479–528, 1990.

[Zho08] S. Zhong. The growth in time of higher Sobolev norms of solutions to Schrödinger equations on compact Riemannian manifolds. *J. Differential Equations*, 245(2):359–376, 2008.